

Forecasting dengue transmission in Thailand by time and space

Yalda Jafari

St. Peter's College

University of Oxford



A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2026

Acknowledgements

This thesis was made possible through the guidance and support of many individuals and institutions. I am especially thankful to my supervisors, Richard Maude, Oliver Brady, and Chawarat Rotejanaprasert, for their mentorship and intellectual guidance throughout this journey. Their expertise, patience, and encouragement have shaped both this thesis and my development as a scholar. I also extend sincere thanks to Moritz Kraemer for hosting me at Oxford and providing mentorship that greatly enriched my research perspective and experience.

I am grateful to my collaborators who contributed their expertise to different aspects of this work. Ahyoung Lim, Wynn Zaw, Tuyen Huynh Ngoc, and Katherine Tiley provided assistance with the systematic review. I thank Marc Choisy for his support with wavelet analysis, Bernardo Gutiérrez-Granja and Rhys Inward for providing guidance and thoughtful feedback on phylogenetic analyses, and Rory Gibb and Adam Howes for freely sharing their expertise on INLA. Joseph Tsui's technical support with the server infrastructure was essential to the computational aspects of this research. I am particularly grateful to Sumali Bajaj, Cathal Mills, Iago Perez Fernandez, Kulchada Pongsoipetch, Prathyush Sambaturu, and Qian Wang for their support which greatly strengthened the analytical foundations of this work. I thank Monnaphat Jongdeepaisal, Suwanna Mukem, and Massaya Sirimatayanant for providing support in understanding the Thai healthcare system and with translation when needed.

This research was made possible through the generous financial support of the Tropical Health Network, the Clarendon Scholarship, the Goodger and Schorstein Award, and the Nuffield Department of Medicine. I am privileged to have received this support, which allowed me to pursue this work.

I am deeply grateful to the Thai Ministry of Public Health for providing the dengue surveillance that formed the basis of this thesis. Without their commitment to open data sharing and public health transparency, this work would not have been possible. I also thank Facebook Data for Good for providing mobility data.

I thank Pawadee Boonyakanjanapon and students and staff at Mahidol-Oxford Tropical Medicine Research Unit for making me feel welcome in Thailand. I thank the Kraemer group for the insightful discussions.

Finally, I am forever grateful to my family, my parents Maryam and Ahmad, my sisters Sara and Mitra, and my partner Bas for their unconditional love, patience, and encouragement throughout this endeavour.

This thesis is dedicated to my grandmothers, whose strength and wisdom shaped the woman I am today. Their resilience and determination in the face of barriers that limited women of their generation inspire me to make the most of the opportunities they helped make possible.

Attribution

Chapter 1: I wrote this chapter independently.

Chapter 2: I led the study design, wrote the protocol, ran the searches, and completed screening, data extraction, analysis, and write-up. Ah-Young Lim, Tuyen Huynh Ngoc, and Katherine Tiley, Wynn Zaw contributed to screening and data extraction.

Chapters 3–7: I designed the study, wrote the ethics application, processed and analysed the data, built and ran the models, produced the figures, conducted literature searches, and wrote all sections.

All supervisors (Pof. Richard Maude, Prof. Oliver Brady, Dr. Chawarat Rotejanaprasert) provided guidance on analyses and edited the thesis draft.

.

Papers in preparation

This thesis will result in three publications. The first, '*Systematic Review of Temporal Forecasting of Dengue Transmission*,' will be based on Chapter 2. The second, '*Descriptive Analysis of Dengue Transmission in Thailand, 2013–2022*,' will present the epidemiological and phylogenetic analyses from Chapter 3. The third, '*Spatiotemporal Forecasting of Dengue Transmission in Thailand to Inform Public Health Response*,' will synthesise Chapters 4-6, focusing on spatiotemporal forecasting to support public health decision-making. Titles are provisional and may change prior to publication.

Papers published during DPhil

During my DPhil, I contributed as second author to a systematic review, '*A Systematic Review of the Data, Methods and Environmental Covariates Used to Map Aedes-Borne Arbovirus Transmission Risk*', where I carried out the literature search, screening, and data extraction. I also published, as first author, '*Could Prophylactic Antivirals Reduce Dengue Incidence in a High-Prevalence Endemic Area?*', a study I initiated prior to my DPhil and completed during my doctoral study.

Abstract

Dengue causes substantial morbidity and mortality in Thailand annually. Forecasting when and where cases will increase can help public health officials plan vector control and health sector responses. A systematic review of dengue forecasting literature revealed that most studies used fixed effects models that ignore temporal and spatial correlation, with evaluation focused primarily on point predictions rather than probabilistic forecasts.

Using Thai surveillance data (2013-2022), this thesis characterized dengue dynamics and developed forecasting models. Spatial and temporal patterns showed strong heterogeneity. Wavelet analysis revealed annual seasonality and multi-year interannual cycles, likely driven by different mechanisms. Serotype data indicated serotype 1 dominance, but major gaps in serotype and genomic sampling highlighted the need for more systematic surveillance.

Bayesian spatio-temporal models were developed to forecast dengue cases using an expanding window framework. Climate variables (temperature, Standardized Precipitation Evapotranspiration Index, wind speed), environmental variables (Enhanced Vegetation Index), at no lag up to 6-week lags, and genetic diversity and mobility data were tested individually to identify optimal lags, then incorporated into autoregressive and seasonal baseline models. Models were evaluated at 1–24-week horizons using the Continuous Ranked Probability Score.

Forecasting skill decreased with longer horizons. The autoregressive model with seasonality performed best overall. Covariates improved only the autoregressive

baseline, suggesting the seasonal component already captured covariate-associated information. All models performed poorly during outbreak periods compared to endemic periods.

This work demonstrates that seasonal models using surveillance data alone can inform planning during endemic periods. However, improving outbreak forecasting remains a critical challenge requiring further methodological development to better support public health decision-making during high-transmission periods.

Table of Contents

Acknowledgements	i
Attribution	iii
Papers in preparation:	iv
Papers published during DPhil:	iv
Abstract	v
List of Figures	xiv
List of Tables.....	xxii
1 Introduction	1
1.1 Global epidemiology	1
1.2 Clinical presentation	3
1.3 Clinical management.....	4
1.4 Diagnosis	5
1.5 Surveillance	9
1.6 Phylogenetics.....	10
1.7 Vector control	11
1.8 Drivers of Dengue Transmission	14
1.8.1 Climatic variables	14
1.8.1.1 Temperature	14
1.8.1.2 Precipitation.....	17
1.8.1.3 Humidity	19
1.8.1.4 Wind speed	21
1.8.2 Environmental variables	22

1.8.2.1	Vegetation index	22
1.8.3	Other variables	23
1.8.3.1	Socioeconomic variables	23
1.8.3.2	Mobility	24
1.9	Forecasting	26
1.9.1	Bayesian forecasting	29
1.10	Dengue in Thailand	31
1.11	Objectives	35
2	Systematic Review of Temporal Forecasting of Dengue Transmission.....	37
2.1	Introduction	37
2.2	Methods	40
2.2.1	Information sources	40
2.2.2	Search strategy	41
2.2.3	Selection process	41
2.2.4	Data collection process	42
2.2.5	Data items	42
2.2.6	Risk of bias assessment	42
2.3	Results.....	43
2.3.1	Prediction target	45
2.3.2	Prediction resolution	46
2.3.3	Prediction period	48
2.3.4	Covariates	49
2.3.5	Methods used	50
2.3.6	Random effects	54
2.3.7	Evaluation metrics:	55

2.3.8	Comparator	57
2.3.9	Risk of bias assessment	57
2.4	Discussion	57
3	Descriptive Chapter	62
3.1	Introduction	62
3.2	Methods	65
3.2.1	Surveillance data	65
3.2.2	Population data	66
3.2.3	Serotype data from surveillance	66
3.2.4	Wavelet analysis	68
3.2.5	Phylogenetic analysis.....	70
3.3	Results:.....	72
3.3.1	Surveillance data:	72
3.3.1.1	National trends.....	72
3.3.1.2	Provincial trends.....	73
3.3.1.3	District trends	75
3.3.2	Serotype results.....	77
3.3.3	Wavelet analysis	79
3.3.3.1	Reconstructed time-series.....	81
3.3.3.2	Phase difference	84
3.3.4	Phylogenetic analysis.....	85
3.3.4.1	Data availability.....	85
3.3.5	Phylogenetic trees	86
3.3.6	Genetic diversity	87
3.4	Discussion	92

4	Baseline Model Analysis	96
4.1	Introduction	96
4.2	Methods	98
4.2.1	Data	98
4.2.2	Models	99
4.2.2.1	Model structure	99
4.2.2.2	Priors	100
4.2.3	Forecast evaluation	102
4.2.3.1	Outbreak prediction evaluation	103
4.3	Results:.....	105
4.3.1	Forecasting evaluation by forecast horizon	105
4.3.1.1	CRPS	105
4.3.1.2	Sharpness	107
4.3.1.3	Calibration	108
4.3.1.4	Bias	110
4.3.2	Forecasting evaluation by target week.....	111
4.3.2.1	CRPS	111
4.3.2.2	Sharpness	113
4.3.2.3	Calibration	113
4.3.2.4	Bias	114
4.3.3	Forecasting evaluation by space	115
4.3.4	Outbreak prediction evaluation	117
4.3.4.1	Outbreak prediction evaluation by horizon	117
4.3.4.2	Outbreak prediction by target week	118
4.3.4.3	Outbreak prediction by space	120
4.4	Discussion	123

5	Univariate Analysis.....	126
5.1	Introduction	126
5.2	Methods.....	127
5.2.1	Covariates	127
5.2.1.1	Temperature	127
5.2.1.2	Standardized Precipitation Evapotranspiration Index (SPEI)	128
5.2.1.3	Wind Speed.....	128
5.2.1.4	Enhanced Vegetation Index (EVI)	129
5.2.1.5	Genetic Diversity	129
5.2.1.6	Mobility Data	129
5.2.2	Relationship with dengue cases.....	130
5.2.3	Forecasting dengue cases with covariates	133
5.3	Results.....	134
5.3.1	Relationship between covariates and dengue cases	134
5.3.1.1	Temperature	134
5.3.1.2	SPEI	141
5.3.1.3	Wind Speed.....	143
5.3.1.4	EVI	145
5.3.1.5	Genetic diversity.....	147
5.3.1.6	Mobility	147
5.3.2	Forecasting evaluation	149
5.3.2.1	Forecasting with temperature	149
5.3.2.2	Forecasting with SPEI	150
5.3.2.3	Forecasting with wind speed	151
5.3.2.4	Forecasting with EVI	152
5.3.2.5	Forecasting with genetic diversity	153

5.3.2.6	Forecasting with mobility	154
5.4	Discussion	156
5.4.1	Summary of evidence.....	156
5.4.1.1	Climate variables	156
5.4.1.2	Environmental variables	160
5.4.1.3	Mobility	162
5.4.2	Limitations	162
6	Multivariable Analysis	164
6.1	Introduction	164
6.2	Methods.....	168
6.2.1	Sub-analyses.....	170
6.3	Results.....	170
6.3.1	All years.....	170
6.3.1.1	Forecast horizon.....	170
6.3.1.2	Target week.....	172
6.3.2	Seasonal status	172
6.3.3	Outbreak status	175
6.3.4	Effect estimates.....	177
6.4	Discussion	182
7	Discussion	185
7.1	Summary of findings	185
7.2	Implications for our understanding of dengue epidemiology.....	189
7.3	Implications for dengue forecast model development.....	190
7.4	Public health utility of findings	191

7.5	Limitations	192
7.5.1	Data	192
7.5.2	Methodology.....	194
7.6	Future directions	195
7.7	Conclusions	197
8	Appendices.....	198
9	References	218

List of Figures

Figure 1.1-Timeline of dengue biomarker appearance in patients experiencing primary and secondary infection. NS1 and virus are detectable from onset in primary infection, with IgM appearing around day 3 and IgG toward the end of the acute period. Secondary infections show early IgG and shorter NS1/virus detection duration. Reproduced from Muller et al. (2017) with permission[6].	6
Figure 1.2-Map of Thailand by region, province, and district.	31
Figure 2.1-PRISMA flow diagram of study selection.	44
Figure 2.2-Number of studies published, by year and region.	45
Figure 2.3-Spatial scope and resolution of included studies.	47
Figure 2.4-Temporal resolution of included studies.	48
Figure 2.5-Frequency of methods used.	53
Figure 2.6-Proportion of methods used, by year.	53
Figure 3.1-Sentinel hospitals for serotyping, 2017-2019. Population for each province is the average of 2017 - 2019.	68
Figure 3.2- Incidence per 100,000 people	72
Figure 3.3-Annual cases with outbreak threshold in red.	73
Figure 3.4-Monthly cases with outbreak threshold in red.	73
Figure 3.5-Incidence per 100,000 people, by province.	74
Figure 3.6-Incidence per 100,000 people, by district.	76
Figure 3.7 a) Proportion of dengue cases by type, tested for serotype. b) Number of cases by serotype. c) Number of serotypes by type of dengue disease.	78
Figure 3.8 a) Log of proportion of dengue cases tested for serotype by province and year. b) Proportion of serotype cases per province by year.	79
Figure 3.9-Power wavelet analysis of weekly cases in Thailand	80
Figure 3.10-National average power of weekly cases in Thailand	81

Figure 3.11-Reconstructed time-series, national level, 2013-2022	82
Figure 3.12-Reconstructed time-series, by district, annual and multi-annual periods	83
Figure 3.13-Phase difference, by district, by annual and multi-annual periods ...	85
Figure 3.14-Number of serotypes sampled in Thailand, 2013 - 2020.....	86
Figure 3.15-Phylogenetic trees of dengue serotypes in Thailand, 1985 - 2020. Coloured branches represent samples collected in Thailand while grey branches represent samples from other South Asian countries.	87
Figure 3.16-a) Proportion of genotypes by serotype and b) overall.....	89
Figure 3.17-Number of lineages and samples over time, by serotype	90
Figure 3.18-Nucleotide diversity (π) and number of sequences (n) used to calculate diversity, by serotype and year. Lines correspond to nucleotide diversity (π) (left axis) and bars correspond to number of sequences (n) (right axis).....	92
Figure 4.1-Mean national level CRPS by forecast horizon and model. Lower values indicate better performance. Lower CRPS values indicate better forecast skill.	106
Figure 4.2-Mean district level CRPS by forecast horizon and model. Lower CRPS values indicate better forecast skill.	106
Figure 4.3-Mean national sharpness, as measured by NMAD, by forecast horizon and model. Lower NMAD values indicate narrower (sharper) prediction intervals.	107
Figure 4.4-Mean district level sharpness, as measured by NMAD, by forecast horizon and model. Lower NMAD values indicate narrower (sharper) prediction intervals.....	108
Figure 4.5-Mean national empirical coverage, 50% and 90%, by forecast horizon and model. Higher empirical coverage indicates more observed values fell within the prediction intervals.	109

Figure 4.6-Mean district level empirical coverage, 50% and 90%, by forecast horizon and model. Higher empirical coverage indicates more observed values fell within the prediction intervals.....	109
Figure 4.7-Mean national bias, by forecast horizon and model. Lower bias indicates that on average, forecasts are closer to the observed values.	110
Figure 4.8-Mean district level bias, by forecast horizon and model. Lower bias indicates that on average, forecasts are closer to the observed values.	111
Figure 4.9-Mean national CRPS by target week and forecast horizon. Lower CRPS values indicated better forecasting skill.	112
Figure 4.10-Mean national NMAD by target week and forecast horizon. Lower NMAD values indicate narrower (sharper) prediction intervals.....	113
Figure 4.11-Mean national empirical coverage by target week and forecast horizon. Higher empirical coverage indicates more observed values fell within the prediction intervals.....	114
Figure 4.12-National mean bias by target week, model and forecast horizon. Lower bias indicates that on average, forecasts are closer to the observed values. ...	115
Figure 4.13-District CRPS by model and forecast horizon, coloured by region. Lower CRPS indicates better forecasting skill.	116
Figure 4.14-District level difference in CRPS by model and forecast horizon. ..	116
Figure 4.15-AUC by model and forecast horizon.....	117
Figure 4.16-Brier score by forecast horizon (weeks) and model, higher values indicating better performance.	118
Figure 4.17-Agreement between observed and predicted outbreaks, by target week, model, and forecast horizon.	119
Figure 4.18-Brier Score by target week, model, and forecast horizon.....	120
Figure 4.19-Difference in agreement between no seasonality and seasonality model, by district and forecast horizon (weeks).....	121

Figure 4.20-Brier Score by district and model, coloured by region.....	122
Figure 4.21-Difference in brier score between models and significance of difference by district and forecast horizon.	123
Figure 5.1-Relative Risk of dengue, by mean temperature and lag	135
Figure 5.2-Relationship between relative risk of dengue and mean temperature, by lag	135
Figure 5.3-Errors of residuals for modelling dengue and mean temperature, by region, national and regional model.....	136
Figure 5.4-Relative Risk of dengue, by minimum temperature and lag	137
Figure 5.5-Relationship between relative risk of dengue and minimum temperature, by lag	138
Figure 5.6-Errors of residuals for modelling dengue and minimum temperature, by region.	138
Figure 5.7-Relative Risk of dengue, by maximum temperature and lag	139
Figure 5.8-Relationship between relative risk of dengue and maximum temperature, by lag	140
Figure 5.9-Errors of residuals for modelling dengue and maximum temperature, by region.	140
Figure 5.10-Relative Risk of dengue, by SPEI and lag.....	141
Figure 5.11-Relationship between relative risk of dengue and SPEI, by lag	142
Figure 5.12-Errors of residuals for modelling dengue and SPEI, by region.....	142
Figure 5.13-Relative Risk of dengue, by wind speed and lag.....	143
Figure 5.14-Relationship between relative risk of dengue and wind speed, by lag	144
Figure 5.15-Errors of residuals for modelling dengue and wind speed, by region	144
Figure 5.16-Relative risk of dengue by EVI and lag	145

Figure 5.17-Relationship between relative risk of dengue and EVI, by lag	146
Figure 5.18-Errors of residuals for modelling dengue and EVI, by region	146
Figure 5.19-Correlation between genetic diversity and dengue cases, by serotype	147
Figure 5.20-Incoming visits to districts.....	148
Figure 5.21-Relationship between national dengue incidence and correlation between district-level incidence and incoming population flow, 2013-2022. Blue is the trend line and grey represents 95% CI.	149
Figure 5.22-Mean CRPS of forecasts by mean, maximum, and minimum temperature, by baseline model, and lag. Lower CRPS indicates better forecasting skill.	150
Figure 5.23-Mean CRPS of forecasts by SPEI, baseline model, and lag. Lower CRPS values indicate better forecasting skill.	151
Figure 5.24-Mean CRPS of forecasts by wind, baseline model, and lag. Lower CRPS indicates better forecasting skill.	152
Figure 5.25-Mean CRPS of forecasts by EVI, baseline model, and lag. Lower CRPS indicates better forecasting skill.	153
Figure 5.26-Mean CRPS of forecasting with diversity by forecast horizon. Lower CRPS values indicate better forecasting skill.	154
Figure 5.27-National CRPS results of mobility and baseline models, by forecast horizon. Lower CRPS values indicate better forecasting skill.....	155
Figure 5.28-District level CRPS difference between base and mobility models, by forecast horizon.	155
Figure 6.1-CRPS by forecast horizon and model. The left panel shows the models with a no seasonality base and the right panel shows the models with a seasonality base.	171
Figure 6.2-CRPS by target week, forecast horizon, and model.....	172

Figure 6.3-CRPS by forecast horizon, season, and model.....	174
Figure 6.4-Difference in CRPS between baseline model and model with all covariates, by rainy season, and forecast horizon (1, 4, and 24 weeks).	174
Figure 6.5-CRPS by forecast horizon, outbreak status, and model.....	176
Figure 6.6-CRPS difference between baseline model and model with formula 1, by forecast horizon, outbreak status, and baseline model. Dark grey indicates districts that did not experience outbreaks.	176
Figure 6.7-Effect estimates by formula, for a no seasonality base model. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, pi_Dengue_1 is genetic diversity for serotype 1, pi_Dengue_2 is genetic diversity for serotype 2, pi_Dengue_3 is genetic diversity for serotype 3, pi_Dengue_4 is genetic diversity for serotype 4, EVI_4:pi_Dengue_1 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 1, EVI_4:pi_Dengue_2 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 2, EVI_4:pi_Dengue_3 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 3, EVI_4:pi_Dengue_4 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 4, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, and TMin_6:Wind_Speed_6 is the interaction term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.	178
Figure 6.8- Effect estimates by formula, for a seasonality base model. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, pi_Dengue_1 is genetic diversity for serotype 1,	

pi_Dengue_2 is genetic diversity for serotype 2, pi_Dengue_3 is genetic diversity for serotype 3, pi_Dengue_4 is genetic diversity for serotype 4, EVI_4:pi_Denugue_1 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 1, EVI_4:pi_Dengue_2 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 2, EVI_4:pi_Dengue_3 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 3, EVI_4:pi_Dengue_4 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 4, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, and TMin_6:Wind_Speed_6 is the interaction term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.179

Figure 6.9-Effect estimates by formula, without diversity related estimates, for no seasonality base models. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, and TMin_6:Wind_Speed_6 is the interaction term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.180

Figure 6.10- Effect estimates with confidence intervals by formula, without diversity related estimates, for seasonality base models. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, an TMin_6:Wind_Speed_6 is the interaction

term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.181

List of Tables

Table 1.1- Criteria for DF, DHF, and DSS[1].	4
Table 2.1-Prediction target type.	46
Table 2.2-Prediction period of included studies (n=162).	49
Table 2.3-Covariates investigated in included studies (n=162).	50
Table 2.4-Methods used in included studies. (Abbreviations: ADL, Autoregressive Distributed Lag; ANN, Artificial Neural Network; AR, Autoregressive; ARIMA, Autoregressive Integrated Moving Average; ARIMA-ANN, hybrid ARIMA-Artificial Neural Network; ARIMA-LSTM, hybrid ARIMA-Long Short-Term Memory; ARIMA-NNAR, hybrid ARIMA-Neural Network Autoregression; ARIMAX, Autoregressive Integrated Moving Average with Exogenous Variables; BATS, Box-Cox transform, ARMA errors, Trend and Seasonal components; BME, Bayesian Maximum Entropy; DLNM, Distributed Lag Non-linear Model; EPM, Epidemic Prediction Markets; EPRA, Evolutionary Polynomial Regression Algorithm; EpiEstim, Epidemic Estimation; ETS, Error, Trend, Seasonal; GAM, Generalized Additive Model; GARP, Genetic Algorithm for Rule-set Prediction; GLARMA, Generalized Linear Autoregressive Moving Average; GLM, Generalized Linear Model; GLMM, Generalized Linear Mixed Model; GP, Gaussian Process; KDE, Kernel Density Estimation; LASSO, Least Absolute Shrinkage and Selection Operator; LSTM, Long Short-Term Memory; MCP, Minimax Concave Penalty; NNAR, Neural Network Autoregression; PDLM, Polynomial Distributed Lag Model; SARIMA, Seasonal Autoregressive Integrated Moving Average; SARIMAX, Seasonal Autoregressive Integrated Moving Average with Exogenous Variables; SCAD, Smoothly Clipped Absolute Deviation; SSTAR, Spatial-Spatio-Temporal Autoregressive; STAR, Spatio-Temporal Autoregressive; STCFM, Spatio-Temporal Convolutional Forecasting Model; StructTS, Structural Time Series;	

TBATS, Trigonometric seasonality, Box-Cox transform, ARMA errors, Trend and Seasonal components; VAR, Vector Autoregression.	52
Table 2.5-Evaluation methods used in included studies. (Abbreviations: AUC, Area Under the Curve; CRPS, Continuous Ranked Probability Score; DSS, Dawid-Sebastiani Score; GMRAE, Geometric Mean Relative Absolute Error; MAD, Mean Absolute Deviation; MAE, Mean Absolute Error; MAPE, Mean Absolute Percentage Error; MASE, Mean Absolute Scaled Error; MCC, Matthews Correlation Coefficient; NPV, Negative Predictive Value; PPV, Positive Predictive Value; RMAE, Relative Mean Absolute Error; RMAPE, Root Mean Absolute Percentage Error; RMSE, Root Mean Square Error; RMSPE, Root Mean Squared Percentage Error; RMSWE, Root Mean Squared Weighted Error; SRMSE, Standardized Root Mean Square Error; UMBRAE, Unscaled Mean Bounded Relative Absolute Error; VAF, Variance Accounted For)	56
Table 3.1-Kendall's tau between number of lineages and number of samples, by serotype.	91
Table 3.2-Kendall's tau between genetic diversity and number of samples.	92
Table 5.1-Summary of findings and variables to be included in multivariable models.	161

List of abbreviations

ADL – Autoregressive Distributed Lag
Ae – Aedes
ANN – Artificial Neural Network
AR – Autoregressive
ARIMA – Autoregressive Integrated Moving Average
ARIMA-ANN – Hybrid ARIMA–Artificial Neural Network
ARIMA-LSTM – Hybrid ARIMA–Long Short-Term Memory
ARIMA-NNAR – Hybrid ARIMA–Neural Network Autoregression
ARIMAX – Autoregressive Integrated Moving Average with Exogenous Variables
AUC – Area Under the Curve
BATS – Box-Cox transform, ARMA errors, Trend and Seasonal components
BME – Bayesian Maximum Entropy
BoE – Bureau of Epidemiology
CDC – Centers for Disease Control and Prevention
CI – Confidence Interval
CRPS – Continuous Ranked Probability Score
DALYs – Disability-Adjusted Life Years
DENV – Dengue virus
DF – Dengue Fever
DHF – Dengue Haemorrhagic Fever
DLNM – Distributed Lag Non-linear Model
DOPA – Department of Population Administration
DSS – Dengue Shock Syndrome
DSS – Dawid–Sebastiani Score (*context-dependent; both appear*)
EPM – Epidemic Prediction Markets
EPRA – Evolutionary Polynomial Regression Algorithm
EpiEstim – Epidemic Estimation
ETS – Error, Trend, Seasonal
EVI – Enhanced Vegetation Index
EWS – Early Warning System
GAM – Generalized Additive Model
GARP – Genetic Algorithm for Rule-set Prediction
GBD – Global Burden of Disease
GLARMA – Generalized Linear Autoregressive Moving Average
GLM – Generalized Linear Model
GLMM – Generalized Linear Mixed Model
GMRAE – Geometric Mean Relative Absolute Error
GP – Gaussian Process
GEE – Generalized Estimating Equations
IgG – Immunoglobulin G
IgM – Immunoglobulin M
INLA – Integrated Nested Laplace Approximation
KDE – Kernel Density Estimation

LASSO – Least Absolute Shrinkage and Selection Operator
LSTM – Long Short-Term Memory
MAD – Mean Absolute Deviation
MAE – Mean Absolute Error
MAPE – Mean Absolute Percentage Error
MASE – Mean Absolute Scaled Error
MCC – Matthews Correlation Coefficient
MCP – Minimax Concave Penalty
MEDLINE – Medical Literature Analysis and Retrieval System Online
MoPH – Ministry of Public Health
NASA – National Aeronautics and Space Administration
NDVI – Normalized Difference Vegetation Index
NMAD – Normalized Median Absolute Deviation
NNAR – Neural Network Autoregression
NPV – Negative Predictive Value
NS1 – Non-structural protein 1
PDLM – Polynomial Distributed Lag Model
PPV – Positive Predictive Value
PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses
R₀ (R0) – Basic Reproduction Number
RMAE – Relative Mean Absolute Error
RMAPE – Root Mean Absolute Percentage Error
RMSE – Root Mean Square Error
RMSPE – Root Mean Squared Percentage Error
RMSWE – Root Mean Squared Weighted Error
RR – Relative Risk
RT-PCR – Reverse Transcriptase Polymerase Chain Reaction
SARIMA – Seasonal Autoregressive Integrated Moving Average
SARIMAX – Seasonal Autoregressive Integrated Moving Average with Exogenous Variables
SCAD – Smoothly Clipped Absolute Deviation
SPEI – Standardized Precipitation Evapotranspiration Index
SRMSE – Standardized Root Mean Square Error
SSTAR – Spatial–Spatio-Temporal Autoregressive
STAR – Spatio-Temporal Autoregressive
STCFM – Spatio-Temporal Convolutional Forecasting Model
StructTS – Structural Time Series
TBATS – Trigonometric seasonality, Box-Cox transform, ARMA errors, Trend and Seasonal components
TCN – Temporal Convolutional Network
TimesFM – Time-Series Foundation Model
UMBRAE – Unscaled Mean Bounded Relative Absolute Error
VAR – Vector Autoregression
VAF – Variance Accounted For
WHO – World Health Organization

XGBoost – Extreme Gradient Boosting

1 Introduction

1.1 Global epidemiology

Dengue is a vector-borne disease transmitted by *Aedes aegypti* and *Aedes albopictus* mosquitoes[10] and is found in tropical and sub-tropical regions worldwide[11], with South Asia, Southeast Asia, and tropical Latin America most affected[12]. According to the World Health Organization (WHO), dengue is currently endemic in more than 100 countries[13].

It has been estimated that there are approximately 390 million cases per year, 96 million of which are symptomatic[14]. Based on the Institute for Health Metrics and Evaluation's Global Burden of Disease (GBD) study in 2021 which collects data from a variety of sources, including vital registrations, verbal autopsies, police records, scientific literature, household survey data, epidemiological surveillance data, disease registry data, and clinical informatics data, there were an estimated 58.96 million reported cases in 2021[15]. This substantial difference reflects the well-documented underreporting of dengue in health surveillance systems globally, with studies estimating that only 8-13% of symptomatic dengue cases are captured by passive surveillance systems[15-17]. In 2024, the mortality rate was estimated at 4.81 deaths per million, and a case-fatality rate of 0.07%[18].

Globally, there has been a change over time in the age distribution of Disability-Adjusted Life Years (DALYs) due to dengue. In 1990, the highest burden was amongst those under 5 years of age, while in 2019, the largest proportion of burden was amongst those aged 15 – 49 years[19]. Considering specifically those under the age of 20, there was an estimated annual percentage increase of incidence of

1.39% (95% confidence interval [CI], 1.13%–1.65%) from 1990 to 2021, and an estimated annual percentage increase of DALYs of 0.67% (95% CI, 0.40%–0.95%)[20]. The overall age-standardized incidence rate of dengue increased by 0.56% per year between 1990 and 2021[15]. Age-standardized DALYS increased by 0.28% (95% UI: -0.38–0.92) per year, from 21.63 per 100,000 (95% UI: 15.09–26.92) in 1990 to 27.76 (95% UI: 14.21– 41.65) in 2021[15], though this increase was not statistically significant.

Dengue also has negative economic impacts. There are costs due to dengue infection at various levels, from direct costs such as cost of healthcare visits, hospitalizations, and treatments, and indirect costs such as loss of productivity and caregiving responsibilities[21-23]. A systematic review found that the cost of dengue illness varies largely across countries, from as low as 15 USD in an outpatient setting of Burkina Faso to as high as 9,386 USD for intensive care unit patients in Mexico[24]. Shepard et al.[25] analysed the results from the GBD Study 2013 and found that the total annual global cost of dengue illness was 8.9 billion USD (95% UI 3.7 billion–19.7 billion). At the country level, in Thailand and Brazil, Marczell et al.[22] found that in 2019, dengue resulted in loss of 22.4 million workdays and reduced the national GDP by 0.05% or by 876 million USD. Dengue can also impact the tourism industry. Marczell et al.[22] found that in Thailand, dengue resulted in loss of tourism equivalent to about 1.81 billion USD or 0.33% of the national GDP in 2019.

Dengue outbreaks can have disproportionate impacts on the burden of disease compared to underlying endemic cases. A large proportion of cases in endemic areas are a result of seasonal outbreaks as described for example by Mills et al.[26] in Peru and Oneda et al.[27] in Brazil. Outbreaks can overwhelm the health

system with a sudden increase in patients resulting in shortage of resources [4, 23]. This demand on resources can also negatively impact those without dengue[4]. In addition, dengue outbreaks can exacerbate socioeconomic inequalities, disproportionately affecting those with limited access to healthcare and appropriate sanitation services [4, 21].

1.2 Clinical presentation

Symptomatic dengue is often mild, where symptoms include fever, headache, pain behind the eyes, and muscle and joint pain[28]. Some people develop severe dengue, whose symptoms include bleeding gums or nose, persistent vomiting, severe abdominal pain and/or shock amongst others. Individuals with severe dengue need immediate care to reduce the risk of a fatal outcome[28]. Symptomatic dengue is widely classed into three syndromes, according to the 1997 WHO guidelines[1]: dengue fever (DF), dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS). DHF and DSS are forms of severe dengue. The criteria for these are shown in Table 1.1. Although this classification has been criticized as being difficult to apply in clinical situations and an increase in severe cases which did not fulfil its strict criteria, leading to a revised WHO classification in 2009[11], it is still widely used for case reporting and surveillance, including in Thailand.

Dengue fever	Dengue Haemorrhagic Fever	Dengue Shock Syndrome
Acute febrile illness with two or more of the following: Headache Retro-orbital pain Myalgia Leukopenia Arthralgia Rash Hemorrhagic manifestations Supportive serology or occurrence at the same location and time as other confirmed cases of dengue fever	All of the following must be present: Fever or history of acute fever, lasting 2–7 days, occasionally biphasic. Hemorrhagic manifestations: Positive tourniquet test; Petechiae, ecchymosis, purpura or bleeding from mucosa, gastrointestinal tract, injection sites or other locations; or hematemesis/melena. Thrombocytopenia ($<100,000$ platelets for mm^3) Evidence of plasma leakage due to increased vascular permeability	DHF with hypotension for age or narrow pulse pressure (>20 mmHg), plus one of the following: Rapid and weak pulse Cold, clammy skin, restlessness

1.3 Clinical management

There are four serotypes of dengue virus (DENV), DENV-1 to DENV-4[29]. Infection with one serotype results in lifelong immunity to that serotype, and immunity for up to three months to other serotypes[29]. However, after this short time period, infection with another serotype results in severe dengue, known as antibody-dependent enhancement (ADE)[28]. There are currently no medications available to treat dengue, with vector control remaining the main method of reducing transmission[28]. Reducing mortality of infected individuals is achieved by high quality supportive care, including appropriate fluid management. There have been challenges in developing a vaccine that can protect against all serotypes and prevent ADE[30, 31]. Dengvaxia was the first dengue vaccine developed. However, its use is only limited to individuals with previous dengue infection[31] and the manufacturer, Sanofi, has recently decided to discontinue its manufacture in the third quarter of 2026[32]. Qdenga is a newer vaccine

manufactured by Takeda, for individuals aged 4 years and up regardless of prior infection status that has shown favourable benefit and risk profiles in a series of trials in the Americas and Southeast Asia, including Thailand[33]. Qdenga received marketing authorization from the European Medicines Agency in May 2022[34] and has been available for use in Thailand since 2023[35]. There are other new candidates in development, which provide protection against all four serotypes[30, 31], as well new antivirals[31].

1.4 Diagnosis

Initial steps of diagnosis include clinical suspicion based on signs and symptoms[6, 36-38]. Early signs as outlined in Table 1.1 include fever, nausea, and headache. As these signs are very non-specific, understanding the local context is crucial. During an outbreak, cases with early signs may be presumed to be dengue in the absence of laboratory testing and managed accordingly[36].

The appropriate biomarker to test for depends on days since symptom onset and patient's history of previous dengue infection. This is illustrated in Figure 1.1, from Muller et al[6].

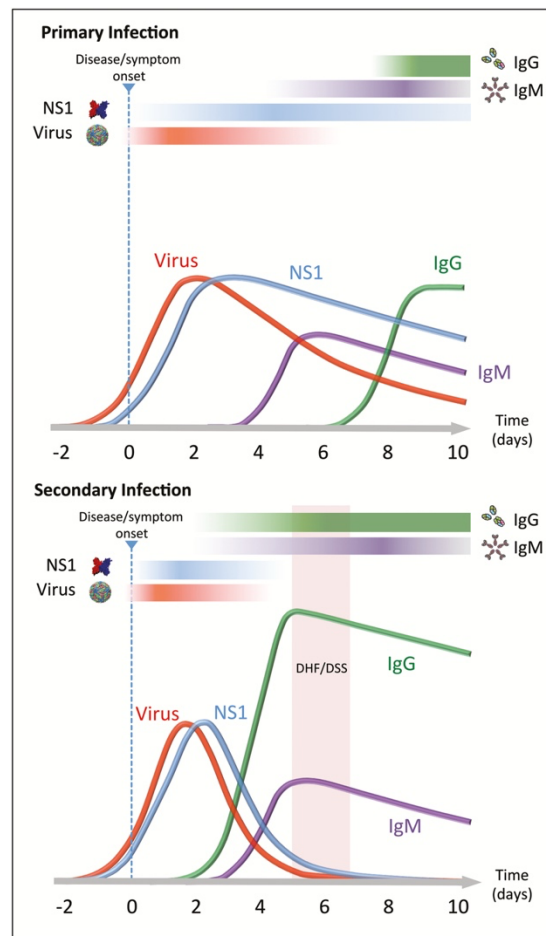


Figure 1.1-Timeline of dengue biomarker appearance in patients experiencing primary and secondary infection. NS1 and virus are detectable from onset in primary infection, with IgM appearing around day 3 and IgG toward the end of the acute period. Secondary infections show early IgG and shorter NS1/virus detection duration. Reproduced from Muller et al. (2017) with permission[6].

Traditionally, virus isolation has been the main method of identifying dengue infection. Typically, blood, serum, or plasma taken during the first 5 days after symptom onset is cultured in cells lines of mosquitos or mammals or in live mosquitoes. In patients with secondary infections, this method is more difficult to use due to the production of immune complexes of cross-reactive antibodies and the virus[6, 37] or more rapid rate of viral clearance[37]. Viremia can also be detected 2 or 3 days before symptoms onset[37]. Despite its certainty in identifying

dengue virus, this method has been replaced by reverse transcript PCR (RT-PCR) or NS1 antigen-capture enzyme-linked immunosorbent assays (ELISAs) as virus isolation requires high technical expertise and can take days or weeks to perform[6, 29, 37].

Detection of viral dengue RNA in blood, serum, or plasma using RT-PCR is a highly sensitive and specific confirmatory test. It must be noted that RT-PCR is highly sensitive when used within 8 days of symptoms onset in primary infections[6, 37] and within 4 days in secondary infections[6], coinciding with the high viraemic period. This method also allows for the quantification of viral titers which are useful in studying pathogenesis[29]. Despite their fast turn-around and accuracy, they require specialized laboratories and staff, and therefore their use is limited in low-resource settings[6, 29, 37].

Another type of test is those that detect dengue NS1 antigen. NS1 is a highly conserved glycoprotein that is secreted from infected cells. NS1 can be detected in blood and plasma and has been shown to be correlated with viremia and disease severity. It can be found in blood circulation for up to 9 days or longer after illness onset in primary infections[6, 37] and up to 4 days in secondary infections[6]. Utility of NS1 detection in secondary infections is reduced compared to primary infections due to rapid formation of immune complexes of cross-reactive antibodies and NS1[6]. NS1 is detected using antigen-capture ELISA in a laboratory setting or using lateral-flow tests that can be used at the point of care. ELISA tests can provide quantitative results that can shed light on disease progression while lateral flow tests can only provide positive or negative results. Even though ELISA tests are more accurate, they take at least a day for results to be available while lateral-flow test results are available within minutes [6, 29]. Lateral-flow tests have

drastically improved case detection particularly in low-resource settings and have thus contributed to more timely treatment and control efforts[6].

In response to the dengue virus, the body produces anti-dengue antibodies, IgG and IgM. The detection of these antibodies is another method to identify dengue infection. IgG and IgM antibody-capture ELISA and point-of-care lateral flow tests can be used to identify these antibodies. ELISA tests have higher sensitivity, but point-of-care tests are more practical. In both primary and secondary infections, IgM can be detected from 3 to 5 days after symptom onset, with levels increasing up to approximately 2 weeks and persisting up to 3 months[39]. Early anti-dengue IgM can be cross-reactive to other flaviviruses such as Zika, Japanese encephalitis or yellow fever so understanding the epidemiology of other flaviviruses in the area is important. In cases of primary infection, IgG can be detected from day 6 to 10 days after illness onset, while in secondary infections there is a rapid increase of IgG 2 to 4 days after symptoms onset. IgG can persist lifelong. During the acute phase of the disease, the ratio between IgG and IgM can be used to identify whether an infection is primary or secondary [6, 29, 37].

Given the complex biokinetics of markers of dengue infection over the course of infection and by type of infection, a combination of tests, such as NS1 antigen test and antibody tests is best. Previous research [40-46] has shown that combining antigen and antibody tests have improved acute diagnosis of dengue, particularly by increasing sensitivity.

1.5 Surveillance

National dengue surveillance systems are the primary source of information for monitoring dengue fever cases and outbreaks. However, they are imperfect and are affected by many biases. Most endemic countries rely on passive surveillance of dengue cases. When an individual is feeling sick, they visit a health facility such as a hospital or a health centre. When dengue is suspected or confirmed, health care providers then report the case to regional or national health departments[29]. The criteria for a suspected or confirmed case vary between countries. Many reported cases are suspected and are not confirmed through a confirmatory diagnostic test. Confirmation generally requires a positive dengue PCR and/or NS1 and/or rise in antibody titers. Passive surveillance systems often capture more serious cases that require treatment and miss milder or asymptomatic cases, as many of those cases do not come in contact with the health system. Those that do, may present with non-specific symptoms and may not be reported as dengue. As referred to in Table 1.1, clinical criteria for severe forms of dengue are more specific and less likely to be misreported as another illness. Additionally, many countries only report cases from the government health sector, missing cases which present to private sector healthcare providers. As a result, burden estimates based on reported cases only are a gross underestimate of total infections[47]. A systematic review of asymptomatic dengue infection rates in 50 studies found this to be around 54%[48], indicating that symptomatic cases account for only about half of true infections.

Additionally, there are reporting delays that result in right truncation, where most recent observations in the surveillance time series are incomplete because many case or outcomes of cases have not yet entered the dataset, resulting in an under-

estimation of events in real time. Reporting delays can arise at multiple points in the reporting process, such as from the time symptoms develop to healthcare visit, from clinical suspicion to laboratory diagnosis, or from diagnosis to administrative processing. For example, in Brazil, 9 weeks are required for 90% of cases to be reported while less than 50% of cases are reported in the first week[49]. In Thailand, one study found that 75% of cases were reported within 6 weeks while 50% of cases were reported within 5 weeks[50]. These delays are country-dependent and can also change over time and during outbreak and non-outbreak periods of both dengue and other infectious diseases. Outbreaks can impact the workload of clinicians, laboratories, and public health staff. Although many countries have mandatory reporting by specific timelines in place, these policies are often not enough to eliminate delays[29, 50, 51]. As a result, nowcasting methods are increasingly used to adjust for reporting delays and improve accuracy of real-time dengue surveillance [51, 52].

1.6 Phylogenetics

Phylogenetics is used to study the evolution and genetic diversity of DENV genes and genomes[53, 54]. One way to do this is to build a phylogenetic tree from sampled virus genomes. Branches of the tree represent evolutionary lineages, with closely related sequences indicating recent divergence from a common ancestor[55, 56]. These analyses help shine light on evolutionary characteristics and how transmission within and across regions may be connected[55].

Dengue virus is an enveloped virus with a positive-sense RNA genome of approximately 11 kbp in length. There are three structural and seven non-structural coding genes, all translated as a single polyprotein and surrounded by two

untranslated or non-coding regions (UTR) at the 3' and 5' ends[53, 54]. The evolutionary rate of dengue virus is estimated to be on average 7.6×10^4 substitutions per site per year[53]. There is 30-35% amino acid diversity between serotypes and about 3% within each serotype[53]. Nucleotide divergence of genotypes within serotypes does not exceed 6%[53, 57]. Certain genotypes have demonstrated their epidemic potential and fitness advantages. For example, the DENV-2 Asian-American genotype has been spreading in Americas since the 1980s, displacing the local American genotype[58]. Recently, the Cosmopolitan genotype has been found to be spreading in the Americas[59-64]. These replacements are often accompanied by more severe disease[65] and outbreaks [53, 58, 59, 65]. Understanding these evolutionary patterns and tracking circulating genotypes can inform surveillance strategies and forecasting models.

1.7 Vector control

Vector control activities target either mosquito vectors to reduce their numbers or human and vector contact to reduce infectious bites.

In urban centres, *Aedes (Ae.) aegypti* is responsible for most transmission of dengue. *Ae. aegypti* proliferates in standing water because females lay their eggs in still water, where larvae can hatch and grow[66]. Around the home, items such as household plants or household containers used for water storage provide opportunities for oviposition or egg-laying. Other possible sites include blocked gutters and storm drains, tires, buildings under construction, and discarded containers. Control of *Ae. aegypti* is mainly achieved through removing these potential mosquito breeding sites in standing water. WHO recommends regularly emptying and cleaning standing water containers[29]. If water needs to be stored

in containers, they should either be covered with tight lids or nets to prevent mosquitoes from entering. Larvae can also be targeted using insecticides. Another approach includes biological agents including certain species of predatory fish, copepods[29], and tadpoles[67] which eat the larvae.

To reduce the population of adult mosquitoes, indoor residual spraying of insecticide is used. In this method, insecticides are applied to indoor surfaces where adult *Ae. aegypti* may rest such as walls or furniture. Mosquitoes that rest on these surfaces for extended periods will then be killed[68]. Insecticide-treated materials such as screens for windows and doors, in addition to acting as barriers, can also kill mosquitoes upon contact[69].

There are also methods, which rather than killing mosquitos, aim to reduce their ability to transmit dengue. This is achieved through releasing mosquitoes infected with *Wolbachia*. *Wolbachia* is a bacteria found naturally to infect about 50% of insects. Infecting mosquitoes with this bacterium does not harm the mosquito but affects the ability of female mosquitoes to process dengue virus when they bite an infected individual. As a result, the female mosquito will not be infectious. Certain *Wolbachia* strains, such as *wMelPop*, can reduce the adult lifespan of *Ae. aegypti* mosquitoes. Mosquitoes infected with *Wolbachia* pass on the bacteria to their offspring, propagating the bacteria through the population, thus also propagating the ability of female mosquitoes to stop development of dengue virus inside the mosquito[70, 71]. Most notably, a cluster-randomized trial of *Wolbachia* infected *Ae. aegypti* in Yogyakarta, Indonesia, found that the protective efficacy of the intervention was 77.1% (95% CI, 65.3 to 84.9) when comparing intervention and non-intervention clusters[72]. In addition to direct effects, Chow et al. [73] also found indirect effects of spillover, which contributed to 37.7% (95% CI, 36.4 – 38.5)

case reduction in adjacent non-intervention areas. Currently, the World Mosquito Program, a non-profit that works to protect the global community from mosquito-borne diseases, have released *Wolbachia* mosquitoes in 14 countries to protect over 13.3 million people[74].

During outbreaks, vector control measures are intensified and there is a more targeted effort focused on killing adult mosquitoes rather than larvae. This is most often done through fogging with the most commonly used insecticide formulations containing pyrethroids[75, 76]. There are two types of fogging techniques. Basic thermal spray or thermal foggers are used to spray inside and outside premises that contain potential breeding sites for *Aedes* mosquitos. The mosquitoes die within 24 hours of encountering spray, reducing virus replication and transmission. Another kind of fogging, Ultra Low Volume (ULV) sprays or cold foggers convert the liquid to a fine mist, dispersing it into the air at high velocity. ULV fogging can cover a wider range of areas and can have more longer lasting effects than thermal foggers[29, 75]. Fogging every 2 to 3 days for 10 days helps rapid reduction of vector density. Further applications once or twice a week then help keep the population of adult mosquitoes low [29].

With continued and increased use of insecticides, resistance is also a possibility. Therefore, routine surveillance and testing of mosquito larvae and adults is necessary to ensure that insecticides remain useful[29]. Many countries have already reported various forms of resistance in *Aedes* populations[77-81], such as Bangladesh[82], Brazil[83, 84], Indonesia[77, 85], Jamaica[86], Malaysia[77, 87, 88], Mauritania[89], Singapore[77], Senegal[90], Sri Lanka[91], and Thailand[77, 78].

1.8 Drivers of Dengue Transmission

Dengue transmission has been associated with several factors, including climatic variables, environmental variables, and mobility. A mechanistic framework that links these covariates to the basic reproduction number (R_0) is helpful to understand how these covariates influence dengue transmission. The Ross-MacDonald model, originally developed for malaria, is widely also applied to dengue and other mosquito-borne infections[92]. For dengue, R_0 can be expressed as

$$R_0 = \frac{ma^2bce^{-\mu\tau}}{\gamma\mu} \quad (1)$$

where m is the ratio of mosquitoes to humans, a is the mosquito biting rate, b is the probability of transmission from mosquito to human per bite, c is the probability of transmission of infection from an infectious human to mosquito per bite, μ is the mosquito mortality rate, τ is the extrinsic incubation period (EIP) (time from infection to infectiousness within the mosquito), and γ is the human recovery rate[92-94]. Covariates affect transmission by modulating one or more of these parameters.

1.8.1 Climatic variables

1.8.1.1 Temperature

Temperature is an important driver of dengue transmission impacting R_0 through its influence on mosquito density (m), biting rate(a), mosquito mortality (μ), EIP (τ), and transmission probabilities (b, c). A meta-analysis of 49 studies found that temperature alone is sufficient to explain development rate variability in *Ae. aegypti*

[95]. They found that the relationship between temperature and development rate of both hatch to pupation and hatch to emergence was positive and linear within the median range of 14-31°C. However, this effect was found to be heterogeneous across studies with significant variation in the temperature at which development stops and the cumulative effect of temperature on development rate. Faster development rates at higher temperature increases the progression from immature to adult stages, increasing adult mosquito density.

Brady et al. [96] conducted a meta-analysis modelling adult *Aedes aegypti* and *Aedes albopictus* survival at different temperatures in both laboratory and field. The models revealed the optimal survival of *Ae. aegypti* was at 21°C with a median longevity of 38 days, while *Ae. albopictus* had a wider range of optimal temperatures, 20-30°C, with a median longevity over 60 days. Temperature-dependent mortality directly affects R_0 , with extreme temperatures increasing mortality and reducing transmission potential.

Temperature also affects mosquito behaviour, which is critical for transmission. Female *Ae. aegypti* were able to fly between 10 °C and 35 °C with sustainable flight between 15 °C and 32 °C, and with flight performance better below 27°C [97]. Feeding rate increased between 26-35 °C [97, 98]. In *Ae. albopictus*, biting occurred between 16-37 °C, peaking at 27°C [99]. These temperature dependent changes in feeding directly modulate biting rate(a), impacting transmission potential.

Temperature also impacts EIP, the period between when a mosquito acquires the dengue virus and to when the mosquito can transmit the virus to humans. Some time is required for the virus to replicate, exit the gut, and reach the salivary glands

[100]. A meta-analysis by Chan and Johansson[101] reviewing EIP values estimated that at 25°C, 95% of EIPs ranged between 5 and 33 days, with a mean of 15 days, while at 30°C, the range was narrower, between 2 and 15 days with a mean of 6.5 days, demonstrating that higher temperatures shorten the time required for a mosquito to become infectious. In the R_0 formula, a shorter EIP (τ) is associated with a higher transmission rate.

Epidemiologically, temperature has widely been shown to be associated with dengue transmission through a non-linear, inverted U-shaped relationship, with peak dengue risk observed between approximately 27-31°C across multiple geographic settings and declining risk at both lower and higher temperatures[102-106]. However, the effect estimates of this relationship vary by geographic context. A meta-analysis found that mean temperatures between 23.2–27.7 °C showed the strongest association with dengue incidence (OR = 1.35 per 1 °C increase, 95% CI: 1.18–1.52), with risk increasing steeply from 22 °C to 29 °C and declining below and above this range [107]. A more recent meta-analysis found a 13% increase in risk of dengue infection (RR = 1.13, 95% CI: 1.11-1.16) for each 1°C increase in high temperatures, with greater effects observed in tropical monsoon climate zones (RR = 1.29, 95% CI: 1.11–1.51) and humid subtropical climate zones (RR = 1.20, 95% CI: 1.15–1.25)[108]. Additionally, minimum temperature ($r = 0.28$, 95% CI = 0.05, 0.48) may be more strongly correlated with dengue transmission than mean temperature ($r = 0.07$, 95% CI = -0.1, 0.24), with observed lags ranging from 1 to 4 months, depending on geographic location [109].

1.8.1.2 Precipitation

Precipitation influences dengue transmission primarily through its effect on mosquito density (m), by creating or eliminating larval breeding habitats, though this relationship is context dependent.

Ae. aegypti eggs can survive dry periods and hatch en-masse after a rainfall [110]. Alto et al.[111] conducted an experiment at 22, 26, and 30°C, and three precipitation regimes: low fluctuation, high fluctuation, and drying. They found that greater temperature and absence of drying resulted in higher adult populations. Montgomery et al. [112] found that across six cities in Central Africa, both *Ae. aegypti* and *Ae. albopictus* abundance increased with rainfall. Other studies have found positive correlations with precipitation and *Aedes* populations[113-116].

In a study in Northern Italy, accumulated rainfall 1-4 weeks prior was negatively correlated with *Ae. albopictus* abundance, possibly due to rainfall flushing larvae from breeding sites[117]. Other studies have also found a negative effect of heavy rains on *Aedes* populations [118-120].

A modelling study, investigating the impact of different rainfall patterns on *Ae. aegypti* populations found that concentrating the same total rainfall into fewer, larger events rather than spreading it evenly resulted in reduced mosquito abundance[121]. Another modelling study[122] in Taiwan also found complex effects of rainfall. In northern cities with rainy winters, *Ae. aegypti* was absent because winter rainfall triggered egg hatching when temperatures were too cold for larvae to survive, while in southern cities with dry winters, the species thrived. They also found that moderate dry season length increased survival through

persistence of eggs, while very long droughts caused extinction as eggs became non-viable.

Another factor that can impact the relationship between precipitation and mosquito populations is that during periods of low or no precipitation, human-made water sources have been found to support mosquito populations[123]. Production of *Ae. aegypti* has been found to even trigger dengue outbreaks during dry seasons[124]. These water sources include potted plants as well as water storage containers, especially in areas with limited access to piped water[125].

Multiple reviews have established the positive correlation between precipitation and dengue incidence. A meta-analysis by Li et al.[126] found that precipitation had a statistically significant risk of association with dengue fever (RR = 1.01, 95% CI: 1.01, 1.01). Another meta-analysis found that precipitation between 83-15 mm had a correlation of 0.38 (95% CI = 0.31, 0.45), at 1 to 5 months lag [109]. In contrast, a study investigating the effect of climatic factors on two Brazilian hotspots found that total precipitation was negatively associated with dengue cases[127].

Given the described complex relationship between precipitation and mosquito population and dengue transmission, standardized indices have been developed that capture both drought and flood conditions across multiple timescales. Standardized Precipitation Index (SPI) standardizes precipitation data relative to historical precipitation data, with negative values indicating drought and positive values indicating wet conditions[128]. SPI can be calculated from a range of periods, allowing for identification of both short to long term values corresponding to meteorological, agricultural and hydrological droughts to understand lagged

effects on dengue incidence[129]. Since SPI only includes precipitation, another index, Standardized Precipitation Evapotranspiration Index (SPEI), was developed which incorporates both precipitation and potential evapotranspiration, accounting for the effect of temperature variability[130]. SPEI has been shown to be correlated with dengue cases[131, 132]. Other drought indices, such as the Palmer Drought Severity Index (PDSI), have also been developed and found to be correlated with dengue cases[133, 134]. However, SPEI is generally thought to be a better index because it can be calculated at multiple timescales, like SPI, while maintaining PDSI's sensitivity to temperature effects[130, 135].

1.8.1.3 Humidity

Humidity influences dengue transmission by affecting mosquito survival (μ), biting behaviour (a), and viral replication affecting transmission probability (b).

In a pooled survival analysis, Schmidt et al.[136] found that mortality of female *Ae. aegypti* increased non-linearly with decreasing humidity, because of desiccation stress. For *Ae. albopictus*, associations were observed in stratified models but were not statistically significant. Data were limited at low humidities and temperature extremes for both species. A later study[137] found a similar result with higher humidity associated with longer survival of adult mosquitoes, therefore decreasing μ and increasing the period during which mosquitoes can transmit dengue.

A study examining the effects of temperature and humidity on reproductive activity of *Ae. aegypti* found that at a high temperature of 35°C, relative humidity (RH) of 80%, 28% of mosquitoes exhibited inhibited oviposition compared to 45% at 60%

RH[138]. *Ae. aegypti* survival rate was found to decrease at humidity levels decreased. Blood feeding rates showed a peak at 50% RH but were the same at 30% and 80% RH[139]. Similarly, Manzano-Alvarez et al.[140] investigated the impact of RH shock and found that decreased RH was associated with higher mortality and blood-feeding rates, potentially reflecting compensatory feeding behaviour under stress conditions. This finding was supported by Canyon et al.[141], using *Ae. aegypti* sourced from arid conditions. They found that in low humidity conditions, host-biting rates were six times higher than expected. These results suggest that while low humidity may increase mortality (μ), it may simultaneously increase biting rate (a) among surviving mosquitoes. The net effect on transmission potential depends on the balance between these two parameters.

Humidity also affects dengue virus replication within the mosquito vector. Thu et al.[142] found that virus propagation within the mosquito was higher at higher RH. Enhanced viral replication could increase transmission probability (b) or decrease EIP (τ), both of which would increase R_0

Epidemiologically, humidity has also been shown to be associated with dengue cases. In their meta-analysis, Abdullah et al.[109] found that relative humidity between 60.5% and 88.7% was positively correlated with dengue incidence ($r = 0.30$, 95% CI = 0.23, 0.37). A spatiotemporal study in Indonesia found that RH was one of the most influential covariates associated with dengue incidence rates[143]. Several studies have also investigated the lag effects between humidity and dengue cases. In India, a 19-year retrospective analysis in East Delhi[144] found RH with lag of 2 months had the highest correlation with dengue cases, while another study in Pune[145] found the same lag most associated with dengue mortality. A study in Thailand found a lag of 7 months was associated with dengue

cases[146] while a study in Sri Lanka found positive correlation at lag of 1 month but a negative correlation at lag of 9 months[147].

1.8.1.4 Wind speed

Wind speed influences dengue transmission by impacting mosquito density (m) through dispersal and modified oviposition behaviour.

Costa-da-Silva et al.[148] found that wind speed was positively associated with egg abundance, egg density index, and ovitrap index for both *Ae. aegypti* and *Ae. albopictus* in Miami and Charleston in the United States of America. Similarly, a study in Kenya[149] found a positive correlation between wind speed and number of eggs laid. In a Peru, a study found high wind speeds was associated with higher capture rates of adult mosquitos in homes[150]. The mechanism for this phenomenon, however, was not well understood. Wind can either facilitate adult mosquito dispersal to new areas or impair flight and force gravid females to oviposit in nearby traps rather than dispersing eggs across multiple sites[148-150]. In fact, a study investigating the habitat suitability of *Ae. albopictus* in Europe, found that wind speed was negatively correlated with suitability[151].

Epidemiological studies have also found mixed associations between wind speed and dengue cases. Studies in Bangladesh[152], Malaysia[153, 154], and Paraguay[155] found that wind speed was negatively associated with incidence while a study in Chinda[156] found that wind speed was not a significant driver in the highest-burden province. On the other hand, a study in Dhaka, Bangladesh found wind speed to be positively correlated with dengue cases[157]. Another study, also in Bangladesh, found a positive linear relationship with wind speed and

dengue cases until a threshold of 7m/s, beyond which risk increased non-linearly[158].

1.8.2 Environmental variables

1.8.2.1 Vegetation index

Vegetation index influences dengue transmission mainly through its effect on mosquito density (m) and spatial patterns of mosquito-human contact.

Normalized Difference Vegetation Index (NDVI) measures the amount of vegetation over an area, using satellites. Values range from -1 to 1, with low values meaning barren areas such as exposed soil, while higher values mean greener vegetation, including forests, croplands, and wetlands[159].

A study in central Mexican highlands found NDVI to be one of the determinants of a suitable habitat for *Ae. aegypti*[160]. In Cordoba city, Argentina, variability in vegetation was found to be associated with *Ae. aegypti* oviposition temporal patterns[161], therefore affecting local mosquito density (m). A study in an urban Brazilian neighbourhood found that higher vegetation cover was negatively associated with *Ae. aegypti* possibly because these mosquitoes prefer urban structures[127], .

Epidemiological studies have also shown an inverse relationship between NDVI and dengue incidence, where lower vegetation cover is associated with higher number of cases. In Bhutan[162] and Lao PDR[163], spatio-temporal analyses found that NDVI was negatively associated with higher transmission. Higher NDVI values correspond to areas with lower population density, where there is a lower

chance of contact between mosquitoes and humans[163]. A similar trend was found in Brazil, though this effect was modified by socio-economic vulnerability, with negative association in most vulnerable areas, while a positive association was observed in least vulnerable areas. In the Pearl River Delta area in China, risk of DF declined nonlinearly with increasing NDVI until approximately 0.65, after which the response increased[164].

1.8.3 Other variables

1.8.3.1 Socioeconomic variables

Even though it might be expected that lower socioeconomic status would be correlated with *Ae. aegypti* and increased burden of dengue infections, a recent review[165] has found highly variable and location-specific associations. Another review also found inconsistent relationship between socioeconomic status and distribution and abundance of *Ae. aegypti*[166]. Socioeconomic variables considered were income, education, housing quality, employment status, social class, and community services. These inconsistent findings reflect the fact that socioeconomic variables impact mosquito density (m) through multiple pathways that can operate in opposing directions. Lower socioeconomic status may be associated with poorer housing quality and inadequate water structure but may also be associated with different building materials or other factors that reduce mosquito survival or habitat suitability

Population density has a more consistent relationship with mosquito abundance and dengue incidence. A study in Tucson, Arizona found a significant positive association between number of residents in home and odds of finding *Ae. aegypti* positive containers[125]. In Sao Paulo, Brazil, the authors found that female *Ae.*

aegypti density was proportional to number of residents[167], leading to an increased R_0 . In Vietnam, the relationship between dengue cases and population density was modulated by access to piped water. Areas without access showed peak risk at moderate densities (around 190 people/100m), while areas with access showed highest risk at very low densities with decreasing risk at higher densities[168]. In China, it was found that areas with 30,000-40,000 people per km squared were at a lower risk compared to areas with lower or higher densities[164].

1.8.3.2 Mobility

Human mobility is one of the drivers of dengue, since it extends the spatial scale of virus dispersal beyond the limited flight range of mosquitoes. Ross-Macdonald model assumes a homogenous mixing of mosquitos and humans[169]. Introducing human movement allows for the incorporation of spatial heterogeneity and connectivity that links infected and susceptible populations across space.

Mosquitoes have a limited flight range of around 100 meters[170]. As a result, human movement is the main method of dispersing the virus. At the global scale, Wilder-Smith et al.[171] reviewed the role of international travel on the spread of dengue and provided evidence that travel introduced dengue to areas previously dengue free. A case-control study in Iquitos, Peru, found that individual's house-to-house movements played a key role in defining infection risk and local patterns of incidence [172].

Using high-resolution phone data to inform mobility patterns, Massaro et al. found that mobility was a major factor in the spread of dengue in Singapore[173]. Mobile phone data was also found to improve dengue forecasting in Pakistan and

Thailand[174, 175]. Despite the advantages of including mobile phone data to capture human mobility, these models may have limited generalizability. Mobile phone availability and representatives can vary across settings, with coverage often sparse outside of urban areas and dependent on individual service providers. Mobile phone services may not be representative of the entire population as mobile phone ownership and usage may vary across demographics or socioeconomic groups. Additionally, high-resolution mobile phone data can increase model complexity, although this can be addressed through aggregation.

In Sri Lanka[176], using a travel survey, authors found that human mobility was a stronger indicator of dengue outbreaks than land-use or climate variables. As demonstrated, even though travel surveys are an informative source of mobility data, their implementation requires additional resources, and this is not generally feasible for use in routine surveillance.

Previous studies found that including travel networks better explained dengue expansion and relevance of risk factors in Vietnam[177], Bangkok, Thailand[178], and improved forecasting in Fortaleza, Brazil[179].

In a simulation meta-population model based on the region of Caldas, Colombia, authors evaluated how altering patch-to-patch movement rates influenced spatial spread of dengue[180]. They found that restricting mobility, specially from highly connected, high incidence municipalities, reduced overall incidence.

It is also important to consider the impact of illness on mobility. Schaber et al.[181] used a simulation agent-based model to better understand the impact of mobility on disease dynamics. They modified the mobility matrix based on previously

studied changes to movement during the symptomatic period. Those with symptomatic dengue were more likely to spend more time at home than away from home, consequently leading to increased contact rates at home and decreased contact rates away from home. This approach highlighted an important mechanistic consideration that infection status and outbreak intensity can change movement patterns. However, due to lack of empirical social networks, this model was not accurately parametrized.

Finally, climate and environment can impact mobility. For example, increased rainfall can lead to flooding which temporarily reduces mobility or can cause population displacement[182, 183].

1.9 Forecasting

Accurate forecasting can provide lead time for public health officials to prepare for dengue outbreaks, enabling targeted interventions and more efficient resource allocation before outbreaks or epidemic peaks occur. Early prediction of dengue incidence support health system preparedness by informing resource mobilisation, public awareness campaigns, and policy decisions that facilitate prioritization of interventions, such as vector control[184]. Accurate forecasts also help anticipate increased healthcare demands and promote preparedness[185], allowing health services to allocate human and material resources more effectively, alert clinicians to increase their diagnostic suspicion of dengue, and strengthen community engagement in prevention and control activities[11]. This proactive approach contrasts with traditional reactive responses, which are often too slow to prevent outbreaks unless triggered sufficiently early. Additionally, forecasts can guide strategic timing of vector control interventions, including interventions that provide

longer-term control, such as community-driven source reduction[186, 187]. In resource-constrained settings, where there are often competing interests, forecasting tools play an import role in facilitating decision making, by helping policy makers optimize resource use, implement timely control measures, and reduce impact of dengue outbreaks[188]. As such, forecast models serve as decision-support tools for public health authorities, contributing to early warning systems, scenario evaluations, and evidence-based resource allocation, strengthening immediate outbreak response and long-term strategic planning capacity [189].

If local-level epidemiology and the complex relationships between vector populations, climate variables, socioeconomic factors, and human movement can be adequately captured, models can achieve useful predictive performance. [184, 190]. There are also logistical issues accessing and combining the data needed to build such models. Raw data and covariates should be available in the desired resolution and in high quality at the appropriate time frame, though this is often not the case [50, 190]. Given these challenges, there is currently no real-time model used by the Ministry of Public Health (MOPH) in Thailand to identify increased transmission or predict outbreaks. Through personal communication with MOPH, they have expressed interest in adopting such models for forecasting to improve timeliness of prevention and intervention activities.

Dengue forecasting models can be broadly categorized into four approaches, namely statistical, mechanistic or mathematical, machine learning, and ensemble methods. Statistical methods such as Autoregressive Integrated Moving Average (ARIMA), seasonal ARIMA (SARIMA), and seasonal decomposition techniques like seasonal trend decomposition using Loess (STL) have been widely used for

infectious disease forecasting and are particularly effective at capturing relationships and seasonal patterns in time-series data[188, 191-196]. Other statistical models used for dengue prediction include linear regression[197], Poisson regression[198, 199], and generalized additive models (GAM)[200]. Statistical methods that include random effects allow for inclusion of temporal and spatial correlations. Integrated nested laplace approximation (INLA), one such a method, was used to predict epidemic bands for dengue cases in Brazil[201]. Mechanistic models, based on compartments, such as Susceptible-Exposed-Infectious-Recovered (SEIR), capture dengue transmission dynamics through mathematical representation of host-vector interactions. Even though these models can be used for forecasting, they are most often employed to assess intervention outcomes. In another category of methods, machine learning techniques such as random forest [202], neural networks[203], gradient boosting[204], support vector algorithms[205], and Long Short-Term Memory (LSTM)[189, 206] have been used to predict dengue. Finally, ensemble methods which combine multiple approaches, such as ARIMA and LSTM, show great improvements over individual models[188]. Integrating machine learning approaches with compartmental models can capture both complex non-linear patterns and mechanistic disease dynamics, which are difficult to model with a single approach[207]. A systematic review of dengue outbreak prediction models found that modelling techniques used were broadly categorized into statistical models, comprising approximately 60% of models, and machine learning approaches, comprising approximately 40%[184].

Evaluating forecasts using out of sample data is also an essential part of developing and deploying forecasting models[208]. Such an approach provides an understanding of how a model will perform when it is faced with data it is not trained

on, which is essentially the use case of a forecasting model. There are a variety of approaches to evaluating forecasting models. These are broadly categorized into point-estimate based approaches, such as Mean Absolute Error and Root Mean Square Error, contingency table-based approaches such as sensitivity and specificity, and probabilistic based approaches, such as Weighted Interval Scores and Continuous Ranked Probability Score (CRPS). Point-estimate based approaches compare the forecasted values to the observed values, and do not include the probabilistic nature of forecasts[209, 210]. Contingency-based measures are similar to point-estimate based approaches in that they also do not take advantage of probabilistic forecasts and are mostly used to evaluate the model's ability to correctly identify outbreak and no-outbreak periods[211, 212]. Probabilistic approaches are superior in that they include the uncertainty of predictions, which better reflects reality[213].

1.9.1 Bayesian forecasting

Bayesian modelling is built upon the Bayes theorem, which states that the posterior distribution is dependent on the likelihood and prior distribution. Joint posterior distributions are then used to make inferences and forecasts[214, 215].

Bayesian modelling offers several advantages for infectious diseases forecasting [216, 217]. Hierarchical structures are easy to implement, allowing for variation at different levels such as between and within provinces[218]. The Bayesian approach is advantageous in that it can account for spatial and temporal random effects. Spatial random effects account for spatial heterogeneity through unstructured effects and spatial autocorrelation through structured effects, by allowing areas to borrow strengths from neighbouring areas to produce smoothed

risk estimates[219, 220]. Temporal random effects capture temporal heterogeneity through unstructured effects while serial autocorrelation is captured through structured effects[220].

When the joint posterior distribution cannot be calculated analytically, as is often the case, samples are drawn from it. This is commonly done using Markov chain Monte Carlo (MCMC) methods, such as the Metropolis-Hastings algorithm and Gibbs sampling. Metropolis–Hastings generates samples from the joint posterior distribution by proposing new parameter values that are accepted or rejected probabilistically[215, 217]. Gibbs sampling, a specific case of Metropolis-Hastings, updates one parameter at a time.

Although flexible, MCMC methods can be computationally intensive, time-consuming, and technically complex to implement, particularly for high-dimensional models. To address these challenges, an alternative computational approach for Bayesian inference, known as the Integrated Nested Laplace Approximation (INLA), was proposed[221]. Rather than estimating the full joint posterior distribution, INLA focuses on approximating marginal posterior distributions of model parameters using Laplace approximations. This approach substantially reduces computational burden while maintaining high accuracy[222].

INLA is designed for models that can be expressed as latent Gaussian Markov random fields (GMRF)[215, 222]. Although this represents a restriction on the class of models that can be fitted, it is still applicable to a wide range of commonly used models in spatial and spatiotemporal epidemiology[222]. To facilitate implementation, the R-INLA package has been further developed, which enables efficient and accessible model fitting[223].

Given advantages of Bayesian analysis through the implementation of INLA, many studies have employed this approach to forecast a range of infectious diseases such as COVID-19[224, 225], dengue[226-229], Lassa fever[230], and West Nile[231]. Based on this evidence, I used INLA to conduct Bayesian forecasting of dengue transmission in Thailand in Chapters 4 – 6.

1.10 Dengue in Thailand

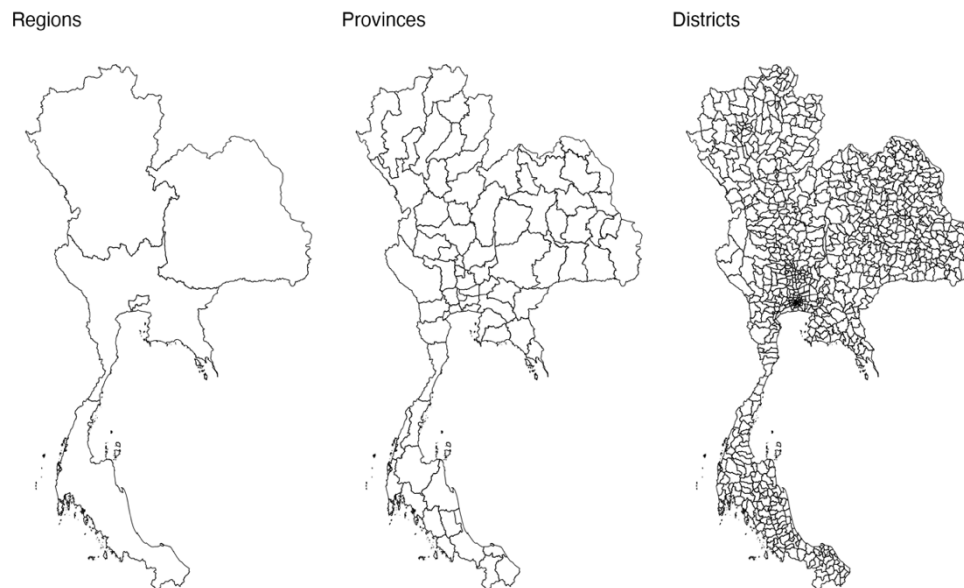


Figure 1.2-Map of Thailand by region, province, and district.

Thailand is a tropical country, with 5 regions, 77 provinces, and 928 districts. These areas are presented in Figure 1.2. Dengue is an endemic disease in Thailand where all four serotypes circulate [232-234]. There is spatial variation in dengue risk across urban and rural settings in Thailand with the central and southern regions having been identified as the highest risk areas [235]. Spatial clustering of dengue cases has been observed in urban areas near Khon Kaen city, while also

in rural areas of Khon Kaen province[234]. Thailand also hosts many refugees, including those housed in nine camps on the western side of Thailand, with a total population 43,872 as of 2022[236], representing potentially vulnerable populations for dengue transmission with a different risk profile than the general Thai population.

Dengue in Thailand exhibits strong seasonal patterns with cases increasing during the rainy season, from May to October [233, 234, 237]. Peak transmission usually occurs in July and August[238]. Multi-year epidemic cycles are also evident in Thailand, with intense peaks occurring approximately every 3 years[239]. In southern provinces of Songkhla, Pattani, Yala, and Narathiwat, distinct peaks were observed in 2010, 2013, and 2016[240]. Nationally, between 2011 and 2018, the highest incidence of all types of dengue combined was observed in 2013 at 241 cases per 100,000 population, followed by 2015 [233]. Historical surveillance data from 2003-2015 shows an overall increasing trend in dengue burden across the country, with interannual variability driven by complex interactions between viral serotype circulation, population immunity, and climatic factors[233, 235].

The national dengue surveillance in Thailand is conducted through the Bureau of Epidemiology (BoE), Ministry of Public Health (MoPH), which collects case reports from all 77 provinces through the R506 reporting system[241]. The surveillance framework captures clinically suspected and laboratory confirmed cases from both public and private healthcare facilities, with data reported through the national Weekly Epidemiological Surveillance Report and also available online in the annual epidemiological surveillance reports[242]. As many other surveillance systems, this system also faces challenges, such as reporting delays,

underreporting of cases, and variable diagnostic practices cross provinces[50, 243]. Despite these limitations, Thailand has a robust surveillance system that has provided data on the burden of the dengue in the country since 1974[240, 242]. District-level dengue surveillance data became publicly available through MoPH's online surveillance portal in mid-2025, whereas previously only province-level data was accessible, enabling finer spatial resolution analysis for public health planning and research[244].

There is extensive literature on risk factors associated with dengue in Thailand. Research in northeastern Thailand on determinant of vector abundance found that abundance of adult female *Ae. aegypti* was significantly associated with ecological and social determinants such as low education level, crowded housing, high numbers of wet containers, and poor premise conditions[245]. An earlier study in Chachoengsao province, located in east of Thailand, found that households in residential areas mixed with commercial activity and densely populated residential areas, had higher risk of dengue than other urban settings[246]. The relationship between urbanization and dengue risk is not straightforward, however. A study on spatio-temporal patterns of dengue incidence in northeastern Thailand found clustering of cases in urbanized areas of Khon Kaen city, as well as in rural areas of the province. This suggests that other environmental risks may be at play[234].

Climate factors have also been shown to be associated with dengue risk in Thailand. A recent study in Bangkok, using cases from 2015 – 2024, found a non-linear and lagged effects of temperature, humidity, rainfall and wind speed with dengue risk, though specific effects were heterogenous across districts of Bangkok[247]. A study found that an increase of 1°C in weekly mean temperatures, up to a 29°C threshold, increased dengue risk by 18.9% in two

provinces in Thailand along the Mekong river[248]. A national study conducted at provincial resolution using case data from 2001– 2014, found precipitation to be the most influential weather variable[249].

Human mobility patterns also play a significant role in Thailand, as was demonstrated for Bangkok[178]. By incorporating public transport networks into epidemiological models, the authors were able to show that accounting for daily human movement patterns improved statistical association between socio-economic variables and dengue incidence. Kiang et al.[175] found that using mobile phone data to model mobility improved forecasting at the provincial level.

Multiple approaches have been used for forecasting dengue in Thailand to support public health activities, with varying performance across spatial scales and forecast horizons. Traditional time series approaches, such as ARIMA and SARIMA models have been used for dengue forecasting in Thailand. These statistical methods take advantage of historical patterns to generate forecasts and can include other covariates [241, 249-256]. Other statistical approaches include Poisson regression [50], which outperformed seasonal baseline in majority of the provinces at up to 2 months horizon, while quantifying the substantial impact of reporting delays by employing a real-time approach. A generalized additive model [257] was able to outperform baseline models in forecasting outbreaks by province. Recent advances in integrating additional data streams by Kiang et al.[175] demonstrated that incorporating mobility using mobile phone data improved forecasts when compared to SARIMA, while Jain et al.[258] used a machine-learning approach with meteorological and socioeconomic data to forecast district-level outbreaks up to one month ahead. Despite this impressive body of work on forecasting in Thailand, there has been no work yet on forecasting at a resolution

of district and week and probabilistic evaluation, representing a critical gap in the literature. This resolution allows for better focalisation of resources aimed at prevention and intervention, such as more targeted vector control measures and public messaging. To achieve results at such a resolution requires surveillance data and covariates at a high resolution. Such data is not often available and if it is, its processing can be resource intensive. Additionally, probabilistic forecasts are essential for capturing uncertainty, enabling public health professionals to implement actions proportional to a likelihood of an event.

1.11 Objectives

The objective of this thesis is to determine whether spatiotemporal models can forecast dengue infections in Thailand, by district and week, for the years 2013 – 2022. This will be done through a sequence of steps with the following aims:

1. Review and summarize methods used to forecast dengue transmission, provide an overview of covariates used, and methods used to evaluate these models to inform methods and covariate selection for subsequent chapters
2. Describe patterns in dengue case data in Thailand from 2013 – 2022
3. Develop baseline forecasting models using hierarchical Bayesian statistical models implemented using INLA, and evaluate models using probabilistic-based techniques

4. Evaluate the relationship of climatic (temperature (minimum, maximum, mean), SPEI, wind speed), environmental (Enhanced Vegetation Index (EVI)), and mobility covariates with dengue incidence, and evaluate the forecasting ability of individual covariates
5. Build a multivariable model based on the findings of the previous chapter and evaluate its' forecasting performance based on factors such as forecast horizon, seasonality, and outbreak status

2 Systematic Review of Temporal Forecasting of Dengue Transmission

2.1 Introduction

Dengue is a vector-borne disease transmitted by *Aedes aegypti* and *Aedes albopictus* mosquitoes and is found in tropical and sub-tropical areas[10]. It has been estimated that there are approximately 390 million cases per year, 96 million of which are symptomatic[14]. Dengue infections have been found to increase in conditions of higher temperature, precipitation, and relative humidity[109]. With increasing climate change, there has been an increase in magnitude of case numbers and spread of the disease to areas where it was previously not seen[259, 260]. Even though it might be expected that lower socioeconomic status be correlated with increased burden of dengue infections, a recent review[165] has found highly variable and location-specific associations. Several studies have shown human mobility to be linked to dengue outbreaks[171, 174, 175, 181].

Forecasting dengue transmission is important in helping national and international actors better prepare for responding to increased incidence and outbreaks. By having insight into the magnitude of future cases and the timing of incidence peaks at the appropriate spatial resolution, public health professionals can allocate resources more effectively, implement targeted vector control interventions, and strengthen surveillance systems in anticipation of increased disease burden.

Forecasting dengue transmission is a work in progress. The complex relationship between vector populations, climate variables, socioeconomic factors, and human movement, amongst other covariates, have made it challenging to build models with useful prediction skills [184, 190]. There are also logistical issues, such as

access to suitable data. Raw data and covariates should be available in the desired resolution and in high quality at the appropriate time frame, though this is often not the case [50, 190].

Previous systematic reviews have examined temporal forecasting models. However, they have each reviewed a smaller subset of the literature to examine a specific aspect such as Bayesian methods only or models for outbreak prediction.

Sylvestre et al.[261] considered models using big data and found that in studies predicting incidence, regression-based models showed the highest performance while in studies predicting outbreaks, time series and decision trees showed best performances. They also recommended better utilisation of novel data streams, such as from internet search engines and social media.

De Lima et al.'s [262] review of prediction models for arboviruses transmitted by *Ae. aegypti* found that moving average models were most common. They found that even though deep learning models had good results, the high demand of training and resources required to apply them in real world settings makes them less useful than moving average models, which are easier and quicker to apply. Regarding models that forecasted outbreaks and risk, they found that there was not a comprehensive comparative approach in evaluating different predictor configurations. The review also found that the included studies lacked clarity and uniformity in evaluation metrics, making them difficult to compare.

Aswi et al.[216] focused exclusively on Bayesian methods and found that the most common model was a generalized linear mixed model with spatial random effects.

They noted Bayesian models' advantages of better estimating uncertainty compared to frequentists methods and considering spatial structures. They found that including a spatially structured component using a CAR prior often improved fit but noted that choice of neighbour structure could impact significance of covariates. Therefore, they recommend other smoothing priors should be investigated.

Baharom et al.[186] conducted a systematic review of early warning systems (EWS) for predicting dengue outbreaks and found that machine learning approaches was the most common approach. They reported limited use of probabilistic evaluation methods (1 out of 17 studies) and recommended using novel data source such as social media.

Leung et al.[184] examined methods used to predict dengue outbreaks and found that about 61% of models were statistical while 40% were machine learning models. Only 5.2% of the included studies included external validation, while 20.2% reported no form of validation. This lack of transparent reporting of modelling procedures and performance metrics hindered meaningful comparison across studies.

In contrast, this systematic review aimed to include any type of study that performed temporal forecasting of dengue transmission regardless of covariates or methods. Transmission was defined broadly to include such targets as incidence, outbreaks and reproduction number. This systematic review is also unique because it was limited to studies that had conducted out-of-sample validation of their predictions, an important criterion in determining prediction

accuracy because it assesses the model's ability to predict using data it has not been trained on.

The aim of this work was to review the research conducted on temporal forecasting of dengue transmission. I sought to understand the diversity in the methods used, identify the most common evaluation criteria, and summarize covariates incorporated in these models. A key component of this review was the systematic quality assessment of included studies using the EPIFORGE (Evaluating Performance In FORecasting GuidancE) framework[208], which evaluates the rigor and transparency of forecasting methodology and reporting. I highlight gaps in the current forecasting practices, suggest areas for improvement, and propose future research directions to advance dengue forecasting science.

2.2 Methods

The protocol for this review was registered with PROSPERO, with ID CRD42022358144 and adheres to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [28].

2.2.1 Information sources

Ovid Embase (1974 to present), CAB Global Health (CAGF), Ovid MEDLINE (1946 to present), and Web of Science Core Collection databases were searched (Yalda Jafari (YJ)) without any time limit on January 11, 2023.

2.2.2 Search strategy

The search strategy combined terms related to dengue (dengue OR denv) with prediction-related terms (predict* OR forecast* OR anticipat* OR expect*). A wildcard symbol was used in front of prediction related terms to broaden the search. For MEDLINE, Embase, and Web of Science, animal studies were excluded unless they also involved human subjects using database-specific filters. The complete search strategies are presented in Appendix A.

2.2.3 Selection process

Duplicates were removed, and the records were uploaded to Swift ActiveScreeener[29] software for screening. Swift ActiveScreeener is a machine learning web application that helps with the screening stage of systematic reviews by reducing the number of records required to screen. The app functions with two models: one model prioritizes which records should be screened next and the other model estimates recall. For this review, recall estimate was set to 99%. Each title and abstract were reviewed by two reviewers (YJ plus one of Ahyoung Lim (AL), Tuyen Huynh Ngoc(TN), and Katherine Tiley(KT), Wynn Zaw (WZ)). Disagreements were resolved through consensus. Reference lists of other systematic reviews conducted on the topic were also searched to identify relevant records to include through a snowball search.

Full text articles were reviewed by two reviewers (YJ, AL, TN, KT, WZ) to identify articles for inclusion. Inclusion criteria were studies that predicted dengue transmission in humans, temporally, with at least one of incidence, incidence rate, outbreak or R0. Studies must have predicted dengue transmission using independent data from the data used to build the model. Only studies with

evaluated predictions, either visually or numerically, were included. Studies not in English were excluded.

2.2.4 Data collection process

Data extraction was completed by two reviewers (YJ plus one of AL, TN, KT, WZ). Conflict was resolved through consensus.

2.2.5 Data items

The data extraction form was developed based on a modification of EPIFORGE, a guideline for reporting epidemic forecasting research [208]. The EPIFORGE checklist includes items on study goals, data description, model characteristics, model evaluation, and translation of results for public health practice, interpretability and generalizability. Criteria from EPIFORGE that were relevant to this systematic review were incorporated into the data extraction form.

Items extracted included article identifiers (title, authors, journal, year of publication), study population details (location and date of study), methodological details (forecasting methods, covariates used, spatial and temporal random effects), and model evaluation metrics (methods of evaluating predictions).

2.2.6 Risk of bias assessment

Risk of bias was assessed using six items adapted from the EPIFORGE reporting guidelines for epidemic forecasting studies[208]. These items evaluated whether studies: (1) defined forecasting targets and study objectives; (2) documented model specifications and assumptions with appropriate references; (3) described

and justified forecast accuracy evaluation metrics; (4) specified and justified the forecast horizon; (5) compared results to benchmark models; and (6) quantified prediction uncertainty. A score of 1 was given for each item satisfied, and the total was calculated for each study.

2.3 Results

Overall, 12,466 records were identified from Embase, Global Health, Medline and Web of Science databases, as described in Figure 2.1. After de-duplication, 5,431 records were screened by their titles and abstracts. 519 records were assessed by further eligibility by reviewing their full-texts in addition to 7 records identified by a snowball search. After this review, 162 full-text articles were included in the study. Characteristics of included studies are provided in Appendix B.

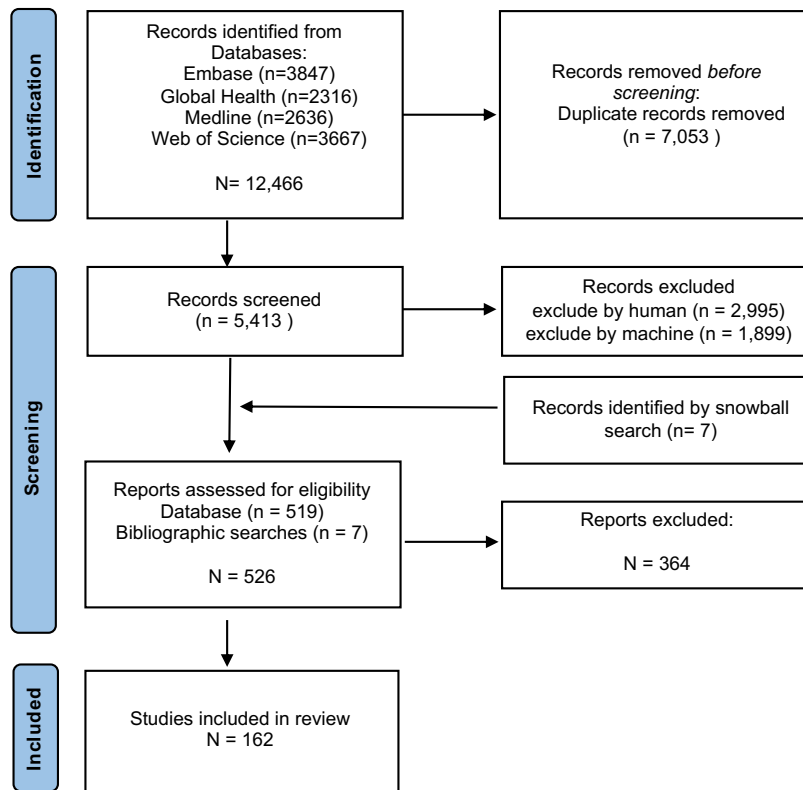


Figure 2.1-PRISMA flow diagram of study selection.

Of the 162 papers included in this study, the majority (120/162, 74.1%) made temporal predictions only while the rest (42/162 25.9%) made spatio-temporal predictions. Papers were published between the years 1991 and 2023, with the highest number published in 2020 and 2022, (22/162, 13.6% in both), Figure 2.2.

Geographical distribution of study populations by year is presented in Figure 2.2. Most of the studies were conducted on data from the Western Pacific (58/162, 35.8%) followed by the Americas (48/162, 29.6%) and South-East Asia (45/162, 27.8%) WHO regions[263]. There was only one study in each of the African and Eastern Mediterranean WHO regions, highlighting limited research in these

regions. Nine studies (9/162, 5.6%) were conducted on data from countries from multiple regions. Geographic diversity has increased over time. Earlier studies focused on the Western Pacific followed by South-East Asia and the Americas. From 2017 onwards, studies with global coverage have been increasing. Brazil was the most studied country (28/162, 17.3%), followed by Thailand (23/162, 14.2%) and China (19/162, 11.7%).

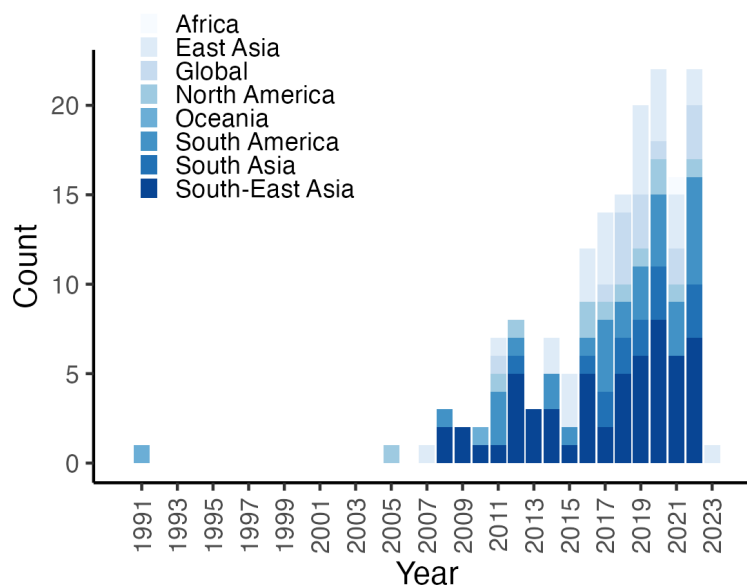


Figure 2.2-Number of studies published, by year and region.

2.3.1 Prediction target

Prediction targets can be grouped into two categories: case related and time related (Table 2.1, Appendix C). Case related measures included incidence related measures: peak incidence, prediction of outbreaks (as defined by authors), and occurrences of disease denoted by presence or absence, as well as reproduction number. Time related targets included forecasting of timing of peak incidence. The most common prediction target was related to incidence, 87.0% (141/162), including incidence and incidence rate, followed by forecasting of outbreaks at 19.1% (31/162). Time of peak incidence was a target for 3.7% (6/163) of studies.

Target	Frequency (%)	Includes
Case related		
Incidence related	141 (87.0%)	Incidence, Incidence Rate, Peak Incidence
Outbreak	31 (19.1%)	
Peak incidence	6 (3.7%)	
Occurrence	4 (2.5%)	Presence, chance of presence, Number of areas with cases
R	3 (1.9%)	R0, Rt
Time related		
Peak time	6 (3.7%)	

Table 2.1-Prediction target type.

2.3.2 Prediction resolution

Spatial resolution of model output was categorized into area or point estimates, Figure 2.3, Appendix C. Within area estimates, spatial resolution was summarized as resolution at country level, region (sub-national) or province level, and any higher resolutions including district, city, or village boundary polygons. Point estimates were defined as predictions for specific geographic coordinates (latitude/longitude), rather than aggregated to administrative boundaries. If a study included models with more than one resolution, the highest resolution was included in this analysis. Spatial scope was summarized at global, national, or sub-national levels. If a study included models with more than one spatial scope, the largest scope was included as part of this analysis. A cross-tabulation of spatial resolution and spatial scope is presented in Figure 2.3. The majority of studies, 54.3%

(88/162), included models with a district or city level resolution at a sub-national level scope.

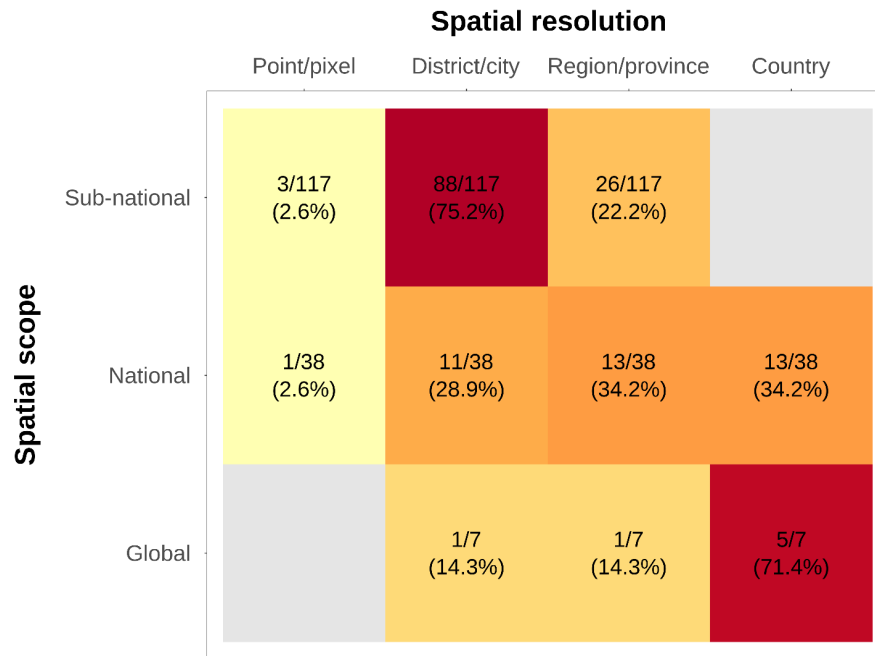


Figure 2.3-Spatial scope and resolution of included studies.

Similar to spatial resolution analysis, if a study investigated multiple models with different temporal resolutions, the highest resolution was included for analysis, Figure 2.4. The most common temporal resolutions of the models were monthly, 48.1% (78/162) followed by weekly, 42.0% (68/162).

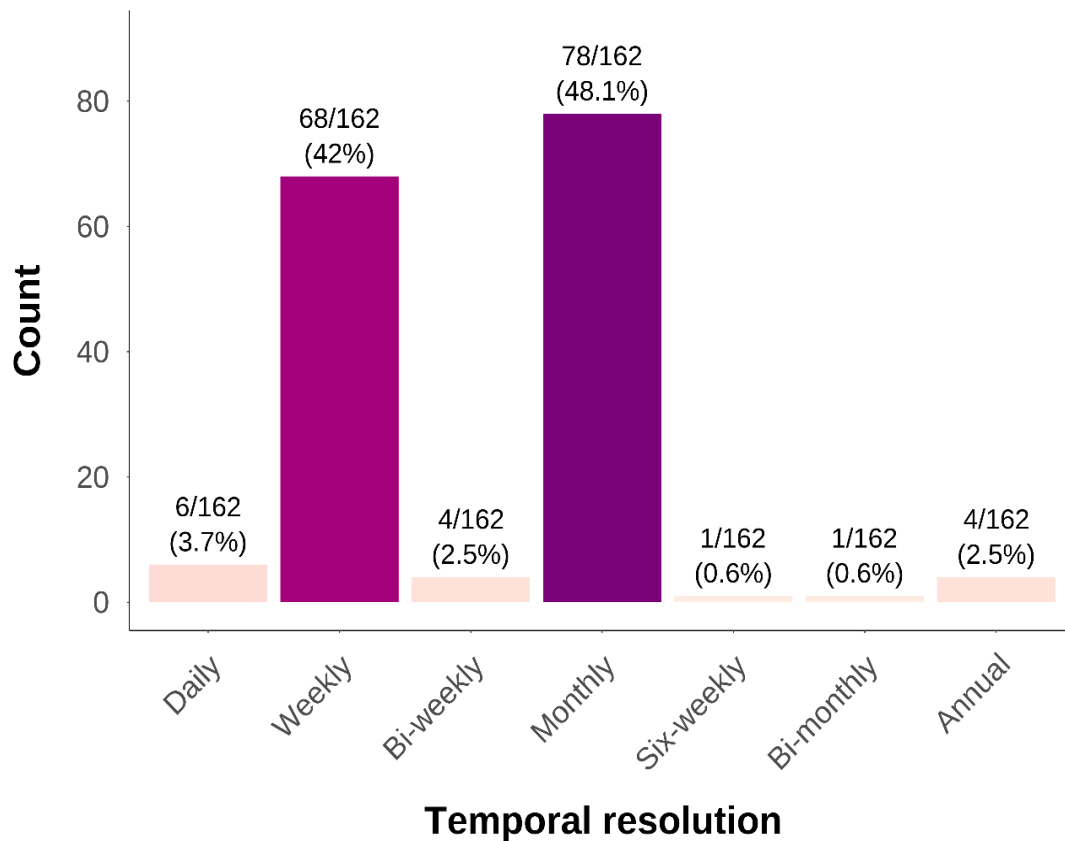


Figure 2.4-Temporal resolution of included studies.

2.3.3 Prediction period

For the prediction period, the spans of dates for forecasts were considered and are presented in Table 2.2 . The earliest prediction period was between 1 January 1982 and 31 December 2005[264], while the latest prediction period was between 1 January 2020 and 31 December 2021[241]. The most common prediction period was from 6 to 24 months ahead, used in 59.3% (96/162) of studies. The shortest prediction period was one month ahead, for the period 1 June 2014 to 30 June 2014[227], while the longest period was 24 years, between 1 January 1982 and 31 December 2005[264]. Three studies did not give a clear prediction period.

Category	Period	Count (n=162)	Percentage
----------	--------	---------------	------------

Very short-term	< 1 month	0	0
Short-term	1 to < 6 months	9	5.6%
Medium-term	6 to ≤ 24 months	96	59.3%
Long-term	> 2 years to ≤ 5 years	36	22.2%
Very long-term	> 5 years	18	11.1%
Unclear		3	1.9%

Table 2.2-Prediction period of included studies (n=162).

2.3.4 Covariates

Covariates tested by studies are presented in Table 2.3 and Appendix C, and included temperature, precipitation, other climatic factors, macroclimatic factors, environmental, demographic, vector related, and other categories.

Category	Type	Publications (n=162)	Notes
Temperature	Mean	79	
	Maximum	57	
	Minimum	56	
	Range	6	
	Other	8	average temperature anomalies, relative median, deviation, number of hours of high temperature, Number of days with TMIN < 18°C, Number of days with TMAX > 32°C, heatwave, temperature suitability index
	Unclear	3	
Precipitation	Total	50	
	Mean	17	
	Maximum	6	
	Minimum	0	
	Number of days	8	

	Other	2	average precipitation anomalies, extremely high rainfall
	Unclear	31	
Other climatic	Humidity	67	
	Pressure	10	
	Sea level pressure index	1	
	Wind	15	
Macroclimatic	ENSO MEI	2	
	Dipole Mode Index (DMI)	1	
	SOI	5	
	SST anomalies	7	
Environmental	NDVI/EVI	5	
	Land cover	2	
	Elevation	2	
	Altitude	1	
Demographic	Population density	3	
	Morbidity	1	
Vector	Breteau Index	3	
Other	Mobility	4	
	Digital data	11	Baidu searches, google searches, tweets

Table 2.3-Covariates investigated in included studies (n=162).

2.3.5 Methods used

Methods used in studies were broadly in six categories: statistical fixed effect models, statistical mixed effect models, machine learning, mathematical models, ensembles, and other, Table 2.4 and box 1.

The most common group of methods used in the included studies was autoregressive models within the statistical fixed effect category, shown in Table 2.4 and Figure 2.5. Within the statistical mixed effect category, other methods, and withing machine learning category neural networks methods were most common.

Mathematical models, ensemble models, and otherwise not specified models were each combined into their own category.

Over time, the proportion of fixed effect statistical models decreased while the proportions of other model types, particularly machine learning models, increased, Figure 2.6.

Methods used	Publications (n=162)	Notes
Statistical fixed effect		
AR	62	ADL, ARIMA, ARIMAX, SARIMA, SARIMAX, VAR
GLM	39	
GAM	16	
LASSO	13	
Other	15	BATS, DLNM, elasticity, ElasticNet, ETS, exponential smoothing methods, GLARMA, historical likelihood, Kalman filter, MCP, multi-dimensional method of analogues, null model, PDLM, Ridge, SCAD, Spatial-Temporal Hotspot Analysis, TBATS
Statistical mixed effect		
AR	2	
GLMM	7	
GP	2	
Other	13	Bayesian weighted outbreaks, BME, EpiEstim, EpiFilter, KDE, SSTAR, STAR, StructTS
Machine learning		
Neural networks	19	
LSTM	13	
Support vector	16	

Random Forest	10	
Gradient boost	8	
Fuzzy	2	
Other	6	k-Nearest Neighbour, graph-embedding algorithm, extreme learning machine, EPRA, STCFM, transformer
Mathematical models	18	
Ensemble	10	
Other	7	ARIMA-ANN, ARIMA-LSTM, ARIMA-NNAR, Danger theory algorithm, EPM, GARP, multi-dimensional method of analogues, prediction market system, unable to ascertain specifics

Table 2.4-Methods used in included studies. (Abbreviations: ADL, Autoregressive Distributed Lag; ANN, Artificial Neural Network; AR, Autoregressive; ARIMA, Autoregressive Integrated Moving Average; ARIMA-ANN, hybrid ARIMA-Artificial Neural Network; ARIMA-LSTM, hybrid ARIMA-Long Short-Term Memory; ARIMA-NNAR, hybrid ARIMA-Neural Network Autoregression; ARIMAX, Autoregressive Integrated Moving Average with Exogenous Variables; BATS, Box-Cox transform, ARMA errors, Trend and Seasonal components; BME, Bayesian Maximum Entropy; DLNM, Distributed Lag Non-linear Model; EPM, Epidemic Prediction Markets; EPRA, Evolutionary Polynomial Regression Algorithm; EpiEstim, Epidemic Estimation; ETS, Error, Trend, Seasonal; GAM, Generalized Additive Model; GARP, Genetic Algorithm for Rule-set Prediction; GLARMA, Generalized Linear Autoregressive Moving Average; GLM, Generalized Linear Model; GLMM, Generalized Linear Mixed Model; GP, Gaussian Process; KDE, Kernel Density Estimation; LASSO, Least Absolute Shrinkage and Selection Operator; LSTM, Long Short-Term Memory; MCP, Minimax Concave Penalty; NNAR, Neural Network Autoregression; PDLM, Polynomial Distributed Lag Model; SARIMA, Seasonal Autoregressive Integrated Moving Average; SARIMAX, Seasonal Autoregressive Integrated Moving Average with Exogenous Variables; SCAD, Smoothly Clipped Absolute Deviation; SSTAR, Spatial-Spatio-Temporal Autoregressive; STAR, Spatio-Temporal Autoregressive; STCFM, Spatio-Temporal Convolutional Forecasting Model; StructTS, Structural Time Series; TBATS, Trigonometric seasonality, Box-Cox transform, ARMA errors, Trend and Seasonal components; VAR, Vector Autoregression.

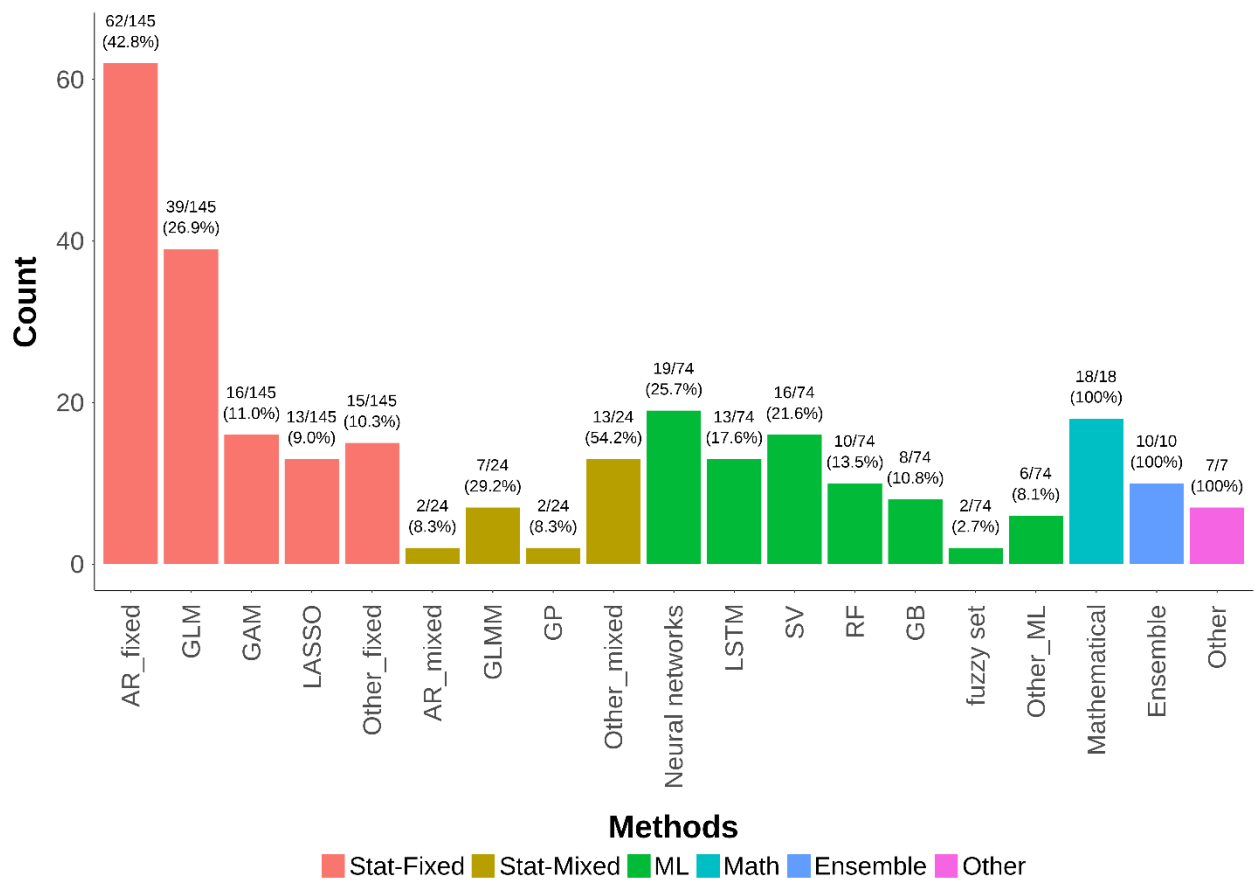


Figure 2.5-Frequency of methods used.



Figure 2.6-Proportion of methods used, by year.

Box 1. Definitions of categories of methods:

- Statistical models – fixed effects: Models that assume unobserved heterogeneity is constant over time [2].
- Statistical models – mixed effects: Models that combine fixed effects, which account for time-invariant heterogeneity, and random effects, which capture time-variant heterogeneity. This structure allows for modeling both within-group and between-group variations while accounting for correlation across different hierarchical levels [2, 3].
- Machine learning models: Models that learn patterns from data and make decisions or predictions without being specifically programmed [4, 5].
- Mathematical models: Models that represent real-worlds systems using mathematical frameworks, such as equations to simulate and predict outcomes [7, 8].

Ensembles models: Models that combine multiple individual models to improve performance by leveraging strengths of each individual model [9].

2.3.6 Random effects

Random effects are important as they account for unobserved heterogeneity[265].

Structured spatial random effects were incorporated into models in 6.8% (11/162) of studies. Eleven studies included structured random effects comprising Conditional Autoregressive (CAR) models (5/11), spatiotemporal random field (S/TRF) (2/11), Spatial Autoregressive models (SAR) (1/11, 9.1%), neighbourhood-level random effects (1/11, 9.1%), empirical Bayes smoothing, bivariate local Moran's index and ConvLSTM, a deep learning model to handle

spatial-temporal dependencies (1/11, 9.1%). Five studies (5/12, 41.7%) included unstructured random effects, while 33.3% (4/12) used both.

8.6% (14/162) of studies incorporated structured temporal random effects into their models. Five studies, 35.7% (5/14), included seasonal effects, one, 20.0% (1/5), of which used year as a random effect and four, 80% (4/5), used an auto-correlated monthly effect. Two, 14.3% (2/14), included autoregression dengue cases, two, 14.3% (2/14) studies use spatiotemporal random field (S/TRF) and two, 14.3% (2/14), studies used kernels. One, 7.1% (1/14), study included an auto-regressive moving average (ARMA) noise process, and one, 7.1% (1/14), included time-varying coefficients.

2.3.7 Evaluation metrics:

Evaluation methods are categorized into point-based and probabilistic evaluations. A single study can include more than one kind of evaluation method. For this analysis, each category is counted only once per study, to denote the use of the specific method at least once.

Only 4.3% (7/162) of studies used a probability-based evaluation method. Probability based evaluation methods included log scores and Continuous Ranked Probability Score (CPRS).

Within point evaluation methods, the most used methods for evaluation were point difference methods at 63.0% (102/162). Some examples of this category of methods were root mean square deviation (RMSE) and mean absolute error (MAE). The second most used category within point evaluation was categorical

evaluation at 25.3% (41/162). Examples of this kind of evaluation included Matthew's correlation coefficient, sensitivity and specificity. Finally, correlation methods were used to compare predicted points with observations in 23.5% (38/162) of studies. Examples included coefficient of determination and Pearson correlation coefficient. 14.2% (23/162) used only a qualitative approach to evaluation.

Evaluation method	Publications (n=162)	Includes
Probabilistic Evaluation	7	Brier score, coverage rates, CRPS, DSS, log score
Point Evaluation		
Categorical	41	AUC, chi-square, detection rate, hit rate, MCC, NPV, PPV, sensitivity, specificity, Theil's U
Correlation	38	Coefficient of determination, Pearson correlation coefficient
Point difference	102	Average error, GMRAE, MAD, MAE, MAPE, MASE, mean-squared difference, RMAE, RMAPE, RMSE, RMSPE, RMSWE, SRMSE, UMBRAE, VAF
Qualitative only	23	Visual comparison

Table 2.5-Evaluation methods used in included studies. (Abbreviations: AUC, Area Under the Curve; CRPS, Continuous Ranked Probability Score; DSS, Dawid-Sebastiani Score; GMRAE, Geometric Mean Relative Absolute Error; MAD, Mean Absolute Deviation; MAE, Mean Absolute Error; MAPE, Mean Absolute Percentage Error; MASE, Mean Absolute Scaled Error; MCC, Matthews Correlation Coefficient; NPV, Negative Predictive Value; PPV, Positive Predictive Value; RMAE, Relative Mean Absolute Error; RMAPE, Root Mean Absolute Percentage Error; RMSE, Root Mean Square Error; RMSPE, Root Mean Squared Percentage Error; RMSWE, Root Mean Squared Weighted Error; SRMSE, Standardized Root Mean Square Error; UMBRAE, Unscaled Mean Bounded Relative Absolute Error; VAF, Variance Accounted For)

2.3.8 Comparator

About 61.7% (100/162) of studies compared forecasts between more than one model. Specifically, 39% (39/100) of studies compared the models to each other, 46% (46/100) compared models to a baseline, and 15% (15/100) compared the model under investigation to other models. A baseline and model under investigation were defined by authors of each study.

2.3.9 Risk of bias assessment

The average risk of bias score of included studies was 4.9 out of 6, ranging from 3 to 6. The lowest scored category was whether studies had included uncertainty in their forecasting results. Only 27.8% (45/162) of studies reported uncertainty in their predictions.

2.4 Discussion

Based on a search of the literature without any time limits, 162 studies were identified that conducted temporal forecasting with an out-of-sample validation. The first study was from 1991 while the last included study was from 2023, the majority (74.1%) of which made temporal predictions while 25.9% made spatio-temporal predictions. Most studies were conducted in the Western Pacific (35.8%), Americas (29.6%), and South-East Asia (27.8%) WHO regions, with only one study each from the African and Eastern Mediterranean regions. Brazil was the most studied country (17.3%), followed by Thailand (14.2%) and China (11.7%). Geographic diversity increased over time, with studies achieving global coverage from 2017 onwards. Sylvestre et al. found that populations from South-East Asia (37%) were the most common studied region, though this study also included

surveillance studies. Baharom et al.[186] who conducted a systematic review of EWSs, found that most studies were conducted in Asia. Other reviews did not specify geographical spread of the studied populations[184, 216, 262]. Overall, these findings align with global burden of dengue. Based on the most recent global epidemiological overview by WHO for the year 2024, the region of Americas accounted for 90% of dengue cases, followed by South-East Asia and Western Pacific[266].

Most models forecasted at a sub-national spatial scope and district/city resolution and (54.3%), while the most common temporal resolution was monthly (48.1%). Data availability most likely constrains model resolution and scope. Additionally, computational demands may increase with higher spatiotemporal resolution and scope, possibly limiting feasibility in real-world settings. The district/city and monthly resolutions may align well with public health measure intervention scales and timelines.

The most common prediction target was incidence-related measures (87.0%), with medium-term prediction periods of 6-24 months being most frequent (59.9%). Incidence-related measures provide more direct evidence than outbreak status for resource allocation as knowing the expected number of cases allows for specific resource allocation[267]. While studies report that predictions of shorter period are more accurate [202, 226, 268], the preference for 6-24 months prediction periods may reflect the lead time required for adequate planning[269].

The most common methods were autoregressive models (statistical fixed effects), with the proportion of fixed effect models decreasing over time while machine learning approaches, particularly neural networks, increased. De Lima et al.[262], looking at case prediction models for arboviruses found a similar finding of a

majority of moving average models. Leung et al.[184] found that for dengue outbreak predictions, statistical models, particularly times series/autoregressive models were most common.

There is still a large focus on fixed effects models, which discount the temporal and spatial correlation structures of data. Only 6.8% of studies incorporated structured spatial random effects (primarily CAR models) and 8.6% included structured temporal random effects. Even though fixed effect models might be easier to build, run, and explain to non-experts, they may underperform in forecasting as they fail to account for correlations and unobserved heterogeneity. For example, Lowe et al.[270] found that a generalized linear mixed model (GLMM) with temporal and spatial random effects improved dengue forecasting as compared to a generalized linear model (GLM). Models with random effects account for overdispersion, spatial and temporal correlation, and unmeasured confounding factors[216], though they may be more complicated to estimate.

Temperature, precipitation, and humidity were the most frequently included predictors. This finding is similar to other systematic reviews examining dengue forecasting models[184, 186, 216, 261] and arboviral disease predictions more broadly[262].

For model evaluation, 63.0% used point difference methods (e.g., RMSE, MAE), while only 4.3% employed probabilistic evaluation methods such as continuous ranked probability scores. Baharom et al.[186] also found a very low use of probabilistic evaluations, with only one study among 17 included studies using CRPS. Point difference and correlation methods offer advantages as they are faster to implement, easier to compute, and easier to communicate. However, they

have important limitations. By focusing on average accuracy or a general agreement between forecasts and observed outcomes, they ignore measures of uncertainty associated with the forecasts. Probabilistic evaluation methods, such as probabilistic skill scores, provide a more comprehensive view of model performance by reflecting both accuracy and uncertainty of forecasts[271].

About 61.7% of studies compared multiple models. Model comparison is important to demonstrate the advantages of proposed models over alternative approaches, particularly any models already in operational use by public health agencies.

This review had some limitations. This study only focused on English language publications, which may have resulted in the exclusion of relevant studies. Dengue is mostly endemic in non-English speaking countries, and as such, this restriction may have led to an underrepresentation of studies led by local researchers. Additionally, this language bias may limit the generalizability of the findings. Some studies included multiple modelling approaches and due to time restrictions, covariates extracted were covariates used at the study level rather than at the individual model level. This may have led to limited granularity in synthesis of evidence, as differences between covariates by model could not be explored. Additionally, effect size estimates of covariates were not captured, for the same above reason, and because the effect size estimates are highly context dependent and may have limited value in informing future modelling work.

A valuable aspect to consider in future research is the real-world impact of forecasting models on public health practice. While a forecasting model may indicate a good predictive performance, it may not be effective if it is cumbersome to implement, difficult to interpret, or requires resources beyond what is feasible in

practice. Therefore, future evaluation criteria should go beyond traditional metrics like accuracy and precision and incorporate user experience as a vital component. This can include assessing how easy it is for public health officials to use the model, how well it integrates with existing systems, and whether it provides actionable insights in a timely and clear manner. By considering these practical aspects, a more holistic evaluation of forecasting models can be achieved and will help ensure forecasting models are not only scientifically robust but also valuable and useable in real-world public health decision making and increase their impact in preventing and controlling diseases like dengue.

This systematic review was important in guiding the subsequent research in this thesis. The limited use of spatio-temporal random effects and probabilistic evaluation identified in the literature motivated the focus on forecasting through a Bayesian hierarchical approach with proper scoring rules. In addition, the review found that most studies used climate and environmental predictors, with few studies incorporating viral genetic diversity or human mobility data. These findings further informed this research, and subsequent chapters examine whether genetic diversity and human mobility improve forecast skill. Overall, this systematic review's findings shaped the methodological framework as well as the selection of covariates used in this thesis.

3 Descriptive Chapter

3.1 Introduction

In this chapter, I aim to provide a descriptive overview of the dengue surveillance and serotype data provided by MoPH in Thailand. This will be done by analysing the temporal and spatial distribution of dengue cases within Thailand and identifying patterns such as seasonal trends and synchrony of outbreaks. I will review dengue incidence at provincial and district spatial scales and at weekly, monthly, and yearly temporal scales to identify spatio-temporal patterns. I will also review data on dengue serotypes.

Previous analyses of dengue data in Thailand have covered varied temporal and spatial scales. National-level studies have predominantly analysed province-level data from the early 1980s to recent years. Longer-term national studies using province-level data have covered periods from 1981 to 2017[272] and 1985 to 2005[273] across Thailand's provinces. Bekoe et al.[235] analysed province-level monthly data from all provinces from 2003 to 2015, while Saita et al.[241] examined monthly reported dengue cases from all provinces from 2011 to 2021. Garcia-Carreras et al. conducted one of the longest national studies, analysing monthly province-level dengue case data from 1968 to 2018[274]. Finer-scale district-level analyses have typically been restricted to specific provinces or regions rather than national coverage. Phantichat et al.[234] analysed monthly sub-district level data from 199 sub-districts in Khon Kaen Province, northeastern Thailand, from 2006 to 2016, while Jeefoo et al.[275] examined monthly village-level data from Chachoengsao Province from 1999 to 2007. More recent district-level work

has focused on specific regions, such as the analysis of 377 sub-districts in the four southernmost provinces from 2008 to 2020[240].

These studies have revealed consistent patterns in dengue transmission across Thailand. National analyses have shown that dengue incidence peaks in the rainy season, around May to August, with increasing trends in reported cases from 2003 to 2015[235]. The peak is often in June, though it was delayed to July during COVID-19 pandemic[241].

Temporal analyses have identified cyclic patterns, with epidemic peaks around every 2 to 5 years[276]. Garcia-Carreras et al.[274] found similar results of interannual cycles of 1.5 and 4.5 years, that varied over time and space. They also found periodic synchronization events where dengue outbreaks became synchronised across provinces, coinciding with El Niño (lower than normal temperatures) and La Niña (higher than normal temperatures) events ((1972 to 1973, 1982 to 1983, 1997 to 1998, and 2014 to 2016).

Dengue waves in Thailand are non-stationary, following seasonal patterns[105, 274]. However, there may be dynamic changes in this non-stationarity as a result of changing environmental factors, human mobility, and changing immune landscape[274, 277]. Wavelet analysis helps to disaggregate these influences by revealing how timing and strength of dengue cycles vary across scales (inter and intraseasonal) and over time[278].

In a seminal study, Grenfell et al.[278] analysed measles epidemics in the UK between 1944 and 1994, using a variety of tools including wavelet analysis. They demonstrated that London acted as a hub for infection spread and showed that

both distance from London and population size significantly influenced the timing or lag of outbreaks across the UK. In a study by Xu et al.[105], looking at dengue incidence across Thailand, Bangkok was treated as the epidemic hub and the authors found that peak dengue activity in Bangkok often predicted dengue incidence elsewhere. More specifically, a study by Cummings et al.[279] found traveling waves of DHF that spread outward from Bangkok at roughly 148 km per month.

In this chapter, I apply wavelet analysis to disaggregate cases into inter and intraseasonal components and explain the role of seasonal and non-seasonal forces in shaping epidemics. I also compare timings in waves between Bangkok and other districts at inter and intraseasonal temporal scales to assess the direction and coherence of epidemic spread.

Phylogenetic analysis will be used to explore the role of genetic diversity on disease distribution across Thailand. Genetic differences between and within serotypes can influence transmissibility, immune response, cross-protection and epidemic potential[280-283]. Co-circulation and changing dominance of various serotypes and genotypes may contribute to spatial and temporal heterogeneity of dengue incidence. By examining the evolutionary relationship between publicly available dengue sequences, I aim to identify whether certain lineages are more dominant than others in Thailand and whether there is an interplay between this dominance and epidemic spread.

The objectives of this chapter are to visualise and analyse Thailand's dengue surveillance data to better understand patterns and trends that will inform the development of forecasting models in chapters 4-6. The temporal and spatial

patterns identified through wavelet analysis will guide the specification of the time series models that will incorporate epidemiological, climatic, environmental, and mobility factors. Serotype and genetic data are included to assess representativeness and spatio-temporal variability as potential covariates for multivariable forecasting models.

3.2 Methods

3.2.1 Surveillance data

Dengue fever is a notifiable disease by law in Thailand. The BoE, MoPH, first initiated the national surveillance system for DHF in 1972, becoming fully operational in 1974 [242, 284]. In 1994, DF was added to the surveillance system [242].

Surveillance is passive. Case definitions follows the 1997 WHO case definitions, classifying dengue into three categories of DF, DHF, and DSS [1, 233, 285]. A detailed definition of these categories is presented in Table 1.1 in Chapter 1. Briefly, dengue virus infection can be asymptomatic or symptomatic. Symptomatic infections present as undifferentiated fever (viral syndrome), DF syndrome (which can occur with or without haemorrhage), or DHF (characterized by plasma leakage), which may progress to DSS [1, 286]. Cases include probable without any laboratory confirmation as well as laboratory confirmed cases [1, 243]. Within 24 hours of diagnosis, public and private hospitals report cases to district surveillance centres. District surveillance centres then report cases up to provincial epidemiological units, which then report to the BoE [243]. Even though not all suspected cases are tested which may result in a lower specificity, passive surveillance is a useful tool for monitoring dengue trends over long time

periods [287]. Based on Thai guidelines, even if a commercial NS1 test is negative, dengue cannot be ruled out[285].

We obtained surveillance data from BoE, MoPH. Ethical approval was received from the Ethical Committee of Faculty of Tropical Medicine, Mahidol University. Ethical approval from OxTREC was not required as only anonymised routine surveillance data was used. The data covers the years 2013 to 2022 and is available for all 928 districts and at weekly time intervals. The data is also disaggregated by severity into DF, DHF, and DSS. For our analysis, we combined all types of dengue disease.

National and provincial level cases at monthly level were also accessed online through MoPH's public-facing website available at the following reference [288].

Outbreak threshold was defined as median of number of cases in the past five years for the same spatio-temporal resolution, following MoPH's methodology.

3.2.2 Population data

Population data at the district level was downloaded from Department of Provincial Administration (DoPA) for years 2013 to 2022[289].

3.2.3 Serotype data from surveillance

Dengue serotype testing is provided free of charge for patients that meet specific criteria including suspected severe cases or fatalities. For patients clinically suspected of dengue virus infection, the following criteria are used to determine eligibility for dengue serotype testing[290]:

- A positive test results from a dengue rapid diagnostic test
- Fever lasting no more than five days, and
- A complete blood count showing white blood cell count of less than 5,000 cells/mL, predominantly composed of lymphocytes or atypical lymphocytes.

Detection of dengue virus serotype is performed using real time polymerase chain reaction (RT-PCR). Serotype data is reported to the national disease surveillance database through two pathways. In the first pathway, the data flows from each of the 12 Offices of Disease Prevention and Control (ODPC) and Institute for Urban Disease Control and Prevention in the Bangkok area, via sentinel hospitals within each ODPC region. The data is then reported to the Department of Disease Control (DDC), and then the Division of Vector-Borne Diseases (DVBD) who then report to the national disease surveillance database. A list and location of sentinel hospitals is provided in Figure 3.1. The second pathway includes reporting of dengue serotype testing data directly by the hospitals to the national disease surveillance database.

For this thesis, serotype data was provided by MoPH from 2014 to 2022. No serotype data was available for 2013. The data was an anonymised line list of tests and included date of collection, dengue disease type (DF, DHF, DSS), epidemiological week of data collection, year, and serotype.

The data was analysed by year, province, serotype and dengue disease type. Test coverage by dengue disease type was calculated by dividing the number of serotype tests by the number of dengue disease cases identified through surveillance.

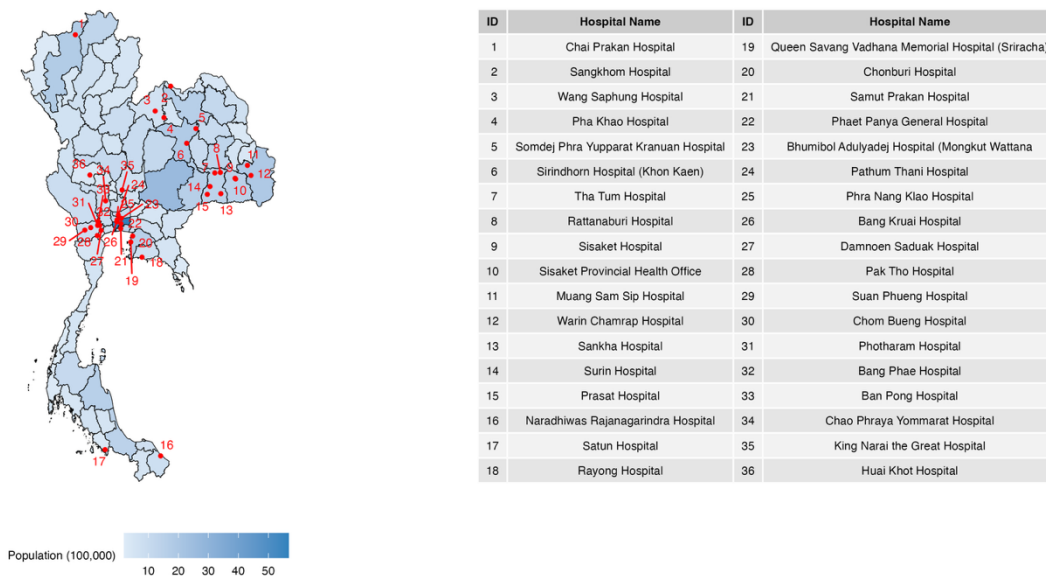


Figure 3.1-Sentinel hospitals for serotyping, 2017-2019. Population for each province is the average of 2017 - 2019.

3.2.4 Wavelet analysis

I applied a wavelet analysis to dengue data in Thailand to understand the spatial and temporal patterns of the epidemic.

Epidemic waves in each district were compared to the aggregated epidemic wave in Bangkok province as reference given Bangkok's status as the capital and largest metropolitan area in the country.

Incidence was calculated at a rate per 100,000 population. Time series were logged after adding a constant of 1.

WaveletComp[291] was used to complete the wavelet analysis. A loess span of 0.75 within wavelet.coherency function was used to detrend the time series. This is the default in the package and has been applied in previous research using time-

series data[292]. A loess span of 0.75 indicates moderate smoothing using 75% of the data points for local polynomial fitting, which is a good balance between removing long-term trends (bias) and preserving periodic components of interest (variance)[293].

Annual periods were defined as periods equal to or greater than 0.5 years and less than 1.5 years and multi-annual periods were defined as periods equal to and greater than 1.5 years and equal to and less than 5 years. This was based on previous studies that found that dengue in Thailand follows annual and multi-annual cycles of two to four years[273, 294].

Time series were reconstructed by filtering periods to annual and multi-annual periods.

Phase differences between districts and Bangkok were calculated for annual periods and multi-annual periods. To calculate this, first the cases in 50 districts of Bangkok province were aggregated into one time series. Then the wavelet coherencies between each district and Bangkok province were calculated. Using only significant coherencies at 0.05 level, the phase differences angles were extracted and converted to time difference for each week and period within the annual $[0, 1.5)$ and multi-annual $[1.5, 5]$ period ranges to obtain an average weekly phase difference.

The relationship between travel time and phase difference between each district and Bangkok was investigated. To determine travel time, I extracted the centroids of each district and determined the duration in minutes to travel from each district's

centroid to Bangkok province's centroid, using the OSRM package[295] in R. The default of using a car for travel was used.

3.2.5 Phylogenetic analysis

Publicly available genetic sequences and their metadata for DENV-1, DENV-2, DENV-3, and DENV-4 for Thailand and the South East Asian region were downloaded from GenBank[296], with reported sample collection dates from January 1, 2010 to December 31, 2022. Sequences with missing collection date or location were excluded from the analysis. Sequences with partially missing data on collection date (i.e., where only the collection year or month and year was available), were assigned to the mid-point of their respective sampling times.

The arbovirus typing tool provided by Dengue-GLUE[297] was used to accurately identify the serotype, genotype, and sub-lineage of each sequence. Sequences where the Dengue-GLUE typing mismatched the reported serotype labels were removed. I aligned sequences to serotype-specific references (DENV-1 = NC_001477.1, DENV-2 = NC_001474.2, DENV-3 = NC_001475.2, DENV-4 = NC_002640.1) using NextAlign[298]. All sequences were trimmed to include only the E gene. Sequences exceeding 31% missing or ambiguous sites were excluded, based on pre-established sequencing thresholds followed by Inward et al.[299].

I included all Thai sequences for each serotype and added sequences from the rest of Southern and Southeast Asia (SEA) to allow for the identification of Thailand-specific lineages. An evenly weighted sampling approach was used to ensure a selection of sequences proportional to the available number of

sequences across different locations and time periods. For the total of available sequences from a specific geographical location collected within the same year, I assigned a weight to each individual sequence, defined as $1/x_{ij}$, where x_{ij} represents the number of sequences from location i collected during time period j . This weighting inversely correlates with the probability of selecting a sequence from the same location each year. Consequently, locations with fewer sequences within a particular year were given higher selection probabilities, ensuring a balanced representation of geographical locations and reducing the impact of sampling biases.

Following initial data processing, maximum-likelihood (ML) trees were constructed for each serotype, inferred using IQ-TREE2 under a HKY+F+I substitution model. The ML tree outputs were time-calibrated with TreeTime, informed by tip sampling dates. TreeTime was also used to assess the temporal signal and remove outliers from the sub-sampled datasets. To improve clarity and ensure interpretable results, I selected only genotypes with over 10 sequences from Thailand.

Pairwise nucleotide diversity, π , (hereby referred to as genetic diversity) for each serotype was completed using the pegas package[300] in R. Checks of co-linearity were done using vifcor function from the R package usdm[301] to test whether genetic diversity was related to number of samples.

This analysis used code adapted from Inward et al[299].

3.3 Results:

3.3.1 Surveillance data:

3.3.1.1 National trends

Incidence per 100,000 population per week from 2013 to 2022 is presented in Figure 3.2. From this figure, it can be observed that dengue follows both inter and intra-annual cycles. Increase in cases are observed every two to four years.

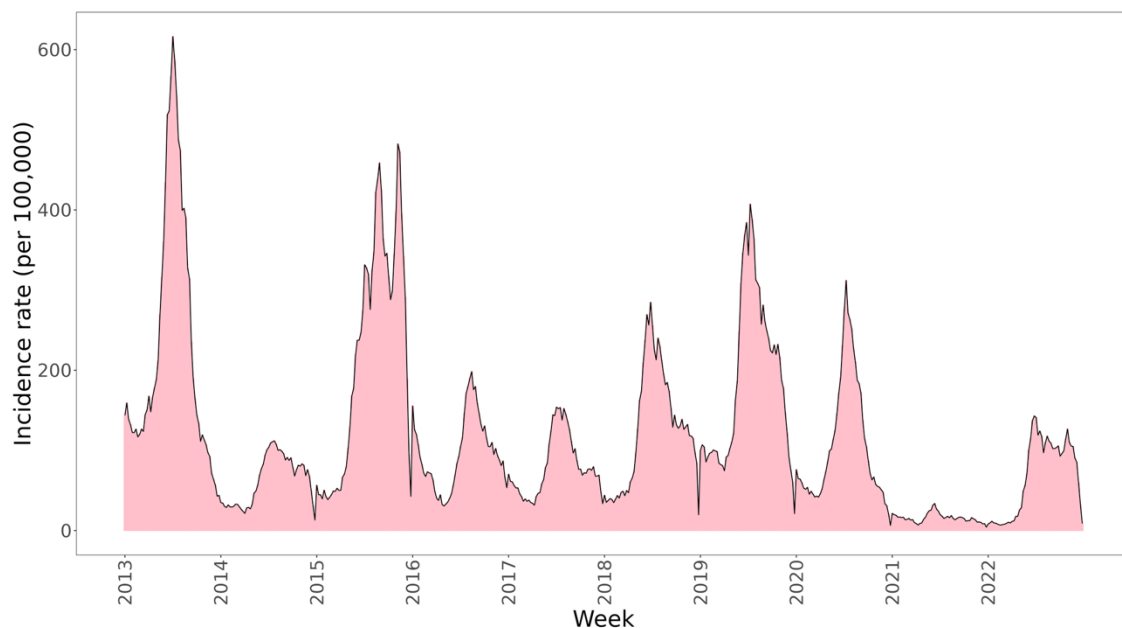


Figure 3.2- Incidence per 100,000 people

Figure 3.3 shows the annual number of cases for the same period. Based on the MoPH's defined outbreak threshold, which is the median of the past five years, there were outbreaks in years 2013, 2015, 2018 and 2019. Outbreaks at monthly level are depicted in Figure 3.4. From some years, outbreaks are observed in certain months, such as in 2016, while in other years, outbreaks were observed in every month, such as in 2015.

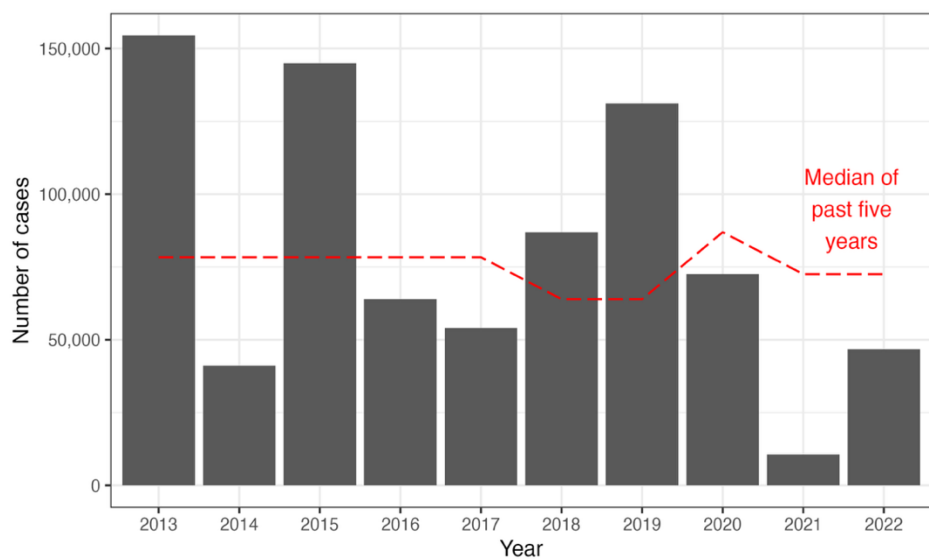


Figure 3.3-Annual cases with outbreak threshold in red.

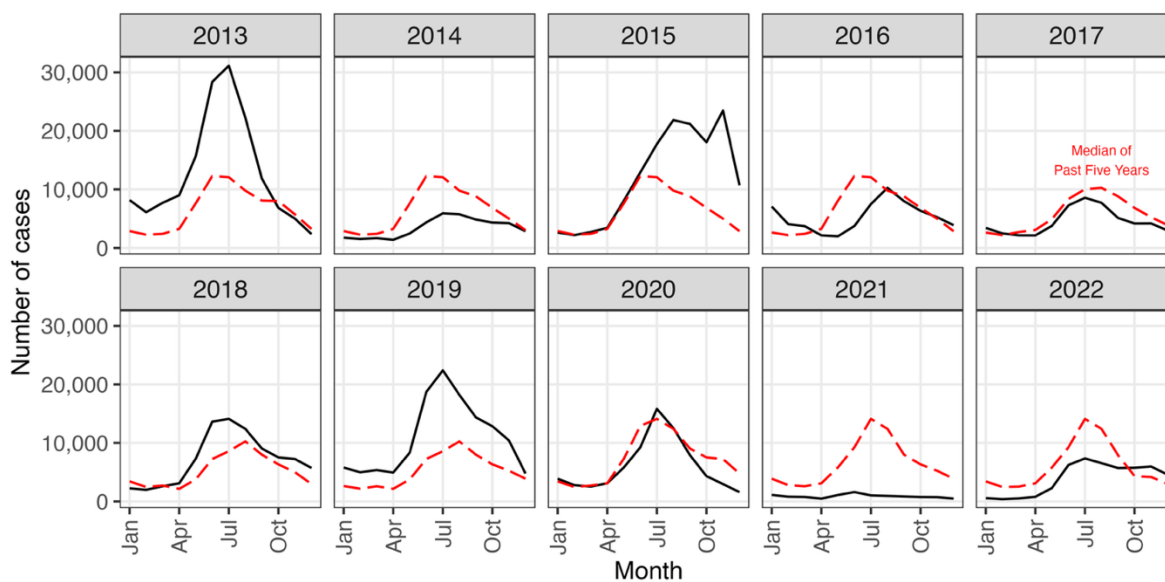


Figure 3.4-Monthly cases with outbreak threshold in red.

3.3.1.2 Provincial trends

Weekly incidence by province is presented in Figure 3.5. Heterogeneity in disease patterns can be observed between provinces.

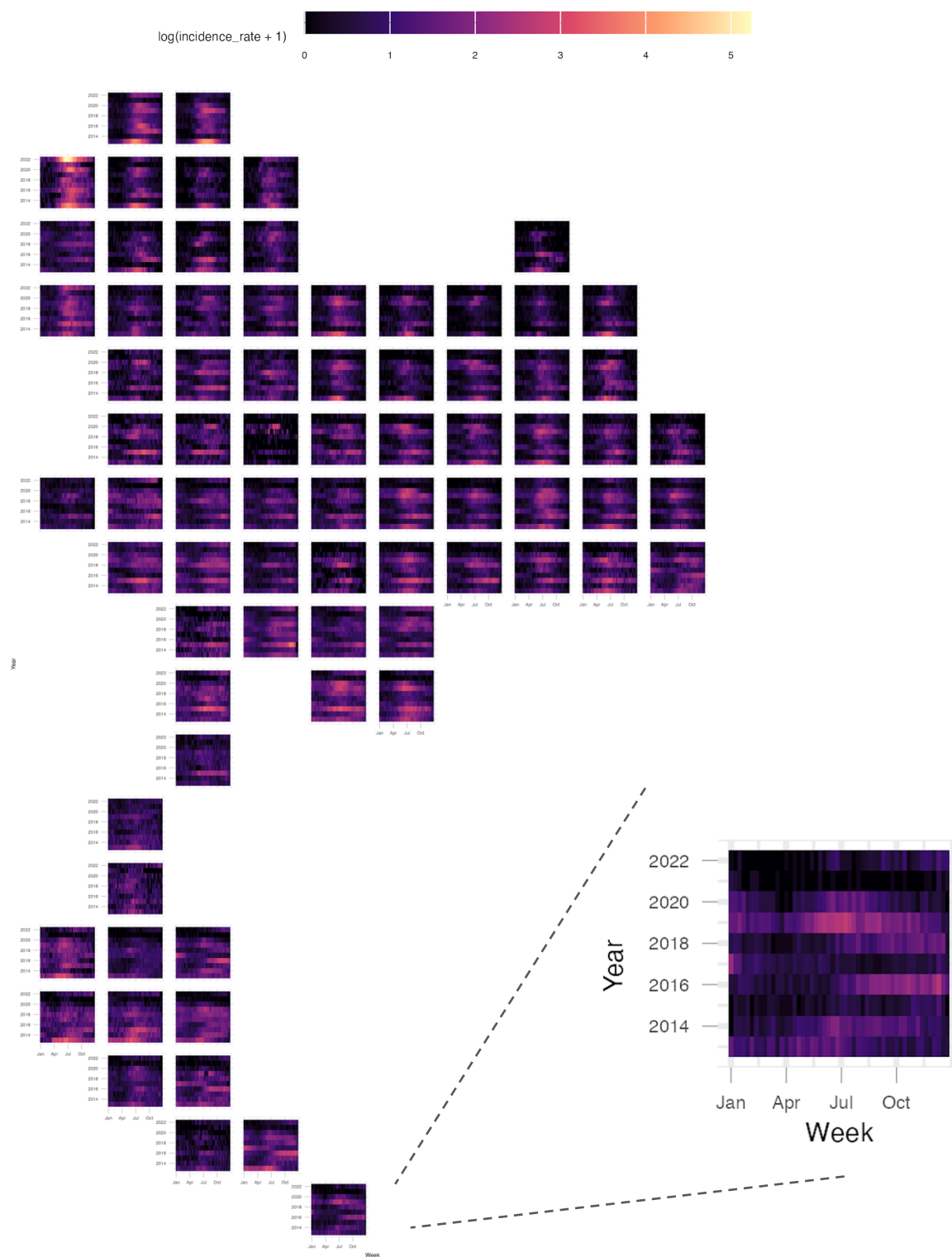


Figure 3.5-Incidence per 100,000 people, by province.

3.3.1.3 District trends

Weekly incidence of cases by district are presented in Figure 3.6, with districts ordered from south to north. Brighter yellow colour indicates higher incidence while darker blue indicates lower incidence. The temporal synchrony between districts is generally strong, with most areas experiencing elevated incidence (yellow) simultaneously during transmission seasons. There is, however, some variation in season lengths across the years. Some years, such as 2016 and 2017, show more concentrated peaks of around 2 to 3 months, while others such as 2013, 2015, and 2019 show longer transmission seasons of around 4 to 5 months. There are some outliers. A drastic drop in incidence is observed in 2021 across all districts, coinciding with the COVID-19 pandemic and the subsequent lockdowns and reduction in health-seeking behaviour. There is some recovery of the incidence rates in 2022. Across all years, within the same year, there are districts with more bright yellow bands than others, such as districts in two-thirds down the plot (district in and around Bangkok) in 2015 and most southern districts in 2016. Indeed, the bright band in 2015 corresponds to a an unprecedented dengue outbreak that occurred that year[239]. Additionally, certain districts consistently demonstrate higher baseline and peak incidence throughout the study period, such as districts two-thirds down the plot and northern districts. These observations indicate spatial heterogeneity of transmission across districts.

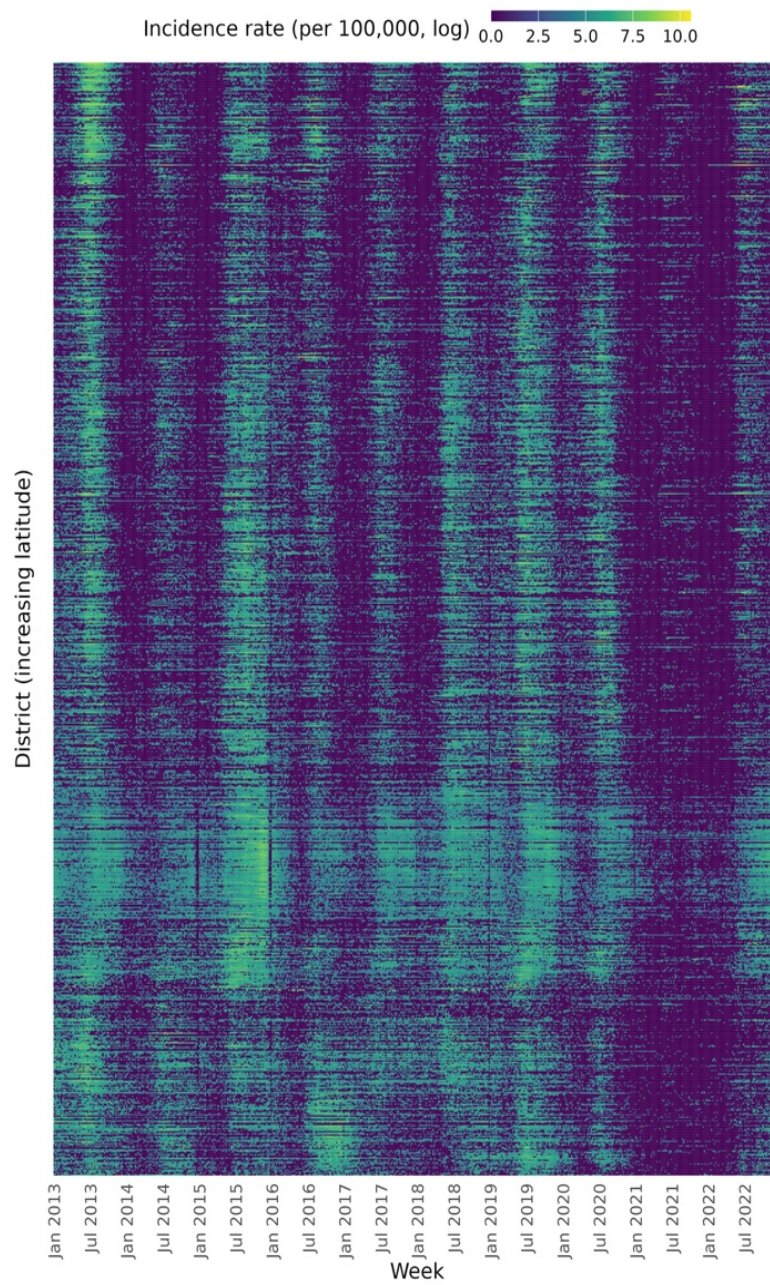


Figure 3.6-Incidence per 100,000 people, by district

3.3.2 Serotype results

Between 2013 and 2022, there were 3,163 serotype tests conducted, only 0.02% of all cases. The highest proportion, 34.0%, was done in 2019 followed by 27.9% in 2020. Comparing serotype testing by disease type, individuals with DSS followed by DHF were more likely to be tested for serotype, as expected from serotype testing eligibility criteria, Figure 3.7a. Looking at serotypes results, DENV1 was the most frequent dengue serotype, 65.0%, followed by DENV2, 28.5%, DENV4, 3.9% and DENV3, 2.7%. This relative pattern was consistent across the years from 2014 to 2022, Figure 3.7b. This pattern was also mainly consistent across disease types, Figure 3.7c, though there were some differences. For example, between 2014 and 2016, DENV3 was responsible for a higher proportion of cases of DF as compared to DHF and DSS, though this difference may be an artefact of the sampling strategy.

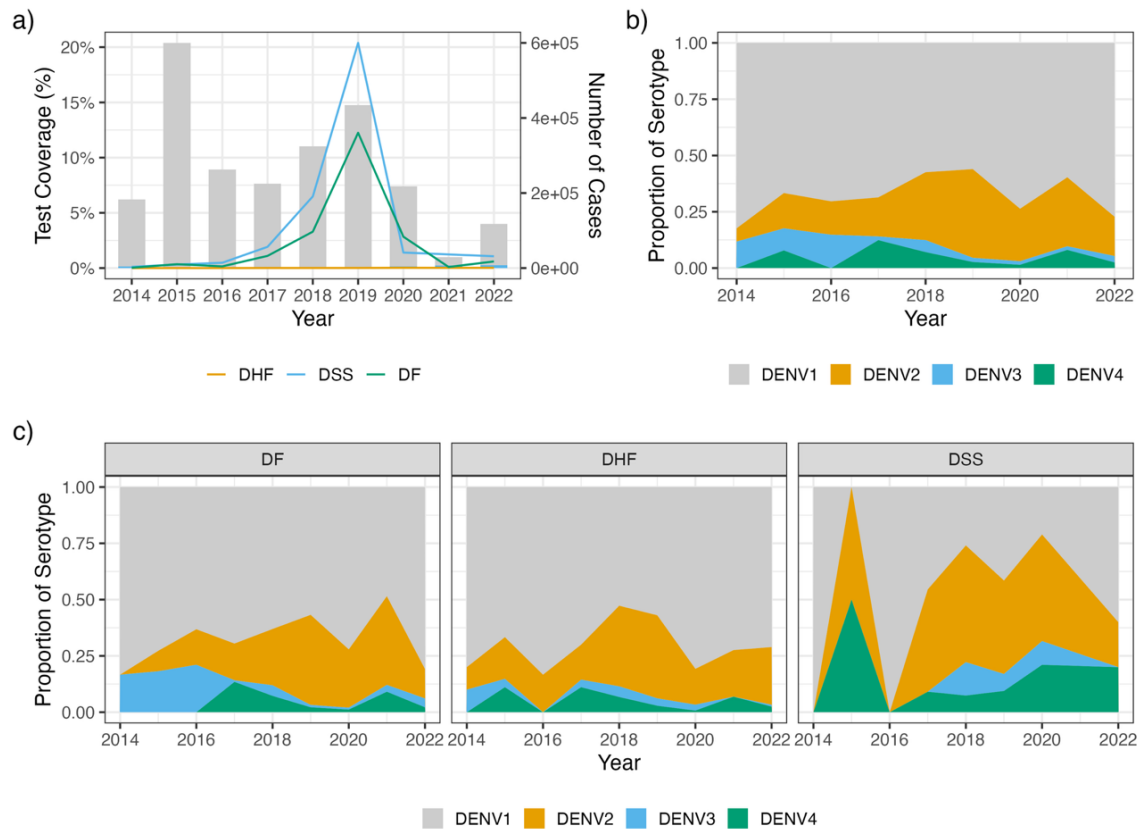
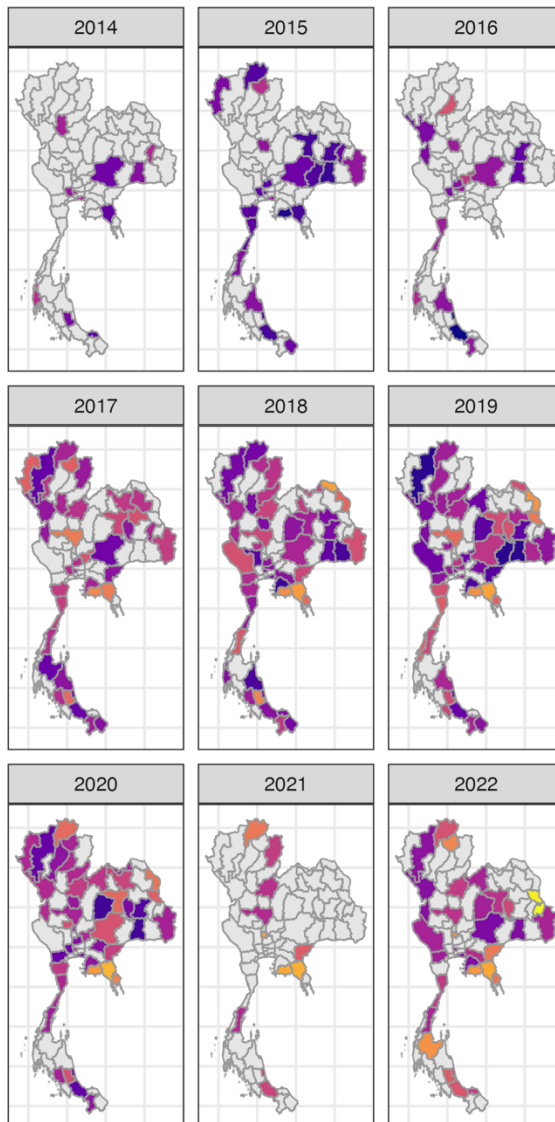


Figure 3.7 a) Proportion of dengue cases by type, tested for serotype. b) Number of cases by serotype. c) Number of serotypes by type of dengue disease.

Figure 3.8 illustrates the provincial distribution of serotype testing in Thailand from 2014 to 2022. Panel a) shows the proportion of dengue cases that were tested for serotype, while panel b) presents the proportion of all serotype tests conducted per province. Each panel is organized by year. In panel a), higher test rates (indicated by yellow shades) are concentrated in a few provinces, while many others had little to no serotype testing. Panel b) reveals that a small number of provinces consistently account for a large share of national testing, suggesting a geographic concentration of testing capacity or activity. Notably, Rayong province appears as the leading contributor to serotype testing from 2017 to 2022. Together, the panels highlight persistent disparities in serotype testing coverage and distribution across provinces and over time.

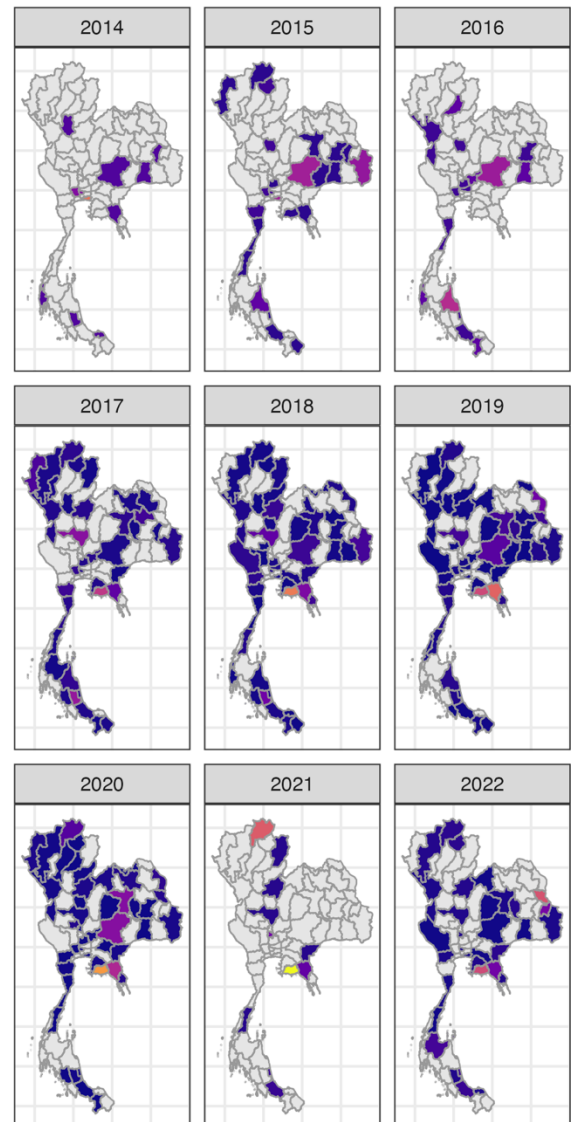
a) Proportion of cases tested for serotype, by province



Serotype test rate (log)

-8 -6 -4 -2

b) proportion of all serotype tests, by province



Proportion of tests

0.1 0.2 0.3 0.4

Figure 3.8 a) Log of proportion of dengue cases tested for serotype by province and year. b) Proportion of serotype cases per province by year.

3.3.3 Wavelet analysis

The wavelet spectrum of weekly dengue time series in Thailand is presented in Figure 3.9 while the average power is presented in Figure 3.10.

In Figure 3.9, there is a high band of power at around the one-year period, which is present every year, as indicated by the black ridge line. This annual periodicity shows the highest average wavelet power (approximately 29) across all time points, Figure 3.10. There is also a high-powered region observed at periods between approximately two and four years, with a period of two years between 2013 and 2016, after which the period increases up to four years, peaking at an average power of approximately 10. These high-power regions are enclosed within the white contours, indicating they are statistically significant at the 0.05 level. Caution must be taken when analysing these results as the observations inside the cone of influence (more greyed-out colour at the edges of the plot) are less reliable due to an edge effect.

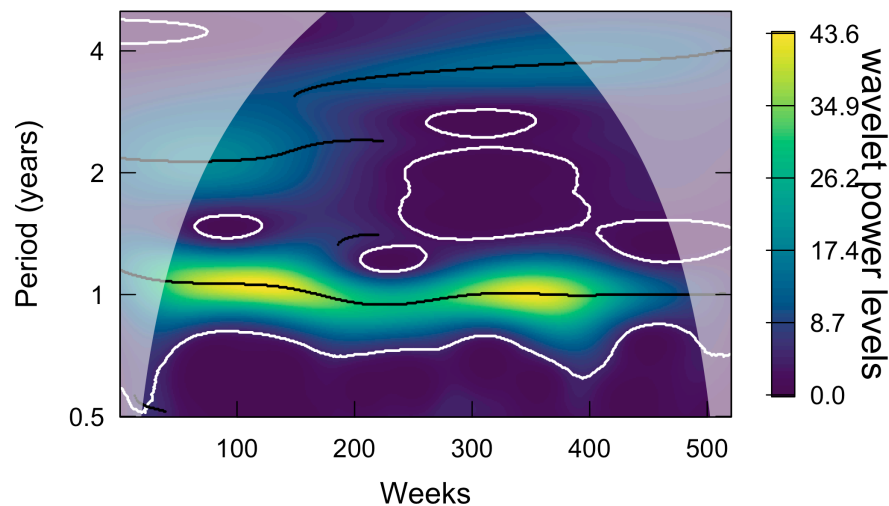


Figure 3.9-Power wavelet analysis of weekly cases in Thailand

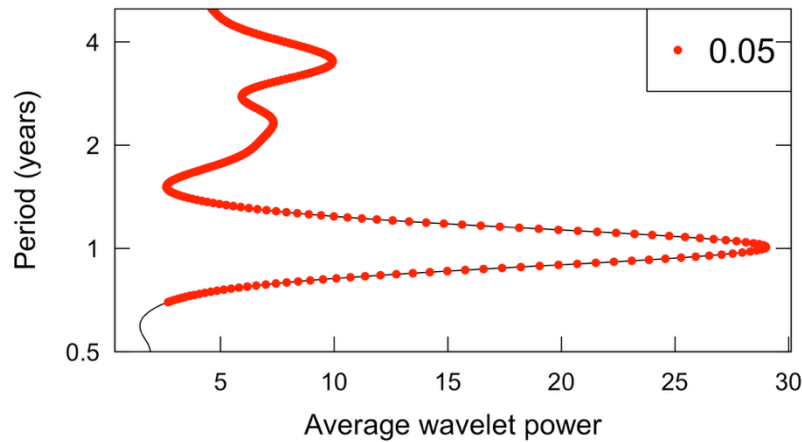


Figure 3.10-National average power of weekly cases in Thailand

3.3.3.1 Reconstructed time-series

The original time series along with reconstructed time series, filtering the annual and multi-annual periods, are plotted in Figure 3.11. As previously described, annual periods are defined as periods between and including 0.5 and up to 1.5 years, and multi-annual periods include periods between and including 1.5 and 5 years. These reconstructed time series are only limited to significant periods and observations outside of the cone of influence. This plot shows the component of incidence where incidence occurs with annual and multi-annual periodicity. In almost all weeks, annual and multi-annual reconstructed time series, the peaks and the troughs align except in year 2014, where the peak of the reconstructed annual time series coincides with the trough of the reconstructed multi-annual series, which may have contributed to the much lower incidence in 2014 compared to 2013.

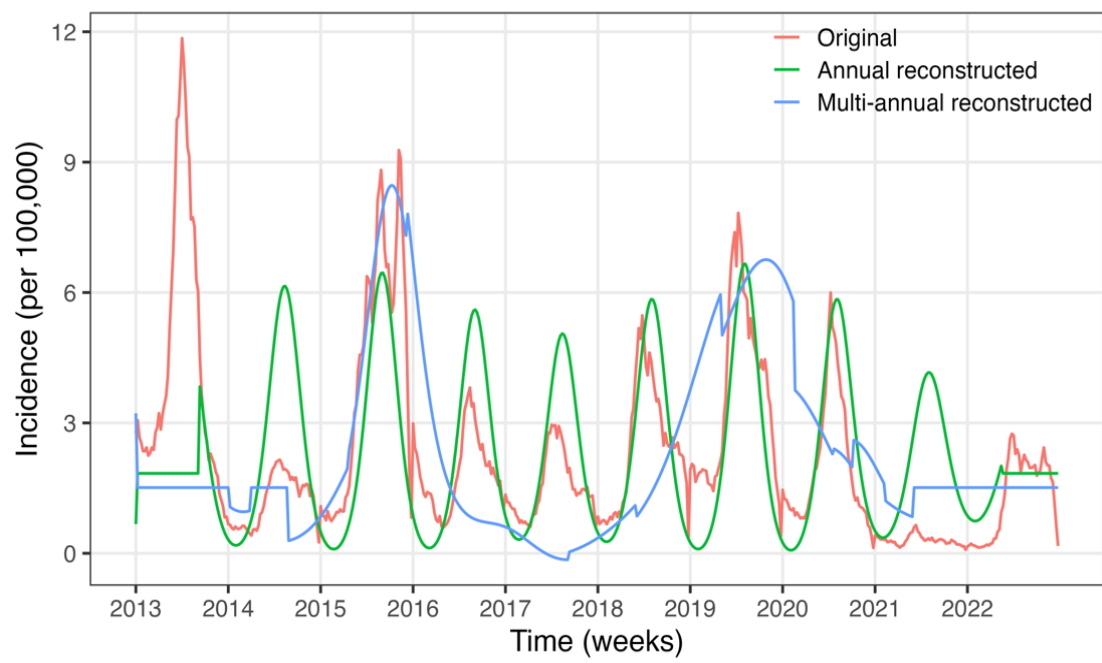


Figure 3.11-Reconstructed time-series, national level, 2013-2022

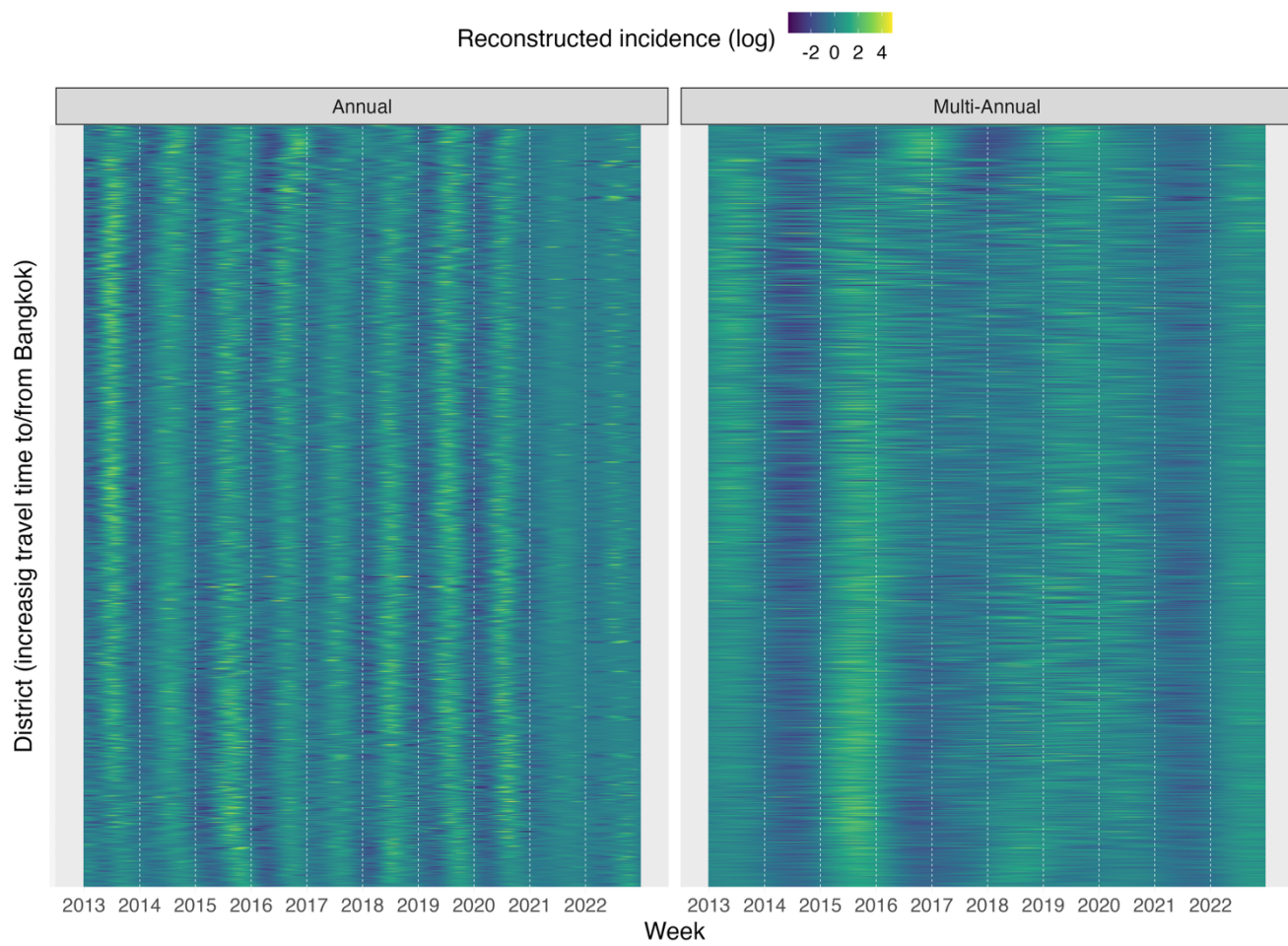


Figure 3.12-Reconstructed time-series, by district, annual and multi-annual periods

3.3.3.2 *Phase difference*

Spatio-temporal pattern of dengue waves at the district level were investigated by calculating the phase difference between each district and the province of Bangkok, using the methodology described in methods. This analysis was completed separately for annual and multi-annual periods and is presented in Figure 3.13. The districts are arranged by increasing travel time from each district to Bangkok. Negative values indicate that the wave occurred first in Bangkok while positive values indicate that the wave was first observed in the district. Grey values indicate weeks where wavelet coherences were not significant.

In the annual period, for significant coherencies, 90.8% of district level phase difference by week were larger than zero, meaning that for those districts and weeks, the wave of cases in those districts occurred before Bangkok.

This trend is quite different for the waves with multi-annual periods. About 57.9% of phase differences of districts by week were more than zero. As the time of travel to Bangkok increases, the phase difference becomes more negative, with cases first occurring in Bangkok.

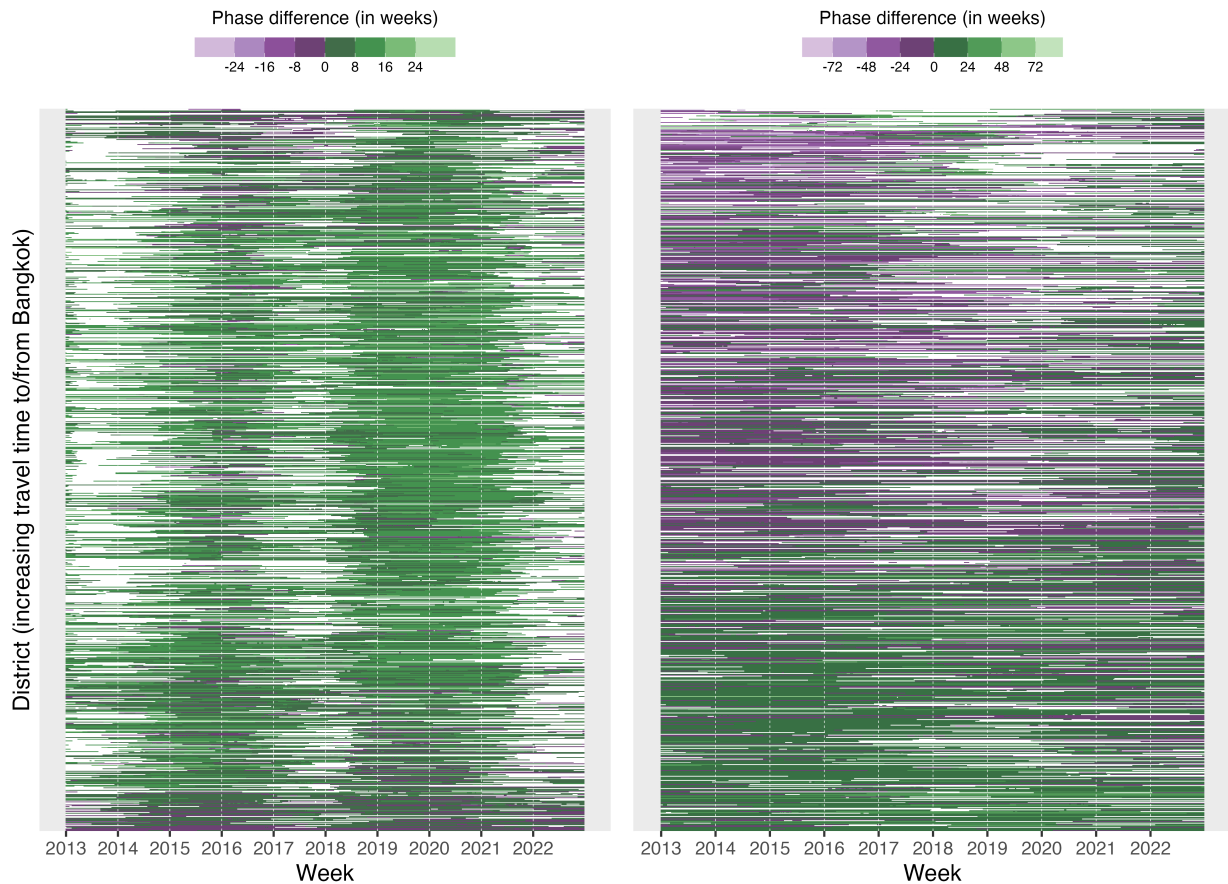


Figure 3.13-Phase difference, by district, by annual and multi-annual periods

3.3.4 Phylogenetic analysis

3.3.4.1 Data availability

From 2010 to 2022, 413 samples from Thailand were included in the analysis. Due to limited geographical information associated with each sample, this analysis was conducted at the national level.

All four serotypes were present every year except in year 2019 when there was no serotype 3 observed, Figure 3.14, with highest number of sequences available for serotype 1 (n=149), followed by serotype 2 (n=138), serotype 3 (n=75), and serotype 4 (n=51).



Figure 3.14-Number of serotypes sampled in Thailand, 2013 - 2020

3.3.5 Phylogenetic trees

Phylogenetic trees are presented in Figure 3.15 and show the circulation of DENV in Thailand and the South East Asian (SEA) region from 1985 to 2020, the year of the most recent available sample. Trees are truncated to start from 1985 to focus on the more recent patterns. The trees show multiple independent introductions of each serotype into Thailand over this period, as indicated by the Thai sequences (coloured branches) being interspersed among the sequences from SEA region (grey branches). This pattern shows ongoing viral exchange between Thailand and countries in the region, rather than evolution from a single lineage. This also shows Thailand's role as both a recipient and source of regional viral dispersal.

Temporally, the figure also shows how the dominance of each serotype changes. Dengue 1 shows distinct clusters of Thai sequences particularly between 2000 and

2005, and 2010 to 2015. Dengue 2 displays relatively continuous circulation with Thai sequences distributed throughout the tree. Dengue 3 displays more sporadic clustering of Thai sequences, consistent with less frequent circulation in the region. Dengue 4 shows increased diversity in more recent years, 2010 to 2020, with multiple Thai lineages co-circulating.

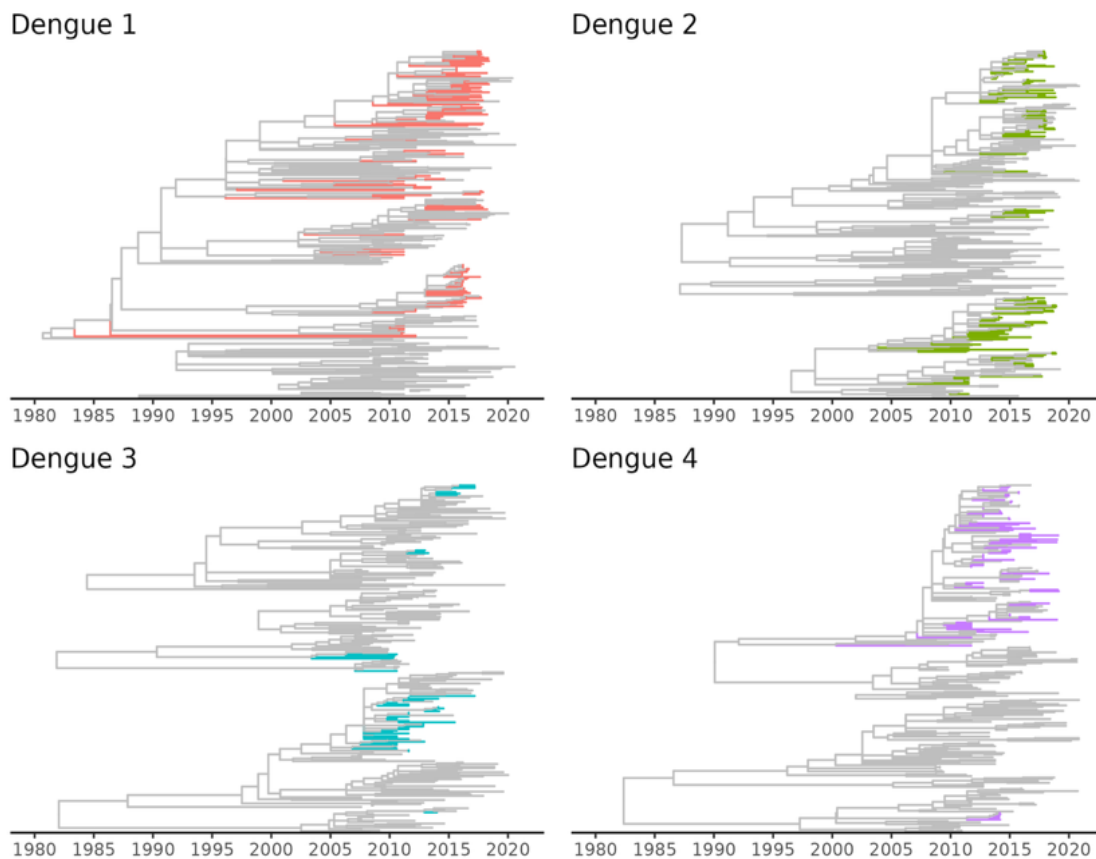


Figure 3.15-Phylogenetic trees of dengue serotypes in Thailand, 1985 - 2020. Coloured branches represent samples collected in Thailand while grey branches represent samples from other South Asian countries.

3.3.6 Genetic diversity

Genotypes overall and within each serotype are presented in Figure 3.16. As illustrated in panel a), between 2013 and 2020, ten genotypes were identified. Overall genetic diversity has sporadic high points, such as in 2013, 2015 and 2016,

with reduction later. Eight genotypes were circulating in 2013, compared to 2020 where six genotypes were circulating.

As shown in panel b), three genotypes were identified within serotypes 1 and 3 while two genotypes were identified within serotypes 2 and 4.

Within serotype 1, genotype 2 was most common, except an increase of an unassigned genotype in 2015. Within serotype 2, genotype Asian I(AI) was first observed as the only genotype in 2013 with presence of genotype Cosmopolitan (C) in 2014 and increasing in proportion over time. Within serotype 3, genotype III was most common, with genotype I increasing in proportion from 2014 onwards. Within serotype 4, genotype I was the most common except in 2013 and 2015 with genotype II being most common. By the year 2020, only serotype 2 shows some meaningful co-circulation of genotypes, while in all other serotypes, a single genotype seems to be dominant.

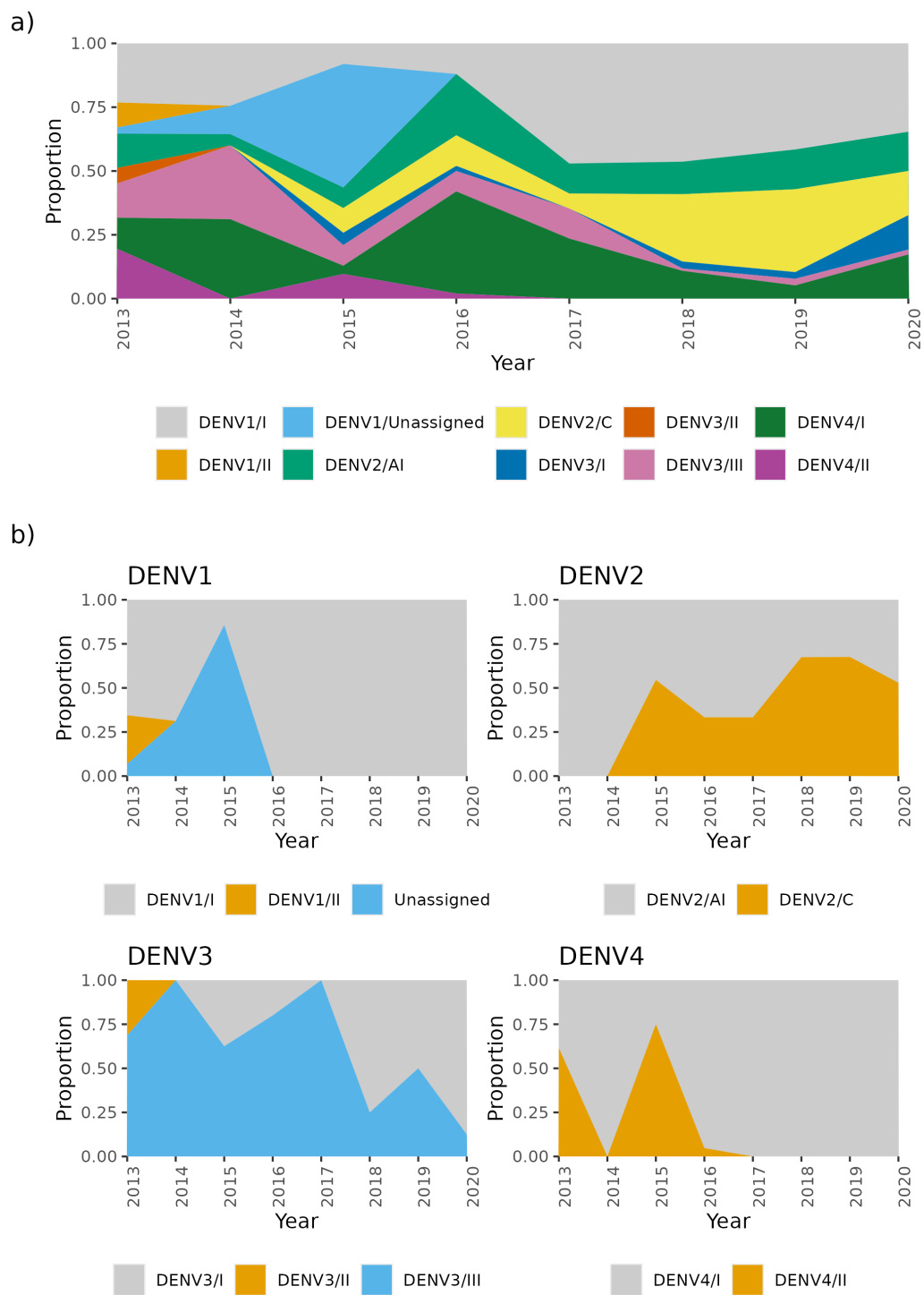


Figure 3.16-a) Proportion of genotypes by serotype and b) overall.

Number of lineages and number of samples per time per serotype is presented in Figure 3.17. Around 2013 for DENV3, 2016 for DENV4, and 2018 for DENV1 and DENV4, lineage diversity decreases suddenly even though the number of samples continue to increase.

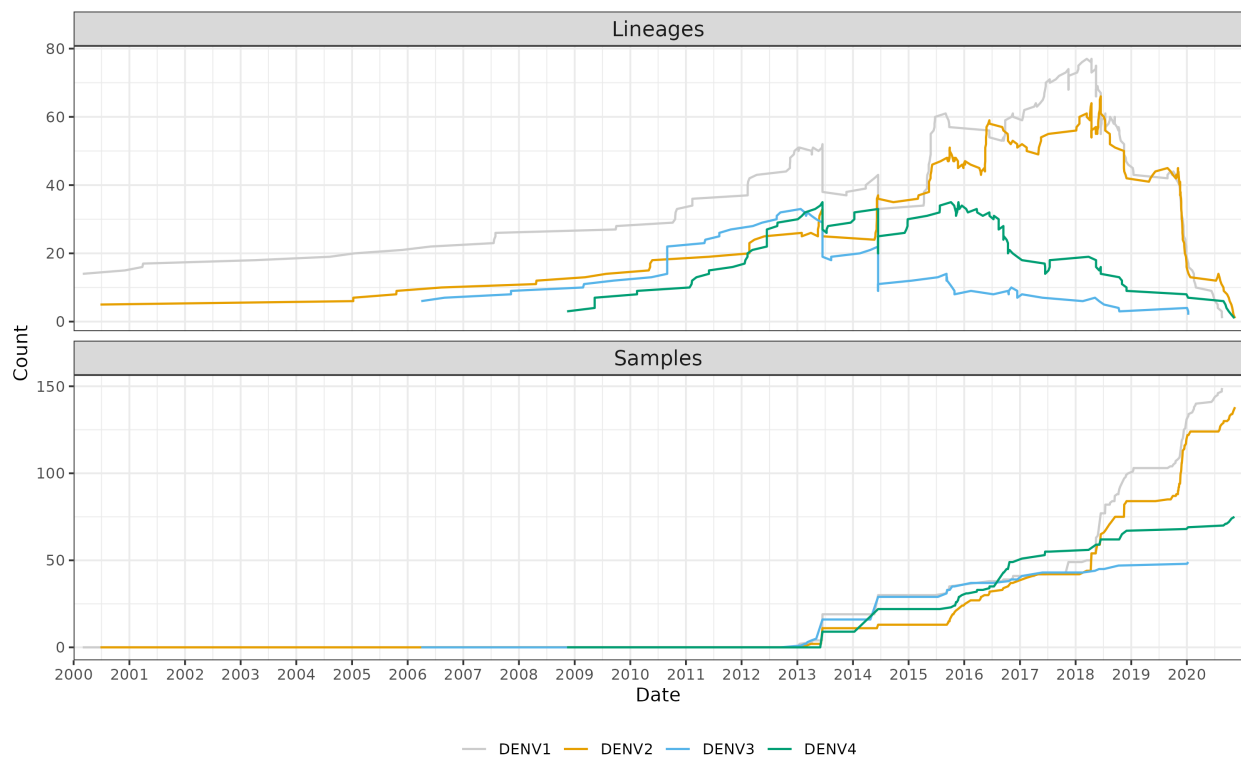


Figure 3.17-Number of lineages and samples over time, by serotype

To assess whether the change in number of lineages is related to number of samples, Kendall's correlation was performed and the results are presented in

. For DENV1 and DENV2, a weakly positive but significant correlation was observed, suggesting that as number of samples increased, the number of

detected lineages also increased. In contrast, DENV3 showed a moderate negative correlation while DENV4 showed a weak negative correlation. Both were significant. Negative correlations indicate that more samples do not correspond to more lineages, which could be a result of a dominate lineage during the sampling period or uneven sampling.

Serotype	Kendall's tau	p-value
DENV1	0.105	< 0.05
DENV2	0.117	< 0.05
DENV3	-0.419	< 0.001
DENV4	-0.164	< 0.01

Table 3.1-Kendall's tau between number of lineages and number of samples, by serotype.

Genetic diversity results and number of sequences used for calculations are presented in Figure 3.18. Kendall's tau correlation was used to test whether nucleotide diversity increased with the number of sequences. Results are presented in Table 3.2. For serotype 2, there was a significant moderate negative correlation, consistent with the plot. Diversity was highest when sequence counts were lower (e.g. 2014, 2015, 2020), suggesting co-circulation of multiple lineages and later dominance of a single strain (2018, 2019). For serotypes 1, 3, and 4, correlations were not statistically significant. For serotype 1, the plot suggests a possibility of lineage dominance during years of high sampling (2018, 2019). Serotype 4 showed a moderate positive trend, but this was not significant and likely a result of small sample size.

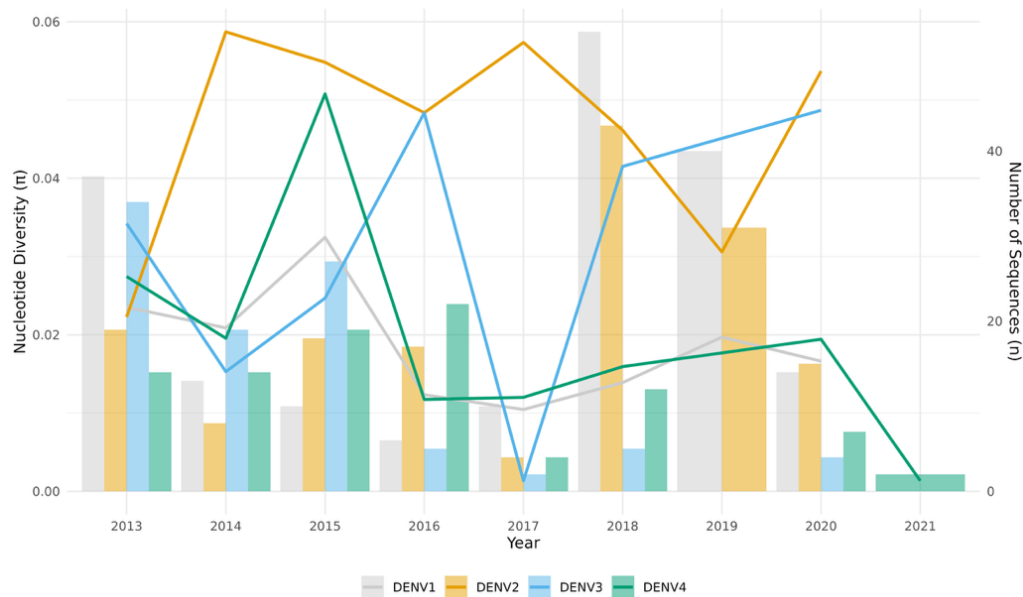


Figure 3.18-Nucleotide diversity (π) and number of sequences (n) used to calculate diversity, by serotype and year. Lines correspond to nucleotide diversity (92π) (left axis) and bars correspond to number of sequences (n) (right axis).

Serotype	Kendall's Tau	p-value
DENV1	0.0364	> 0.05
DENV2	-0.571	< 0.05
DENV3	-0.0976	> 0.05
DENV4	0.473	> 0.05

Table 3.2-Kendall's tau between genetic diversity and number of samples.

3.4 Discussion

Reviewing dengue incidence data at the provincial scale, Figure 3.5, revealed considerable heterogeneity in the incidence pattern between provinces over the years. For example, in 2016, northwestern provinces experienced peak incidence

between April and August while in southern provinces, higher incidence was observed between December 2016 and March 2017. At the district level, a clearer seasonal pattern can be observed with most area experiencing peak between May and July.

Serotype data provided by MoPH showed that DENV1 is the dominant circulating serotype, though data was limited. As shown in Figure 3.8, there is a substantial disparity across the country, both considering the proportion of cases that were tested for serotype within provinces and the distribution of serotypes across provinces. There are also many provinces without any serotype testing at all, making it difficult to understand serotype diversity across the country. More routine and equitable serotype surveillance is needed to better capture the spatial and temporal dynamics of circulating viral diversity.

Wavelet analysis provided insights on the seasonal and interannual patterns of dengue across Thailand. Seasonal cycle was identified as around between 6 months to 18 months with the strongest power at annual frequency while the interannual pattern was observed at between 1.5 years and 5 years, with peak of power at around 3.5 years followed by a smaller peak at approximately 2.5 years, Figure 3.10. These findings highlight the multi-temporal scale of dengue in Thailand.

To expand on the previous research by Cummings et al.[279] that identified Bangkok as the epidemic start, I compared the phase difference at the seasonal and interannual scales between Bangkok and other districts, and their travel distance. At the seasonal scale, peak in most of the districts occurred before Bangkok while at the interannual scale, the peak mostly occurred in Bangkok first.

These seemingly contradictory findings may reflect the different factors that drive seasonal and interannual cycles. Seasonal epidemics are mostly driven by climate factors such as temperature and rainfall, affecting local vector dynamics which in turn affect local dengue epidemics. Vector control in Thailand is also implemented by local level public health officials and may impact local dynamics. Interannual cycles on the hand, may be driven by human mobility, where virus is introduced to areas further away from large centres such as Bangkok over a longer period. Kang et al.[175] found that using mobile data in Thailand, provinces that were more connected by mobility were more correlated in their dengue incidence than less connected provinces. Interannual cycles may also arise due to accumulation of susceptible populations over a period of few years.

Analysis of lineage and genetic diversity showed both overlapping and divergent trends across serotypes. For serotype 1 and 2, a weak but significant positive correlation between number of samples and detected lineages was observed, suggesting that increased sampling led to the detection of more lineages. In contrast, serotype 3 and 4 showed significant negative correlations, possibly due to the dominance of a single lineage or uneven sampling. When assessing genetic diversity, only serotype 2 showed a statistically significant negative correlation with sequence count, with diversity highest during periods of lower sampling, consistent with co-circulation of multiple lineages followed by dominance of a single strain. For serotype 1, 3, and 4, no significant relationships were found. Lineage diversity and genetic diversity did not always align. A serotype may exhibit multiple closely related lineages with low nucleotide diversity, or few lineages with high genetic divergence. A serotype may display high lineage diversity with low genetic diversity if lineages are closely related or show high genetic diversity despite few lineages if those represent divergent introductions. Lineage diversity is an

epidemiological phenomenon driven by new importations and the propagation of these importations dependent on the local context[61, 302] while genetic diversity is either affected by natural selection, new importations or genetic drift because of chance[303, 304]. These results must be reviewed with caution as the genomic data was not sampled to be nationally representative. They may also not represent the true proportion of serotypes.

This chapter contributes to a more nuanced understanding of dengue dynamics in Thailand between 2013 and 2022, highlighting both spatial variation in epidemic timing and temporal complexity introduced by different cycles. Since annual seasonality was found to be significant, it will be considered in the subsequent modelling chapters through seasonal model with a 52-week period. The spatial variation in epidemic dynamics motivated fitting case data at the province level while modelling seasonality at the region level to balance granularity with computational efficiency. This chapter also highlights some of the shortcomings of the available data, including limited serotype and genomic surveillance data, that constrain my ability to interpret the role of immunity and evolutionary forces shaping dengue spread. Despite these limitations, genetic diversity of each serotype per year was included as a covariate to capture immune dynamics, even if imperfectly. The temporal complexity observed in this chapter suggested that immunity plays an important role in shaping epidemic patterns. The lack of comprehensive genomic surveillance prevented modelling of evolutionary dynamics directly. However, genetic diversity measures were used an indirect indicator of viral population dynamics, which may relate to population immunity levels in subsequent modelling chapters.

4 Baseline Model Analysis

4.1 Introduction

Previous studies have shown that simple models such as autoregressive model and those based on seasonality can sometime outperform more complex models when forecasting infectious diseases including dengue[194]. This is particularly true in context when there are no unexpected changes in circumstances. In particular, autoregressive models have performed well in short-term forecasts[192, 194, 195] while seasonal models can perform better in the long term, predicting the expected seasonal changes in cases[196, 305, 306].

Simpler models also have practical advantages. They are easier to conceptualize, faster to fit and run, more intuitive to interpret, and more amenable to wide-scale implementation. They are also easier to adopt and run by public health authorities, particularly in more resource-limited settings where technical capacity can be severely constrained. This is the case for the Thai MoPH which prefers to have an in-house forecasting platform. These characteristics are important for a rapid and meaningful public health response. Finding and processing additional variables can also be timely and resource intensive, leading to delays in predictions. In some cases, even though a more complex model might offer slightly better performance, for example as measure by lower mean absolute percentage error (MAPE)[192], the difficulty in fitting and running the model may not necessarily be worth the effort.

Forecasting models for dengue generally focus on predicting incidence or outbreaks, each providing different but complementary results. Predicting outbreaks is important because often, public health interventions such as vector

control initiatives and public messaging are triggered by the declaration of outbreak by public health authorities[307-309]. Binary classification is also easier for communication to stakeholders and the public[310]. Predicting incidence on the other hand, can provide a more nuanced view of the situation. With specific numbers of cases predicted, health facilities and hospitals can better plan to ensure they have sufficient tests, medication, hospital beds, and staff[311, 312].

This chapter presents the development and analysis of two baseline models used to forecast dengue incidence and outbreaks in Thailand. These models will provide a baseline for evaluating the performance of more complex models introduced in later chapters.

The first model is a seasonal naïve approach that uses historical patterns in dengue transmission, while the second model, builds upon the first with an addition of a seasonal component. Understanding the predictive accuracy characteristics of these two models will help to better understand the subsequent improvements that may be observed with more complex models.

To formally assess the performance of the models, the forecasted incidence outputs will be evaluated and compared to each other using a range of performance metrics, including Continuous Ranked Probability Score (CRPS), calibration, sharpness, and bias. Each metric measures a different aspect of the predictive performance of the model and should be reviewed together.

CRPS measures the accuracy of the cumulative distribution function (CDF) compared to a true value[213, 313]. In other words, CRPS can be thought of as the distance between a predicted distribution and the actual outcome. CRPS is

considered a strictly proper scoring tool. This means that the score is uniquely minimized when the forecasted probability distribution matches the true distribution of the observed outcomes.

Calibration is used to assess whether the forecast distribution is consistent with the data generating distribution[271, 314]. This is measured by assessing an empirical coverage of specific prediction intervals. Sharpness is measured as the normalized median of the absolute deviation (NMAD) from the median of the predictive samples[271, 314]. Bias is evaluated as the average proportion of posterior draws that are smaller than the observed value[271, 314].

In addition to forecasting incidence, it is important to assess each model's ability to forecast outbreaks. When concerned with a model that can assist in preparation and planning of prevention and intervention strategies led by government-led agencies, understanding the probability of an outbreak can provide a useful threshold for action. Outbreak detection ability of the models is assessed using area under the curve (AUC) and the Brier Score, which measures the accuracy of exceeding a predefined threshold.

4.2 Methods

4.2.1 Data

This chapter used dengue incidence data from the Thai MoPH, as described in Chapter 3.

4.2.2 Models

To forecast dengue cases without covariates, two models were used and compared to each other for their forecasting ability. Forecast horizons of 1 to 24 weeks were chosen because they are most relevant to public health practice.

4.2.2.1 Model structure

Let d index the districts of Thailand.

Let t index the week.

I modelled the number of cases in district d during week t using a zero-inflated Poisson likelihood:

$$y_{d,t} \sim ZIP(\lambda_{d,t}, \pi_{d,t}) \quad (2)$$

This meant that the observed number of cases $y_{d,t}$ in district d at week t followed a zero-inflated Poisson distribution with rate parameter $\lambda_{d,t}$ and zero-inflation probability $\pi_{d,t}$.

The two models investigated are as follows:

Model 1:

$$\log(\lambda_{d,t}) = \alpha + IW_{t,p} + u_d + v_d + \log(pop_d) \quad (3)$$

α was the intercept, $W_{t,p}$ considered effect of week(t) at the province(p) level, using a first order auto-regressive model (AR1). $u_{d,t}$ and $v_{d,t}$ were the spatially structured (conditional autoregressive, u) and unstructured (i.i.d; v) district level random effects, jointly specified with Besag-York-Mollie (BYM2) model, and $\log(pop_d)$ was the log of the population of each district.

Model 2:

$$\log(\lambda_{d,t}) = \alpha + W_{t,p} + Q_{t,r} + u_d + v_d + \log(pop_d) \quad (4)$$

In addition to elements in Model 1, Model 2 included $Q_{t,r}$ which accounted for seasonality effect at the weekly(t) at the region(r) level.

The models were fit using INLA[221, 315].

4.2.2.2 Priors

4.2.2.2.1 Auto-regressive 1 (AR(1)) model

Temporal dependence in the latent process was modelled using a first-order autoregressive (AR(1)) structure

$$x_t = \rho x_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \tau^{-1}) \quad (5)$$

where $\rho \in (-1,1)$ was the autocorrelation parameter and τ the precision.

Penalized complexity (PC) priors[316] were assigned to both hyperparameters. For the precision parameter τ , I used a PC prior defined as $PC(\sigma = 0.5, \alpha = 0.01)$, which corresponded to a 1% prior probability that the marginal standard deviation exceeds 0.5. For the autocorrelation parameter ρ , a $pc.cor0(0.5, 0.5)$ prior was used, implying a 50% chance that the absolute value of ρ exceeds 0.5.

4.2.2.2.2 Seasonality model

Seasonality was modelled using the INLA's seasonal latent effect with a 52 week cycle [215, 317]. Let $\mathbf{x} = (x_1, \dots, x_n)$ denote the latent seasonal effect defined over n weeks, where x_i represented the seasonal effect at week i . The seasonal period was $m = 52$.

The INLA seasonal model was defined by assuming that the total seasonal effect over one complete cycle followed a Gaussian distribution with common precision τ . Specifically, the model assumed that seasonal sums of length m , are independent Gaussian random variables with mean zero and precision τ ,

$$x_i + x_{i+1} + \dots + x_{i+m-1} \sim \mathcal{N}(0, \tau^{-1}) \quad (6)$$

The density for $\mathbf{x}$ was,

$$\pi(\mathbf{x} | \tau) \propto \tau^{\frac{n-m+1}{2}} \exp \left\{ -\frac{\tau}{2} \sum_{i=1}^{n-m+1} (x_i + x_{i+1} + \dots + x_{i+m-1})^2 \right\} \quad (7)$$

A default prior assigned by INLA was used:

$$\text{Log}(\tau) \sim \text{Log} - \text{Gamma}(1, 0.00005) \quad (8)$$

4.2.2.2.3 Besag, York, and Mollié model 2 (BYM(2)) model

In the BYM(2) model, $\mathbf{b}$ was the spatial random effect which combined structured ($\mathbf{u}_*$) and unstructured ($\mathbf{v}_*$) random effects[219, 316].

$$\mathbf{b} = \frac{1}{\sqrt{\tau_b}} (\sqrt{1 - \phi} \mathbf{v}_* + \sqrt{\phi} \mathbf{u}_*) \quad (9)$$

Precision parameter $\tau_b > 0$ controlled the marginal variance contribution of the weighted sum of $\mathbf{v}_*$ and $\mathbf{u}_*$ controlling the strength of the spatial effect. The mixing parameter $0 \leq \phi \leq 1$ measured the proportion of the marginal variance of the spatial effect explained by the structured effect $\mathbf{u}_*$.

The PC prior on τ_b was $pc.prec(U = 0.1, \alpha = 0.01)$. This prior specified a low variance, meaning that low spatial variation was favoured unless otherwise suggested by data. The prior for the mixing parameter, ϕ , was $pc(U = 0.5, \alpha = 0.5)$, meaning that the model did not favour structured or unstructured random effects.

4.2.3 Forecast evaluation

Forecasts were made every three weeks, for 24 weeks of dengue incidence at week t in districts $d = 1, \dots, 928$. Let $N_{d,t}$ represent the number of new cases and $\hat{N}_{d,t+1}$ to $\hat{N}_{d,t+24}$ denote the three-weekly, 24-week ahead forecasts. Model based forecasts were generated for weeks between 29 December 2013 and 31 December 2022, starting on Sundays.

Forecasts were made using all data from 1 January 2013, until immediately up to the start of the week, t , before the forecast started. The training data included all historical data, (i.e. $N_{d,1} \dots N_{d,t}$) to train models, M , to generate probabilistic forecasts, $\hat{N}_{d,t+k}^{m,r}$, where $\hat{N}$ was the forecasted number of new cases where r denotes the draws from the posterior distribution $F_{d,t}^m$ of the model m .

The probabilistic forecasts, $\hat{N}_{d,t+k}^{m,r}$, were then log transformed, and compared to log transformed observed data, $N_{d,t}$. Continuous ranked probability scores (CRPS) were calculated for each model, by week, t and district, d . CRPS calculations were completed with `scoringutils` package[271].

Additional measures of calibration, sharpness, and bias were also calculated.

Evaluations were made by district, forecast horizon, and target week.

4.2.3.1 Outbreak prediction evaluation

Models were assessed on whether they were able to predict outbreaks. To ensure the utility of this analysis, an outbreak was defined using the criteria of the MoPH i.e. when the number of cases is higher than the median of the past five years for a specific place and time period.

A contingency table was created based on the results of the predicted outbreaks by the models and the observed outbreaks and Area Under the Curve (AUC) was evaluated at forecast horizons of 1, 4, and 24 weeks

Additionally, agreement between the predicted and observed outbreak was calculated at each district and week. If both the predicted cases and the observed cases were above the epidemic threshold, or if they were both below the epidemic threshold, then there was agreement between the model and observed cases. Otherwise, there was no agreement. All agreement values were then averaged at the national level to obtain agreement at the target week. To understand spatial distribution of agreement and by model, the difference between agreement of no-seasonality and seasonality models were computed for each district at forecast horizons of 1, 4, and 24 weeks.

For each district and week, outbreak thresholds were calculated and compared to posterior predictions. The probability of cases exceeding the threshold was then calculated. Using these probabilities, Brier scores were calculated for each model and by district and forecast horizon. Brier score is a proper score and is measure as the mean squared difference between the probability, p_i , and the observed outbreak, Y_i (0, 1)[318]. Smaller Brier score indicates a smaller difference between the forecasted probability and the observed outbreak which indicates a model with greater accuracy. A Brier score of 0 indicates perfect accuracy while a score of 1 indicates perfect inaccuracy.

$$Brier\ Score = \frac{1}{N} \sum_{i=1}^N (p_i - Y_i)^2 \quad (10)$$

Bootstrapping using 1000 samples was used to calculate confidence intervals for the difference in brier score between the two models and to identify significance at the 0.05 level.

4.3 Results:

Forecasts of the two models were evaluated using CRPS, sharpness, calibration, and bias.

4.3.1 Forecasting evaluation by forecast horizon

In this section, forecasts are evaluated by forecast horizon, ranging from a horizon of one week, up to 24 weeks. This analysis will help quantify the performance of the model at different lead times.

4.3.1.1 CRPS

Forecasting results of models 1 and 2 which are models without covariates are presented in this section. Model 1 is referred to as the no-seasonality model, while model 2 is referred to as the seasonality model.

CRPS results for each forecast horizon for each district was averaged at the national level and are shown in Figure 4.1. At forecast horizon of one week ahead, the seasonality model had a lower CRPS suggesting that this model performs better than the no-seasonality model. However, at forecast horizon of two weeks and more, up to 24 weeks ahead, the seasonality model had lower CRPS with the gap increasing over length of horizon. A similar trend is found when comparing the distribution of CRPS results at the district level by forecast horizon and model, presented in Figure 4.2.

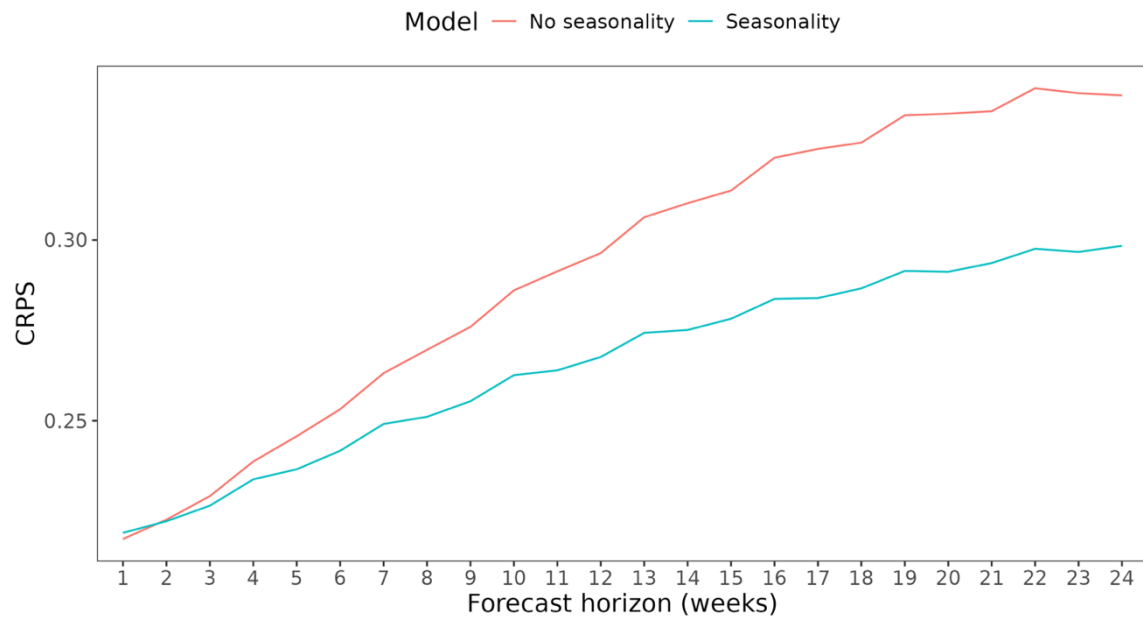


Figure 4.1-Mean national level CRPS by forecast horizon and model. Lower values indicate better performance. Lower CRPS values indicate better forecast skill.

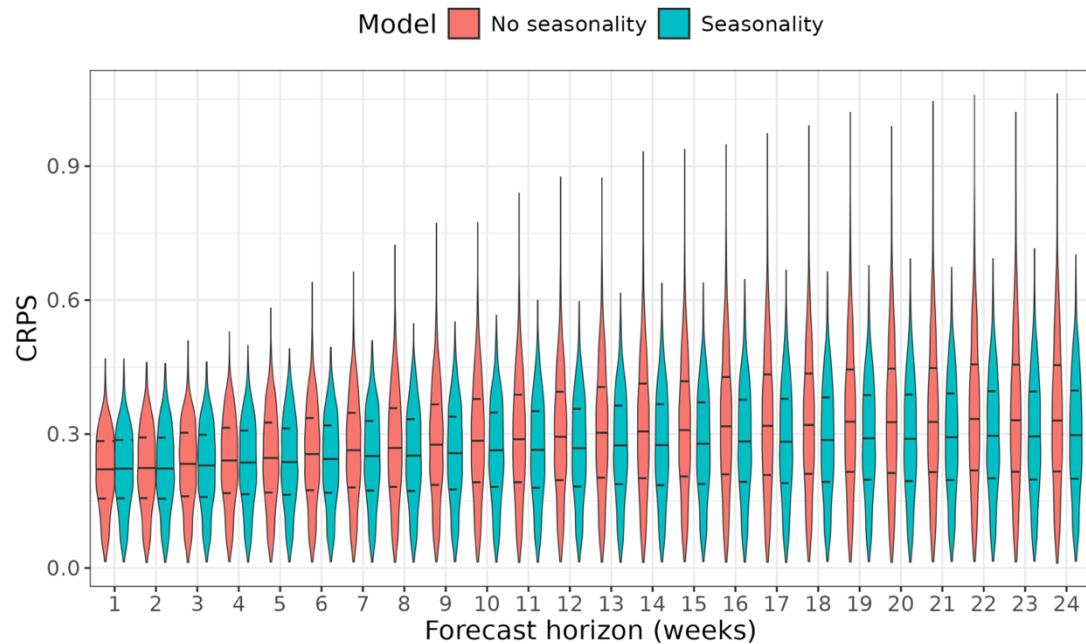


Figure 4.2-Mean district level CRPS by forecast horizon and model. Lower CRPS values indicate better forecast skill.

4.3.1.2 Sharpness

NMAD was used to measure sharpness of forecast outcomes. This measure is based on the model outputs itself and is not a function of comparison to the observed data. As the forecast horizon increased, the NMAD increased, suggesting a decrease in sharpness. There was then a drop in NMAD for the no-seasonality model at around seven weeks of forecast horizon while the NMAD for the seasonality model continued to increase. The additional seasonality term in the seasonality model contributed to wider confidence intervals as the model may be less certain of the seasonality effects at longer horizons. Similar results were observed at the district level, presented in Figure 4.4.

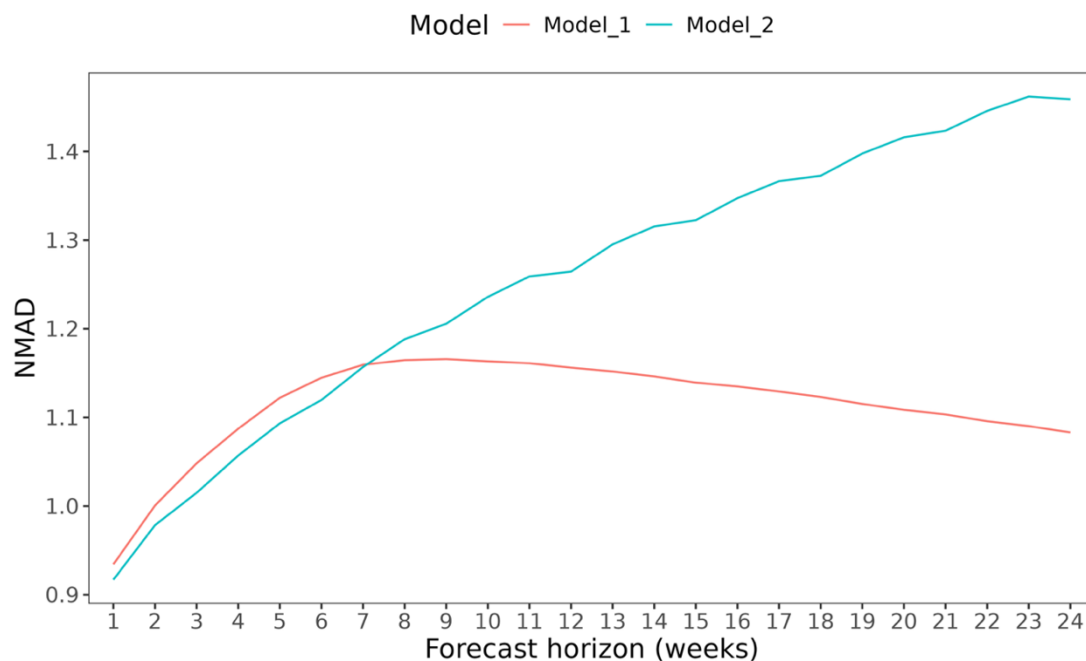


Figure 4.3-Mean national sharpness, as measured by NMAD, by forecast horizon and model. Lower NMAD values indicate narrower (sharper) prediction intervals.

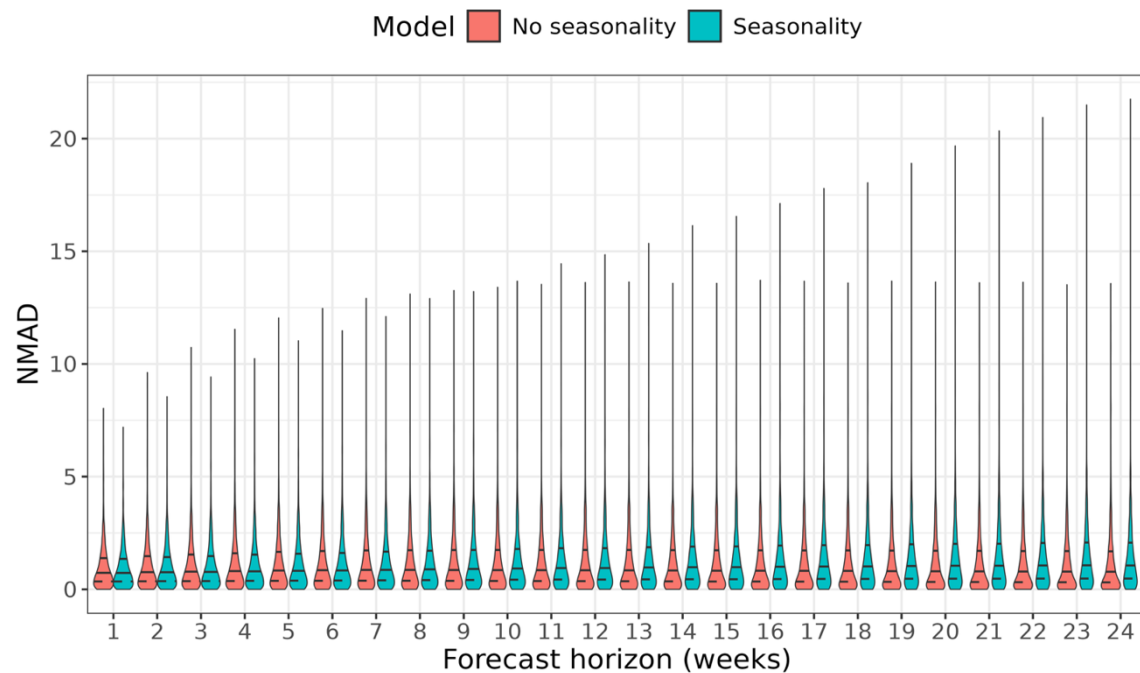


Figure 4.4-Mean district level sharpness, as measured by NMAD, by forecast horizon and model. Lower NMAD values indicate narrower (sharper) prediction intervals.

4.3.1.3 Calibration

Calibration measures the proportion of forecasted events for which the 50% and 90% prediction intervals contained the true value. The 50% prediction interval had lower empirical coverage than the 90% interval, meaning fewer observed values fell within the narrower bounds. This pattern was indeed observed as shown in Figure 4.5 and Figure 4.6. At the national level, the mean of empirical coverage of all forecast events was higher than 50% and 90% (dotted lines), respectively, indicating the model was well calibrated. At the district level, a large proportion of forecasted events included the true value.

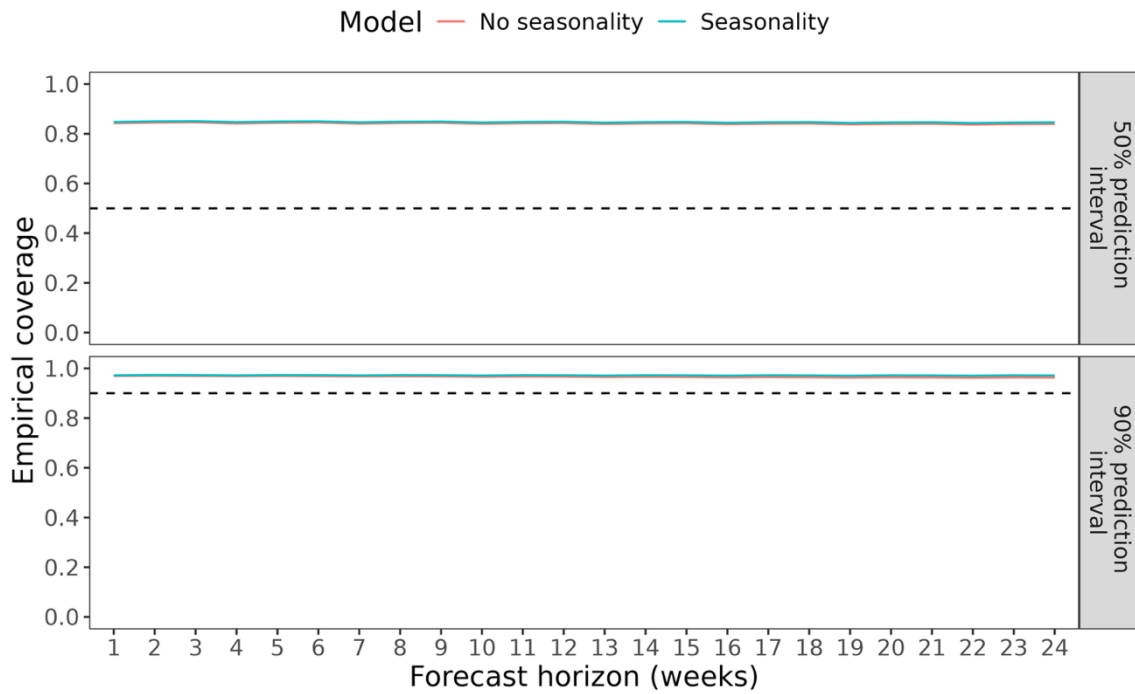


Figure 4.5-Mean national empirical coverage, 50% and 90%, by forecast horizon and model. Higher empirical coverage indicates more observed values fell within the prediction intervals.

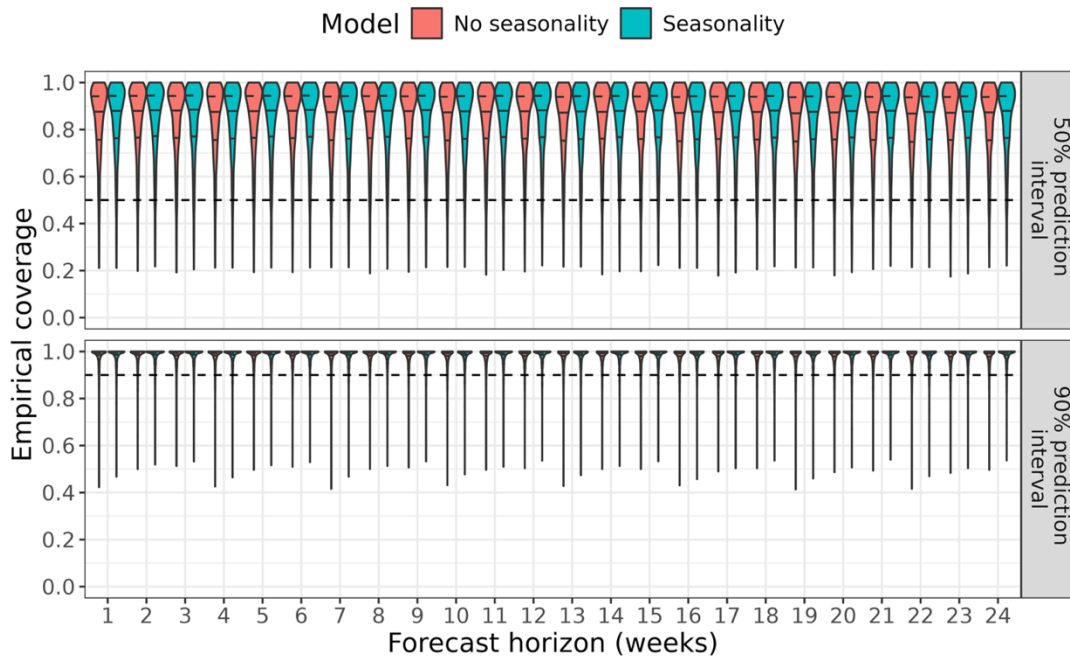


Figure 4.6-Mean district level empirical coverage, 50% and 90%, by forecast horizon and model. Higher empirical coverage indicates more observed values fell within the prediction intervals.

4.3.1.4 Bias

At both the national, Figure 4.7, and district levels, Figure 4.8, bias increased as forecast horizon increased. This increase was at a higher rate for the seasonality model.

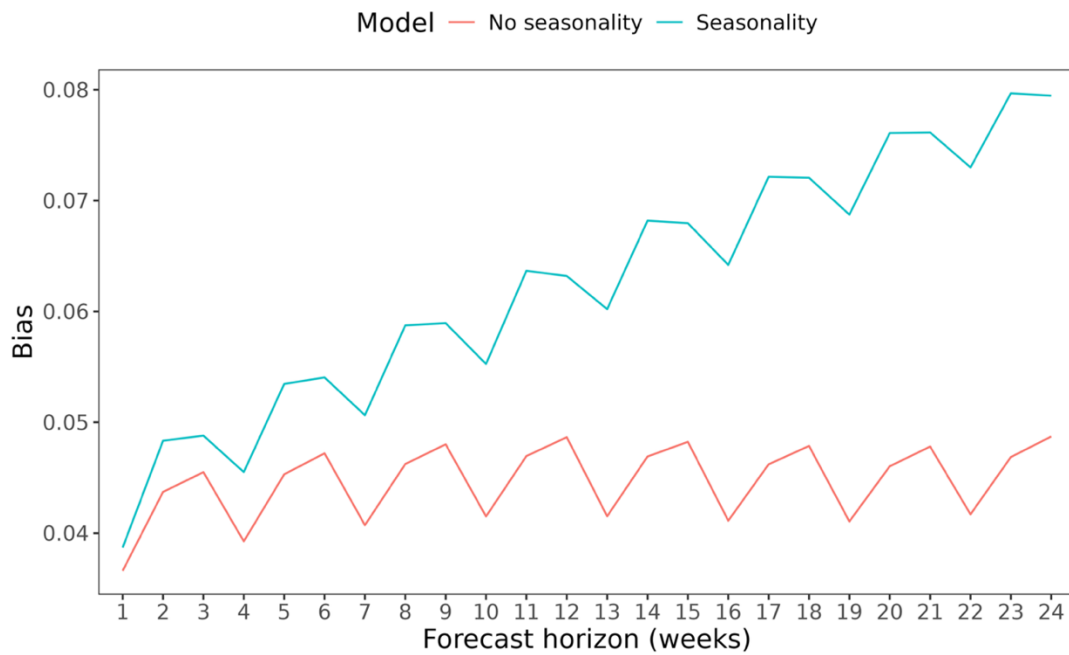


Figure 4.7-Mean national bias, by forecast horizon and model. Lower bias indicates that on average, forecasts are closer to the observed values.

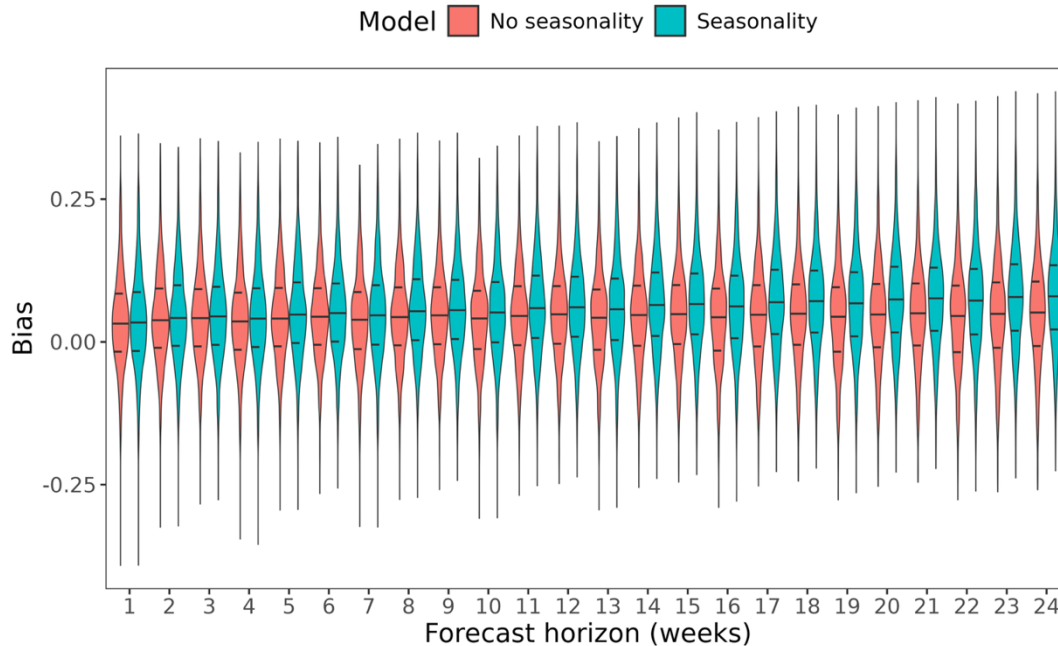


Figure 4.8-Mean district level bias, by forecast horizon and model. Lower bias indicates that on average, forecasts are closer to the observed values.

4.3.2 Forecasting evaluation by target week

In this section, forecasts are evaluated by target week. This analysis will help identify how forecasts may be impacted by a specific week or season.

4.3.2.1 CRPS

CRPS by target week and by forecast horizons of one week, four weeks and 24 weeks is shown in Figure 4.9. The top panel of the image includes the time series of dengue cases, for better understanding of the context of the target week under evaluation.

Both models performed best at one week horizon, as indicated by the smallest CRPS. As forecast horizon increased, results showed higher CRPS and variability. As cases peaked, the performance of both models deteriorated, reflecting the difficulty in predicting variability in cases over time.

The seasonal model generally performed better or equal to the non-seasonal model, especially for longer forecast horizons. The no-seasonality sometimes performed better when there was a large variability between subsequent seasons, for example, as was observed in mid-2016, when there was a large decrease in cases compared to 2015. The seasonal model was expecting a larger number of cases that was observed based on the previous season while the no-seasonality model was better to able to predict the true cases as it was better able to adapt using lagged cases.

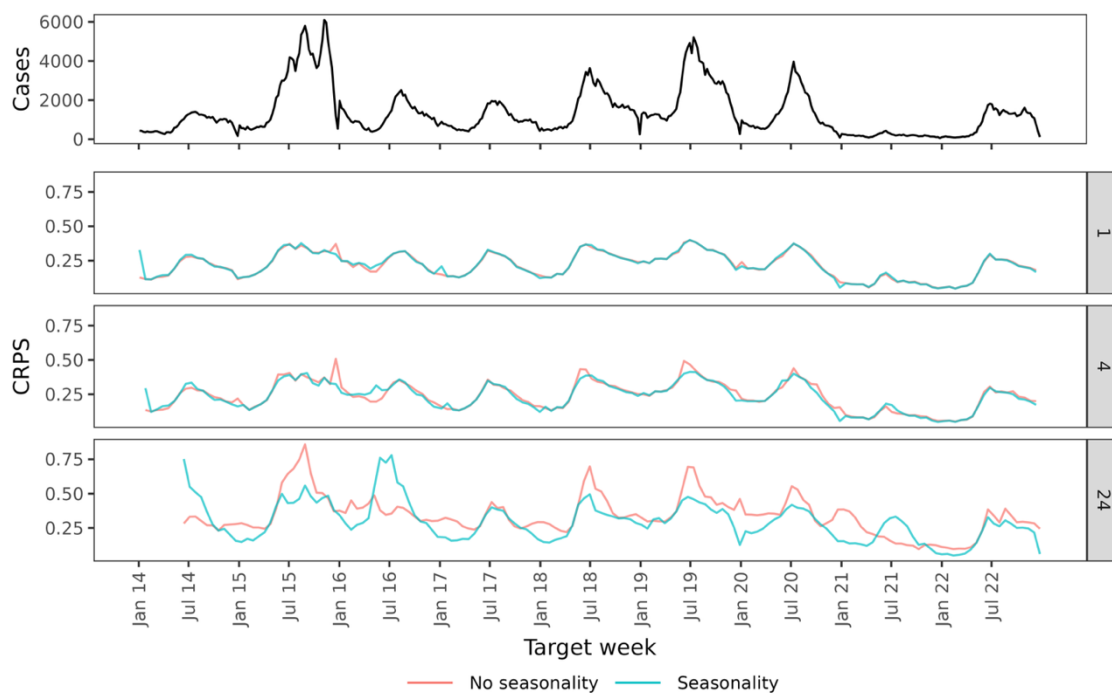


Figure 4.9-Mean national CRPS by target week and forecast horizon. Lower CRPS values indicated better forecasting skill.

4.3.2.2 Sharpness

Sharpness, as measured by NAMD, is shown in Figure 4.10. As forecast horizon increased, the uncertainty of the forecasts increased. Particularly at the 24 weeks horizon, NAMD for the seasonality model was high when there was a change in the seasonal pattern such as in July 2016 when there were fewer cases than the previous season and in July 2019 when there were more cases than the previous season.

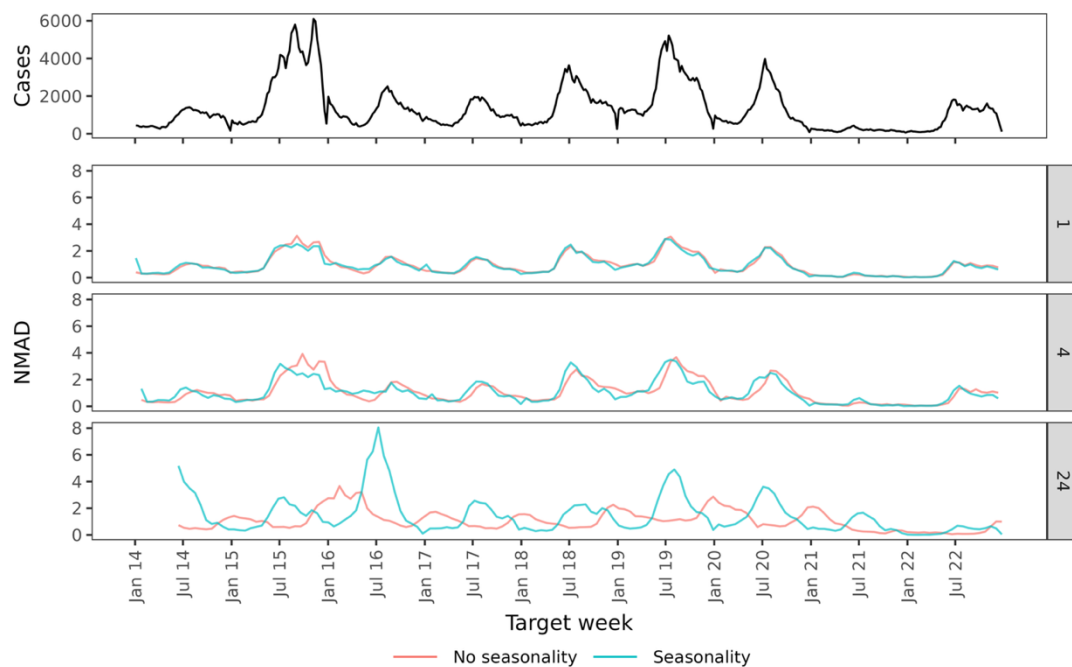


Figure 4.10-Mean national NAMD by target week and forecast horizon. Lower NAMD values indicate narrower (sharper) prediction intervals.

4.3.2.3 Calibration

National empirical coverage is presented in Figure 4.11. As expected, calibration decreased as forecast horizon increased. There was a marked decrease in calibration when there was an increase in cases, such as around July 2015 and July 2019.

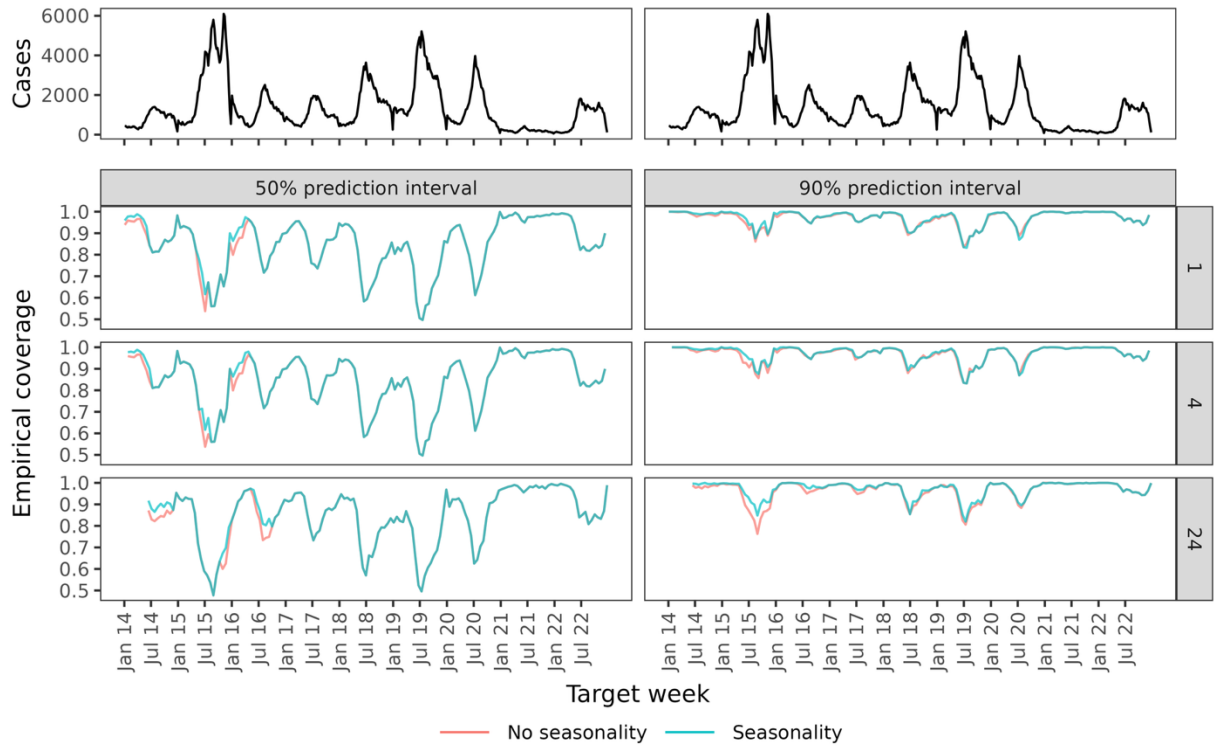


Figure 4.11-Mean national empirical coverage by target week and forecast horizon. Higher empirical coverage indicates more observed values fell within the prediction intervals

4.3.2.4 Bias

As forecast horizon increased, bias deviated more from zero, Figure 4.12. In particular, the performances of the no-seasonality and seasonality models were quite different at the 24 weeks horizon, with the no-seasonality model overall performing worse than the seasonality model.

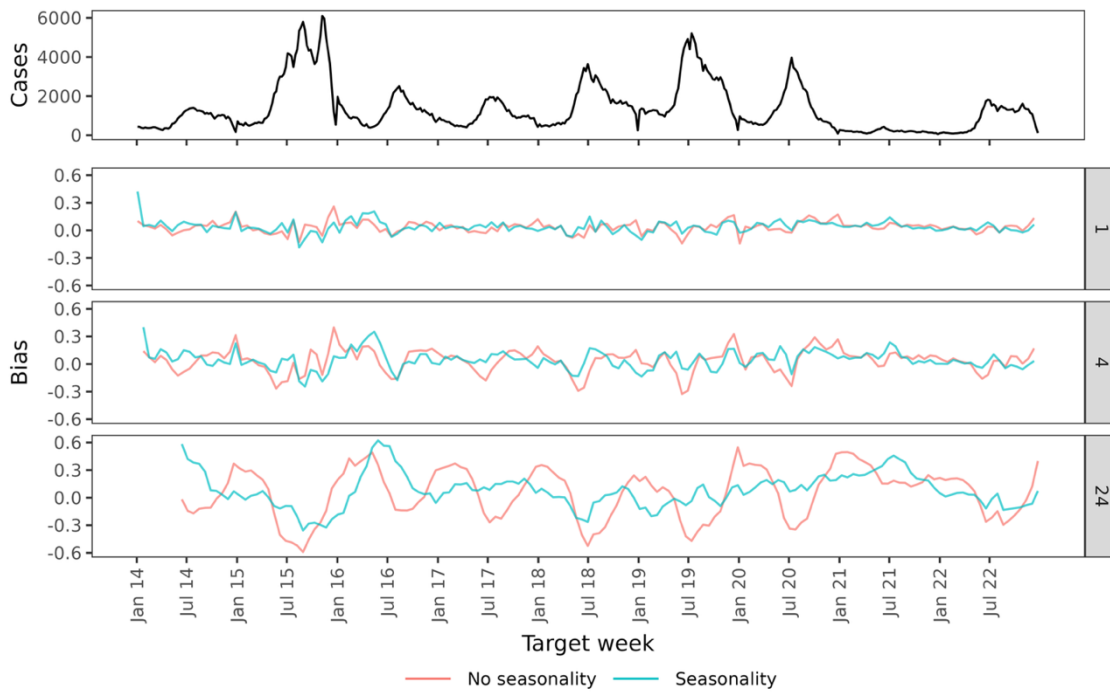


Figure 4.12-National mean bias by target week, model and forecast horizon. Lower bias indicates that on average, forecasts are closer to the observed values.

4.3.3 Forecasting evaluation by space

Performance of forecasts was evaluated over space. In Figure 4.13, district level CRPS results of seasonality and no-seasonality models are plotted against each other with each colour representing regions. As expected, and previously seen, the difference in CRPS between the two models increased as forecast horizon increased. Through this plot, however, it can be observed this difference was not equal across space. The no-seasonality model tended to perform better in northern and northeastern regions while the seasonality model performed better in other regions, Figure 4.14.

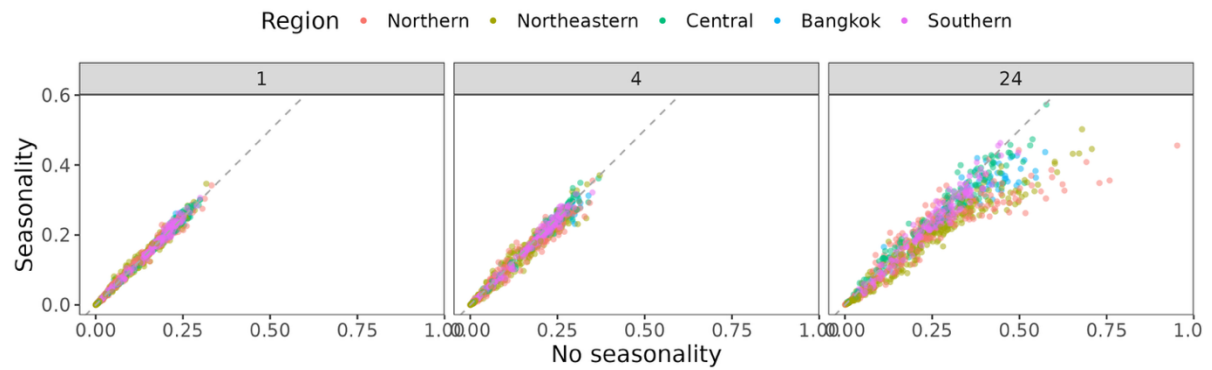


Figure 4.13-District CRPS by model and forecast horizon, coloured by region. Lower CRPS indicates better forecasting skill.

The difference in CRPS between no-seasonality and seasonality models within districts by forecast horizon is shown in Figure 4.14. The more positive and darker blue indicates that the no-seasonality model performed better in that district while a negative and darker orange difference shows that the seasonality model performed better in that district. These results align with Figure 4.13, where the no-seasonality model performed better in the northern and northeastern regions.

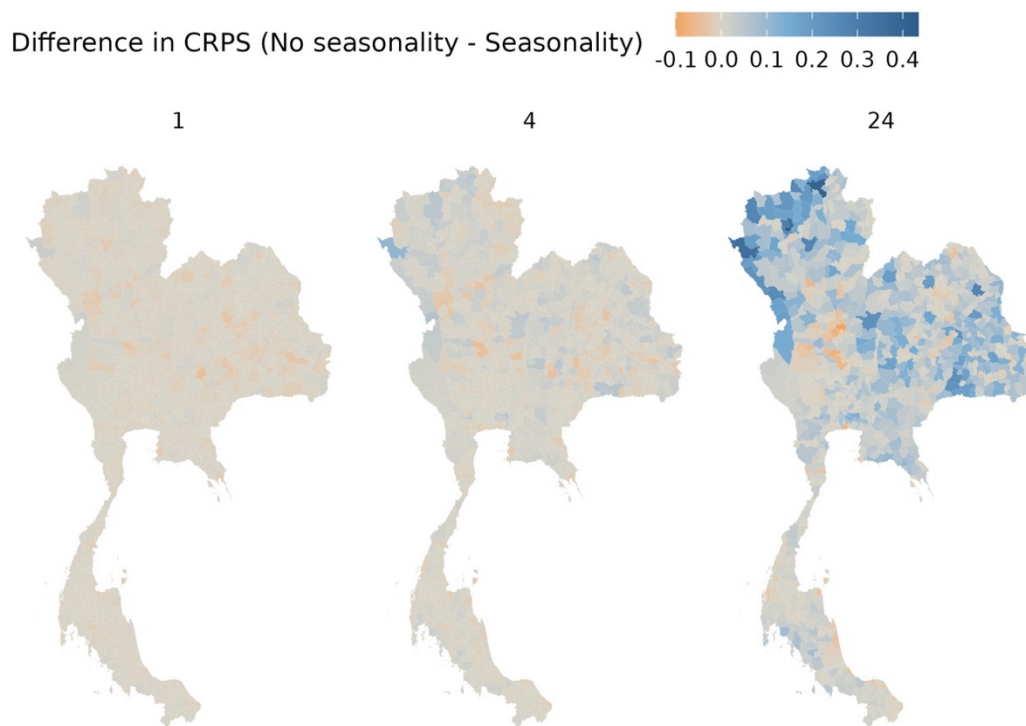


Figure 4.14-District level difference in CRPS by model and forecast horizon.

4.3.4 Outbreak prediction evaluation

4.3.4.1 Outbreak prediction evaluation by horizon

At each forecast horizon, each district was identified as having an outbreak if the observed cases and predicted number of cases exceeded the MoPH criteria. These classifications were then compared using a contingency table and AUC values were computed at forecast horizons of 1, 4, and 24 weeks and are presented in Figure 4.15. As previously observed, forecasting accuracy decreased with increasing forecast horizon, as demonstrated by decreasing AUC values. When comparing the two models of seasonality and no seasonality, the seasonality model had higher accuracy at all forecast horizons.

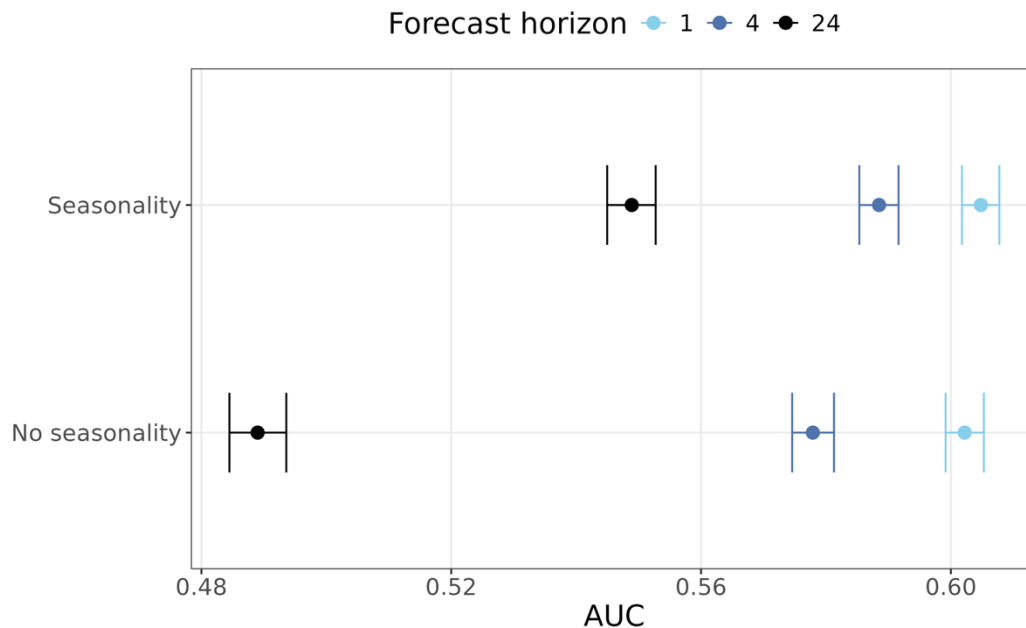


Figure 4.15-AUC by model and forecast horizon.

The distribution of district-level Brier scores by model and forecast horizon is presented in Figure 4.16. Overall, both models performed similarly over forecast horizons. The median was still very close to 0.5, or a random assignment of outbreak status. Importantly, for both models and across all forecast horizons, the distribution of brier scores was wide, indicating that there was a high diversity in the performance of models across districts.

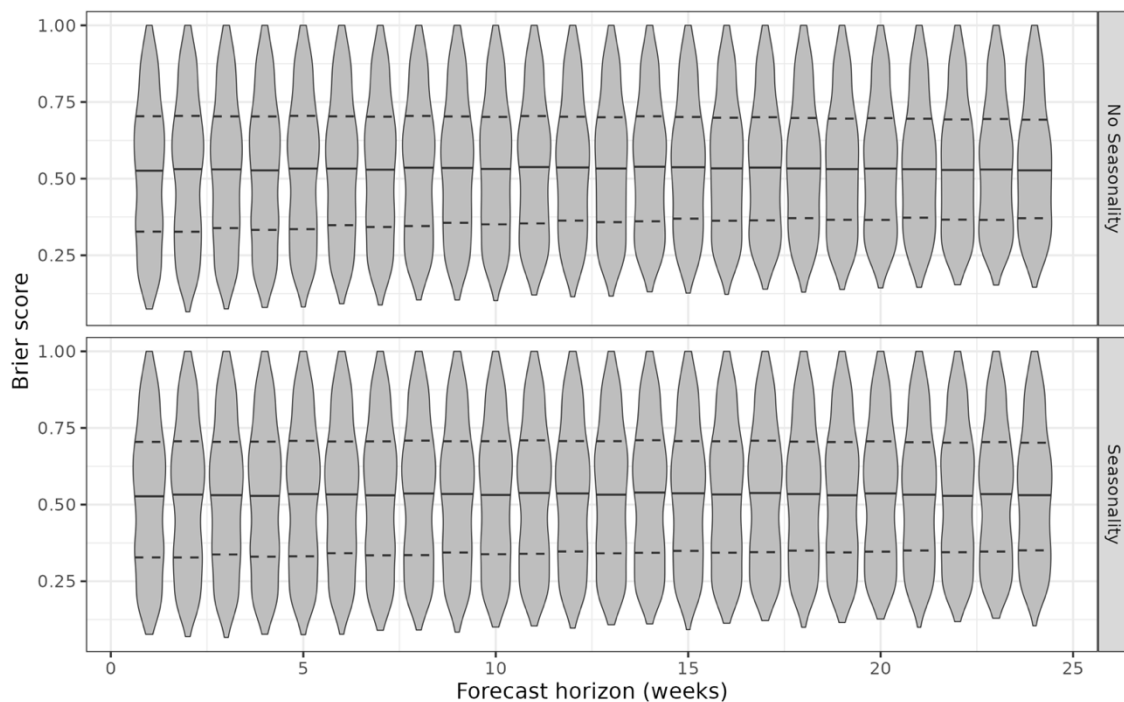


Figure 4.16-Brier score by forecast horizon (weeks) and model, higher values indicating better performance.

4.3.4.2 Outbreak prediction by target week

Agreement between outbreak classification based on observed cases and predicted cases at each district, forecast horizon, and target week was computed as 1 (agreement) or 0 (no agreement). The national level agreement at each target week and forecast horizon was then calculated and is presented in Figure 4.17. As forecast horizon increased, the difference in agreement between the two

models became more apparent. At the forecast horizon of 24 weeks, when there was an outbreak, for example in mid-2019, the seasonality model had a higher agreement with the observed outbreak than when there was no outbreak, for example in mid-2021.

Brier score between by forecast horizon of 1, 4, and 24 weeks was very similar between the two models, Figure 4.18. The difference between the two models increased as horizon increased.

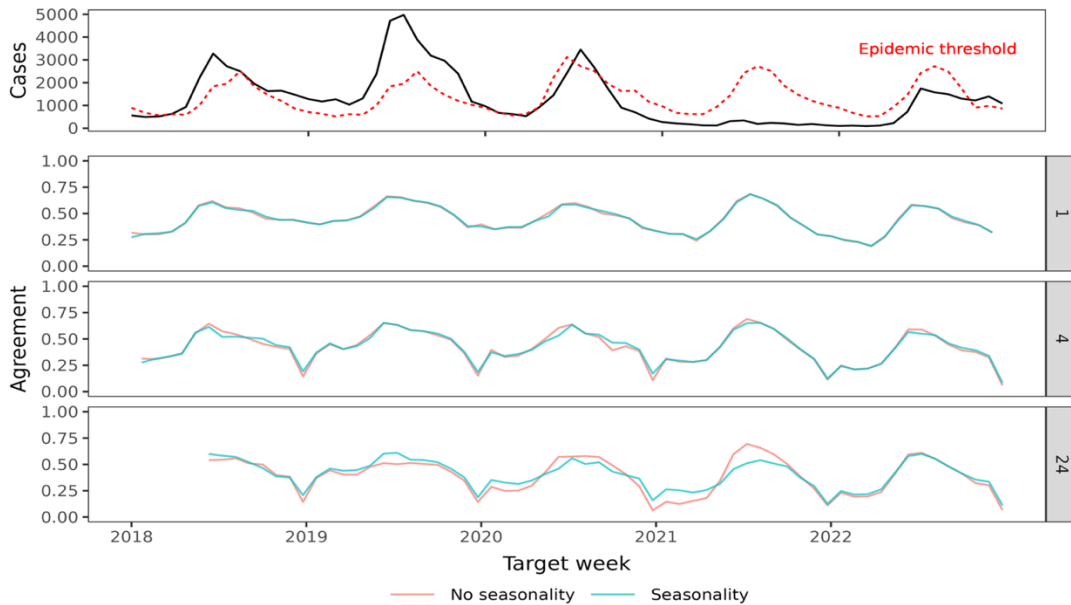


Figure 4.17-Agreement between observed and predicted outbreaks, by target week, model, and forecast horizon.

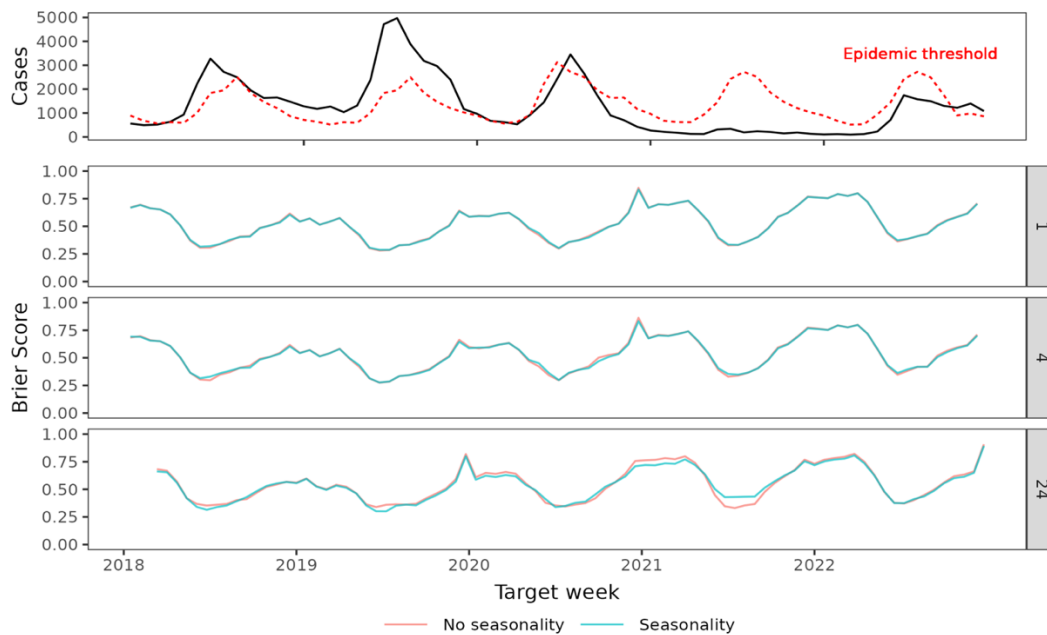


Figure 4.18-Brier Score by target week, model, and forecast horizon.

4.3.4.3 Outbreak prediction by space

Difference in agreement between no seasonality and seasonality models at the district level at forecast horizons of 1, 4, and 24 weeks were calculated and are presented in Figure 4.19. If the values are more positive and bluer, the no seasonality model had a higher agreement with observed outbreaks than seasonality model. Conversely, if the values are more negative and redder, the seasonality model had a higher agreement than the no seasonality model with the observed outbreaks.

As observed in both Figure 4.15 and Figure 4.17, as forecast horizon increases, the difference in agreement between the two models increased. At the 24-week horizon, the difference in agreement ranged from -0.30 to 0.22.

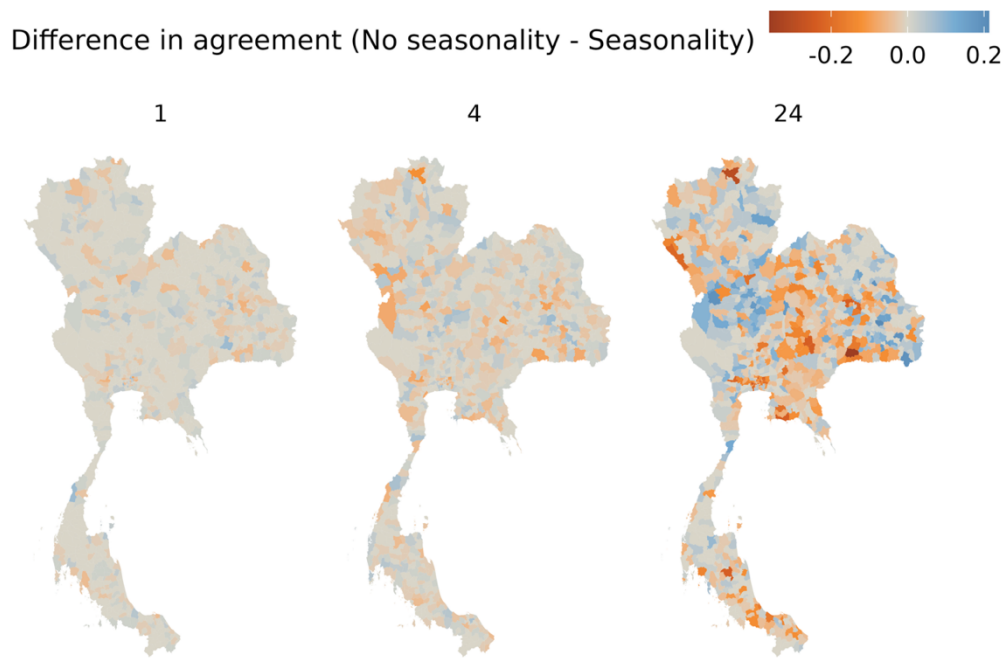


Figure 4.19-Difference in agreement between no seasonality and seasonality model, by district and forecast horizon (weeks).

Brier scores by districts are compared by models and forecast horizon, and coloured by region, presented in Figure 4.20. As forecast horizon increases, the difference in model performance by district increases, with the seasonality model having a lower brier score, indicating better performance. For some districts, particularly in the northeastern region, this difference is more pronounced.

Difference in Brier Score by model and forecast horizon is also plotted on the map of Thailand in Figure 4.21. The lower panel shows the districts where the difference between the two models is significant. As forecast horizon increases, the number of districts that show a significant difference between the two models increases.

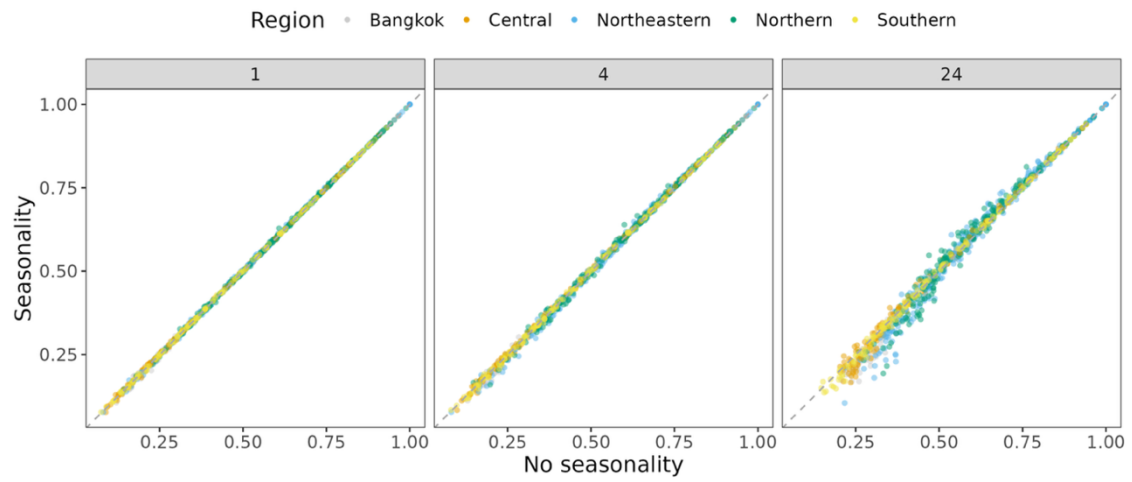


Figure 4.20-Brier Score by district and model, coloured by region.

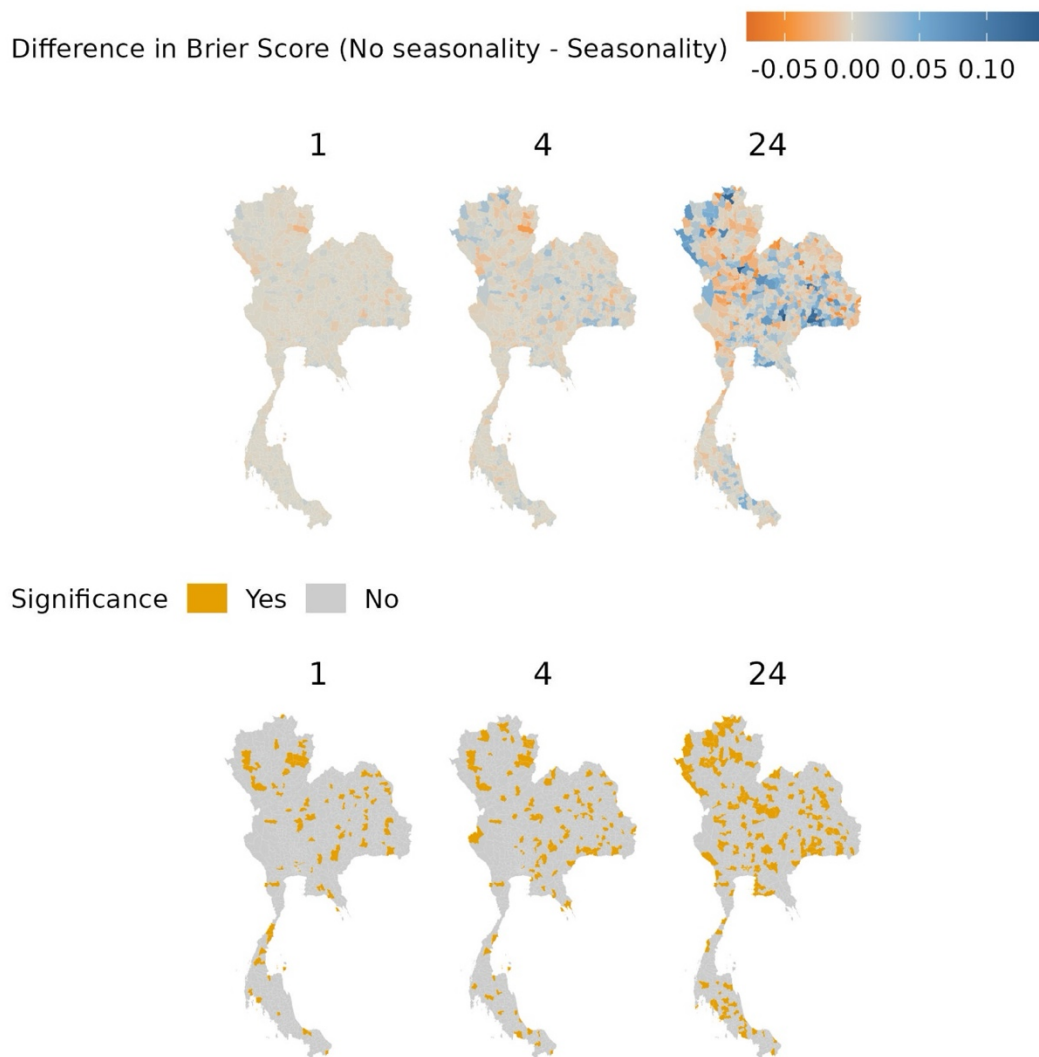


Figure 4.21-Difference in brier score between models and significance of difference by district and forecast horizon.

4.4 Discussion

Two models, one with an autoregressive term accounting for historical dengue cases and another with an additional seasonality term accounting for recurring transmission patterns were used to forecast dengue incidence up to 24 weeks ahead. Forecasted dengue incidence was also used to identity probability of an outbreak against a predefined threshold. Forecasts were evaluated using CRPS, sharpness, calibration, bias, AUC, and the Brier Score.

Seasonality model performed better based on CRPS and AUC, while the no seasonality model performed better based on NMAD and Bias metrics. Both models performed similarly based on empirical coverage and Brier scores. Overall, forecasting accuracy declined as the forecast horizon increased. Both models performed similarly at shorter horizons, suggesting the autoregressive model can effectively capture short-term transmission patterns. However, as the horizon increased, differences in model performance became more pronounced across districts. This variation may reflect the underlying heterogeneity in seasonal dynamics or other district-level factors not captured by the baseline models.

Despite their simplicity, these models provided interpretable results, were able to be run with minimal data input, and were computationally efficient. However, their limited capacity to capture complex environmental relationships may reduce their performance at longer horizons or in dynamic settings.

Seasonality models can overfit expected patterns, for example, expecting a peak during every rainy season every year even if an unexpected event like a pandemic or climate anomaly interrupts that trend. The no-seasonal model, by not making these assumptions, can be more responsive to unexpected shifts, but lacks the structure that helps with consistent long-term accuracy. This is also observed in previous research. Tuan[319, 320] found that in a province in Thailand, even though SRIMAX and LSTM correctly modelled seasonality, they struggled with predicting short-term outbreaks due to overfitting. Sebastianelli et al.[191] used an ensemble machine learning approach to forecast dengue outbreaks across regions in Brazil. They found the model performed best in regions with stable seasonality and performed worse in areas with sudden shifts, irrespective of the season.

With baseline models established, more complex models incorporating additional covariates can now be assessed. These models may offer increased accuracy, particularly at longer-term forecast horizons or in districts where transmission patterns are variable. The following chapters will explore whether and how much complexity in model structure will lead to gains in predictive performance.

5 Univariate Analysis

5.1 Introduction

This chapter will explore the relationships between each covariate of interest and dengue incidence to better understand how they can offer an advantage in forecasting. While the previous chapter focused on establishing baseline models using autoregressive and seasonality terms, this chapter focuses on the role of additional factors such as climate, environmental, genetics, and mobility in predicting transmission dynamics.

Dengue transmission is affected by a range of factors, among which climate variables are the most extensively studied. Increases in temperature and precipitation favour mosquito activity of both larval and adult stages, which lead to increases in dengue cases[95, 98, 107, 112, 126]. These relationships are not always linear. For temperature, transmission is typically higher within an optimal range that supports mosquito survival and viral reproduction[102-106]. Similarly, rainfall provides water necessary for breeding sites[112-116], but excessive rainfall can wash away larvae, reducing mosquito populations[117-120]. Wind speed has shown inconsistent associations with dengue transmission, with studies reporting both positive[157, 158] and negative [152-155] relationships.

There are factors beyond climate that influence dengue transmission. Increasing urbanisation, often reflected in the patterns of land use through vegetation index, has been linked to higher transmission, as dense human population lead to increased human and mosquito contact[162, 163]. Genetic diversity has been found to be correlated with dengue transmission intensity[321]. Human mobility

further amplifies spread by facilitating introduction of virus to new areas and sustaining transmission chains across communities[171, 172].

By examining these covariates one at a time, this chapter provides a univariate perspective on their associations with dengue incidence. This approach provides a better understanding of the individual drivers of disease risk, thus setting the stage for more complex multivariable modelling in subsequent analyses.

5.2 Methods

5.2.1 Covariates

5.2.1.1 Temperature

Hourly 2 m temperature (2 m above the surface) was downloaded from ERA5[322] for years 2012 – 2022, at a spatial resolution of 0.25 degrees (latitude and longitude). The data was restricted to the area defined by the GPS coordinates 20.46/97.34/5.61/105.64 to include only grids relevant to Thailand. The data was then trimmed to the shapefile of Thailand, obtained from Humanitarian Data Exchange[323]. The grids were then mapped to the district level shapefile[323], with a grid square included if its centre was in the polygon of interest. Then, mean temperature across district and hourly time steps was calculated. For weekly mean temperature, the mean temperature for each district was averaged by week. For minimum and maximum temperature, the minimum and maximum temperature of the mean hourly values per district were selected.

For week 53 in each year, the values were averaged with week 52 and then week 53 was removed. This was necessary to ensure each year had 52 weeks so seasonality at the weekly level could be implemented in the models. Lags were

then calculated for each variable. Data for up to 12 weeks at the end of 2012 were included to allow for calculation of lags for weeks in 2013.

5.2.1.2 Standardized Precipitation Evapotranspiration Index (SPEI)

To calculate SPEI, two components were required. Total precipitation at hourly intervals for years 2012 – 2022 was downloaded from ERA5[322] at a spatial resolution of 0.25 degrees. The data was restricted to the area defined by the coordinates 20.46/97.34/5.61/105.64 to include only grids relevant to Thailand. Mean total precipitation at the district and weekly level was calculated and converted from metres to millimetres.

Hourly evapotranspiration at 0.1 degrees grid resolution for years 2012 – 2022 were downloaded from the hPET global dataset[324]. The data was averaged at the district and weekly level.

Water balance was then calculated by subtracting evapotranspiration from precipitation for each district and week. SPEI at scales 0 to 6 were calculated for each district separately using the function `spei` from the package `SPEI`[325].

5.2.1.3 Wind Speed

To obtain wind speed, hourly wind speed in east-west direction (u) and wind speed in north-south direction (v) were downloaded from ERA5[322] for the years 2012 – 2022, at a spatial resolution of 0.25 degrees. As above, the data was restricted to the area defined by the coordinates 20.46/97.34/5.61/105.64 to include only grids relevant to Thailand and trimmed using the Thailand shapefile[323]. The wind speed in each grid was calculated by using the Pythagorean theorem where wind

speed = $\sqrt{v^2 + v^2}$. The average wind speed at the weekly level for each district level was then calculated.

5.2.1.4 Enhanced Vegetation Index (EVI)

EVI was downloaded from AppEEARS [326, 327] for the previously mentioned bounding box. The dataset provided 16-day composite EVI at 1 km spatial resolution[327]. The 16-day composite for each grid was linearly interpolated to estimate daily values, using the `na.approx()` function in the `zoo` package[328]. Daily EVI values were then averaged to obtain values for each epidemiological week. Subsequently, grid-level values were aggregated to the district level by overlaying with district shapefiles and calculating the average EVI per district. Weeks 52 and 53 in certain years were adjusted to maintain consistency with other data. Lags of 0 to 12 weeks were calculated for the EVI values and then log-transformed and scaled.

5.2.1.5 Genetic Diversity

The method for calculating genetic diversity is explained in Chapter 3.

5.2.1.6 Mobility Data

Activity Space Map data for Thailand was obtained from the Data for Good team at Meta through personal communication under an existing agreement between MORU and Meta. Based on this communication, data was provided by Meta at the district level, as typically the data is only available at the grid level. The data corresponds to a reference date of 7 April 2025, and represents an aggregation of users within a district, over the past three weeks, for users that share their location with Facebook, based on their typical location and assigns that district as their home district. The data is further divided into daytime and nighttime[329]. For the

purposes of this analysis, I only used the data for daytime activity as this reflects most closely the activity of *Aedes* mosquitoes.

The average proportion of visits to each district from each other district was calculated and converted into a matrix. There were some missing districts in the Data for Good dataset. For these districts, it was assumed that their connection with other districts was the average of all the other connections for that district.

5.2.2 Relationship with dengue cases

To better understand the relationships between each covariate and dengue transmission, Distributed Lag Non-Linear Models (DLNM) were used to investigate the lagged and non-linear effects of each covariate on dengue incidence rates per 100,000. A cross-basis function was used to simultaneously model these effects, with lags extended up to 12 weeks to explore potential delayed impacts of climate and environmental factors on dengue transmission, consistent with lag periods commonly used in dengue climate studies[330-332]. The resulting cross-basis terms were then added as a covariate to a generalized additive model (GAM). To control for seasonality, yearly effects at the week level were included. A random effect at the province level was also included to account for unaccounted spatial heterogeneity. This model is referred to as the national model.

A sensitivity analysis was performed by modifying the spatial random effect structure. Instead of the random effect at the provincial level, a region random effect and an interaction between region and province was included to account for the hierarchical spatial structure. This model is referred to as the regional model.

For both models, dengue case counts were modelled with quasi-poisson distribution to account for the many zeroes in the case data.

Mean Absolute Error (MAE) and Root-mean-square deviation (RMSE) were calculated for each model, at the regional level, and compared. Confidence intervals for each of MAE and RMSE was calculated by bootstrapping of residuals.

Cross-basis analysis was performed by the `crossbasis()` function from the `dlnm`[333] package in R, and the model was fit using the `bam()` function in the `mgcv`[334] package in R. Residuals were extracted using the `resid()` function from the `stats`[335] package in R.

For the national model;

Let:

- Y_{it} be the weekly dengue cases in district i and week t
- $\mu_{it} = E[Y_{it}]$
- Pop_i be the population of district i
- cb_temp_{it} represent the cross-basis function of covariate
- $Week_t \in [1, \dots, 52]$ be the week of the year
- $Prov_i$ be a province-level random effect

Then the model was;

$$Y_{it} \sim Quasi - Poisson(\mu_{it}, \phi) \quad (11)$$

$$\log(\mu_{it}) = f^1(cb_{t_c}ov_{it}) + f^2(Week_t) + \log(Pop_i) + b_i \quad (12)$$

Where:

- $f^1(cb_{t_c}ov_{it})$ was a smooth function for the non-linear, lagged effect of covariate
- $f^2(Week_t)$ was a cyclic cubic spline capturing seasonal variation over the calendar year
- $b_i \sim N(0, \sigma^2)$ was a random intercept for province i
- $\log(Pop_i)$ was an offset to account for population size

The response was modelled with a quasi-Poisson distribution to account for overdispersion, with dispersion parameter ϕ .

For the regional model;

Let:

- Y_{it} be the weekly dengue cases in district i and week t
- $\mu_{it} = E[Y_{it}]$
- Pop_i be the population of district i
- $cb_{t_c}emp_{it}$ represent the cross-basis function of covariate
- $Week_t \in [1, \dots, 52]$ be the week of the year
- $Region_r$ be the region-level random effect
- $RegProv_{ir}$ be a random effect for province nested within region (i.e., interaction effect)

Then the model was:

$$Y_{it} \sim \text{Quasi-Poisson}(\mu_{it}, \varphi) \quad (13)$$

$$\log(\mu_{it}) = \log(\text{Pop}_i) + f^1(cb_{t_c}ov_{it}) + f^2(\text{Week}_t) + b_r + b_{ir} \quad (14)$$

Where:

- $\log(\text{Pop}_i)$ was an offset accounting for population
- $f^1(cb_{t_c}ov_{it})$ was a smooth function for non-linear, lagged temperature effects
- $f^2(\text{Week}_t)$ was a cyclic cubic spline for seasonality over 52 weeks
- $b_r \sim N(0, \sigma^2_r)$ was a random intercept for region
- $b_{ir} \sim N(0, \sigma^2_{ir})$ was a random intercept for province within region (a hierarchical structure)

The response was modelled with a quasi-Poisson distribution to account for overdispersion, with dispersion parameter φ , which was calculated as part of the model fitting process.

Covariates included temperature, wind speed, SPEI, and EVI.

5.2.3 Forecasting dengue cases with covariates

Climate and environment covariates with lags 1 to 6 weeks were added to each of Model 1 and Model 2 established in the 'Baseline Chapter'. These variables were modelled as random walk of order 2 (RW2)[219].

$$\gamma_j | \gamma_{j-1} \sim N(\gamma_{j-1}, \sigma_\gamma^2) \quad (15)$$

Each step was dependent on the previous step with an addition of some noise, σ_γ^2 . The PC prior for this model was assumed to be `pc.prec(U=1,  $\alpha=0.01$ )`, which was a weak prior, assuming that the standard deviation, σ_γ , of the next step was very likely to be small.

5.3 Results

5.3.1 Relationship between covariates and dengue cases

5.3.1.1 Temperature

5.3.1.1.1 Mean Temperature

5.3.1.1.1.1 Main Model

The national model previously described in section 5.2.1 was used to model the relative risk (RR) of dengue by temperature and lags. The median value of mean temperature of 26.5°C was used as the reference point for calculating RR. The dengue RR by mean temperature and lag is presented in Figure 5.1. RR was lowest at lowest temperature, 11°C and lag 0, while the highest RR was found at the highest temperature, 35°C and highest lag of 12 weeks.

Relative risk by temperature, over different lags is presented in Figure 5.2. At all lags, RR for temperatures below 20°C was below 1, followed by an increase that was more prominent for longer lags. The RR then declined for all lags, with the largest decrease observed at higher lags.

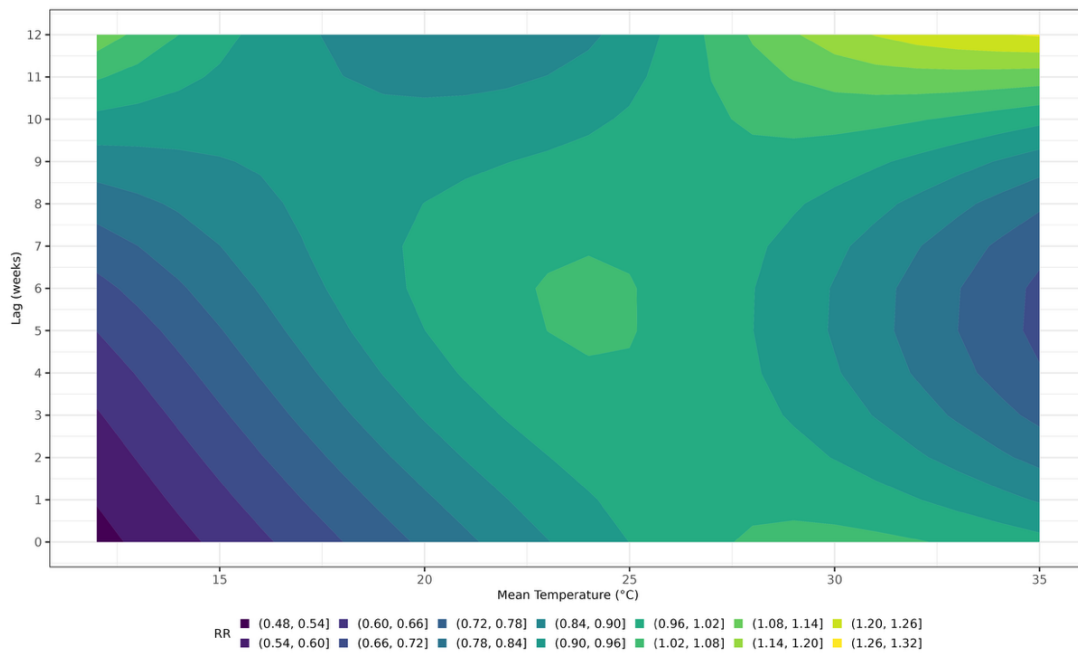


Figure 5.1-Relative Risk of dengue, by mean temperature and lag

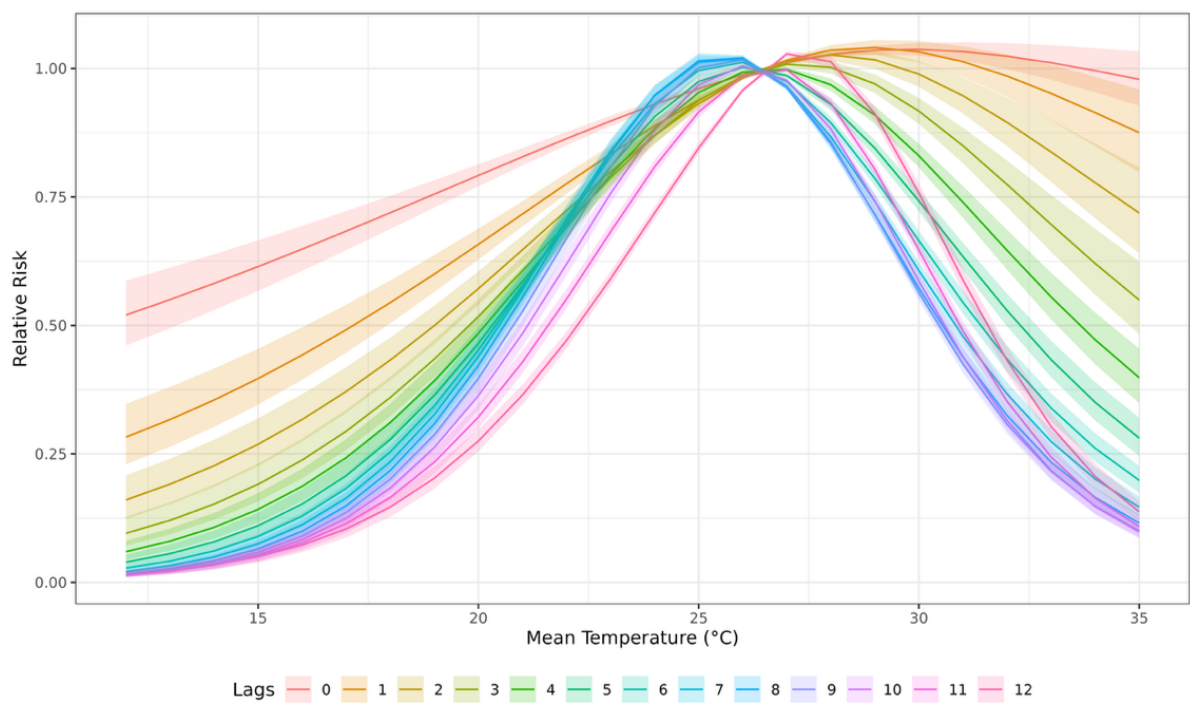


Figure 5.2-Relationship between relative risk of dengue and mean temperature, by lag

5.3.1.1.1.2 Sensitivity analysis

As part of a sensitivity analysis, the national model and regional model described in section 5.2.1 were compared. Results are presented in Figure 5.3. Both MAE and RMSE were lower for the regional model. This is not surprising as increasing the number of parameters increased the flexibility of the model, improving fit.

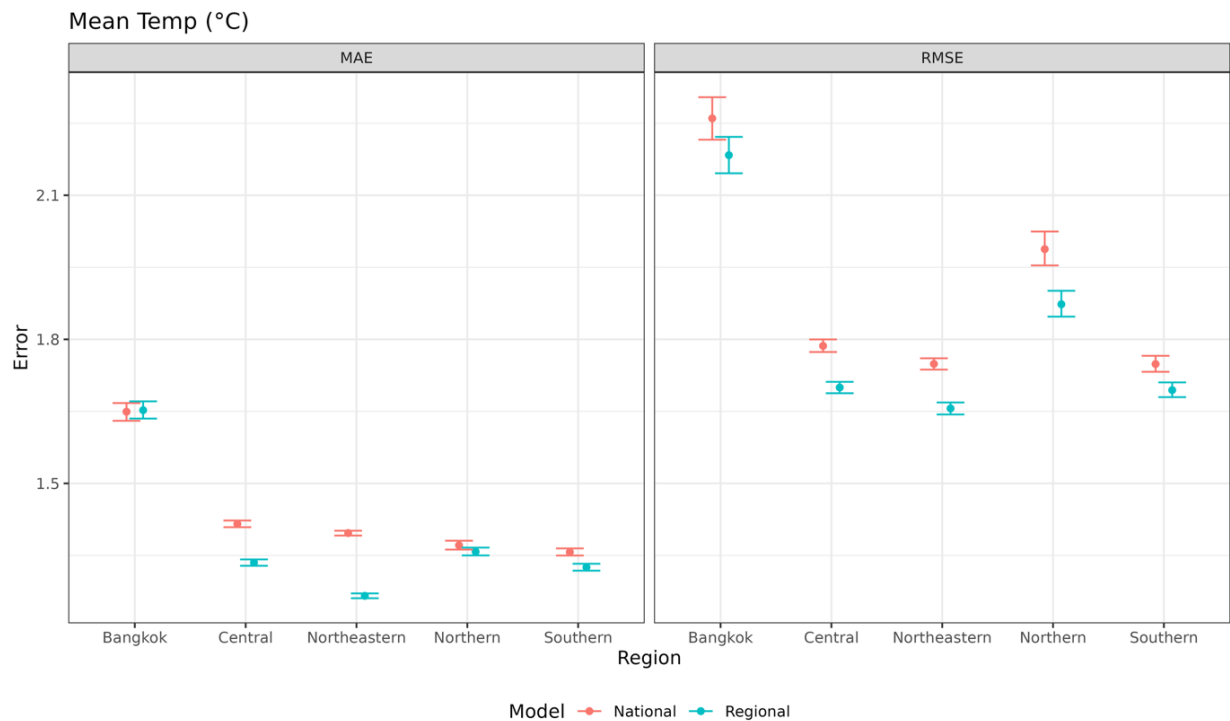


Figure 5.3-Errors of residuals for modelling dengue and mean temperature, by region, national and regional model

5.3.1.1.2 Minimum Temperature

5.3.1.1.2.1 Main Model

A contour plot of RR of dengue by minimum temperature and lags of 0 to 12 weeks calculated for median of minimum temperature of 22.7 °C as reference, Figure 5.4. RR was lowest at lowest temperature, 5°C and lag 0, while the highest RR was found at the highest temperature, 29°C and highest lag, 12 weeks.

Relative risk by temperature, over different lags is presented in Figure 5.5. At all lags, RR below 20°C was below 1, followed an increase and decrease that was more prominent by longer lags.

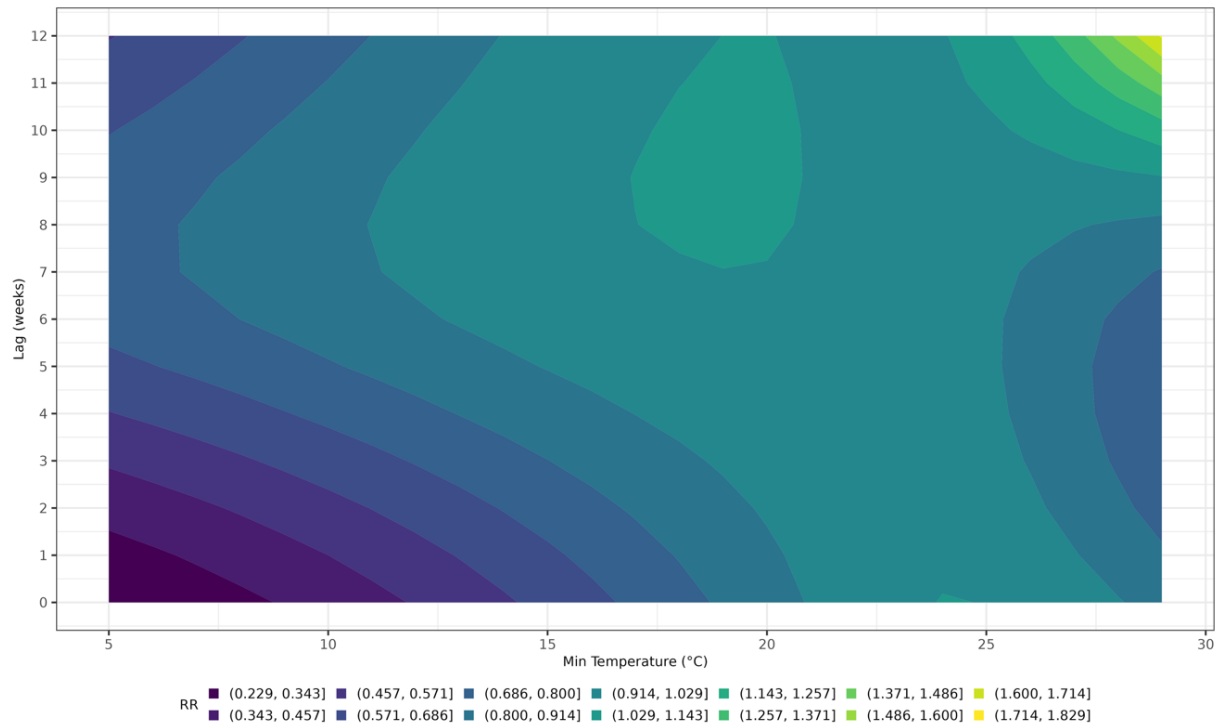


Figure 5.4-Relative Risk of dengue, by minimum temperature and lag

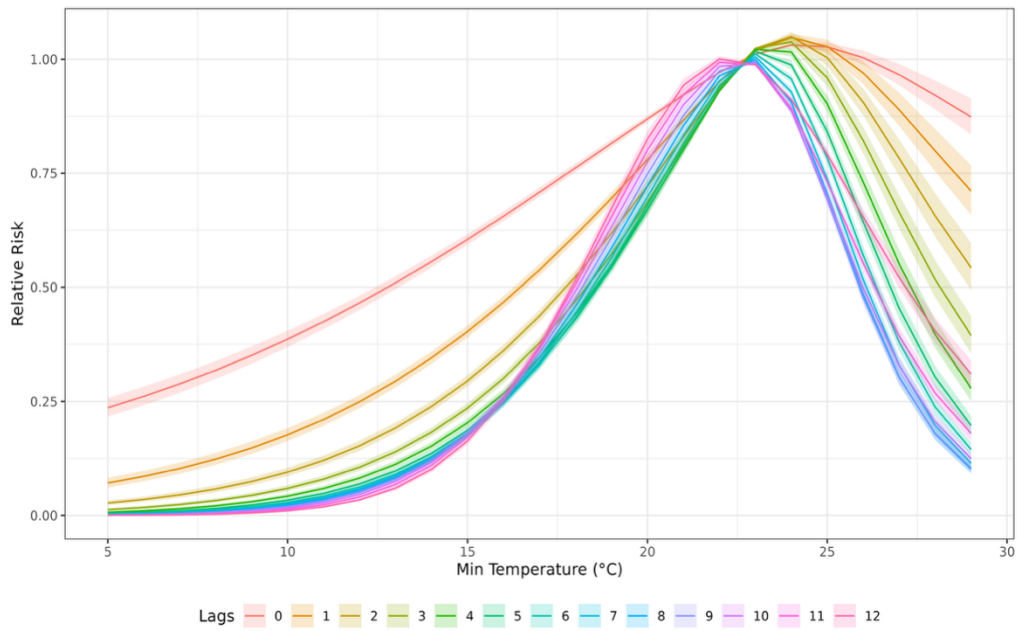


Figure 5.5-Relationship between relative risk of dengue and minimum temperature, by lag

5.3.1.1.2.2 Sensitivity analysis

MAE and RMSE of the national model and regional model were compared and are presented in Figure 5.6. Both MAE and RMSE were lower for the regional model except for Bangkok. In Bangkok, the MAE was higher for the regional model compared to the national model.

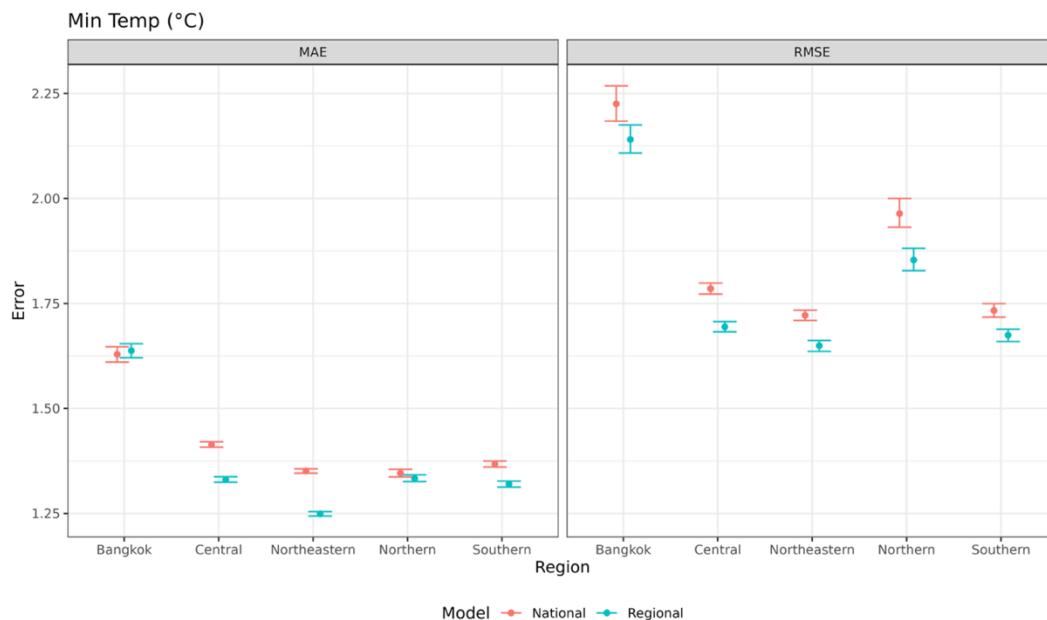


Figure 5.6-Errors of residuals for modelling dengue and minimum temperature, by region.

5.3.1.1.3 Maximum Temperature

5.3.1.1.3.1 Main Model

A contour plot of RR of dengue by minimum temperature and lags of 0 to 12 weeks calculated for median of minimum temperature of 31.9 °C as reference, is presented in Figure 5.7. RR was lowest at highest temperature, 43°C and lags 1 to 6. The highest RR was found at highest lag of 12 and highest temperature, 43°C, and at lowest temperature at 18°C.

Relative risk by temperature, over different lags is presented in Figure 5.8. At all lags, RR below 20°C was below 1, followed an increase and decrease that was more prominent by longer lags.

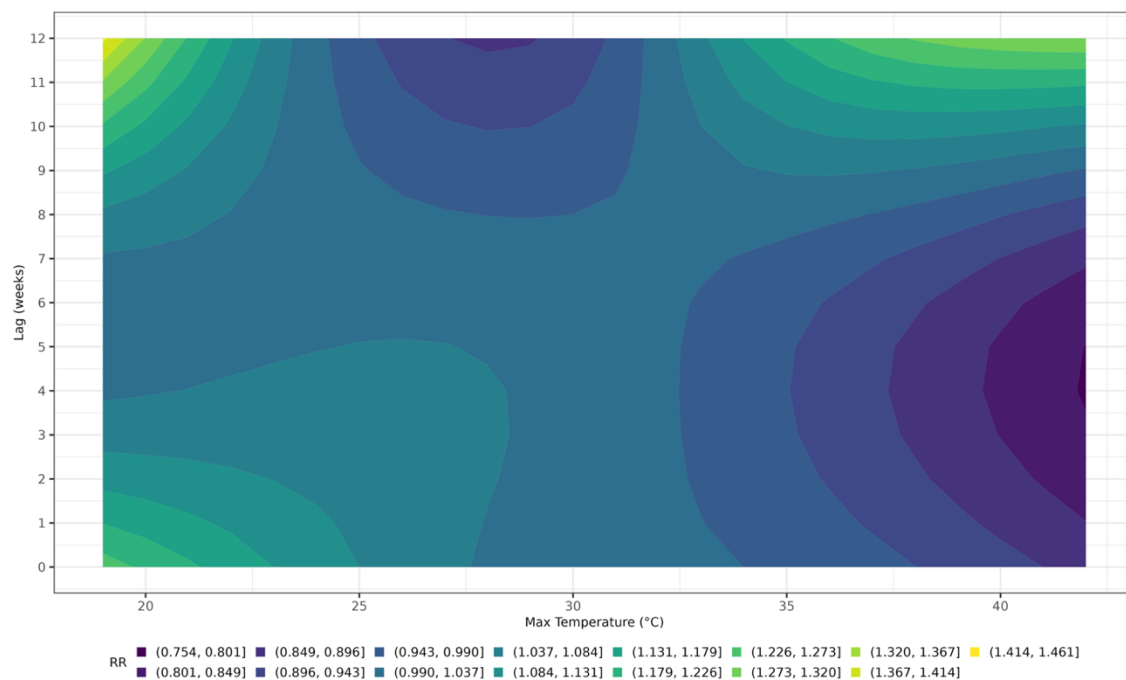


Figure 5.7-Relative Risk of dengue, by maximum temperature and lag

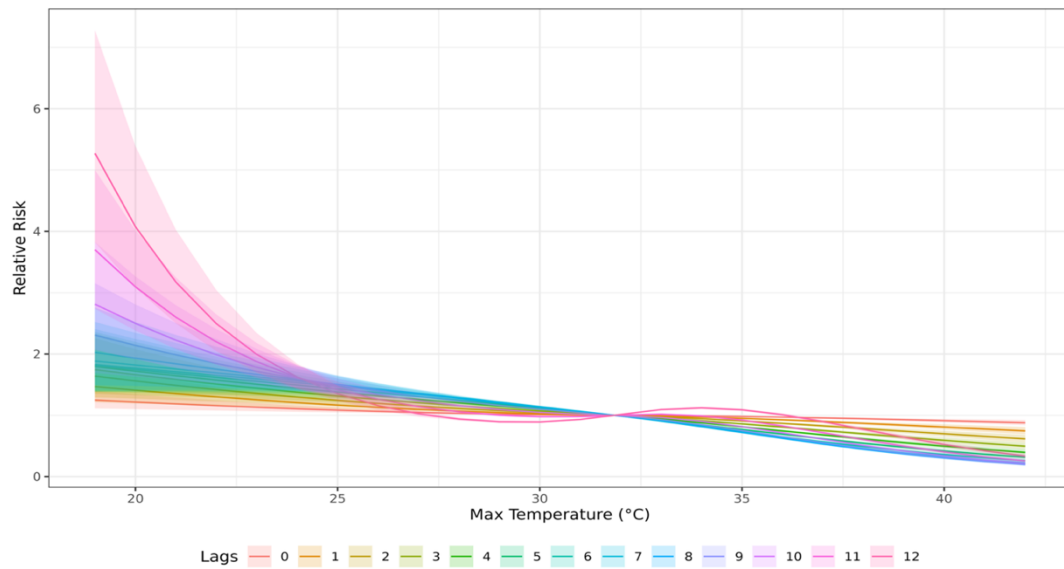


Figure 5.8-Relationship between relative risk of dengue and maximum temperature, by lag

5.3.1.1.4 Sensitivity analysis

MAE and RMSE of the national model and regional model were compared, Figure 5.9. Both MAE and RMSE were lower for the regional model except for Bangkok.

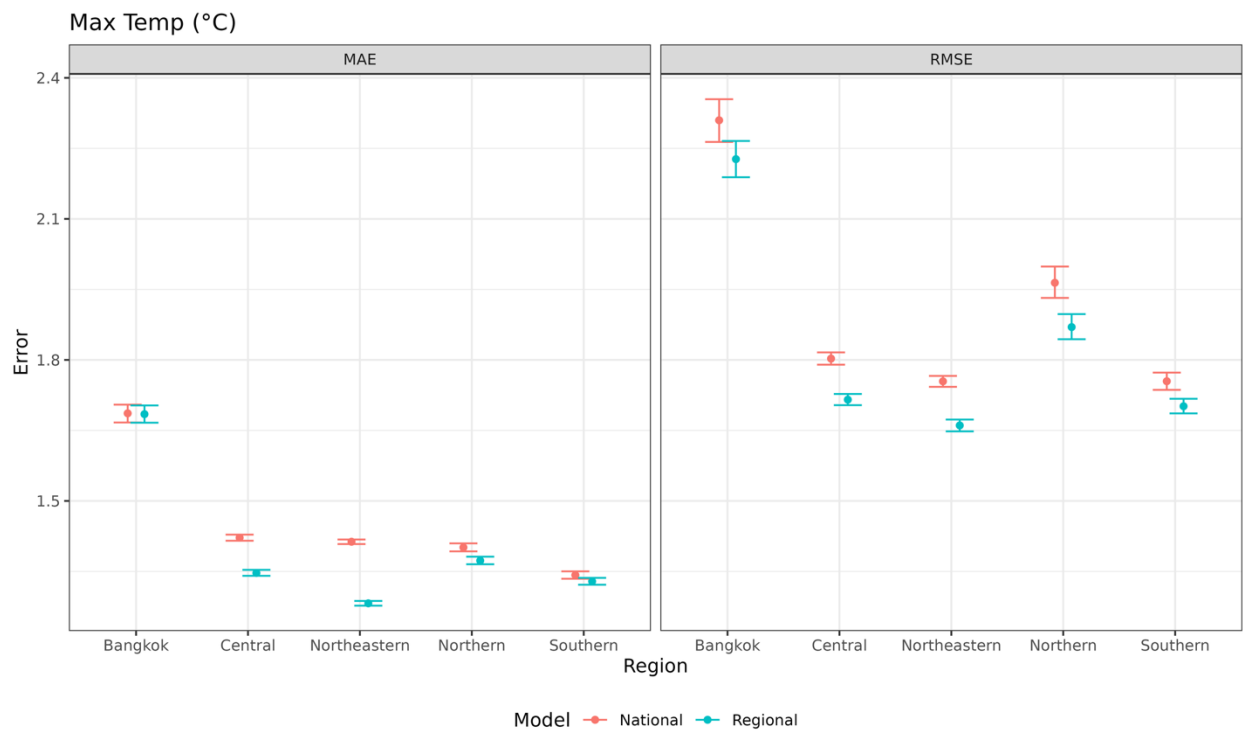


Figure 5.9-Errors of residuals for modelling dengue and maximum temperature, by region.

5.3.1.2 SPEI

5.3.1.2.1 Main Model

RR contour plot by SPEI and lags 0 to 12 weeks is presented in Figure 5.10. Highest RR was observed at lags 5 and 6, and the lowest SPEI. RR decreased at decreasing SPEI and lag of 1 and 12 weeks.

RR by SPEI and different lags is presented in Figure 5.11. At the lowest SPEI, RR increased as las increased. At low SPEI values, short lags of 0 to 2 weeks showed RR below 1, a protective effect against dengue, while long lags of 8 to 12 weeks showed increased risk of dengue. At high SPEI values, the highest lag, 12 weeks, showed the lowest risk of dengue.

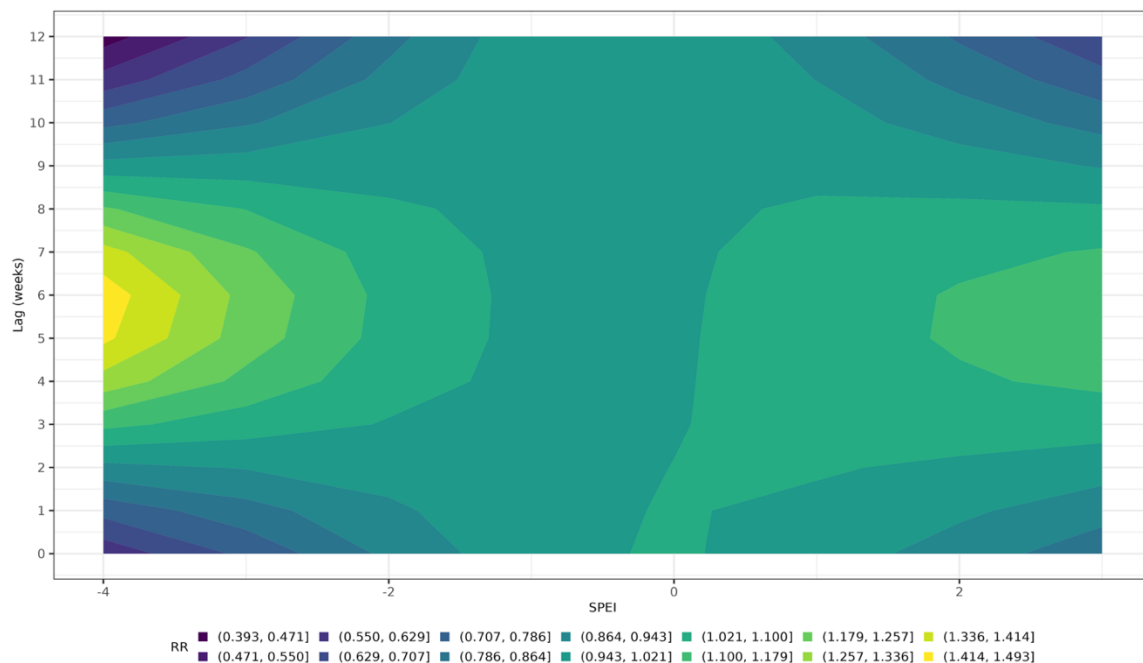


Figure 5.10-Relative Risk of dengue, by SPEI and lag

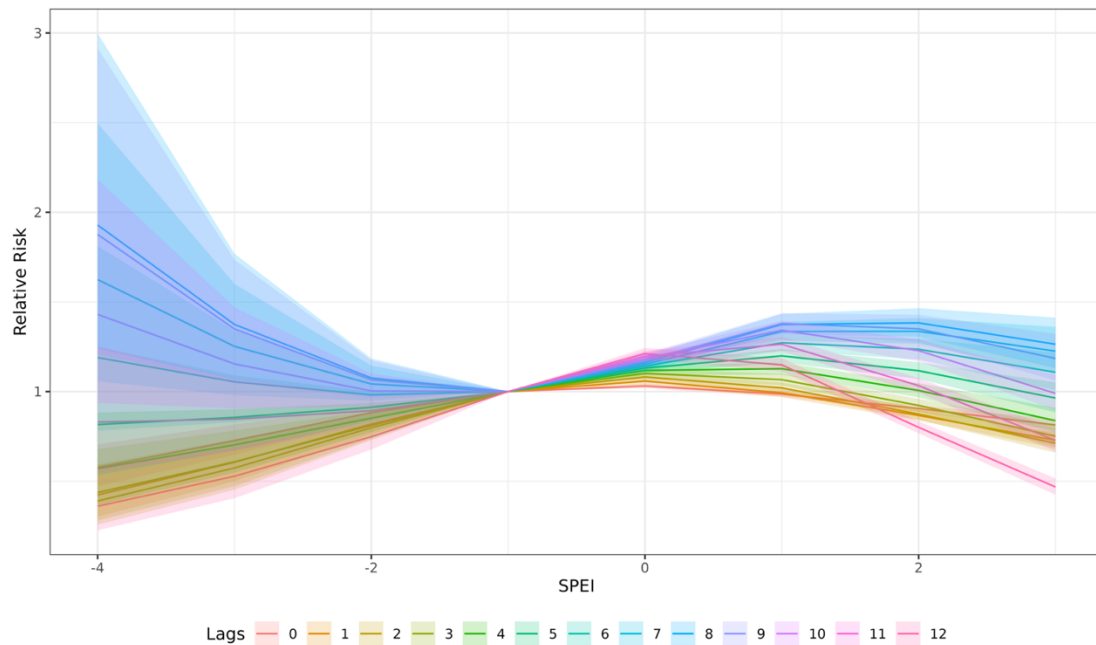


Figure 5.11-Relationship between relative risk of dengue and SPEI, by lag

5.3.1.2.2 Sensitivity analysis

The regional model had lower MAE and RMSE compared to the national model, except in Bangkok for the MAE metric where the regional model performed better, Figure 5.12.

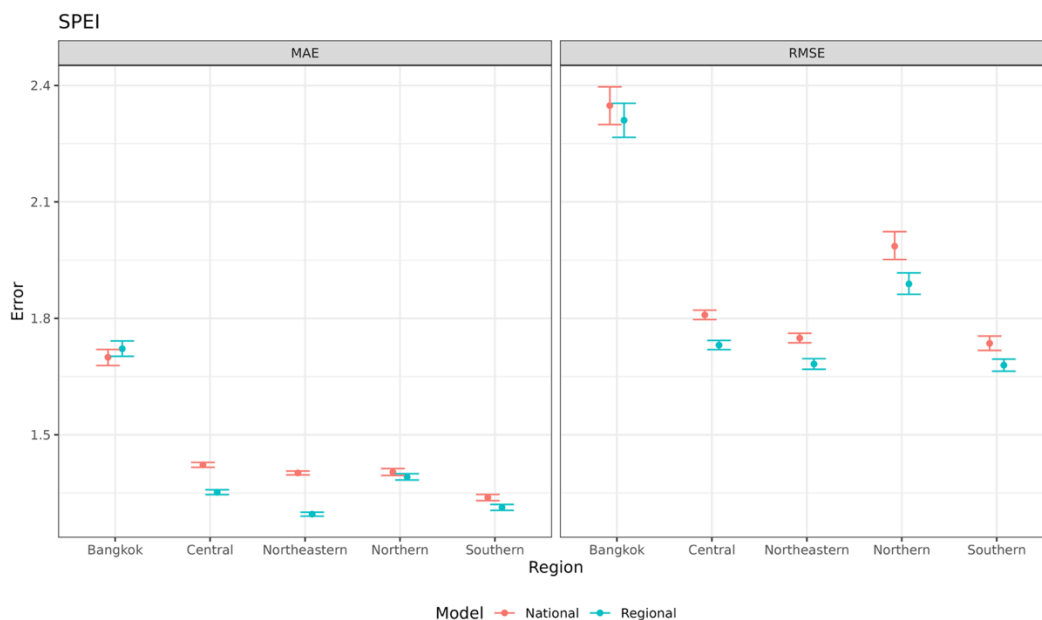


Figure 5.12-Errors of residuals for modelling dengue and SPEI, by region.

5.3.1.3 Wind Speed

5.3.1.3.1 Main Model

A contour plot of RR of dengue by wind speed and lags of 0 to 12 weeks is presented in Figure 5.13. RR was lowest at lowest wind speed and lag 12 while the highest RR was found at lag 0 and lag 12 at highest wind speed.

RR by wind speed, over different lags is presented in Figure 5.14. At all lags, the RR converged to 0 at wind speed of 5 m/s, after which it increased.

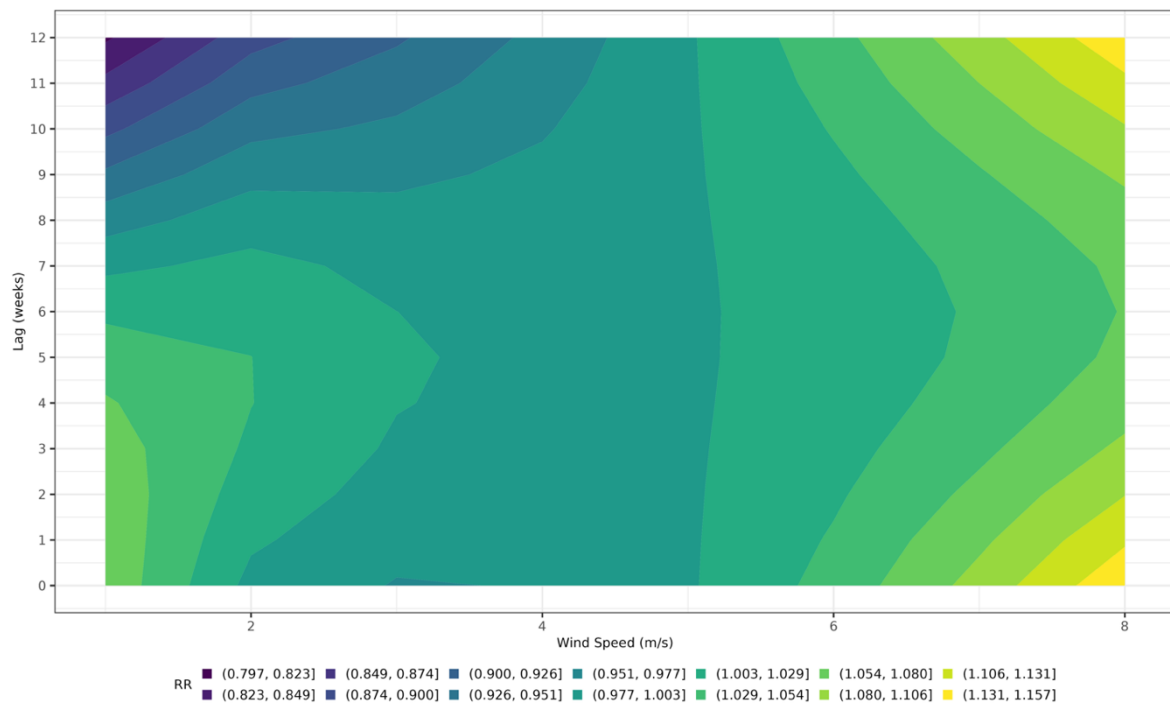


Figure 5.13-Relative Risk of dengue, by wind speed and lag

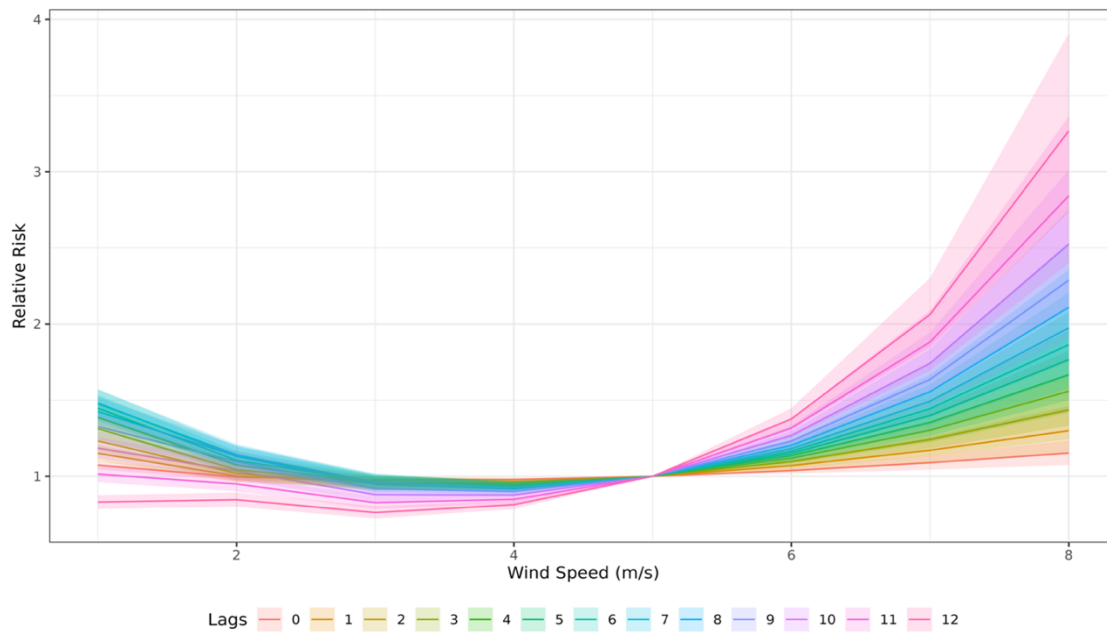


Figure 5.14-Relationship between relative risk of dengue and wind speed, by lag

5.3.1.3.2 Sensitivity analysis

The regional model had lower MAE and RMSE compared to the national model, as presented in Figure 5.15.

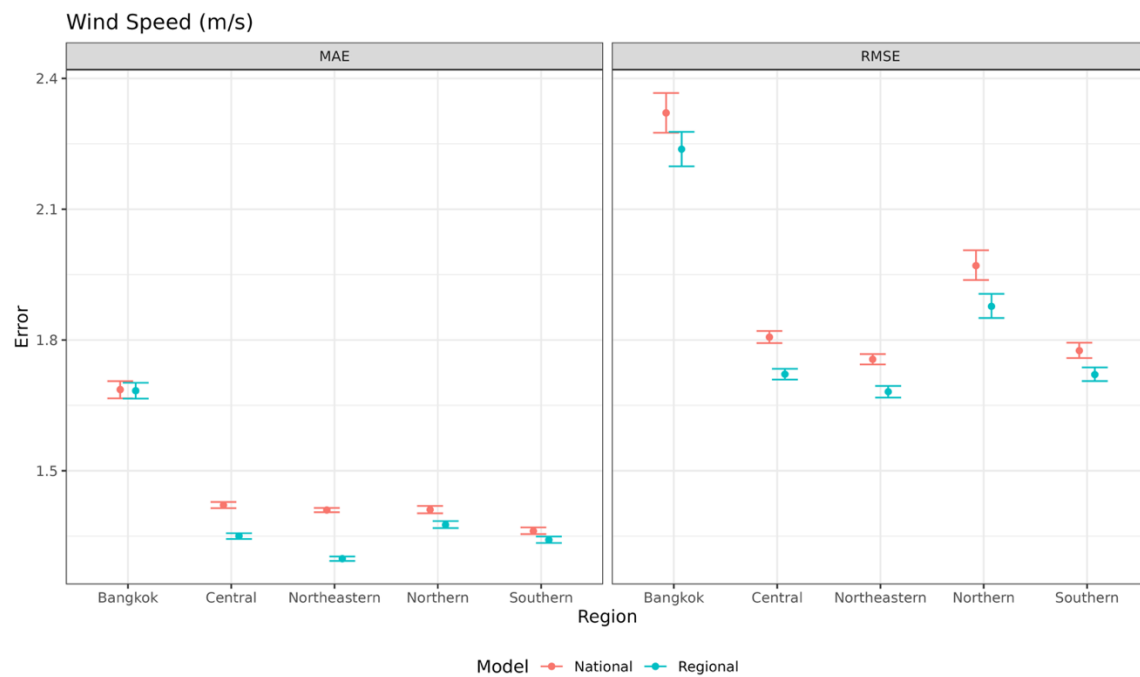


Figure 5.15-Errors of residuals for modelling dengue and wind speed, by region

5.3.1.4 EVI

5.3.1.4.1 Main Model

A contour plot RR by lags of 1 to 12 weeks and EVI is presented in Figure 5.16. The highest RR was observed at lags 0 to 3, at the lowest EVI, while the lowest RR was observed at lag 12 and lower EVI.

RR by EVI, over different lags shows that overall, the RR decreased as EVI increased, Figure 5.17.

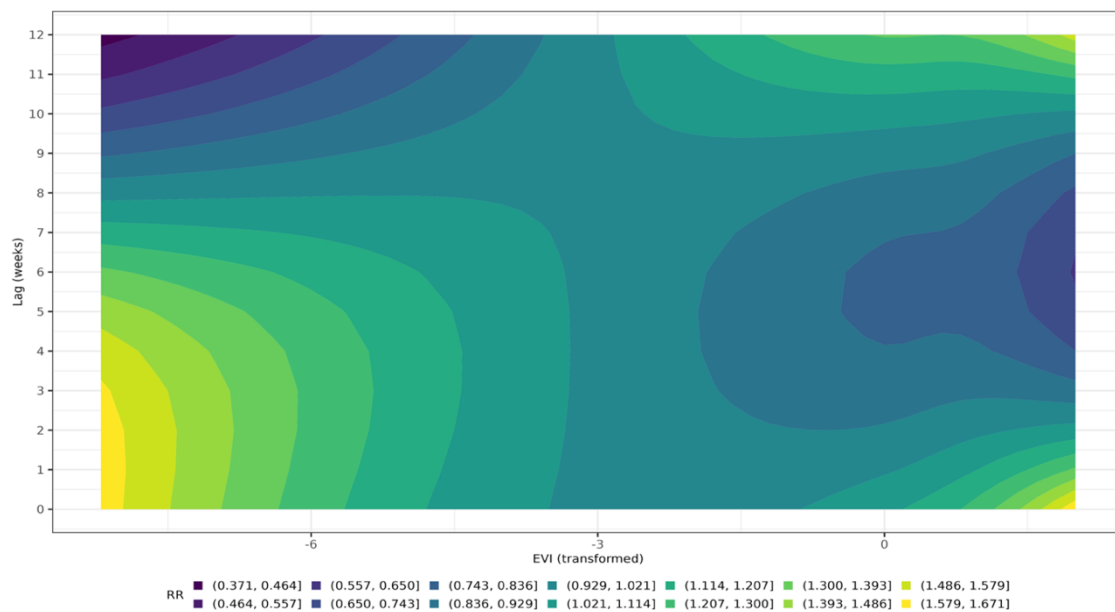


Figure 5.16-Relative risk of dengue by EVI and lag

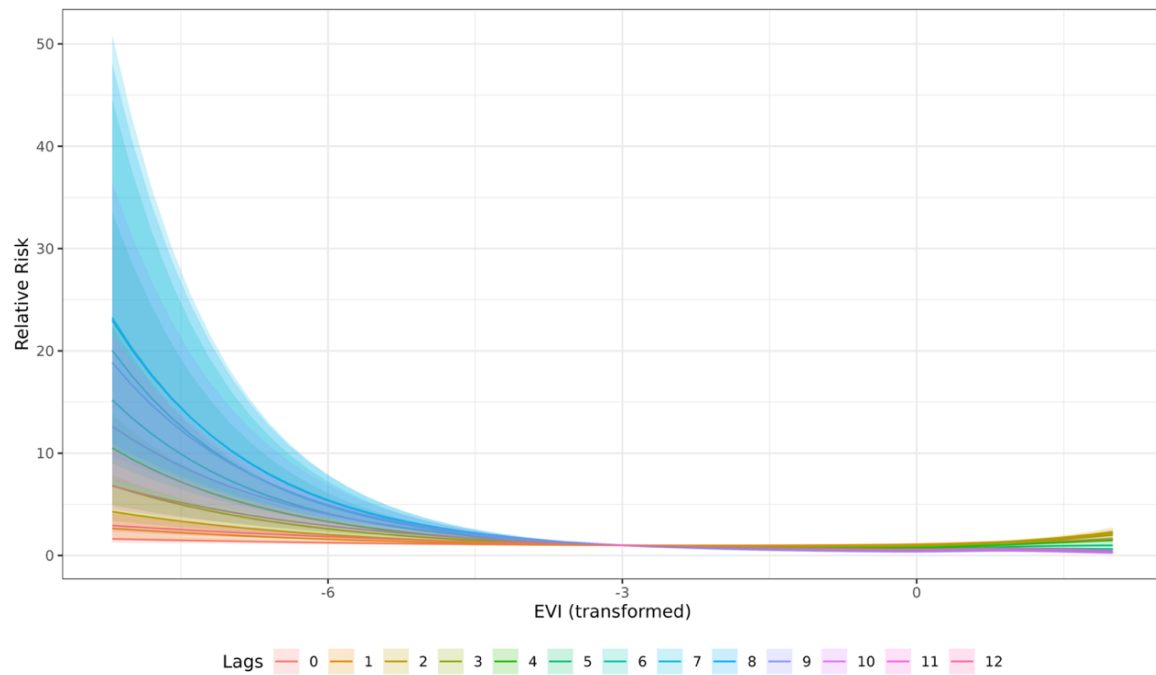


Figure 5.17-Relationship between relative risk of dengue and EVI, by lag

5.3.1.4.2 Sensitivity analysis

The regional model had lower MAE and RMSE compared to the national model, except in Bangkok for the MAE metric where the regional model performed better, Figure 5.18.

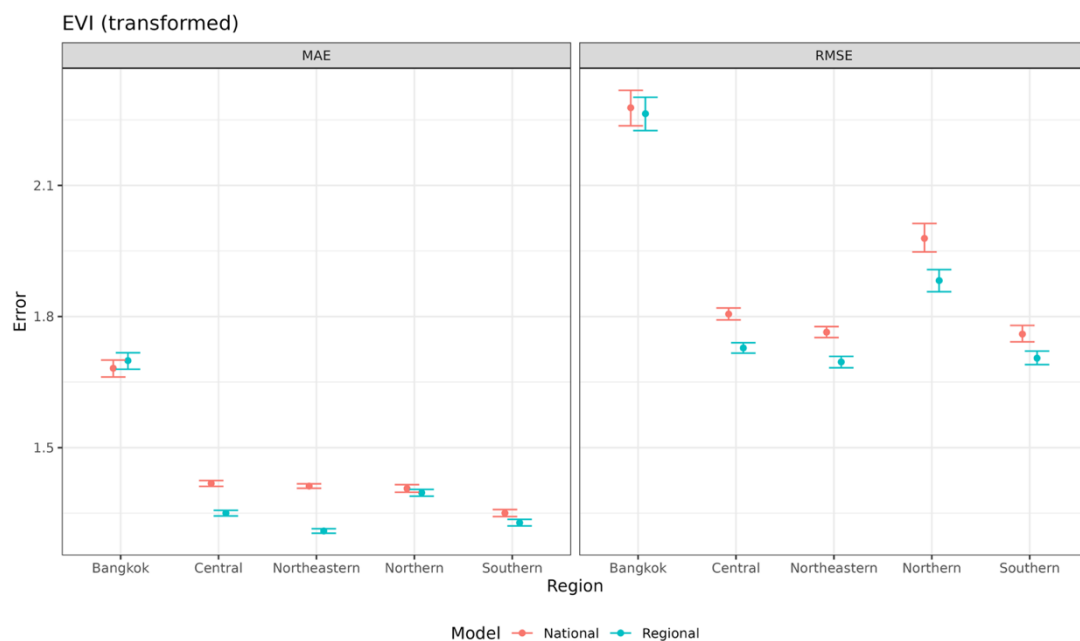


Figure 5.18-Errors of residuals for modelling dengue and EVI, by region

5.3.1.5 Genetic diversity

The correlation between genetic diversity and dengue cases by serotype is shown in Figure 5.19. No relationship was found between current genetic diversity, diversity in the previous year, or in the two years prior and the current annual cases for any serotype.

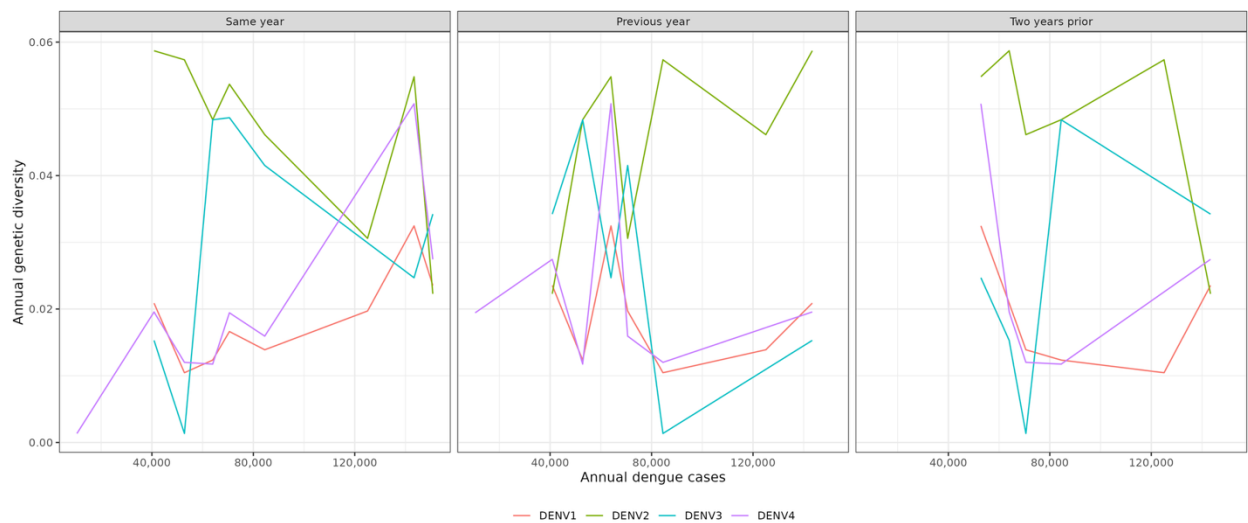


Figure 5.19-Correlation between genetic diversity and dengue cases, by serotype

5.3.1.6 Mobility

Figure 5.20 shows the spatial distribution of total incoming visits at the district level based on Meta's activity space maps. Incoming visits to districts were weighted by the average district population from 2013 to 2022. The individual incoming flows to each district were then aggregated to obtain a total number of incoming visitors to each district.

Districts in and around major urban centres, such as Bangkok, Chiang Mai, Khon Kaen, and Hat Yai, had higher incoming visits, highlighting their role as hubs of social and economic activity. In contrast, more rural districts received lower

volumes of incoming visits. This spatial heterogeneity of mobility underlines the difference in exposure to imported dengue risk, particularly in highly connected districts.

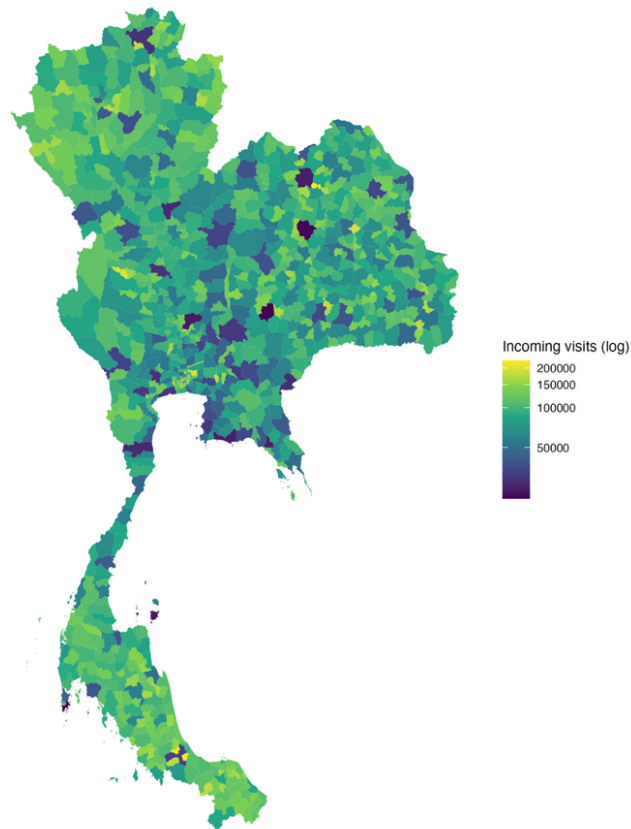


Figure 5.20-Incoming visits to districts

Figure 5.21 depicts the relationship between national incidence and correlation between district-level incidence and incoming flow from 2013 to 2022. Pearson correlations were calculated at the district level, between incidence and incoming flow for each year. The negative trend indicates that as national incidence increased, the association between mobility and dengue incidence decreased. This suggests that during periods of high national transmission, districts may already be experiencing high local transmission and thus the impact of incoming flow on dengue incidence is reduced.

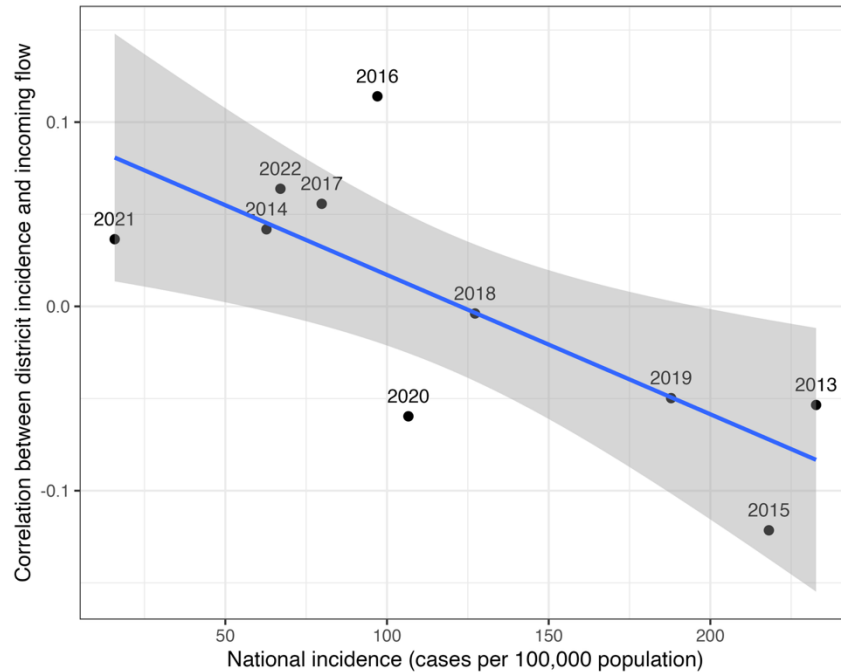


Figure 5.21-Relationship between national dengue incidence and correlation between district-level incidence and incoming population flow, 2013-2022. Blue is the trend line and grey represents 95% CI.

5.3.2 Forecasting evaluation

5.3.2.1 Forecasting with temperature

CRPS of forecasting with mean, maximum, and minimum temperature by no seasonality and seasonality model, lags 1-6 weeks and forecast horizon is presented in Figure 5.22. As forecast horizon increases, CRPS increases for all models. In the no seasonality models, the models with and without temperature covariates performed similarly. In the seasonality models, the baseline model consistently performed the best. Within temperature covariates, minimum temperature with lag 0 performed the best.

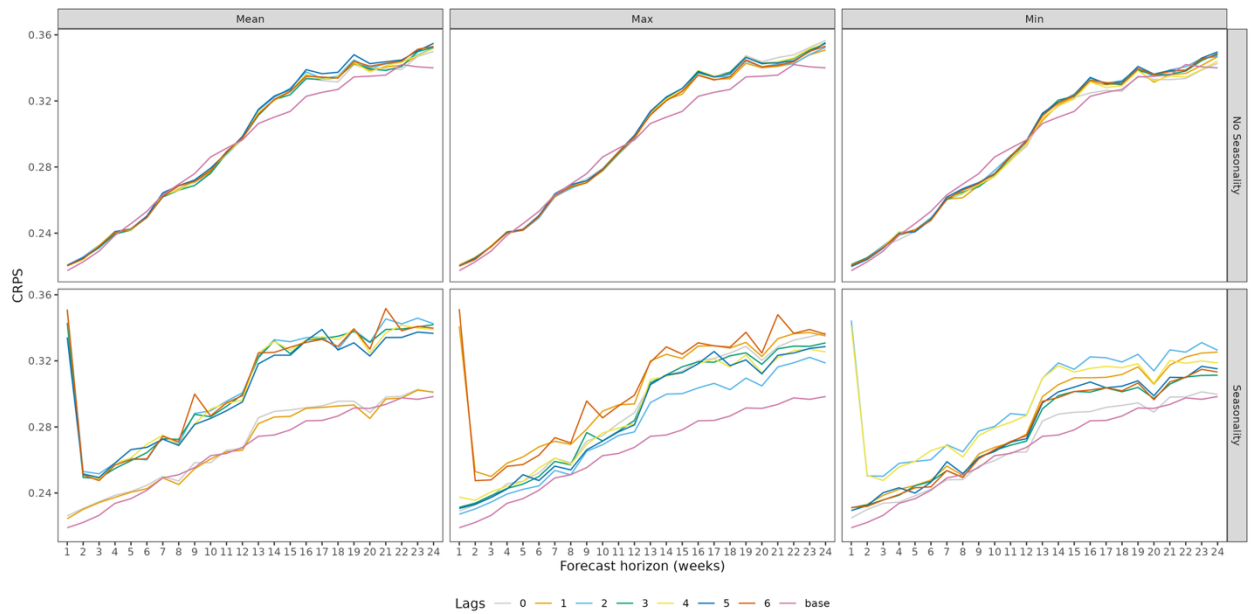


Figure 5.22-Mean CRPS of forecasts by mean, maximum, and minimum temperature, by baseline model, and lag. Lower CRPS indicates better forecasting skill.

5.3.2.2 Forecasting with SPEI

CRPS results of forecasting with SPEI with lags 0 to 6 weeks with a no seasonality and seasonality framework, compared to baseline models, is presented in Figure 5.23. In the no seasonality models, SPEI with lag 6 performed the best, while in the seasonality model and overall, seasonality base line performed best.

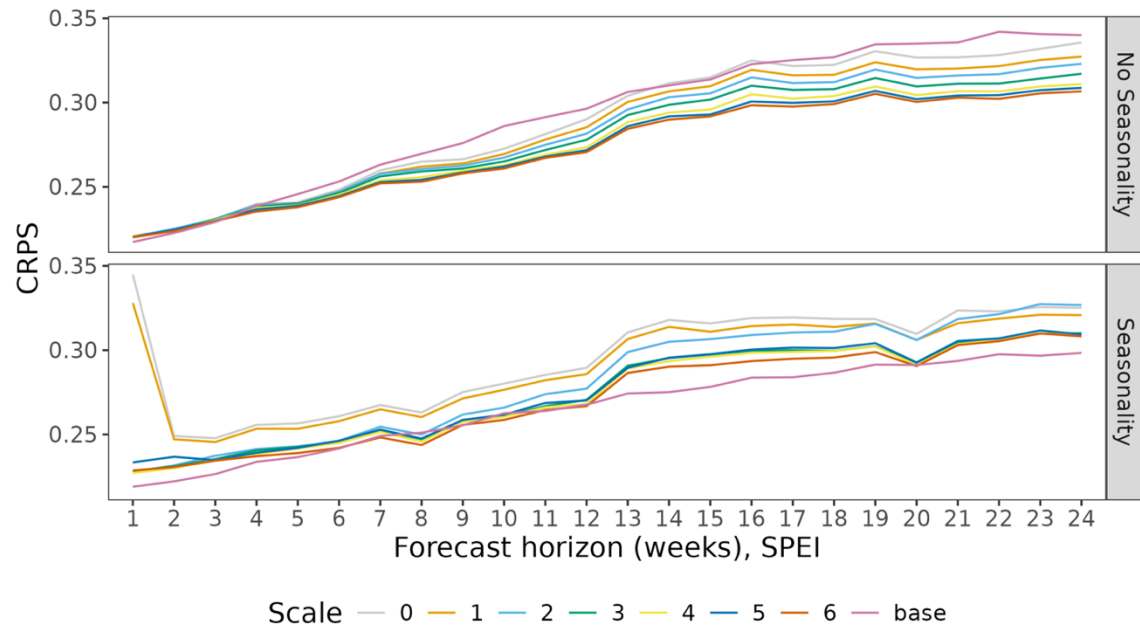


Figure 5.23-Mean CRPS of forecasts by SPEI, baseline model, and lag. Lower CRPS values indicate better forecasting skill.

5.3.2.3 Forecasting with wind speed

CRPS values for forecasting with wind speed at lags 1 to 6 and compared to baseline models of no seasonality and seasonality are presented in Figure 5.24. The no seasonality models performed very similarly, with model with wind speed with lag of 1 performing worst. For the seasonality model, the baseline model performed the best.

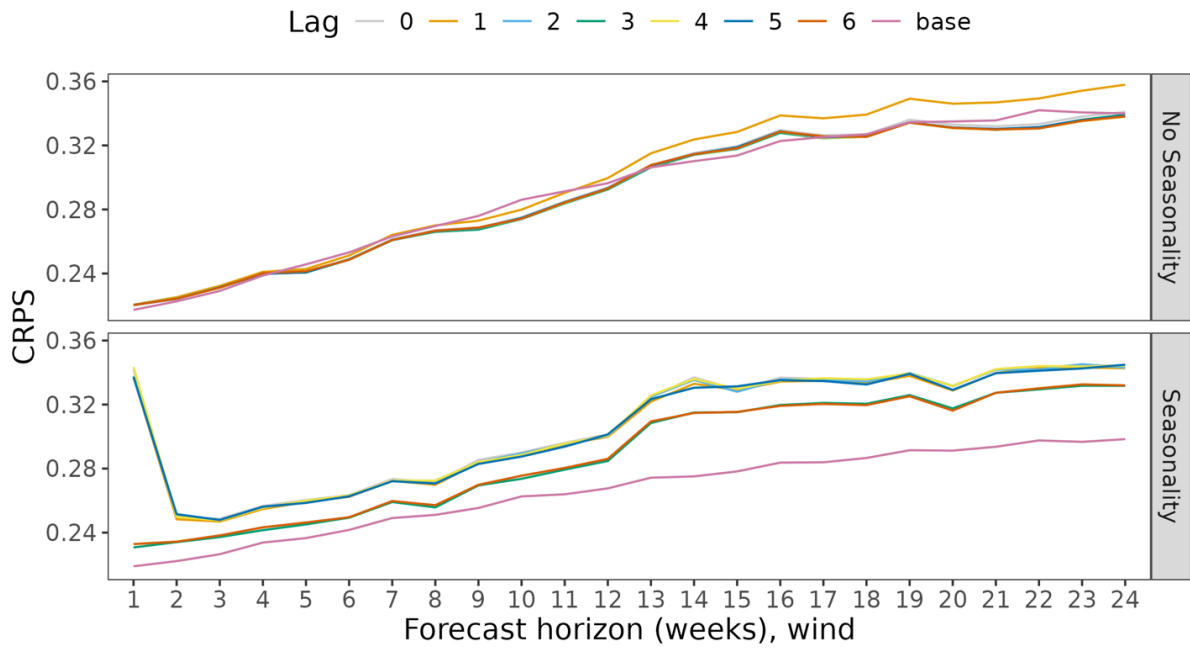


Figure 5.24-Mean CRPS of forecasts by wind, baseline model, and lag. Lower CRPS indicates better forecasting skill.

5.3.2.4 Forecasting with EVI

CRPS results of forecasting with EVI with lags 0 to 6 weeks with a no seasonality and seasonality framework, compared to baseline models, is presented in Figure 5.25. No seasonality baseline performed marginally better than other no seasonality models. Seasonality base line performed better than all other models.

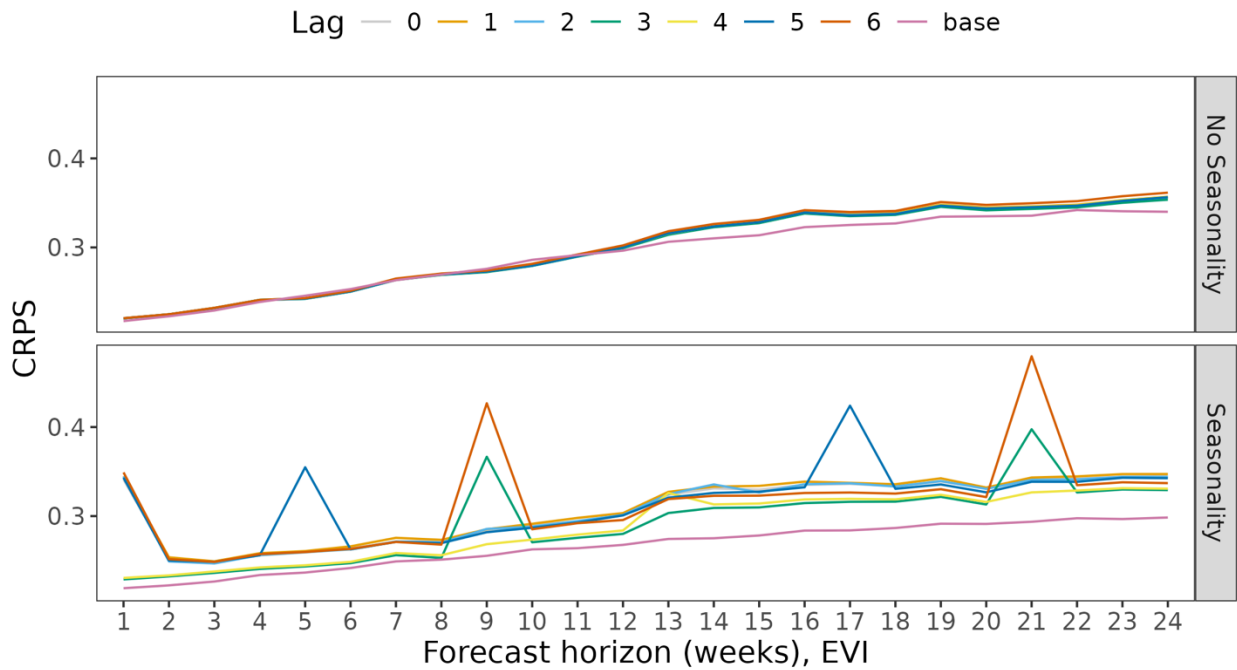


Figure 5.25-Mean CRPS of forecasts by EVI, baseline model, and lag. Lower CRPS indicates better forecasting skill.

5.3.2.5 Forecasting with genetic diversity

As genetic diversity values were only available at the annual level, forecasts were started at the first week of each year with forecast horizon up to 52 weeks. Genetic diversity values did not impact the forecasting ability of the no seasonality and seasonality models as the CRPS with and without this variable were very similar, Figure 5.26.

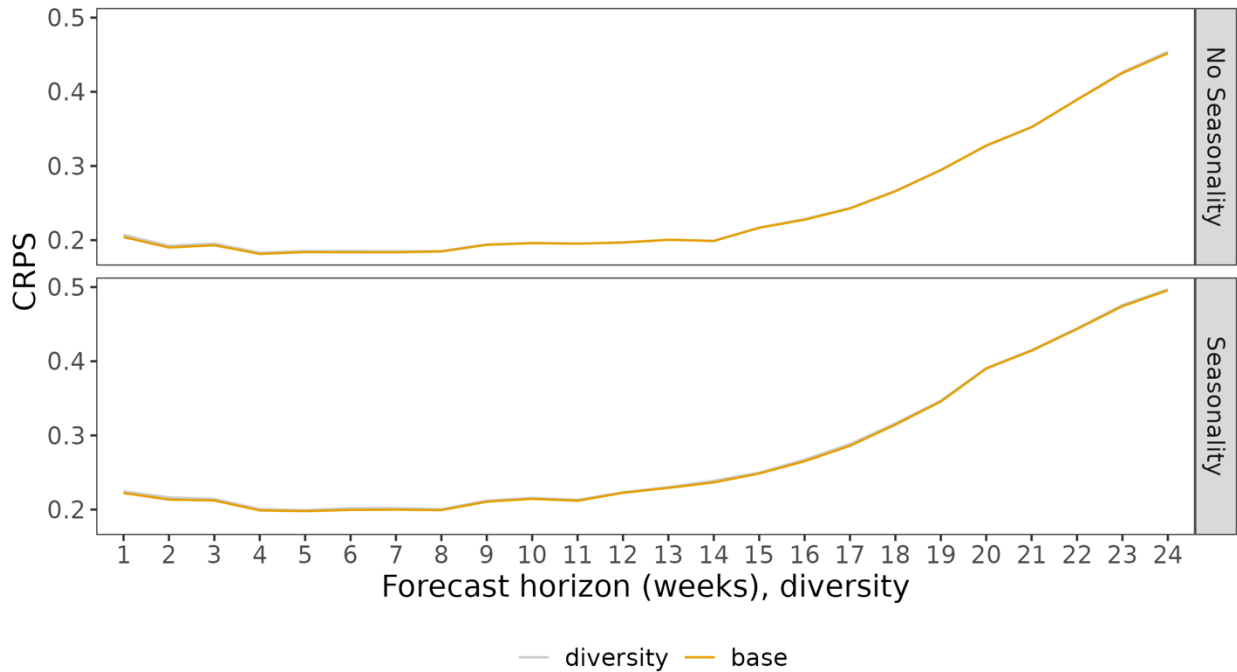


Figure 5.26-Mean CRPS of forecasting with diversity by forecast horizon. Lower CRPS values indicate better forecasting skill.

5.3.2.6 Forecasting with mobility

Including mobility in both the no seasonality and seasonality models did not improve forecasting accuracy as demonstrated by increased CRPS values by forecast horizon, compared to the baseline models, Figure 5.27 and Figure 5.28. By district, mobility increased CRPS values particularly in the northeast, west and southern regions at 24 weeks forecast horizon, Figure 5.28.

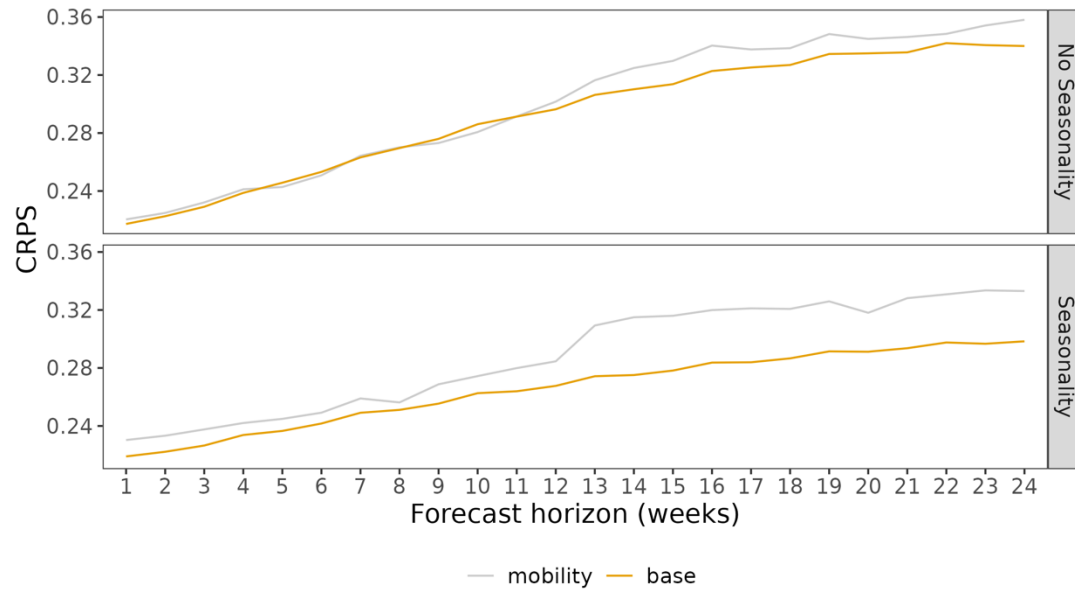


Figure 5.27-National CRPS results of mobility and baseline models, by forecast horizon. Lower CRPS values indicate better forecasting skill.

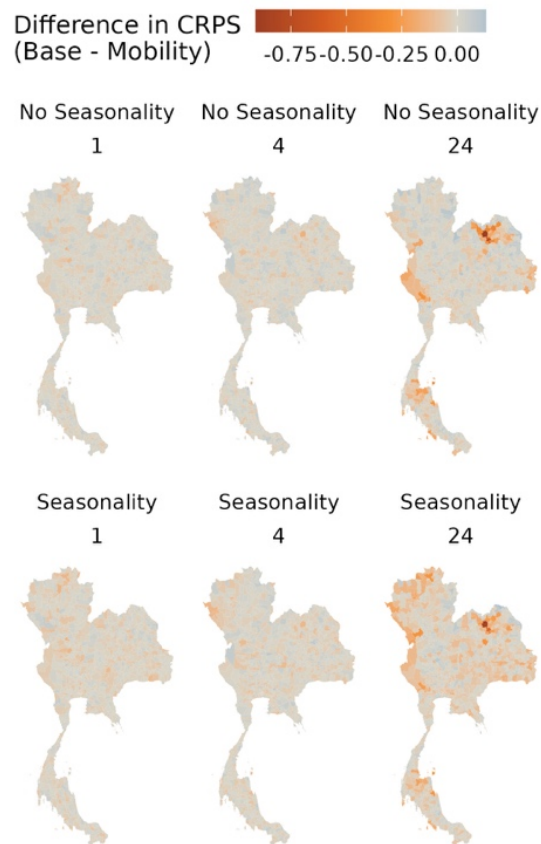


Figure 5.28-District level CRPS difference between base and mobility models, by forecast horizon.

5.4 Discussion

5.4.1 Summary of evidence

5.4.1.1 Climate variables

5.4.1.1.1 Temperature

In the relationships between temperature and RR of dengue incidence, mean and minimum temperatures showed increasing RR with increasing temperature until reaching a maximum RR after which RR decreased. The specific temperature and which maximum RR occurred differed by lag. In contrast for maximum temperature, the RR was highest at the lowest maximum temperature values and decreased with increasing temperature.

A recent systematic review by Kirk et al.[336] investigating impacts of temperature on dengue found that the highest correlation between temperature and dengue occurred around 24°C. Consistent with this result, the contour plot of minimum temperature showed that highest RR was also around 24°C. Other studies found highest association at higher temperatures, such as 27.8°C[337] and between 27°C and 31°C[102]. These findings align more closely with the highest RR.

For mean and minimum temperature, an increasing RR until a maximum RR followed by a decrease in RR by increasing temperature is supported by the literature[104, 193, 247, 336]. This pattern, however, is not the same for maximum temperature. For maximum temperature, RR decreases at temperature increases. This is mostly likely explained by the fact that maximum temperatures exceed the mosquito's optimal thermal range, reducing adult survival, biting activity, amongst other vector competence indicators.

From the forecasting results, none of the temperature variables and their corresponding lags improved the forecasts compared to the no seasonality baseline and all performed similarly. When compared to the seasonality baseline, mean temperature performed best at lags 0 and 1, minimum temperature at lag 0, and maximum temperature at lag 2, but none outperformed the baseline.

The literature on this topic is mixed. A study forecasting dengue incidence in Singapore found a mean temperature with lag 16 showed the best performance[198]. Gharbi et al.[338] found that in Guadeloupe, minimum temperature at lag of 5 weeks was best for forecasting. Using lags of 1 to 4 weeks of median temperature, a study looking at dengue forecasting in Rio de Janeiro, Brazil, found that including lagged variables did no improve forecasting skill[188].

Because CRPS results of this analysis were nearly identical across lags in the no-seasonality models and relatively small differences were observed in the seasonality models, a lag of 6 weeks was selected for inclusion in the multivariable model. This choice is intended to best reflect underlying biological processes and practicality of data availability.

5.4.1.1.2 SPEI

At the lower end of SPEI values, higher lags resulted in higher RR while at higher SPEI values, higher lags resulted in lower RR. At SPEI of -1, all lags resulted in an RR of 1.

This result is very similar to a study conducted in Southeaster China, where authors observed increased risk and both high and low SPEI levels. Looking at

lags of up to 6 months, they found that RR for extremely wet conditions was highest at lag of 1 month while for extremely dry conditions, RR was highest at lag of 6 months[131]. Similarly, Lowe et al.[133] also investigated the correlation between SPEI at lags of up to 6 months and dengue incidence. They found that risk of dengue increased up to 3 months after extremely wet conditions, with the highest RR at lag of 1 month, while RR increased 3 to 5 months after extremely dry conditions, with maximum RR at lag of 4 months.

Lower SPEI values, indicating drought conditions, can lead to water storage practices that create mosquito breeding sites, with effects accumulating over time, explaining higher risk of dengue transmission at higher lags. A study in Bangladesh found that water holding containers were significantly associated with infestation of *Aedes* larvae[339]. Lowe et al. found that in Brazil, following extreme drought, risk of dengue was higher in areas that had higher water supply shortages[133].

In higher SPEI values or very wet conditions, mosquito breeding sites are more likely to be washed away, with protective effects strengthening at longer lags, resulting in lower risk of transmission. In a study in Singapore, authors found a significant decrease in dengue outbreaks 1 to 6 weeks after a flushing event[340]. In another study in Singapore, authors found that pupae decreased by 68.5% after the onset of the monsoon[341]. A study in Pune, India, found that weeks identified as wet were positively correlated with dengue mortality, while weeks identified as flushing weeks, were negatively associated with mortality[145].

No lags improved the forecasting, when compared to the baseline, for both no seasonality and seasonality model.

For purposes of multivariable modelling, SPEI with lag of 6 weeks was chosen as it performed the best in forecasting results in both no-seasonality and seasonality models.

5.4.1.1.3 Wind speed

At higher lags, higher wind speeds were associated with higher RR while lower wind speeds showed protective effects, lower RR. At shorter lags, there was minimal association between wind speed and dengue risk.

This pattern is broadly consistent with the existing literature, which has generally reported an inverse association between wind speed and dengue incidence. A study in Guangzhou, China, found that at lags 0 and 1, wind speed had an inverse relation with dengue incidence[342]. In a study at a municipality in Indonesia, authors also found negative relationship between wind speed and dengue incidence[343]. In a study in Bangkok, authors found that at lag 0, there was a negative association between wind and dengue transmission[247]. Studies in Bangladesh, Malaysia, and Paraguay found a negative relationship while a study in Peru found a positive relationship[156]. A study in Singapore found that at a lag of 0, wind speeds of around 5-7.5 km/h resulted in an increased RR while a lag of 5 weeks resulted in a negative effect on RR[193].

Based on the forecasting results, compared to the no seasonality baseline, all lags performed similarly except lag 1, which had the highest CRPS. In certain forecast horizons, model with covariate performed better than baseline. When compared to the seasonality model, lags 3 and 6 performed better than other lags, but neither outperformed the baseline.

These findings align with the variability in utility of wind speed in forecasting research. In a study in Guangdong and Zhejiang in provinces in China, predicting dengue cases, wind speed with lag 0 and lag 7 months were used, respectively[344]. In Vietnam, wind speed at lag 0 was used for forecasting[226].

For purposes of multivariable modelling, wind speed with lag of 6 weeks was chosen as it performed the best in forecasting results in both no-seasonality and seasonality models.

5.4.1.2 Environmental variables

5.4.1.2.1 EVI

As EVI increased, RR decreased. Higher lags resulted in higher RR at lower EVI values.

This finding aligns with some of the previous research. In studies in Taiwan[345], Belo Horizont [346] and Sao Paolo[347], both in Brazil, the authors found a negative correlation between greenness and dengue cases. A study in Lao PDR investigated the relationship between NDVI at lags 0 to 6 and dengue cases. They found the association was strongest at lag of 1 and as NDVI increased, dengue cases decreased[163]. In a study across 12 provinces in Vietnam, NDVI was found to have both negative and positive association with dengue at lags of 4 and 5 months[348]. A study in central India found that NDVI was positively correlated with dengue cases at a lags of 0 and 1 month, while at longer lags (>4 months), lower NDVI values were associated with higher dengue risk, suggesting an inverse relationship[349].

In forecasting using no seasonality baseline, all EVI lags performed similarly and neither better than the baseline. In the seasonality models, lag 4 performed the best but neither lag was better than the baseline.

Across studies, impact of vegetation on forecast performance is variable. In a study in Colombia, using current and previous 11 weeks of EVI, improved short-term forecasting[202]. A study in Peru created a model to forecast dengue outbreak using both NDVI and EVI values at current and previous 12 weeks[350].

For multivariable analysis, based on the results, EVI with lag 4 was selected.

Variable		Summary of DLNM results	Summary of forecasting results	Variable and lag to be included in multivariable model
Temperature	Mean	Higher lag, high RR	Lag 0, though all did worse than baseline	NA
	Minimum	Lag 6 highest RR	Lag 0, though all did worse than baseline	Lag 6
	Maximum	Higher lag, higher RR but relationship not as straightforward	Lag 2, though all did worse than baseline	NA
SPEI		Lag 8	Lag 6, though all did worse than baseline	Lag 6
Wind		Higher lags, more extreme RR	Lag 6, though all did worse than baseline	Lag 6
EVI		Higher lag, higher RR at lower EVI	Lag 4, though all did worse than baseline	Lag 4

Table 5.1-Summary of findings and variables to be included in multivariable models.

5.4.1.3 Mobility

Including mobility in the model did not improve forecasting, but it may still be useful to include it in a multivariable model to account for confounding.

This result was surprising as previous research has found that including mobility improved forecasting skill. One explanation may be the limited temporal coverage of the mobility data used in this analysis, which may reduce its ability to capture longer-term dynamics. For example, studies using mobile phone data in Pakistan and Thailand found improved forecasting results when mobility was included[174, 175]. Similarly, a study in Peru found that including public transportation data improved forecasting performance[179]. In Brazil, multi-modal transportation data, including air, road, and water travel, was found to improve both dengue incidence and dengue outbreak forecasting[351].

5.4.2 Limitations

There were some limitations with this analysis. This analysis is at a finer spatial scale than previous research which can mean more noise in both dengue data and covariates which can limit the ability of the model to find meaningful relationships. Further work would be helpful to quantify the advantage of using a model at this resolution rather than a lower resolution.

To allow a seasonality component to be added to the formula, years with week 53 were artificially adjusted to have 52 weeks. This may affect the model fit and prediction in real life for the end and beginning of the year, particularly in those years with 53 weeks. In the future, better ways to adjust for seasonality without

modifying the data should be explored such as defining each week in a continuous or fractional time scale rather than categorical.

The forecasting analysis was completed using observed covariates. Only the dengue data was unknown to the model. Therefore, the findings were overestimated. Further work is required to assess the changes in forecasting accuracy when using predicted covariates.

There were also other factors that could impact dengue transmission that were not investigated in this chapter due to lack of suitable available data. These include but are not limited to access to health care, reporting biases, and changes in human behaviour as result of dengue infection itself or changes in climate and environmental variables.

Since each variable was examined individually, the effects of confounders and effect modifiers were not considered. Omitting these factors may lead to misleading estimates of each variable's impact on dengue transmission. The following chapter will investigate multivariable models that adjust for confounders and effect modifiers. Including multiple covariates rather than analysing them in isolation can improve forecasting performance by better capturing spatiotemporal structures, integrating multiple weak signals, and borrowing strength across districts.

To assess the impact of all variables in one model on forecasting accuracy, minimum temperature at lag 6 weeks, wind at lag 6 weeks, SPEI at lag 6 weeks, EVI at lag 4 weeks, and mobility will be assessed in a no seasonality and seasonality framework.

6 Multivariable Analysis

6.1 Introduction

In this chapter, multivariable models will be explored for forecasting dengue transmission. Variables and their associated lags identified in the previous chapter will be used to inform multivariable models. Impact of interactions will also be investigated as they may reveal patterns that could be overlooked when investigating the impact of only one variable.

In univariate models, only one covariate was included at a time. While this approach provided insights into the relationship between each covariate and dengue incidence, the true relationship may be confounded by a third variable that is associated with both the covariate and the outcome but is not part of the causal pathway[352, 353]. This can lead to overestimation or underestimation of the relationship between the first covariate and the outcome[353]. Adjusting for confounders can capture additional information and potentially improve model performance[354].

It is also important to consider effect modifiers. The effect of variables may change at different levels of a third variable, known as an effect modifier[355]. This occurs when there is a biological difference between the effect of the covariate on the outcome, by different levels of the effect modifier. This relationship is captured by including interaction terms between the covariate and effect modifier in the model[356]. This interaction term can improve the predictive outcome of the model by improving the flexibility of the model and allowing for a more precise prediction for subgroups[218].

For the purposes of this thesis, I will consider three interaction terms: interaction between minimum temperature and SPEI, interaction between enhanced vegetation index (EVI) and genetic diversity, and interaction between minimum temperature and wind speed.

From my systematic review, previous research has demonstrated interaction between temperature and precipitation is important to consider for predicting dengue transmission. Warmer temperatures increase biting activity of *Aedes* mosquitoes[357] and shorten EIP[101, 358]. Precipitation can increase availability of larval habitats, leading to increased mosquito abundance[112-116]. However, high precipitation can flush out breeding sites [118-120] while decreased dengue risk during low precipitation periods may be offset by water storage practices that provide habitats for larval growth[133]. The relationship between temperature and precipitation is more nuanced. In a laboratory-based experiment, Alto and Juliano[111] investigated how varying temperatures (22, 26, and 30°C) and precipitation regimes affected *Aedes albopictus*. They used temperature-controlled cages with varying water re-filling schedules to simulate different precipitation scenarios. Their results showed that prevention of drying at higher temperatures led to greatest production of adults while drying led to lower production of adults, because of mortality of aquatic stages. At the lowest temperature, precipitation had no effect on adult production.

Interaction between EVI and genetic diversity can reveal important patterns. *Aedes aegypti*, the main vector of dengue in Thailand[359], thrives in urban areas [359-361]. It has been shown that *Ae. aegypti* abundance[362, 363] and dengue transmission[346] are negatively associated with high values of EVI or greenness. Higher mosquito populations can lead to higher transmission rates, potentially

influencing genetic diversity. In Thailand, Salje et al. [364] found that urban areas had a higher dengue genetic diversity than rural areas. An interaction term between EVI and genetic diversity would examine whether the impact of genetic diversity on transmission is different between urban or low EVI settings and rural or high EVI settings. This could occur if urban environments with higher mosquito densities enhance transmission of circulating viral strains, while in rural areas, the transmission may be lower despite similar viral diversity.

Yet another interaction to consider is between wind speed and temperature. Higher wind speed has been shown to reduce dengue transmission[193, 276, 365-367]. Wind affects dengue transmission by impacting the distribution of the vector. High winds can disturb the flight of the mosquito, making finding a host and oviposition more difficult[148-150]. Moderate winds, on the other hand, may enhance dengue transmission by aiding the dispersal of the vector farther than it can fly[368]. A study of the release and recapture of *Ae. aegypti* found that wind speed impacted recapture frequency[369]. At optimal temperatures, dengue vectors exhibit highest blood feeding rates and female longevity, contributing to higher transmission[370]. However, wind may dampen the positive impacts of temperature by physically disrupting the mosquito's flight behaviour. The interaction between wind and temperature captures whether the impact of temperature on transmission is modified by wind speed.

Mobility was also included in the multivariable models investigated as it can be a confounding factor between climate and environmental variables and dengue transmission. Mobility has been shown to be correlated with dengue transmission. Wesolowski et al.[174] used mobile phone data in addition to climate information to predict the spread of dengue in Pakistan while Kiang et al.[175] found that

mobility correlated with increased dengue cases in Thailand, where provinces that were more strongly connected as shown by mobile phone data had a higher correlation with dengue cases than less connected provinces. As an example of a natural experiment, Chen et al.[371] found COVID-19-related mobility restrictions in Thailand resulted in an aversion of 66% (95% CI: 14-90%) of cases in 2020 whereas mobility restrictions in other countries, such as Belize and Nicaragua, resulted in an increase of cases. Other studies found a decrease in dengue cases in Fortaleza, Brazil [372] and in Sri Lanka[373]. These results, however, must be viewed with caution as there may be biases that are difficult to correct for, such as reduced health-seeking behaviour due to restrictions. Vector control measures such as outdoor spraying were also disrupted, leading to an increase in cases. Cavany et al[374] investigated the impact of the decrease in vector control measures on populations sheltering at home in Iquitos, Peru, and found that this change led to an increase in cases. Based on these results, even though some are conflicting in their conclusions, it is evident that mobility influences dengue transmission.

The aim of this chapter is to assess the predictive ability of multivariable models of dengue incidence and outbreaks by including variables previously investigated in univariate models (Chapter 5) as well as selected interaction terms. Model performance is evaluated by comparing the forecasting ability of each model to a baseline model (Chapter 4), across forecast horizons and by target weeks. Additionally, the effect sizes of each covariate are investigated to better understand the impact of confounding and effect modification in the models.

6.2 Methods

Covariates and interaction terms were added to baseline models of no seasonality and seasonality as fixed effects rather than random effects as was done in the univariate chapter (Chapter 3). This is because as the number of parameters increased, the fitting process became more computationally cumbersome. Therefore, using fixed effects ensured that the models converged in a timely fashion. For example, when minimum temperature at lag of 6 weeks was included as a random effect in the seasonal model, computation time was 6.2 minutes, nearly three times longer than the 2.1 minutes required when the same covariate was specified as a fixed effect.

Four models, each with a no seasonality and seasonality base, were considered. The first model included the covariates of minimum temperature at lag of 6 weeks, SPEI at lag of 6 weeks, wind speed at lag of 6 weeks, EVI at lag of 4 weeks, genetic diversity of dengue serotypes 1, 2, 3 and 4, and finally a mobility precision matrix previously explained in Chapter 3.

The second model was like the first but included an interaction between EVI and genetic diversity measures. The third model was also built upon the first model but included an interaction between minimum temperature and SPEI. The fourth model, also built upon the first model, included an interaction between minimum temperature and wind speed.

Formula 1a

$$\begin{aligned} \log(\lambda) = & \alpha + W_{t,p} + \beta_1 \cdot \text{tmin}_6 + \beta_2 \cdot \text{spei}_{\alpha_4} + \beta_3 \\ & \cdot \text{wind speed}_6 + \beta_4 \cdot \text{evi}_4 + \sum_{i=1} \gamma_i \\ & \cdot \text{pi}_i + f_{\text{mobility}} + u_d + v_d \\ & + \log(\text{pop}) \end{aligned} \quad (16)$$

Where f_{mobility} denotes a random effect modelled using generic0 and a user-defined precision matrix, as previously described in the Descriptive chapter.

Formula 2a

$$\begin{aligned} \log(\lambda) = & \alpha + W_{t,p_4} + \beta_1 \cdot \text{tmin}_6 + \beta_2 \cdot \text{spei}_6 + \beta_3 \cdot \text{wind speed}_6 \\ & + \sum_{i=1} \delta_i \cdot (\text{evi}_4 \cdot \text{pi}_i) + f_{\text{mobility}} + u_d + v_d + \log(\text{pop}) \end{aligned} \quad (17)$$

Formula 3a

$$\begin{aligned} \log(\lambda) = & \alpha + W_{t,p} + \beta \cdot (\text{tmin}_6 \cdot \text{spei}_6) + \beta_3 \\ & \cdot \text{wind speed}_6 + \beta_4 \cdot \text{evi}_4 + \sum_{i=1} \gamma_i \cdot \text{pi}_i \\ & + f_{\text{mobility}} + u_d + v_d + \log(\text{pop}) \end{aligned} \quad (18)$$

Formula 4a

$$\begin{aligned} \log(\lambda) = & \alpha + W_{t,p} + \beta_2 \cdot \text{spei}_6 + \beta \\ & \cdot (\text{tmin}_6 \cdot \text{wind_speed}_6) + \beta_4 \cdot \text{evi}_4 + \sum_{i=1}^4 \gamma_i \\ & \cdot \text{pi}_i + f_{\text{mobility}} + u_d + v_d + \log(\text{pop}) \end{aligned} \quad (19)$$

Formulas 1b to 4b had an additional term of $Q_{t,r}$ which accounts for the seasonality effect at the regional level.

To ensure equal comparison with univariate results, the selected covariates were ran as fixed effects univariately and compared to the coefficients from multivariable models to assess change in values as a result of confounding and effect modifiers. Presence of confounding and effect modification was assessed by comparing effects sizes of covariates between models to identify any statistical difference.

6.2.1 Sub-analyses

Sub-analyses, assessing the forecasting ability of the model by season (rainy or dry) and outbreak status was conducted. These additional analyse will help identify whether the model is generalizable to different periods.

6.3 Results

6.3.1 All years

6.3.1.1 Forecast horizon

CRPS at the national level by forecast horizon is presented in Figure 6.1. In the models with no seasonality, the multivariable models performed better than the

baseline in all forecast horizons except in weeks one to three. This difference however is minimal. In the models with seasonality, the results showed a different pattern, where the baseline performed better than the multivariable models for all forecast horizons except week eight which may be a result of randomness. Comparing both groups of no seasonality and seasonality models, the group of seasonality models performed better and specifically, the seasonality baseline model performed best.

In the no seasonality panel, the covariates may have provided some of the seasonality information that may have been missing, thus resulting in a better performance than the no seasonality baseline. On the other hand, in the seasonality panel, the covariates may have already correlated with the seasonality term and added additional noise, decreasing performance.

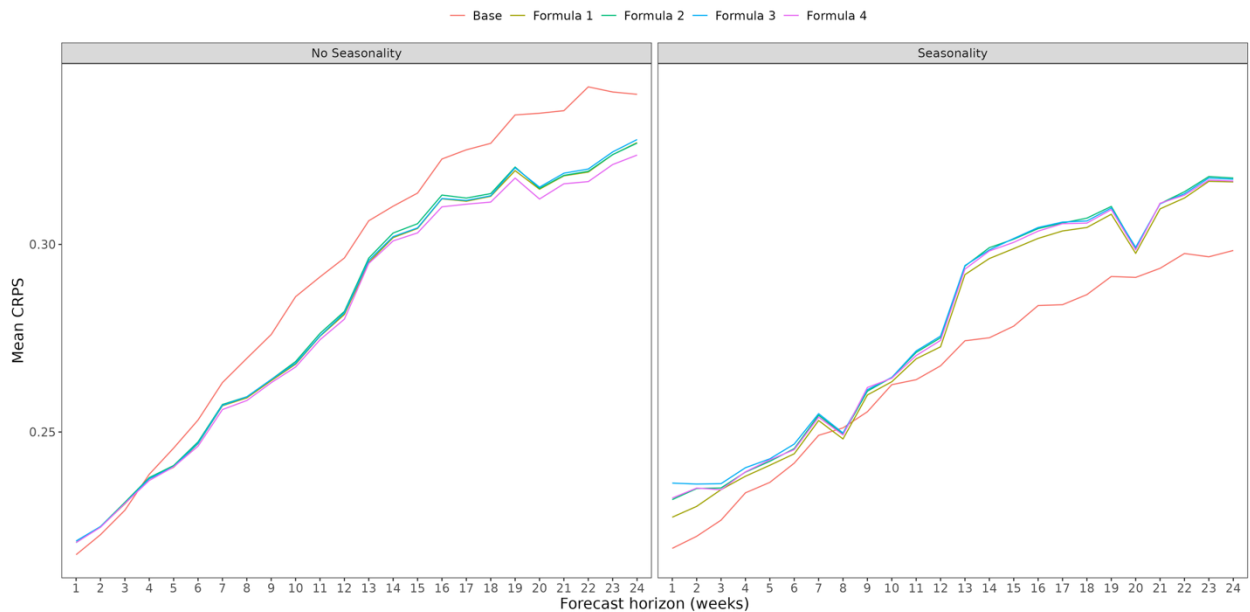


Figure 6.1-CRPS by forecast horizon and model. The left panel shows the models with a no seasonality base and the right panel shows the models with a seasonality base.

6.3.1.2 Target week

Forecasting accuracy by target week and forecast horizon is shown in Figure 6.2. As forecast horizon increased, the difference between the multivariable models and the baselines increased. As also shown in Figure 6.1, for no seasonality models, at 24 weeks forecast horizon, multivariable models performed better than the baseline. In contrast, for the seasonality model at 24 weeks forecast horizon, the multivariable model did not consistently perform better than the baseline model.

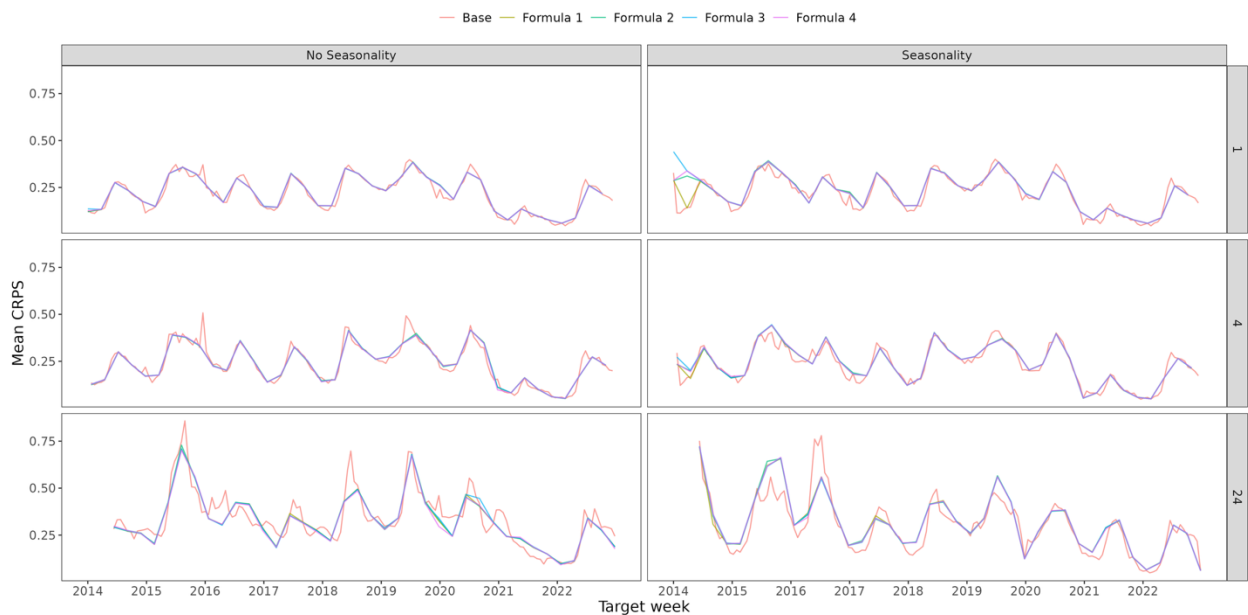


Figure 6.2-CRPS by target week, forecast horizon, and model.

6.3.2 Seasonal status

Rainy season was designated as being from May to October while the dry season was between November and April, inclusive. All models performed better in dry seasons than in rainy seasons (Figure 6.3). In dry seasons, multivariable models performed better than baseline in no seasonality models and performed worse in seasonality models.

In rainy seasons, multivariable models performed similar to the no seasonality and seasonality baselines, with some improvement of the seasonality baseline from week eleven onwards. Overall, models in dry seasons performed better than their rainy season counterparts and the seasonality model in dry seasons performed the best.

As all multivariable models performed similarly, formula 1 was selected as an example for inclusion for the following analysis. Figure 6.4 shows the difference in CRPS between the baseline models and the corresponding formula 1, by district and by dry and rainy seasons. Like Figure 6.3, for the no seasonality models in the dry season, as forecast horizon increases, the difference between baseline model and multivariable model increases, with the multivariable model outperforming the baseline model in 24 weeks of forecast horizon. In the rainy season, the difference between the base and multivariable models by districts is not as pronounced, though there are some districts with strong performance by the baseline model (orange and dark orange in west, south, and northeast areas) and some districts with strong performance by the multivariable model (dark blue in northern areas).

In the seasonality models, the differences between the base and multivariable models increased as forecast horizon increased. At 24 weeks forecast horizon, the difference in the models was most pronounced in the rainy seasons with some districts (orange and dark orange in west, northwest, northeast, south) showing better performance by the baseline model with some areas (blue, central) showing better performance by the multivariable model.



Figure 6.3-CRPS by forecast horizon, season, and model.

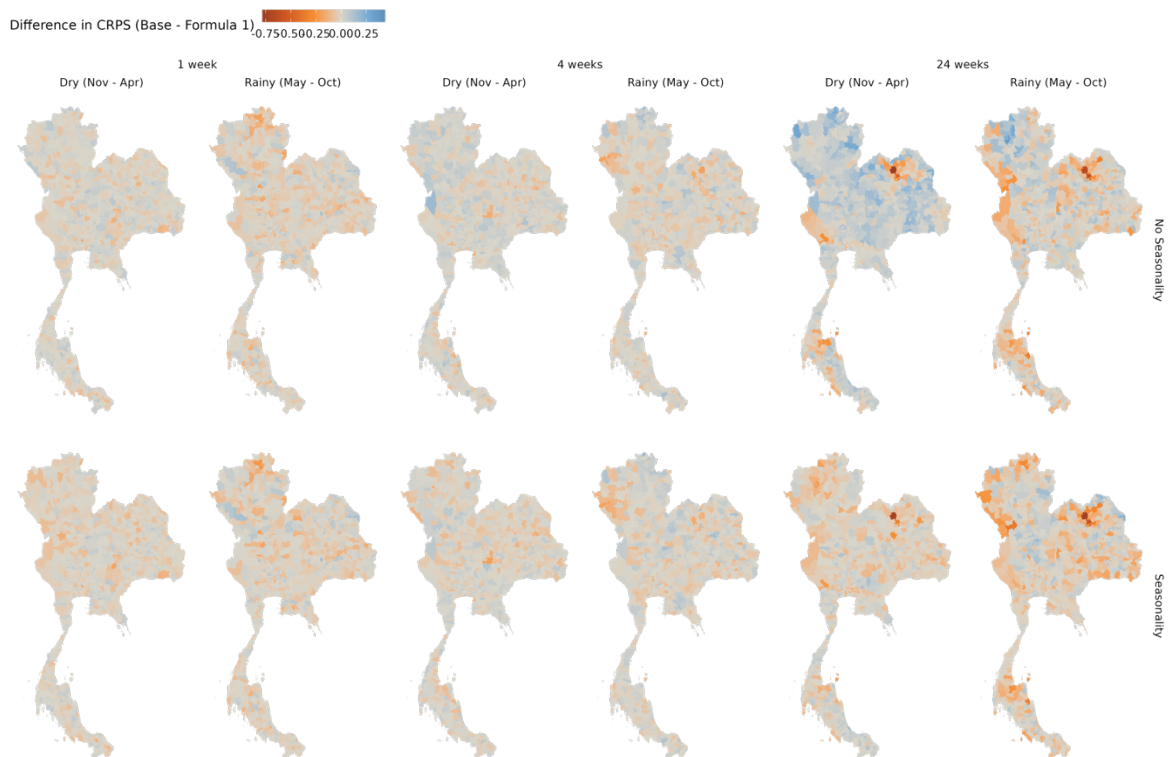


Figure 6.4-Difference in CRPS between baseline model and model with all covariates, by rainy season, and forecast horizon (1, 4, and 24 weeks).

6.3.3 Outbreak status

Each week's outbreak status was determined using the previously described method in Chapter 3. If the number of cases in a week was higher than the median of the past five years for the same week, then that week was considered to have an outbreak. CRPS results were averaged by outbreak status of the target week are presented in Figure 6.5. These results only include target weeks from 2018 onwards as five years of historical data was required to identify outbreaks. CRPS results for all models were lower when there was no outbreak. Comparing the baseline model and multivariable models when there were no outbreaks, the multivariable model performed slightly better than the baseline in the no seasonality models while it performed very similarly to the baseline in the seasonality models. In outbreak settings, the no seasonality models performed very similarly, while in the seasonality models, the baseline model performed better, with difference between the models increasing as forecast horizon increased.

Results of the CRPS difference between baseline and multivariable models, by district and forecast horizon are presented in Figure 6.6. In no-outbreak settings, the difference between the baseline and multivariable models in both the no seasonality and seasonality and for all forecast horizons was minimal, with increase in difference by increase in forecast horizon. In outbreak settings, the difference between the two models by no seasonality and seasonality models also increased by forecast horizon. In the no seasonality models at 24 weeks forecast horizon, there was not a discernible pattern for which districts performed better based on which model. In the seasonality models, the baseline model performed

better in more districts as is evidenced by more orange-coloured districts in Figure 6.6 but the pattern was not clear.

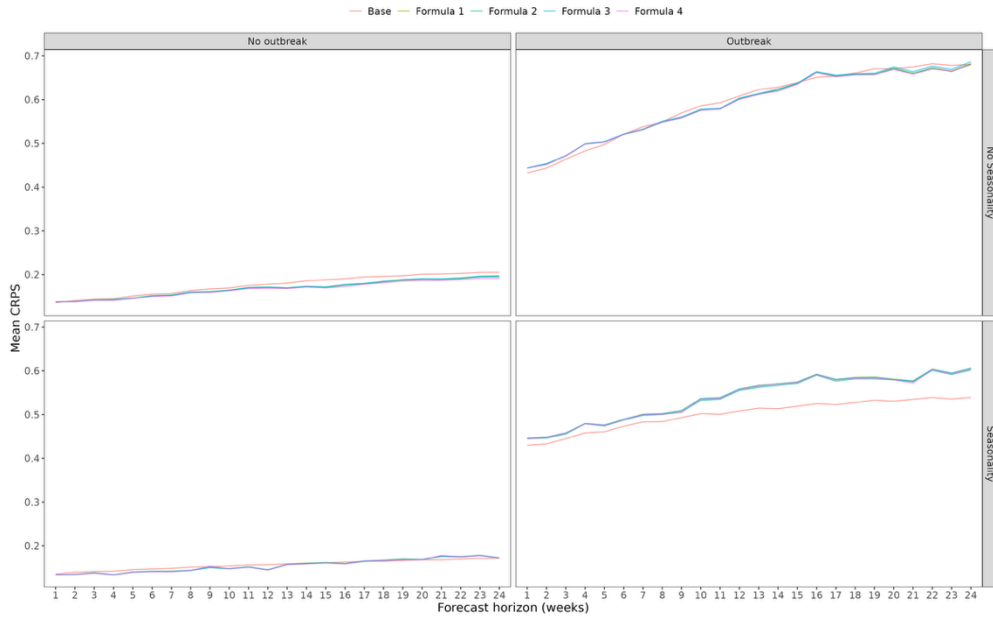


Figure 6.5-CRPS by forecast horizon, outbreak status, and model.



Figure 6.6-CRPS difference between baseline model and model with formula 1, by forecast horizon, outbreak status, and baseline model. Dark grey indicates districts that did not experience outbreaks.

6.3.4 Effect estimates

Effect estimates of variables for the different models are presented below. Figure 6.7 and Figure 6.8 are the effect estimates for all variables for the no seasonality and seasonality baseline models respectively, while Figure 6.9 and Figure 6.10 do not include the genetic diversity related terms. The reason for this was to provide a better view of the non-diversity terms as the estimates of the diversity terms were much larger than the rest of the covariates, making visualisation difficult.

In all sets of formulas (no seasonality and seasonality), genetic diversity for serotype 2 had the largest absolute and the most negative effect. This was followed by genetic diversity for serotype 3 which had the largest positive effect.

Within the climate covariates, SPEI with lag of six weeks had the highest absolute and positive coefficients for both no seasonality and seasonality models. When interaction with minimum temperature with lag of six weeks was included, the SPEI coefficient decreased. Minimum temperature with lag of six weeks for all models had coefficients very close to zero. Wind speed with lag of 6 weeks had small negative effect estimates for all models, while the model with interaction of wind speed and minimum temperature had the largest effect. Addition of this interaction term resulted in a decrease in the effect estimate for wind speed. In the seasonal models, the effect estimates of SPEI and wind speed in formulas 3 and 4 had a larger range than the no seasonality models. This may be due to the collinearity between these variables and the seasonality term, resulting in instability.

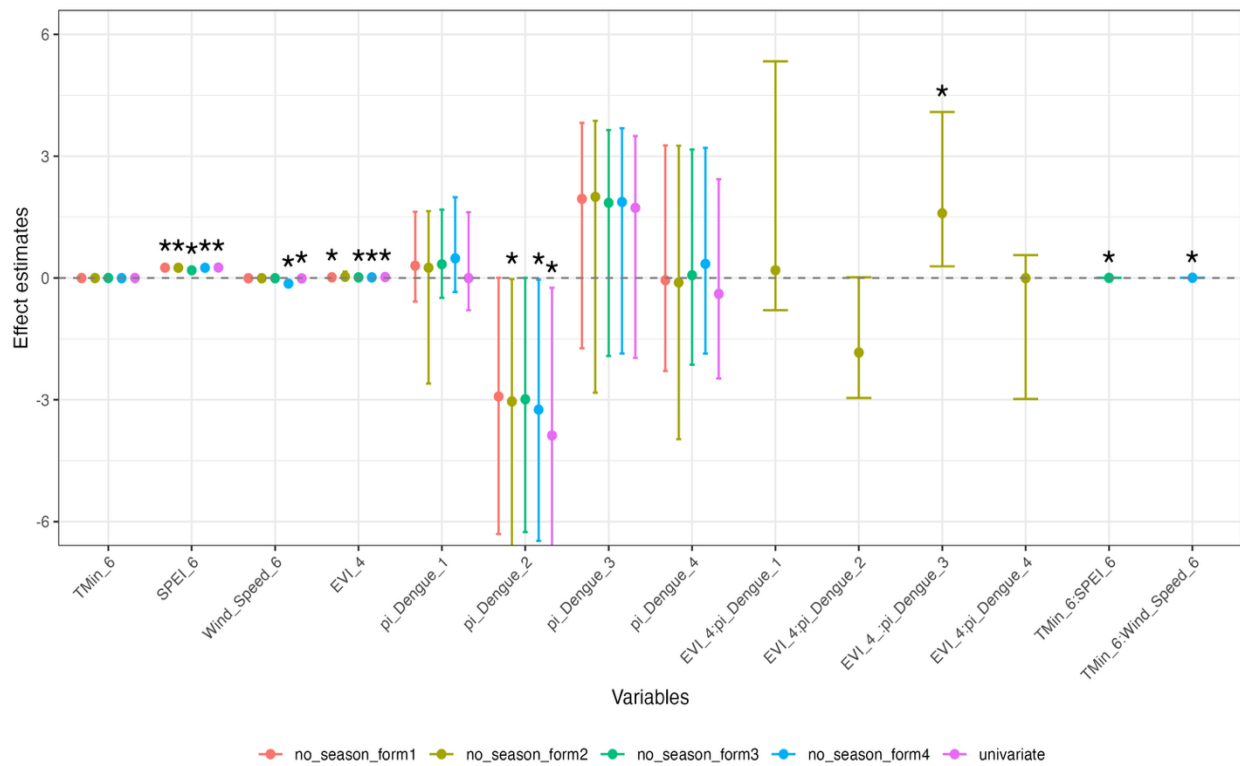


Figure 6.7-Effect estimates by formula, for a no seasonality base model. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, pi_Dengue_1 is genetic diversity for serotype 1, pi_Dengue_2 is genetic diversity for serotype 2, pi_Dengue_3 is genetic diversity for serotype 3, pi_Dengue_4 is genetic diversity for serotype 4, EVI_4:pi_Dengue_1 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 1, EVI_4:pi_Dengue_2 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 2, EVI_4:pi_Dengue_3 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 3, EVI_4:pi_Dengue_4 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 4, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, and TMin_6:Wind_Speed_6 is the interaction term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.

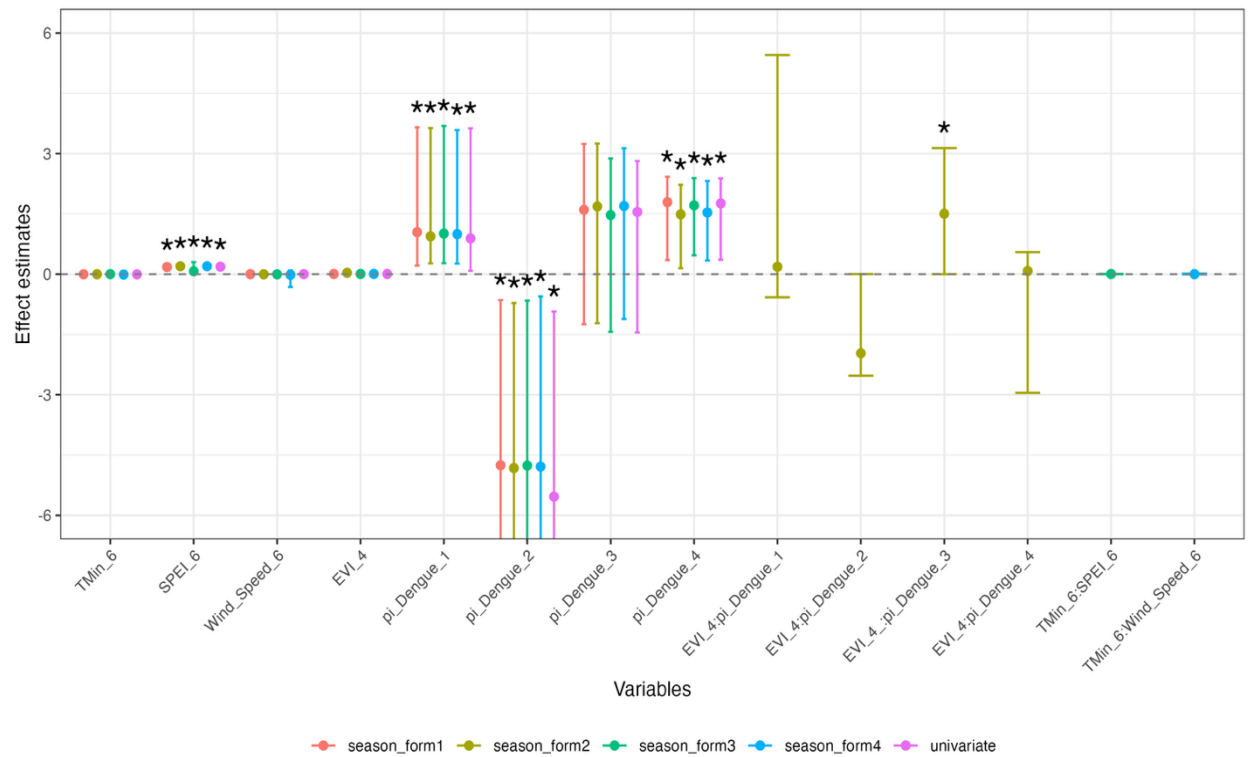


Figure 6.8- Effect estimates by formula, for a seasonality base model. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, pi_Dengue_1 is genetic diversity for serotype 1, pi_Dengue_2 is genetic diversity for serotype 2, pi_Dengue_3 is genetic diversity for serotype 3, pi_Dengue_4 is genetic diversity for serotype 4, EVI_4:pi_Dengue_1 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 1, EVI_4:pi_Dengue_2 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 2, EVI_4:pi_Dengue_3 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 3, EVI_4:pi_Dengue_4 is the interaction term between EVI at lag 4 weeks and genetic diversity of serotype 4, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, and TMin_6:Wind_Speed_6 is the interaction term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.

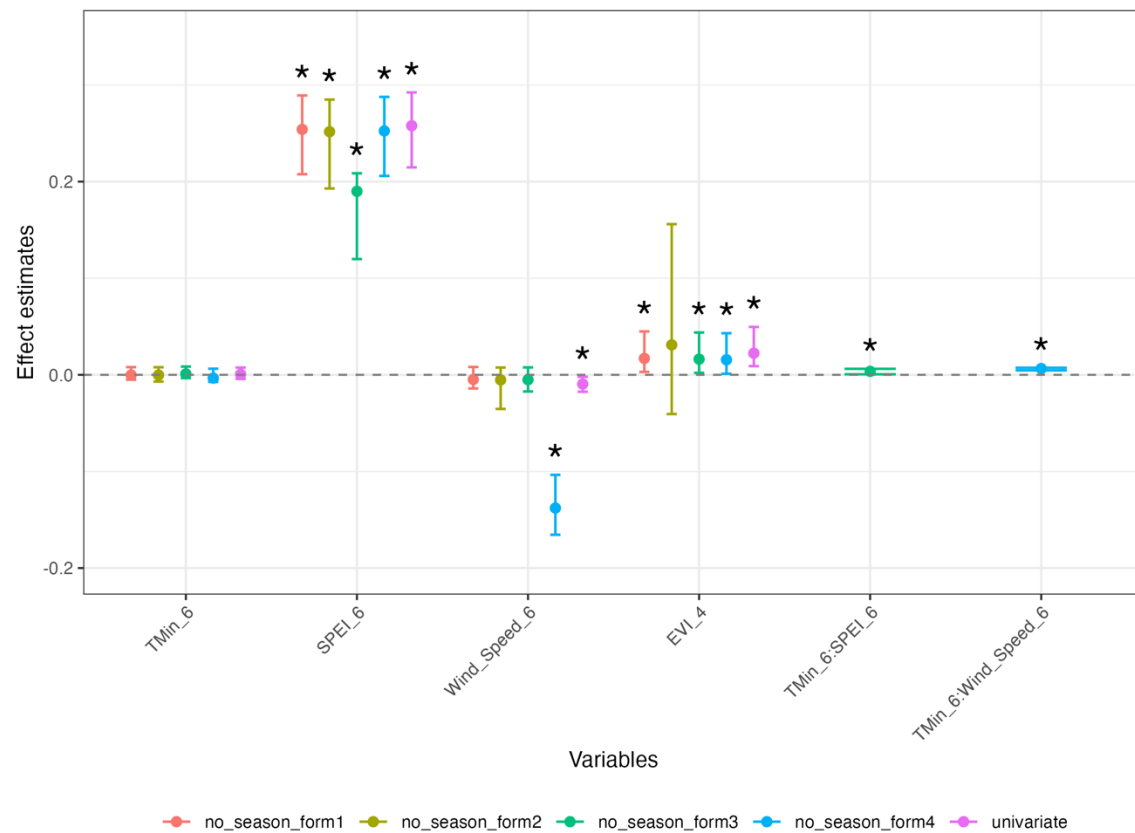


Figure 6.9-Effect estimates by formula, without diversity related estimates, for no seasonality base models. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, and TMin_6:Wind_Speed_6 is the interaction term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.

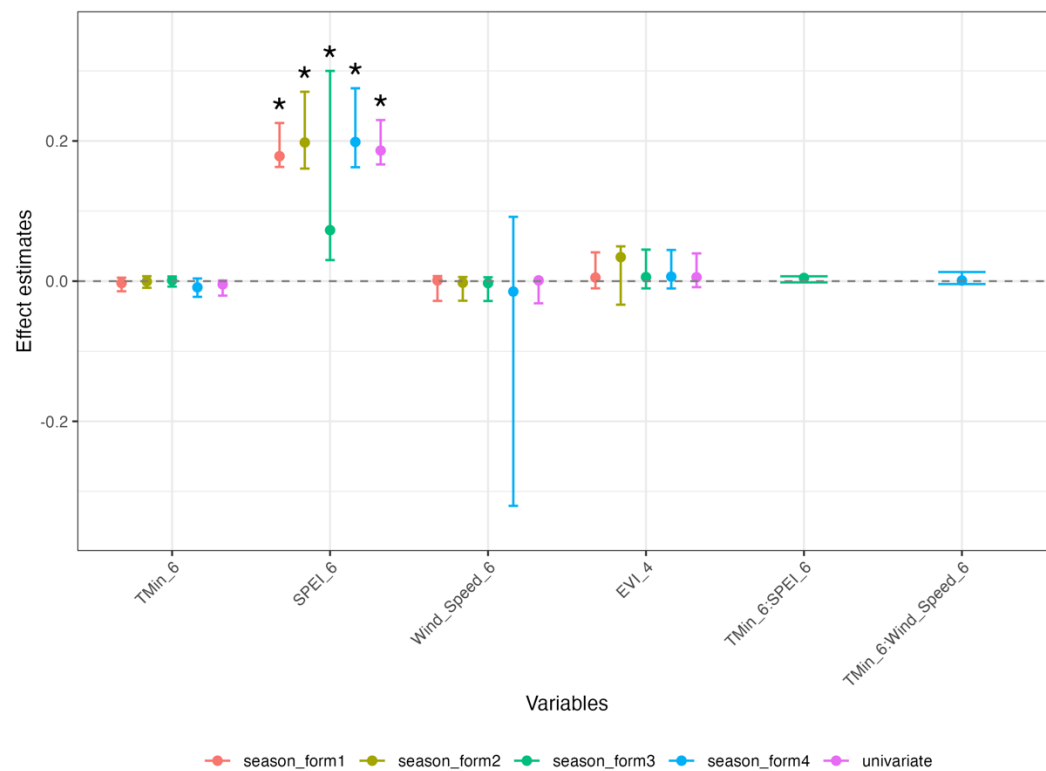


Figure 6.10- Effect estimates with confidence intervals by formula, without diversity related estimates, for seasonality base models. Variables on the x-axis are defined as: TMin_6 is minimum temperature at lag 6 weeks, SPEI_6 is SPEI at lag 6 weeks, Wind_Speed_6 is wind speed at lag 6 weeks, EVI_4 is EVI at lag 4 weeks, TMin_6:SPEI_6 is the interaction term between minimum temperature at lag of 6 weeks and SPEI at lag of 6 weeks, an TMin_6:Wind_Speed_6 is the interaction term between minimum temperature at lag of 6 weeks and wind speed at lag of 6 weeks. Stars (*) indicate significant effects.

6.4 Discussion

Between 2013 and 2022, multivariable models and interactions to adjust for confounding and effect modifiers, compared to no seasonality and seasonality baselines, improved the forecasting accuracy of the no seasonal model while the seasonal baseline performed better in comparison to its respective multivariable models. Overall, the seasonal baseline model was the best performing model as shown by its lowest CRPS results. When the data was disaggregated by season (rainy and dry) and outbreak status, there was an improvement seen with multivariable models compared to the no seasonality baseline. In no-outbreak periods and dry seasons, which are often correlated, the covariates improved forecasts of no seasonality model, particularly at longer horizons. In the seasonality models, the baseline performed better in both rainy and outbreak periods. There was very minimal difference between the performance of the multivariable models. This may be because the covariates are correlated as almost all of them follow seasonal patterns and may be accounting for the same underlying signal.

There may be a range of fundamentally different characteristics between rainy seasons and outbreak periods and dry seasons and no-outbreak periods driving these patterns, many of which were not captured in the data available for my analyses. Due to the influx of cases during rainy seasons, surveillance and quality of reports may be degraded compared to dry seasons. The strain on the healthcare system may also cause a delay in diagnosis and underreporting, particularly of mild cases[17]. Vector control efforts may also be enhanced compared to the dry season[375], introducing a confounding effect not captured in this thesis. There may also be a change in movement patterns such as displacements because of

heavy rains and flooding. Tsuda et al[376] reviewed floods that occurred in Thailand between 2005 and 2019, and found that during my study period of 2013 to 2022, the floods in 2017 resulted in the largest flooded area in northern and southern regions (according to the ratio of inundated area to regional area) which resulted in temporary migration of populations. My model did not capture such movements as only a static mobility matrix could be included. Additionally, after a rainy season and periods of high transmission, population immunity may shift, impacting transmission dynamics in the following seasons[279, 377]. Biologically, environmental covariates may impact dengue differently in different seasons[378], with complex lagged effects on transmission[379]. In rainy seasons, there may also be enhanced vector competition with other viruses such as Zika and chikungunya that rely on the same mosquito hosts for transmission[380, 381], complicating prediction efforts.

Beyond the factors impacting the seasons differently, there are other factors not considered here that may impact dengue transmission such as socioeconomic status, education level, and district level differences in response to dengue outbreaks and implementation of public health measures.

Other researchers have found similar results, where a seasonal model without covariates performed better in forecasting dengue transmission than models with covariates. Forecasting dengue in Mexico, Johansson et al[268] found that climate did not improve the forecasting ability of seasonal autoregressive models, while Lauer et al[257] found that in Thailand, the importance of climate variables was regionally dependent, and past incidence was the most prominent factor in model performance. Colón-González et al[226] found that in Vietnam, a superensemble of models including climate and Earth observations performed better at predicting

dengue cases at 1 to 3 months lead time compared to a seasonal baseline while the baseline outperformed the superensemble at lead times of 4 to 6 months.

The forecasting error between predicted and observed values is a conservative estimate, since the climate data used in the forecasting phase of the analysis was based on observed data rather than predicted values. Therefore, the application of these models and their performance need to be re-assessed using forecasted climate covariates. Different covariates such as higher resolution phylogenetic data and other models such as deep learning methods and ensembles should also be investigated.

These results show that a model based on previous dengue cases and seasonal patterns of dengue had higher forecasting skill than models including covariates in most districts of Thailand. The models also provided details at district levels on where covariate may prove useful. This information is useful to national and local public health officials interested in understanding future burden of dengue in their jurisdiction.

7 Discussion

This thesis developed and evaluated spatiotemporal models to forecast dengue transmission across Thailand over space and time. This was achieved through a systematic review of dengue forecasting methods, epidemiological analysis of dengue data from 2013 to 2022, and the development of Bayesian hierarchical models which included climatic, environmental, genetic diversity, and mobility covariates.

This work addresses key methodological gaps in the dengue forecasting literature, including the limited use of spatial and temporal correlation structures and probabilistic evaluation metrics. By integrating these components, this thesis contributes to improving dengue forecasting performance and interpretability, while emphasizing public health relevance.

7.1 Summary of findings

The systematic review identified an increasing use of machine learning methods for forecasting over time. Most models included in the review failed to account for the temporal and spatial correlation structure of the data. Only 7 of the 162 studies (4.3%) used a probabilistic approach to evaluate their forecasts, with most relying on point estimates.

These findings emphasize the need for advances in, and better use of, methods that explicitly handle spatial and temporal dependencies, as well as increased adoption of forecasting evaluation methods that incorporate uncertainty. This thesis directly addressed this gap by including random effects within a hierarchical

Bayesian framework and evaluating forecasts using probabilistic methods. The review also highlighted that forecasting research should extend beyond accuracy metrics to include implications for decision-making in public health practice, an area that is underexplored[269]. Despite advances in forecasting modelling, evidence of integrating such tools into operational control programs remains scarce, with limited integration of data collection, model development, and interventions [382, 383]. Key gaps include limited alignment between model capabilities and decision-making requirements[227], lack of actionable forecasts at appropriate resolutions[188, 384], and insufficient capacity building of local actors[382].

Between 2013 to 2022, dengue incidence in Thailand showed spatiotemporal heterogeneity. Peaks in northern provinces occurred during late spring and summer, while southern provinces experienced a peak in winter and early spring. DENV1 was the dominate circulating serotype. Wavelet analysis revealed both seasonal and interannual patterns. The seasonal pattern was strongest annually while the interannual pattern was strongest at around 3.5 years followed by 2.5 years. These findings provide important epidemiological context for interpreting forecast performance across districts and forecast horizons. When dengue transmission follows historical norms, the dominant annual seasonal cycle provides predictive skill, allowing models to perform better at forecast horizons of 3 to 6 months [268, 385].

Two baseline models were developed, one with an autoregressive term accounting for historical dengue cases, and another with an additional seasonality term accounting for recurring annual transmission patterns. Baseline models were evaluated over a 24-week forecast horizon at the district level using CRPS,

sharpness, calibration, bias, AUC, and Bier Score. Forecasting accuracy decreased with increasing forecast horizon. At a 1-week forecast horizon, the CRPS of autoregressive-only model was 0.217, compared to 0.219 for the seasonal baseline model. At shorter horizons of approximately four weeks, both models performed similarly, suggesting that autoregressive models alone are adequate for short-term forecasting. However, as the forecast horizon increased, discrepancies between models increased across districts, with the seasonal baseline model demonstrating better performance overall. At a 24-week forecast horizon, the autoregressive-only model had a CRPS of 0.340, while the seasonal baseline model had CRPS of 0.298.

Among climatic and environmental covariates examined, minimum temperature, wind speed, SPEI at 6-week lags and EVI at 4-week lags showed the best forecasting performance. Maximum and mean temperature did not improve forecasts compared with baseline models. Genetic diversity did not improve forecasting while including mobility reduced forecasting skill.

Selected covariates were combined in multivariable models building upon the baseline models to adjust for confounding. Three interaction terms were included to account for effect modification. These included interaction between minimum temperature and SPEI, interaction between EVI and genetic diversity, and interaction between minimum temperature and wind speed. Forecasts were evaluated using CRPS.

Overall, the seasonal baseline model performed best when compared with more complex multivariable models. Adding covariates to the baseline models had minimal impact, suggesting that seasonal patterns already accounted for much of

the predictable variation in dengue incidence. At a 1-week forecast horizon during outbreak periods, the seasonal baseline model achieved CRPS of 0.236, while multivariable seasonal models ranged from 0.268 to 0.269, indicating worse performance. During non-outbreak periods, the seasonal baseline model achieved a CRPS of 0.198, while multivariable seasonal models achieved slightly better performance, with CRPS values ranging from 0.190 to 0.19. At a 24-week forecast horizon, multivariable models outperformed the seasonal baseline during outbreak periods (CRPS 0.264 versus 0.296), but underperformed during non-outbreak periods (CRPS 0.275-0.276 versus 0.251).

Environmental covariates provided greater benefit when added to no-seasonality baseline models, particularly during non-outbreak periods and dry seasons, suggesting that climatic variables play a larger role when seasonal forcing is weaker. At a 1-week forecast horizon during non-outbreak periods, the no-seasonality baseline model achieved a CRPS of 0.198 while multivariable no-seasonality models achieved a CDRPS of 0.194. At a 24-week forecast horizon, the no-seasonality baseline model had a CRPS of 0.305, while some multivariable no-seasonality models showed marginal improvements, with CRPS values ranging from 0.302 to 0.307.

Other studies reported similar findings. In Thailand, Lauer et al. found that climate variables did not improve forecasting performance of an incidence-only model[257]. Chen et al.[386] forecasted multiple infectious disease across four countries, including dengue in Thailand, using LASSO, and found minimal improvement in accuracy when adding climate variables. Johansson et al[268], compared dengue forecasting models in Mexico and found that SARIMA models performed best when compared to models incorporating climate variables. Colon-

Gonzalez et al.[226] developed a seasonal ensemble dengue forecasting model for Vietnam using climate data and found that at forecast horizons beyond 3 months, baseline model predicting same incidence every month performed best.

These results highlight that autoregressive and seasonal components provide strong predictive power for dengue forecasting in Thailand. While climatic and environmental variables are biologically relevant, their contribution to forecasting skill is limited when seasonality is explicitly included in the model. From a public health perspective, short-term forecasts based on autoregressive models may be sufficient for operational planning, while for longer-term forecasts, seasonal models may be more appropriate.

7.2 Implications for our understanding of dengue epidemiology

Beyond forecasting model development, these findings, combined with the descriptive analyses of spatial and temporal patterns, provide insights into the drivers of dengue transmission dynamics in Thailand.

Wavelet analysis revealed that dengue in Thailand operates on multiple temporal scales simultaneously: a seasonal cycle of 6-18 months, with strongest power at an annual frequency and interannual cycles of 1.5-5 years, with peak power at approximately 3.5 years and a smaller peak at 2.5 years. These findings demonstrate that dengue transmission is driven by processes operating at different timescales, which has important implications for forecasting and intervention planning.

The strong annual signal likely reflects consistent environmental forcing through temperature, precipitation, vegetation, and wind patterns associated with

Thailand's seasonal climate [106, 387-390]. The interannual cycles are consistent with susceptible population replenishment following major epidemics, aligning with theoretical dengue dynamic models incorporate serotype-specific immunity and antibody-dependent enhancement[391, 392]. While environmental variables demonstrated clear associations with dengue risk in the DLNM models, individual covariates did not substantially improve forecasting accuracy compared with the seasonality baseline model. This suggests that the seasonality signal in dengue cases in Thailand is better captured by seasonal terms than climatic and environmental variables.

7.3 Implications for dengue forecast model development

The finding that the model with autoregressive correlation and a seasonality component outperformed models incorporating climate and environmental covariates have several important implications for dengue forecast development. The strong performance of relatively simple seasonal autoregressive models suggests that complex models may not be necessary for accurate short-term dengue forecasting in settings with strong seasonal patterns. The strong predictive power of recent case counts compared to external drivers indicates that intrinsic epidemic dynamics dominate short-term forecasting, as was found in other studies[257, 268]. The success of models with seasonal components without other covariates suggests that the annual seasonal cycles already capture the information captured by climate variables. In Thailand, temperature, precipitation, vegetation cover, and wind follow seasonal patterns that drive dengue transmission[106, 387-390]. As a result, these variables indirectly capture seasonality and their inclusion is redundant. However, the risk of such an approach is the application of the model to scenarios when dengue transmission does not

closely follow historical patterns, as was observed during outbreak periods in this thesis. Similar finding was found by Morbey et al.[385] for respiratory syncytial virus in England, where seasonality models performed best during a typical season but were less successful in an atypical season. Additionally, the best-fitting model for typical seasons contained different variables than for an atypical season.

These findings emphasize that there is no best model for dengue forecasting. Best approaches are context dependent, depending on geographic setting, forecast horizon, epidemiological conditions, and where disease follows historical norms. In Thailand's context of strong, predictable annual seasonality, seasonal autoregressive models perform well. However, in regions with weak or variable seasonality, or when forecasting at longer horizons, climate and environmental covariates might provide additional value

7.4 Public health utility of findings

The practical value of a forecasting model does not depend only on its accuracy, but also its feasibility of implementation, timeliness, and alignment with needs of public health decision makers. The finding that the seasonal model had the top forecasting skill when compared to models with covariates is an important finding for implementation in resource-limited settings. Using only historical data will make forecasting more accessible to local health departments, an important criterion of dengue surveillance systems, as described by WHO[29].

The forecast horizons examined in this thesis (1 to 24 weeks), provide sufficient time for public health interventions. Hospitals have adequate time to prepare for increased patient loads by ensuring adequate bed capacity, supplies of

intravenous fluids, blood products, and staff. Health promotion activities can be ramped up to increase care-seeking for dengue symptoms and promote household vector control. Thailand's current dengue response under the 3-3-1 approach is reactive, where local officials are expected to report outbreaks within 3 hours, implement initial control measures within 3 hours of reporting, and conduct mosquito fogging within 1 day[393]. Forecasting could enable a more anticipatory approach, for example, by pre-positioning vector control resources, intensifying surveillance, and prioritising risk communication in areas forecasted to experience elevated transmission.

Forecast horizons examined in this thesis may be insufficient for interventions requiring longer preparation times, such as large-scale community mobilization campaigns or policy changes, highlighting the potential value of longer horizon forecasts (more than 6 months), though the current models showed declining performance at longer horizons.

7.5 Limitations

7.5.1 Data

Surveillance data may be subject to systematic biases in case ascertainment that can affect model performance in complex ways. Case identification is often based on clinical diagnosis without laboratory confirmation. This may result in underdiagnosis, particularly during non-outbreak periods, when clinical suspicion for dengue is low, or overdiagnosis during outbreak periods when clinical suspicion is high.

If a period of underdiagnosis is included in the training data, the model may underestimate future cases. Conversely, if overdiagnosis periods are included, forecasts may overestimate incidence. Given current limitations in understanding temporal variations in diagnostic sensitivity and specificity of Thailand's surveillance system, quantifying and adjusting for these biases remains challenging.

The case data provided by the MoPH are collected through Thailand's national epidemiological surveillance system, a mandatory reporting system that captures cases from public and private health facilities, regardless of a patient's nationality or legal status [394]. However, population denominators used to calculate incidence rates are based on official counts that substantially underestimate the non-Thai population present in the country. According to data released on January 8, 2025, by the Department of Provincial Administration (DOPA), the total non-Thai population was estimated to be 997,549[395], while the United Nations Network on Migration estimated the total non-Thai population to be about 5.345 million as of December 2023[396]. This discrepancy exists because national estimates do not include refugees or illegal migrants including those who live in refugee camps[397]. As a result, incidence rates calculated using DOPA population estimates may result in inflated incidence rates.

The accuracy of NS1 diagnostic tests which are commonly used to diagnose dengue infection, differs by dengue virus serotype. In a systematic review, Da Costa et al[398] compared accuracy of Panbio and Platelia, and found that these tests had the highest accuracy for DENV1 at sensitivity of 81% (95% CI 73–87) for Panbio and 90% (95% CI 81–95) for Platelia and lowest sensitivity for DENV4 at 37% (95% CI 26–50) for Panbio and 58% (95% CI 30–81) for Platelia. In another

systematic review by Shan et al[399] focusing on tests conducted on Asian populations, pooled sensitivity for NS1 tests were 79.5% (95% CI 76.6–82.3) for DENV1, 62.7% (95% CI 59.1–66.2) for DENV2, 73% (95% CI 69.2–76.6) for DENV3, and 46% (95% CI 36–59.3) for DENV4.

These findings demonstrate that NS1 tests have unequal detection capability by serotype, which can directly impact case identification and estimation of disease burden. When serotypes with lower NS1 sensitivity predominantly circulate, case counts based on NS1 testing may lead to underestimation of true incidence. With a wide implementation of NS1 tests in Thailand, it is possible that outbreaks of serotypes other than serotype 1 may have been particularly poorly detected. Despite national guidelines in Thailand stating that a negative NS1 test does not rule out dengue, this may still impact clinicians' diagnostic decision making, particularly in non-outbreak periods[285]. Given the serotype dependence of this bias, it is difficult to adjust for this bias post case detection.

As described in Chapter 5, mobility data were included as a static variable, calculated for a three-week period ending on April 7, 2025, due to limited data availability. Including mobility in the models did not improve forecasting performance. However, this may not have been the most appropriate approach, as a time-varying mobility metric might better capture the dynamic relationship between mobility and dengue transmission.

7.5.2 Methodology

Due to computational limitations, covariates in multivariable models were included as fixed effects. Including variables as random effects would have been preferred

as they would have allowed the model to account for unmeasured temporal variation, such as health-seeking behaviour, vector control practices or environmental factors that could affect both baseline dengue incidence and strength of covariate relationships. Random effects would also allow the model to borrow strength across time periods, improving forecast precision.

This thesis did not explore heterogeneity of covariates and their lags across space but rather included the same covariates and lags across Thailand. It was assumed that biological relationships between dengue incidence and covariates should be constant across space and therefore, lags should be same across space. However, other factors such as interactions with other variables or heterogeneous delay in detection and reporting of cases may lead to apparent spatial heterogeneity of lags across the country, even when underlying biological relationships are constant.

7.6 Future directions

The analysis in this thesis used retrospective dengue surveillance data. However, as is prevalent in almost all surveillance systems, there are well documented reporting delays in Thailand. Reich et al.[50] found that in 2014, reporting delay of dengue cases ranged between 1 to 50 weeks. Not accounting for this delay when using real-time data may result in training data that is underestimating incidence and can thus underpredict future cases. Beesley et al.[400] found that ignoring reporting delays leads to under coverage, ranging from 2 to 60% and higher prediction errors relative to analysis of the validation data. Some of the methods, which can be explored in the future analyses related to this thesis, include

rescaling with estimated reporting factors, multiple imputation approaches, mean model offset methods, or Bayesian nowcasting techniques[243, 400, 401].

Similarly, this thesis used already observed climatic and environmental variables. This gives the models the best-case scenario for forecasting and may result in overestimation of accuracy of the models. To be closer to a real-life scenario, using predicted climatic and environmental variables would have been more appropriate as observed variables would not be available[227]. Therefore, it is recommended that future efforts consider using predicted variables for forecasting.

This thesis focused on hierarchical Bayesian approaches to modelling and forecasting. In the future, using other approaches to compare results would be beneficial. Machine learning methods such as XGBoost [319, 320, 402] and deep learning methods such as Temporal Convolutional Networks (TCN)[403] and Time-Series Foundation Model (TimesFM)[404] have shown competitive performance as they can capture nonlinear relationship and complex dependencies. Recent comparative studies have shown that ensembles of diverse methods for forecasting can improve forecasting accuracy across different scenarios[185, 226, 405].

Allowing for the relationship between dengue incidence and covariates to be spatially heterogeneous could also improve forecasting. Cheng et al.[406] found both statically significant positive and negative associations between dengue incidence and weather conditions in cities in Guangdong province in China while Polrob et al.[247] found different strengths of associations with dengue incidence and climate variables by zones in Bangkok, Thailand.

Additional covariates should be investigated for their usefulness for forecasting. Spatio-temporal information on dengue-infected adult mosquitoes may improve forecasting. Studies found that incorporating female mosquito infection rates were found to improve dengue forecasting in the central region of Thailand[407, 408]. A case-control study in northeastern Thailand found that monitoring dengue infection in mosquitoes was a better indicator of risk than immature entomological indices[409]. Thai Department of Disease Control conducts routine surveys of larval habitats[410], however, there currently is no routine dengue virus detection in mosquitoes[411, 412]. Other relevant variables can include immunity profiles, which can be ascertained through sources such as serological surveys or inferred from historical case data and serotype circulation patterns[247, 274, 413]. Finally, imported case counts, particularly from regions with substantial international travel can serve as indicators of early outbreaks[414, 415].

Due to time limitations, results from this thesis were not evaluated for their ability to inform public health decision making in Thailand. Further work should most certainly address this shortcoming to provide a comprehensive evaluation of models developed and evaluated in this thesis.

7.7 Conclusions

This thesis demonstrates that hierarchical Bayesian spatiotemporal modelling is an effective approach for forecasting dengue transmission in endemic settings such as Thailand. Autoregressive and seasonal components capture much of the predictable variation in dengue incidence across Thailand, while climatic, environmental, genetic diversity, and mobility covariates offer limited, context specific improvements. By integrating probabilistic evaluation metrics and

addressing key methodological gaps, this work improves dengue forecasting and its applicability to public health practice.

8 Appendices

Appendix A: Search Strings

Database	Platform	Coverage	Search Terms	Filters Applied	Date Searched
CAB Global Health	Ovid	Inception to present	(dengue OR denv) AND (predict* OR forecast* OR anticipat* OR expect*)	None	11 January 2023
MEDLINE	Ovid	1946 to present	(dengue OR denv) AND (predict* OR forecast* OR anticipat* OR expect*)	(NOT ((exp animal/ OR nonhuman/ NOT exp human/))	11 January 2023
Embase	Ovid	1974 to present	(dengue OR denv) AND (predict* OR forecast* OR anticipat* OR expect*)	(NOT ((exp animal/ OR nonhuman/ NOT exp human/))	11 January 2023
Web of Science	Core Collection	All years	TS=(dengue OR denv) AND TS=(predict* OR forecast* OR anticipat* OR expect*)	(NOT ((TS=(animal* OR nonhuman*))) NOT TS=(human*)))	11 January 2023

Appendix B: Characteristics of included studies, in order of year

ID	Year	Authors	Title	Country of study	Prediction type
1	1991	Chungue et al.[416]	Dengue surveillance in French Polynesia: An attempt to use the excess number of laboratory requests for confirmation of dengue diagnosis as an indicator of dengue activity	French Polynesia	Temporal
2	2005	Peterson et al.[417]	Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases	Mexico	Spatio-temporal
3	2007	Wu et al.[418]	Weather as an effective predictor for occurrence of dengue fever in Taiwan	Taiwan	Temporal
4	2008	Halide et al.[419]	A predictive model for Dengue Hemorrhagic Fever epidemics	Indonesia	Temporal
5	2008	Luz et al.[195]	Time series analysis of dengue incidence in Rio de Janeiro, Brazil	Brazil	Temporal
6	2008	Silawan et al.[251]	Temporal patterns and forecast of dengue infection in Northeastern Thailand	Thailand	Temporal
7	2009	Tipayamongkholgul et al.[420]	Effects of the El Niño-Southern Oscillation on dengue epidemics in Thailand, 1996-2005	Thailand	Spatio-temporal
8	2009	Wongkoon et al.[421]	Predicting DHF Incidence in Northern Thailand using Time Series Analysis Technique	Thailand	Temporal
9	2010	Aburas et al.[422]	Dengue confirmed-cases prediction: A neural network model	Singapore	Temporal
10	2010	Hu et al.[423]	Dengue fever and El Nino/Southern Oscillation in Queensland, Australia: a time series predictive model	Australia	Temporal
11	2011	Althouse et al.[424]	Prediction of Dengue Incidence Using Search Query Surveillance	Singapore, Thailand	Temporal
12	2011	Chan et al.[425]	Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance	India, Indonesia	Temporal
13	2011	Gerardi et al.[426]	System Identification and Prediction of Dengue Fever Incidence in Rio de Janeiro	Brazil	Temporal
14	2011	Gharbi et al.[338]	Time series analysis of dengue incidence in Guadeloupe French West Indies: forecasting models using climate variables as predictors	Guadeloupe	Temporal

15	2011	Martinez et al.[196]	A SARIMA forecasting model to predict the number of cases of dengue in Campinas State of Sao Paulo Brazil	Brazil	Temporal
16	2011	Martinez et al.[427]	Predicting the number of cases of dengue infection in Ribeirao Preto Sao Paulo State Brazil using a SARIMA model	Brazil	Temporal
17	2011	Yu et al.[428]	A spatio-temporal climate-based model of early dengue fever warning in southern Taiwan	Taiwan	Spatio-temporal
18	2012	Aguiar et al.[264]	How much complexity is needed to describe the fluctuations observed in dengue hemorrhagic fever incidence data?	Thailand	temporal
19	2012	Capistran et al.[429]	Towards uncertainty quantification and inference in the stochastic SIR epidemic model	Mexico	Temporal
20	2012	Earnest et al.[192]	Comparing statistical models to predict dengue fever notifications	Singapore	Temporal
21	2012	Hii et al.[198]	Forecast of dengue incidence using temperature and rainfall	Singapore	Temporal
22	2012	Karim et al.[430]	Climatic factors influencing dengue cases in Dhaka city: a model for dengue prediction	Bangladesh	Temporal
23	2012	Lowe et al.[431]	The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil	Brazil	Spatio-temporal
24	2012	Wongkoon et al.[432]	Assessing the temporal modelling for prediction of dengue infection in northern and north-eastern Thailand	Thailand	Temporal
25	2012	Wongkoon et al.[433]	Development of temporal modeling for prediction of dengue infection in Northeastern Thailand	Thailand	Temporal
26	2013	Dom et al.[434]	Generating temporal model using climate variables for the prediction of dengue cases in Subang Jaya Malaysia	Malaysia	temporal
27	2013	Sitepu et al.[435]	Temporal patterns and a disease forecasting model of dengue hemorrhagic fever in Jakarta based on 10 years of surveillance data	Indonesia	Temporal
28	2013	Wongkoon et al.[252]	Weather factors influencing the occurrence of dengue fever in Nakhon Si Thammarat, Thailand	Thailand	Temporal
29	2014	Buczak et al.[436]	Prediction of High Incidence of Dengue in the Philippines	Philippines	Spatio-temporal
30	2014	Eastin et al.[437]	Intra- and Interseasonal Autoregressive Prediction of Dengue Outbreaks Using Local Weather and Regional Climate for a Tropical Environment in Colombia	Colombia	Temporal
31	2014	Mohsin et al.[438]	Outbreak detection model based on danger theory	Malaysia	Temporal
32	2014	Ongruk et al.[439]	New Key Factors Discovery to Enhance Dengue Fever Forecasting Model	Thailand	Temporal

33	2014	Sang et al.[440]	Predicting local dengue transmission in Guangzhou China through the influence of imported cases mosquito density and climate variability	China	Temporal
34	2014	Torres et al.[441]	Fuzzy model identification of dengue epidemic in Colombia based on multiresolution analysis	Colombia	Temporal
35	2014	Yu et al.[442]	An online spatiotemporal prediction model for dengue fever epidemic in Kaohsiung (Taiwan)	Taiwan	Spatio-temporal
36	2015	Chan et al.[443]	Daily forecast of dengue fever incidents for urban villages in a city	Taiwan	Spatio-temporal
37	2015	Pepin et al.[444]	Utility of mosquito surveillance data for spatial prioritization of vector control against dengue viruses in three Brazilian cities	Brazil	Spatio-temporal
38	2015	Phung et al.[445]	Identification of the prediction model for dengue incidence in Can Tho city, a Mekong Delta area in Vietnam	Vietnam	Temporal
39	2015	Sang et al.[199]	Predicting Unprecedented Dengue Outbreak Using Imported Cases and Climatic Factors in Guangzhou 2014	China	Temporal
40	2015	Tung et al.[446]	Using prediction markets of market scoring rule to forecast infectious diseases: a case study in Taiwan	Taiwan	Temporal
41	2016	Ali et al.[447]	ID-Viewer: a visual analytics architecture for infectious diseases surveillance and response management in Pakistan	Pakistan	temporal
42	2016	Johansson et al.[268]	Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico	Mexico	Spatio-temporal
43	2016	Li et al.[448]	The wisdom of crowds in action: Forecasting epidemic diseases with a web-based prediction market system	Taiwan	Temporal
44	2016	Lowe et al.[227]	Evaluating probabilistic dengue risk forecasts from a prototype early warning system for Brazil	Brazil	Spatio-temporal
45	2016	Mekpanyup et al.[449]	A new approach to detect epidemic of DHF by combining ARIMA model and adjusted Tukey's control chart with interpretation rules	Thailand	Temporal
46	2016	Ramadona et al.[450]	Prediction of dengue outbreaks based on disease surveillance and meteorological data	Indonesia	Temporal
47	2016	Reich et al.[50]	Challenges in Real-Time Prediction of Infectious Disease: A Case Study of Dengue in Thailand	Thailand	Spatio-temporal
48	2016	Shi et al.[451]	Three-Month Real-Time Dengue Forecast Models: An Early Warning System for Outbreak Alerts and Policy Decision Support in Singapore	Singapore	Temporal

49	2016	Wongkoon et al.[452]	Spatio-temporal climate-based model of dengue infection in Southern Thailand	Thailand	Temporal
50	2016	Yamana et al.[453]	Superensemble forecasts of dengue outbreaks	Puerto Rico	Temporal
51	2016	Yu et al.[454]	A spatiotemporal dengue fever early warning model accounting for nonlinear associations with hydrological factors: a Bayesian maximum entropy approach	Taiwan	Spatio-temporal
52	2016	Zhang et al.[455]	Developing a Time Series Predictive Model for Dengue in Zhongshan, China Based on Weather and Guangzhou Dengue Surveillance Data	China	Temporal
53	2017	Chuang et al.[331, 332]	Effects of local and regional climatic fluctuations on dengue outbreaks in southern Taiwan	Taiwan	Temporal
54	2017	de Almeida Marques-Toledo et al.[456]	Dengue prediction by the web: Tweets are a useful tool for estimating and forecasting Dengue at country and city level	Brazil	Temporal
55	2017	Dom et al.[457]	Integration of GIS-Based Model with Epidemiological data as a Tool for Dengue Surveillance	Malaysia	Spatio-temporal
56	2017	Guo et al.[205]	Developing a dengue forecast model using machine learning: a case study in China	China	Spatio-temporal
57	2017	Lee et al.[458]	Early warning signal for dengue outbreaks and identification of high risk areas for dengue fever in Colombia using climate and non-climate datasets	Colombia	Temporal
58	2017	Li et al.[459]	Modeling and projection of dengue fever cases in Guangzhou based on variation of weather factors	China	Temporal
59	2017	Lowe et al.[460]	Climate services for health: predicting the evolution of the 2016 dengue season in Machala Ecuador	Ecuador	Temporal
60	2017	Mahdiana et al.[461]	A Model for Forecasting the Number of Cases and Distribution Pattern of Dengue Hemorrhagic Fever in Indonesia	Indonesia	Spatio-temporal
61	2017	Martinez-Bello et al.[462]	Bayesian dynamic modeling of time series of dengue disease case counts	Colombia	Temporal
62	2017	Ray et al.[463]	Infectious disease prediction with kernel conditional density estimation	Puerto Rico	temporal

63	2017	Sirisena et al.[147]	Effect of Climatic Factors and Population Density on the Distribution of Dengue in Sri Lanka: A GIS Based Evaluation for Prediction of Outbreaks	Sri Lanka	Spatio-temporal
64	2017	Sun et al.[464]	Spatial-temporal distribution of dengue and climate characteristics for two clusters in Sri Lanka from 2012 to 2016	Sri Lanka	Temporal
65	2017	Xu et al.[465]	Climate variation drives dengue dynamics	China	Temporal
66	2017	Yang et al.[466]	Advances in using Internet searches to track dengue	Thailand, Mexico, Brazil, Singapore, Taiwan	Temporal
67	2018	Agarwal et al.[467]	Data mining techniques for predicting dengue outbreak in geospatial domain using weather parameters for New Delhi India	India	Temporal
68	2018	Baquero et al.[200]	Dengue forecasting in São Paulo city with generalized additive models, artificial neural networks and seasonal autoregressive integrated moving average models	Brazil	Temporal
69	2018	Buczak et al.[468]	Ensemble method for dengue prediction	Puerto Rico, Peru	Temporal
70	2018	Carvajal et al.[469]	Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila Philippines	Philippines	Temporal
71	2018	Chen et al.[470]	Neighbourhood level real-time forecasting of dengue cases in tropical urban Singapore	Singapore	Spatio-temporal
72	2018	Chen et al.[386]	The utility of LASSO-based models for real time forecasts of endemic infectious diseases: A cross country comparison	Taiwan, Thailand, Singapore	Temporal
73	2018	Cortes et al.[471]	Time series analysis of dengue surveillance data in two Brazilian cities	Brazil	Temporal
74	2018	Hussain-Alkhateeb et al.[472]	Early warning and response system (EWARS) for dengue outbreaks: Recent advancements towards widespread applications in critical settings	Mexico	Temporal
75	2018	Jing et al.[414]	Imported cases and minimum temperature drive dengue transmission in Guangzhou China: evidence from ARIMAX model	China	temporal
76	2018	Johnson et al.[473]	Phenomenological Forecasting of Disease Incidence Using Heteroskedastic Gaussian Processes: A Dengue Case Study	Puerto Rico, Peru	Temporal
77	2018	Koh et al.[474]	A model comparison algorithm for increased forecast accuracy of dengue fever incidence in Singapore and the auxiliary role of total precipitation information	Singapore	Temporal

78	2018	Lauer et al.[257]	Prospective forecasts of annual dengue hemorrhagic fever incidence in Thailand 2010-2014	Thailand	Spatio-temporal
79	2018	Sanchez-Gonzalez et al.[475]	Prediction of dengue outbreaks in Mexico based on entomological meteorological and demographic data	Mexico	Spatio-temporal
80	2018	Tosepu et al.[476]	Climatic factors influencing dengue hemorrhagic fever in Kolaka district Indonesia	Indonesia	Temporal
81	2018	Withanage et al.[477]	A forecasting model for dengue incidence in the District of Gampaha Sri Lanka	Sri Lanka	temporal
82	2019	Attanayake et al.[478]	Combining Forecasts of Arima and Exponential Smoothing Models	Sri Lanka	Temporal
83	2019	Bett et al.[479]	Spatiotemporal analysis of historical records (2001-2012) on dengue fever in Vietnam and development of a statistical model for forecasting risk	Vietnam	Spatio-temporal
84	2019	Chakraborty et al.[480]	Forecasting dengue epidemics using a hybrid methodology	Peru, Puerto Rico, Philippines	Temporal
85	2019	Chumpu et al.[249]	The association between dengue incidences and provincial-level weather variables in Thailand from 2001 to 2014	Thailand	Temporal
86	2019	Gabriel et al.[481]	Dengue outbreaks: unpredictable incidence time series	Brazil	Temporal
87	2019	Guo et al.[482]	An ensemble forecast model of dengue in Guangzhou, China using climate and social media surveillance data	China	Temporal
88	2019	Jain et al.[258]	Prediction of dengue outbreaks based on disease surveillance meteorological and socio-economic data	Thailand	Spatio-temporal
89	2019	Jayaraj at al.[483]	Developing a dengue prediction model based on climate in Tawau, Malaysia	Malaysia	Temporal
90	2019	Johansson et al.[190]	An open challenge to advance probabilistic forecasting for dengue epidemics	Puerto Rico and Peru	Temporal
91	2019	Li et al.[484]	Climate-driven variation in mosquito density predicts the spatiotemporal dynamics of dengue	China	Spatio-temporal
92	2019	Liu et al.[485]	A dengue fever predicting model based on Baidu search index data and climate data in South China	China	Temporal
93	2019	Marques-Toledo et al.[486]	Probability of Dengue Transmission and Propagation in a Non-Endemic Temperate Area: Conceptual Model and Decision Risk Levels for Early Alert, Prevention and Control	Brazil	Temporal

94	2019	Mincham et al.[487]	Development of a mechanistic dengue simulation model for Guangzhou	China	Temporal
95	2019	Nayak et al.[488]	Prediction of dengue outbreaks in Kerala state using disease surveillance and meteorological data	India	Temporal
96	2019	Pineda-Cortel et al.[489]	Modeling and predicting dengue fever cases in key regions of the Philippines using remote sensing data	Philippines	Temporal
97	2019	Rahayu et al.[490]	Prediction of Dengue Hemorrhagic Fever Cases Based on Weather Parameters Using Back Propagation Neural Networks (Case Study in Pontianak City)	Indonesia	Temporal
98	2019	Rangarajan et al.[491]	Forecasting dengue and influenza incidences using a sparse representation of Google trends electronic health records and time series data	Thailand, Singapore, Taiwan, Brazil, Mexico	Temporal
99	2019	Rypdal and Sugihara[492]	Inter-outbreak stability reflects the size of the susceptible pool and forecasts magnitudes of seasonal epidemics	Puerto Rico	temporal
100	2019	Stolerman et al.[493]	Forecasting dengue fever in Brazil: An assessment of climate conditions	Brazil	Temporal
101	2019	Zhu et al.[494]	Prediction model for dengue fever based on interactive effects between multiple meteorological factors in Guangdong China (2008-2016)	China	Temporal
102	2020	Appice et al.[495]	A Multi-Stage Machine Learning Approach to Predict Dengue Incidence: A Case Study in Mexico	Mexico	Spatio-temporal
103	2020	Bal et al.[496]	Modeling and prediction of dengue occurrences in Kolkata, India, based on climate factors	India	Temporal
104	2020	Benedum et al.[497]	Weekly dengue forecasts in Iquitos Peru; San Juan Puerto Rico; and Singapore	Peru, Puerto Rico, Singapore	Temporal
105	2020	Brett et al.[498]	Dynamical footprints enable detection of disease emergence	Puerto Rico	Temporal
106	2020	Chakraborty et al.[499]	A Robust and Non-parametric Model for Prediction of Dengue Incidence	Singapore	Temporal
107	2020	Chen et al.[500]	Early dengue outbreak detection modeling based on dengue incidences in Singapore during 2012 to 2017	Singapore	Temporal
108	2020	Cheng et al.[501]	Real-time dengue forecast for outbreak alerts in Southern Taiwan	Taiwan	Temporal

109	2020	de Lima et al.[306]	Evaluation of the Models for Forecasting Dengue in Brazil from 2000 to 2017: An Ecological Time-Series Study	Brazil	Temporal
110	2020	Liu et al.[502]	Enhancing fine-grained intra-urban dengue forecasting by integrating spatial interactions of human movements between urban regions	China	Spatio-temporal
111	2020	Majhi et al.[503]	Dynamics of dengue outbreaks in gangetic West Bengal: A trend and time series analysis	India	Temporal
112	2020	Medina et al.[504]	Time Series Forecasting Using a Sarma Model Fitted on Dengue Cases (2010-2017): A Potential Tool for an Early Warning System in Quezon City Philippines	Philippines	Temporal
113	2020	Mussumeci et al.[206]	Large-scale multivariate forecasting models for Dengue - LSTM versus random forest regression	Brazil	Spatio-temporal
114	2020	Mustaffa et al. [505]	Predictive Analysis of Dengue Outbreak Based on an Improved Salp Swarm Algorithm	Indonesia	Temporal
115	2020	Pham et al.[348]	Time-series modelling of dengue incidence in the Mekong Delta region of Viet Nam using remote sensing data	Vietnam	Temporal
116	2020	Polwiang[239]	The time series seasonal patterns of dengue fever and associated weather variables in Bangkok (2003-2017)	Thailand	Temporal
117	2020	Puengpreeda et al.[506]	Weekly Forecasting Model for Dengue Hemorrhagic Fever Outbreak in Thailand	Thailand	Spatio-temporal
118	2020	Stoner et al.[507]	Multivariate hierarchical frameworks for modelling delayed reporting in count data	Brazil	Temporal
119	2020	Tao et al.[508, 509]	Explicit characterization of human population connectivity reveals long run persistence of interregional dengue shocks	Malaysia, Singapore	Spatio-temporal
120	2020	Tsai et al.[510]	Prediction of dengue fever outbreak based on case household locations in southern Taiwan	Taiwan	Spatio-temporal
121	2020	Wagner et al.[511]	Climatological virological and sociological drivers of current and projected dengue fever outbreak dynamics in Sri Lanka	Sri Lanka	Spatio-temporal
122	2020	Xu et al.[512]	Forecast of dengue cases in 20 Chinese cities based on the deep learning method	China	Temporal
123	2020	Zhao et al.[202]	Machine learning and dengue forecasting: Comparing random forests and artificial neural networks for predicting dengue burden at national and sub-national scales in Colombia	Colombia	Spatio-temporal
124	2021	Benkimoun et al.[513]	Dynamic mapping of dengue basic reproduction number	Island of Reunion	Spatio-temporal
125	2021	Caldwell et al.[514]	Climate predicts geographic and temporal variation in mosquito-borne disease dynamics on two continents	Ecuador, Kenya	Temporal

126	2021	Cheng et al.[515]	Extreme weather events and dengue outbreaks in Guangzhou China: a time-series quasi-binomial distributed lag non-linear model	China	Temporal
127	2021	Cheng et al.[516]	Impact of interventions on the incidence of natural focal diseases during the outbreak of COVID-19 in Jiangsu Province China	China	Temporal
128	2021	Colon-Gonzalez et al.[226]	Probabilistic seasonal dengue forecasting in Vietnam: A modelling study using superensembles	Vietnam	Spatio-temporal
129	2021	Ferdousi et al.[517]	A Windowed Correlation-Based Feature Selection Method to Improve Time Series Prediction of Dengue Fever Cases	Brazil	Spatio-temporal
130	2021	Hadi et al.[305]	A seasonal autoregressive integrated moving average model to predict incidence of dengue cases in Kuantan, Malaysia	Malaysia	Temporal
131	2021	Heidrich et al.[518]	Simulation and prediction of dengue outbreaks based on an SIR model with a time-dependent transmission rate including meteorological data. An example for Colombo and Jakarta	Sri Lanka, Indonesia	Temporal
132	2021	Kiang et al.[175]	Incorporating human mobility data improves forecasts of Dengue fever in Thailand	Thailand	Temporal
133	2021	Liu et al.[519]	Facilitating fine-grained intra-urban dengue forecasting by integrating urban environments measured from street-view images	China	Spatio-temporal
134	2021	McGough et al.[520]	A dynamic ensemble learning approach to forecast dengue fever epidemic years in Brazil using weather and population susceptibility cycles	Brazil	Spatio-temporal
135	2021	Nuraini et al.[521]	Climate-based dengue model in Semarang Indonesia: Predictions and descriptive analysis	Indonesia	Temporal
136	2021	Ray et al.[522]	Optimized ensemble deep learning framework for scalable forecasting of dynamics containing extreme events	Puerto Rico	Temporal
137	2021	Srikin et al.[523]	Dengue Transmission Mapping with Weather-Based Predictive Model in Three Southernmost Provinces of Thailand	Thailand	Spatio-temporal
138	2021	Xavier et al.[524]	Analysis of climate factors and dengue incidence in the metropolitan region of Rio de Janeiro Brazil	Brazil	Temporal
139	2021	Yi et al.[525]	SEIR-SEI-EnKF: A New Model for Estimating and Forecasting Dengue Outbreak Dynamics	Taiwan; Singapore; Brazil	Temporal

140	2022	Bracher et al.[526]	Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction	Puerto Rico	Temporal
141	2022	Chen et al.[371]	Measuring the effects of COVID-19-related disruption on dengue transmission in southeast Asia and Latin America: a statistical modelling study	23 countries (16 in Latin America and 7 in southeast Asia)	Spatio-temporal
142	2022	Deb and Deb[527]	An ensemble method for early prediction of dengue outbreak	Puerto Rico Peru	Temporal
143	2022	Heidrich et al.[528]	Prediction of dengue cases based on human mobility and seasonality-An example for the city of Jakarta	Indonesia	Temporal
144	2022	Ismail et al.[529]	The practicality of Malaysia dengue outbreak forecasting model as an early warning system	Malaysia	Temporal
145	2022	Kakarla et al.[530]	Weather integrated multiple machine learning models for prediction of dengue prevalence in India	India	Temporal
146	2022	Koplewitz et al.[531]	Predicting dengue incidence leveraging internet-based data sources. A case study in 20 cities in Brazil	Brazil	Temporal
147	2022	Li[532]	Forecasting Weekly Dengue Cases by Integrating Google Earth Engine-Based Risk Predictor Generation and Google Colab-Based Deep Learning Modeling in Fortaleza and the Federal District Brazil	Brazil	Temporal
148	2022	Li et al.[533]	Improving Dengue Forecasts by Using Geospatial Big Data Analysis in Google Earth Engine and the Historical Dengue Information-Aided Long Short Term Memory Modeling	Brazil	Temporal
149	2022	Li et al.[534]	Dengue Risk Forecast with Mosquito Vector: A Multicomponent Fusion Approach Based on Spatiotemporal Analysis	China	Spatio-temporal
150	2022	Ligue et al.[535]	Deep Learning Approach to Forecasting Dengue Cases in Davao City Using Long Short-term Memory (LSTM)	Philippines	Temporal
151	2022	Mahajan et al.[536]	A Novel Stacking-Based Deterministic Ensemble Model for Infectious Disease Prediction	Hong Kong	Temporal
152	2022	Mukhaiyar et al.[537]	Forecasting of dengue hemorrhages fever cases with autoregression distributed lag model using Poisson regression approach	Indonesia	Temporal
153	2022	Nguyen et al.[203]	Deep learning models for forecasting dengue fever based on climate data in Vietnam	Vietnam	Temporal
154	2022	Ong et al.[538]	Fine-scale estimation of effective reproduction numbers for dengue surveillance	Singapore	Spatio-temporal

155	2022	Othman et al.[539]	Model Forecasting Development for Dengue Fever Incidence in Surabaya City Using Time Series Analysis	Indonesia	Temporal
156	2022	Roster et al.[540]	Machine-Learning–Based Forecasting of Dengue Fever in Brazilian Cities Using Epidemiologic and Meteorological Variables	Brazil	Spatio-temporal
157	2022	Saita et al.[241]	Temporal Variations and Spatial Clusters of Dengue in Thailand: Longitudinal Study before and during the Coronavirus Disease (COVID-19) Pandemic	Thailand	Temporal
158	2022	Sanchez-Gendriz et al.[541]	Data-driven computational intelligence applied to dengue outbreak forecasting: a case study at the scale of the city of Natal, RN-Brazil	Brazil	Temporal
159	2022	Sathish et al.[542]	Autoregressive and moving average models for zero-inflated count time series	India	Temporal
160	2022	Souza et al.[543]	Predicting dengue outbreaks in Brazil with manifold learning on climate data	Brazil	Temporal
161	2022	Surendran et al.[544]	Reduced dengue incidence during the COVID-19 movement restrictions in Sri Lanka from March 2020 to April 2021	Sri Lanka	Spatio-temporal
162	2023	Khan et al.[545]	Dengue transmission under future climate and human population changes in mainland China	China	Spatio-temporal

Appendix C – Prediction target, spatial scope, spatial resolution, and temporal resolution of included studies.

ID	Prediction Target	Spatial Scope	Spatial Resolution	Temporal Resolution
1	Outbreak	National	Country	Monthly
2	Occurrence	Sub-national	Point/Pixel	Monthly
3	Incidence	Sub-national	District/City/Village	Monthly
4	Incidence	Sub-national	District/City/Village	Monthly
5	Incidence	Sub-national	District/City/Village	Monthly
6	Incidence	Sub-national	Regional/Province	Monthly
7	Incidence, Outbreak	Sub-national	Regional/Province	Monthly
8	Incidence	Sub-national	Regional/Province	Monthly
9	Incidence	National	Country	Weekly
10	Occurrence	Sub-national	District/City/Village	Monthly
11	Incidence, Outbreak	Sub-national	Regional/Province	Singapore: Weekly, Thailand: Monthly
12	Incidence	Global	Country	Monthly
13	Incidence	Sub-national	point/pixel	Weekly
14	Incidence	National	Country	Weekly
15	Incidence	Sub-national	District/City/Village	Monthly
16	Incidence	Sub-national	District/City/Village	Monthly
17	Incidence	Sub-national	District/City/Village	Weekly
18	Incidence	Sub-national	Regional/Province	Monthly
19	Incidence	Sub-national	District/City/Village	Acapulco: Monthly, Cuernavaca: Weekly
20	Incidence	National	Country	Weekly
21	Incidence, Outbreak	National	Country	Weekly
22	Incidence, Outbreak	Sub-national	Regional/Province	Monthly
23	Incidence	National	District/City/Village	Monthly
24	Incidence	Sub-national	Regional/Province	Monthly

25	Incidence	Sub-national	Regional/Province	Monthly
26	Incidence	Sub-national	District/City/ Village	Weekly
27	Incidence	Sub-national	District/City/ Village	Monthly
28	Incidence	Sub-national	Regional/Province	Monthly
29	Outbreak	Sub-national	Regional/Province	Weekly
30	Incidence	Sub-national	District/City/ Village	Bi-weekly, Monthly
31	Outbreak	Sub-national	District/City/ Village	Weekly
32	Incidence	Sub-national	District/City/ Village	Bi-monthly
33	Incidence	Sub-national	District/City/ Village	Monthly
34	Incidence	National	Country	Weekly
35	Incidence	Sub-national	District/City/ Village	Weekly
36	Incidence	Sub-national	District/City/ Village	Daily
37	Incidence	Sub-national	District/City/ Village	Weekly
38	Incidence, Outbreak	Sub-national	Regional/Province	Monthly
39	Incidence	Sub-national	District/City/ Village	Monthly
40	Incidence	National	Regional/Province	Weekly
41	Incidence	Sub-national	District/City/ Village	Weekly
42	Incidence	National	Regional/Province	Monthly
43	Incidence	National	Country	Weekly
44	Incidence	National	District/City/ Village	Monthly
45	Outbreak	Sub-national	Regional/Province	Monthly
46	Incidence, Outbreak	Sub-national	District/City/ Village	Monthly
47	Incidence	National	Regional/Province	Bi-weekly
48	Incidence	National	Country	Weekly
49	Incidence	Sub-national	Regional/Province	Monthly
50	Incidence, Peak incidence, Peak time	Sub-national	District/City/ Village	Weekly
51	Incidence	Sub-national	District/City/ Village	Weekly
52	Outbreak	Sub-national	District/City/ Village	Weekly
53	Incidence	Sub-national	District/City/ Village	Weekly

54	Incidence	National	District/City/ Village	Weekly
55	Incidence	Sub-national	District/City/ Village	Monthly
56	Incidence	Sub-national	District/City/ Village	Weekly, Monthly
57	Outbreak	Sub-national	Regional/Province	Monthly
58	Incidence	Sub-national	District/City/ Village	Monthly
59	Incidence	Sub-national	District/City/ Village	Monthly
60	Incidence	Sub-national	District/City/ Village	Monthly
61	Incidence	Sub-national	District/City/ Village	Weekly
62	Incidence, Peak incidence, Peak time	Sub-national	District/City/ Village	Weekly
63	Incidence	Sub-national	District/City/ Village	Monthly
64	Incidence	Sub-national	District/City/ Village	Monthly
65	Incidence, Outbreak	Sub-national	District/City/ Village	Monthly
66	Incidence	Global	Country	Monthly
67	Incidence	Sub-national	District/City/ Village	Monthly
68	Incidence	Sub-national	District/City/ Village	Monthly
69	Peak incidence, Peak time, Incidence	Sub-national	District/City/ Village	Weekly
70	Incidence	Sub-national	Regional/Province	Weekly
71	Incidence	Sub-national	Point/Pixel	Weekly
72	Incidence	Global	Country	Thailand: Monthly, Singapore: Weekly, Taiwan: Weekly
73	Incidence	Sub-national	District/City/ Village	Monthly
74	Outbreak	National	District/City/ Village	Weekly
75	Incidence, Outbreak	Sub-national	Regional/Province	Monthly
76	Peak incidence, Peak time, Incidence	Sub-national	District/City/ Village	Weekly

77	Incidence	National	Country	Weekly
78	Incidence, Outbreak	National	Regional/Province	Yearly
79	Incidence	Sub-national	District/City/Village	Weekly
80	Incidence	Sub-national	District/City/Village	Monthly
81	Incidence	Sub-national	District/City/Village	Monthly
82	Incidence	Sub-national	District/City/Village	Monthly
83	Incidence	National	Regional/Province	Monthly
84	Incidence	Global	District/City/Village	San Juan: Weekly, Iquitos: Weekly; Philippines: Monthly
85	Incidence	Sub-national	Regional/Province	Weekly
86	Incidence	Sub-national	District/City/Village	Monthly
87	Incidence, Outbreak	Sub-national	District/City/Village	Weekly
88	Incidence	Sub-national	District/City/Village	Monthly
89	Outbreak	Sub-national	District/City/Village	Monthly
90	Peak week, Peak incidence, Incidence	Sub-national	District/City/Village	Weekly
91	Incidence	Sub-national	District/City/Village	Bi-weekly
92	Incidence	Sub-national	District/City/Village	Monthly
93	Occurrence	Sub-national	District/City/Village	Weekly
94	Incidence, Outbreak	Sub-national	District/City/Village	Monthly
95	Incidence	Sub-national	Regional/Province	Monthly
96	Incidence	Sub-national	Regional/Province	Monthly
97	Incidence	Sub-national	District/City/Village	Weekly
98	Incidence, Peak incidence, Peak time	Global	Country	Monthly
99	Incidence	Sub-national	District/City/Village	Weekly
100	Incidence	Sub-national	District/City/Village	Yearly
101	Incidence	Sub-national	Regional/Province	Weekly

102	Incidence	National	Regional/Province	Monthly
103	Incidence	Sub-national	District/City/ Village	Monthly
104	Incidence, Outbreak	National	District/City/ Village	Weekly
105	Outbreak	Sub-national	District/City/ Village	Monthly
106	Incidence	National	Country	Weekly
107	Outbreak	National	Country	Weekly
108	Incidence	Sub-national	District/City/ Village	Daily
109	Incidence	National	Regional/Province	Monthly
110	Incidence	Sub-national	District/City/ Village	Weekly
111	Outbreak	Sub-national	District/City/ Village	Weekly
112	Incidence	Sub-national	District/City/ Village	Monthly
113	Incidence	Sub-national	District/City/ Village	Weekly
114	Outbreak	National	Country	Monthly
115	Incidence	Sub-national	Regional/Province	Monthly
116	Incidence	Sub-national	Regional/Province	Monthly
117	Incidence	National	Regional/Province	Weekly
118	Incidence	Sub-national	District/City/ Village	Weekly
119	Incidence	Global	Regional/Province	Weekly
120	Incidence	Sub-national	District/City/ Village	Weekly
121	Incidence	National	District/City/ Village	Monthly
122	Incidence	Sub-national	District/City/ Village	Monthly
123	Incidence	National	Regional/Province	Weekly
124	R0	National	District/City/ Village	Weekly
125	Incidence	Sub-national	District/City/ Village	Monthly
126	Outbreak	Sub-national	District/City/ Village	Daily
127	Incidence	Sub-national	Regional/Province	Monthly
128	Incidence, Outbreak	National	Regional/Province	Monthly
129	Incidence	Sub-national	District/City/ Village	Weekly
130	Incidence	Sub-national	District/City/ Village	Monthly
131	Incidence	Sub-national	District/City/ Village	Weekly

132	Incidence	National	Regional/Province	Monthly
133	Incidence	Sub-national	District/City/ Village	Weekly
134	Outbreak	National	District/City/ Village	Yearly
135	Incidence	Sub-national	District/City/ Village	Daily
136	Incidence	Sub-national	District/City/ Village	Weekly
137	Incidence	Sub-national	District/City/ Village	Monthly
138	Incidence	Sub-national	District/City/ Village	Monthly
139	Incidence	Sub-national	District/City/ Village	Weekly
140	Incidence	Sub-national	District/City/ Village	Weekly
141	Incidence	Global	Country	Weekly
142	Outbreak	Sub-national	District/City/ Village	6-weekly
143	Incidence	Sub-national	District/City/ Village	Weekly
144	Incidence	Sub-national	District/City/ Village	Weekly
145	Incidence	Sub-national	Regional/Province	Monthly
146	Incidence, Outbreak	Sub-national	District/City/ Village	Weekly
147	Incidence	Sub-national	Regional/Province	Weekly
148	Incidence	Sub-national	Regional/Province	Bi-weekly
149	Incidence	Sub-national	District/City/ Village	Weekly
150	Incidence	Sub-national	District/City/ Village	Monthly
151	Incidence	National	Country	Monthly
152	Incidence	Sub-national	District/City/ Village	Monthly
153	Incidence	National	Regional/Province	Monthly
154	R0, Incidence	National	point/pixel	Daily
155	Incidence	Sub-national	District/City/ Village	Monthly
156	Incidence	National	District/City/ Village	Monthly
157	Incidence	National	Regional/Province	Monthly
158	Incidence	Sub-national	District/City/ Village	Weekly
159	Incidence, Occurrence	Sub-national	District/City/ Village	Weekly
160	Outbreak	Sub-national	District/City/ Village	Daily
161	Incidence	National	District/City/ Village	Monthly

162	R0	National	District/City/ Village	Yearly
------------	----	----------	---------------------------	--------

9 References

1. World Health Organization, *Dengue haemorrhagic fever: diagnosis, treatment and control*. 2nd ed. 1997: Geneva.
2. Gardiner, J.C., Z. Luo, and L.A. Roman, *Fixed effects, random effects and GEE: what are the differences?* Stat Med, 2009. **28**(2): p. 221-39.
3. Laird, N.M. and J.H. Ware, *Random-effects models for longitudinal data*. Biometrics, 1982. **38**(4): p. 963-74.
4. Sarker, I.H., *Machine Learning: Algorithms, Real-World Applications and Research Directions*. SN Comput Sci, 2021. **2**(3): p. 160.
5. Kraemer, M.U.G., et al., *Artificial intelligence for modelling infectious disease epidemics*. Nature, 2025. **638**(8051): p. 623-635.
6. Muller, D.A., A.C. Depelsenaire, and P.R. Young, *Clinical and Laboratory Diagnosis of Dengue Virus Infection*. J Infect Dis, 2017. **215**(suppl_2): p. S89-S95.
7. Porgo, T.V., et al., *The use of mathematical modeling studies for evidence synthesis and guideline development: A glossary*. Res Synth Methods, 2019. **10**(1): p. 125-133.
8. Berry, J.S. and K. Houston, *Mathematical modelling*. Modular mathematics series. 1995, London: E. Arnold. x, 142 p.
9. Wu, H. and D. Levinson, *The ensemble approach to forecasting: A review and synthesis*. Transportation Research Part C: Emerging Technologies, 2021. **132**: p. 103357.
10. Centers for Disease Control and Prevention. *Dengue: Transmission*. 2019; Available from: <https://www.cdc.gov/dengue/transmission/index.html>.
11. World Health Organization, *Dengue Guidelines for Diagnosis, Treatment, Prevention and Control*. 2009, World Health Organization: Geneva.
12. Zhang, W.X., et al., *Assessing the global dengue burden: Incidence, mortality, and disability trends over three decades*. PLoS Negl Trop Dis, 2025. **19**(3): p. e0012932.
13. World Health Organization. *Dengue*. 2025; Available from: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>.
14. Bhatt, S., et al., *The global distribution and burden of dengue*. Nature, 2013. **496**(7446): p. 504-7.
15. Zheng, J., et al., *Global burden of dengue from 1990 to 2021: a systematic analysis from the Global Burden of Disease study 2021*. Infect Dis Poverty, 2025. **14**(1): p. 105.
16. Stanaway, J.D., et al., *The global burden of dengue: an analysis from the Global Burden of Disease Study 2013*. Lancet Infect Dis, 2016. **16**(6): p. 712-723.
17. Undurraga, E.A., Y.A. Halasa, and D.S. Shepard, *Use of expansion factors to estimate the burden of dengue in Southeast Asia: a systematic analysis*. PLoS Negl Trop Dis, 2013. **7**(2): p. e2056.
18. Haider, N., et al., *Global dengue epidemic worsens with record 14 million cases and 9000 deaths reported in 2024*. Int J Infect Dis, 2025. **158**: p. 107940.
19. Yang, X., et al., *Global burden for dengue and the evolving pattern in the past 30 years*. J Travel Med, 2021. **28**(8).

20. Deng, J., et al., *Global, regional, and national burden of dengue infection in children and adolescents: an analysis of the Global Burden of Disease Study 2021*. EClinicalMedicine, 2024. **78**: p. 102943.
21. Silburn, A. and J. Arndell, *The impact of dengue viruses: Surveillance, response, and public health implications in Queensland, Australia*. Public Health Pract (Oxf), 2024. **8**: p. 100529.
22. Marczell, K., et al., *The macroeconomic impact of a dengue outbreak: Case studies from Thailand and Brazil*. PLoS Negl Trop Dis, 2024. **18**(6): p. e0012201.
23. Malavige, G.N., et al., *Facing the escalating burden of dengue: Challenges and perspectives*. PLOS Glob Public Health, 2023. **3**(12): p. e0002598.
24. Leelavanich, D., I. Dorigatti, and H.C. Turner, *The Economic Burden of Dengue: A Systematic Literature Review of Cost-of-Illness Studies*. medRxiv, 2025: p. 2025.08.21.25334162.
25. Shepard, D.S., et al., *The global economic burden of dengue: a systematic analysis*. Lancet Infect Dis, 2016. **16**(8): p. 935-41.
26. Mills, C., et al., *Multi-model approach to understand and predict past and future dengue epidemic dynamics*. R Soc Open Sci, 2025. **12**(11): p. 241870.
27. Oneda, R.M., et al., *Epidemiological profile of dengue in Brazil between the years 2014 and 2019*. Rev Assoc Med Bras (1992), 2021. **67**(5): p. 731-735.
28. World Health Organization. *Dengue and severe dengue*. 2023; Available from: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>.
29. World Health Organization, S.P.f.R.a.T.i.T.D., *Dengue: guidelines for diagnosis, treatment, prevention and control - New edition*. 2009.
30. Torres-Flores, J.M., A. Reyes-Sandoval, and M.I. Salazar, *Dengue Vaccines: An Update*. BioDrugs, 2022. **36**(3): p. 325-336.
31. Palanichamy Kala, M., A.L. St John, and A.P.S. Rathore, *Dengue: Update on Clinically Relevant Therapeutic Strategies and Vaccines*. Curr Treat Options Infect Dis, 2023. **15**(2): p. 27-52.
32. Pecci, A. *Dengue is spreading and Sanofi is pulling the US market's only vaccine. What's next?* 2024.
33. Takeda Pharmaceutical Company Limited. *Dengue vaccine (News Release)*. 2025; Available from: <https://www.takeda.com/newsroom/newsreleases/2025/dengue-vaccine/>.
34. European Medicines Agency. *Qdenga (TAK-003)*. 2022 2024; Available from: <https://www.ema.europa.eu/en/medicines/human/EPAR/qdenga>.
35. Thai Travel Clinic (TravM Dr & Dr Putri). *Dengue vaccine for foreigners/travelers in Thailand: Should I get it? UPDATE 2023*. 2023; Available from: <https://www.thaitravelclinic.com/blog/vaccineinfo/dengue-vaccine-for-foreigners-travelers-in-thailand-should-i-get-it-update-2023.html>.
36. World Health Organization (WHO). *Dengue outbreak toolbox*. 2024; Available from: <https://www.who.int/emergencies/outbreak-toolkit/disease-outbreak-toolboxes/dengue-outbreak-toolbox>.
37. Tang, K.F. and E.E. Ooi, *Diagnosis of dengue: an update*. Expert Rev Anti Infect Ther, 2012. **10**(8): p. 895-907.

38. Sánchez-González, L.A., Laura E.; Paz-Bailey, Gabriela, *Dengue*, in *CDC Yellow Book: Health Information for International Travel*. 2026, Centers for Disease Control and Prevention: Atlanta, GA.
39. Centers for Disease Control and Prevention. *Serologic Tests for Dengue Virus*. 2025; Available from: <https://www.cdc.gov/dengue/hcp/diagnosis-testing/serologic-tests-for-dengue-virus.html>.
40. Blacksell, S.D., et al., *Evaluation of six commercial point-of-care tests for diagnosis of acute dengue infections: the need for combining NSI antigen and IgM/IgG antibody detection to achieve acceptable levels of accuracy*. Clin Vaccine Immunol, 2011. **18**(12): p. 2095-101.
41. Bin Suaydan, R.A., et al., *Combining non-structural protein 1 and IgM tests for acute dengue diagnosis in Saudi Arabia*. Int Health, 2025.
42. Fry, S.R., et al., *The diagnostic sensitivity of dengue rapid test assays is significantly enhanced by using a combined antigen and antibody testing approach*. PLoS Negl Trop Dis, 2011. **5**(6): p. e1199.
43. Teoh, B.T., et al., *The Use of NSI Rapid Diagnostic Test and qRT-PCR to Complement IgM ELISA for Improved Dengue Diagnosis from Single Specimen*. Sci Rep, 2016. **6**: p. 27663.
44. Versiani, A.F., et al., *Performance of VIDAS((R)) Diagnostic Tests for the Automated Detection of Dengue Virus NSI Antigen and of Anti-Dengue Virus IgM and IgG Antibodies: A Multicentre, International Study*. Diagnostics (Basel), 2023. **13**(6).
45. Zhang, H., et al., *NSI-based tests with diagnostic utility for confirming dengue infection: a meta-analysis*. Int J Infect Dis, 2014. **26**: p. 57-66.
46. Chan, H.B.Y., C.H. How, and C.W. Ng, *Definitive tests for dengue fever: when and which should I use?* Singapore Med J, 2017. **58**(11): p. 632-635.
47. Shepard, D.S., E.A. Undurraga, and Y.A. Halasa, *Economic and disease burden of dengue in Southeast Asia*. PLoS Negl Trop Dis, 2013. **7**(2): p. e2055.
48. De Santis, O., N. Bouscaren, and A. Flahault, *Asymptomatic dengue infection rate: A systematic literature review*. Heliyon, 2023. **9**(9): p. e20069.
49. Xiao, Y., et al., *Dengue nowcasting in Brazil by combining official surveillance data and Google Trends information*. PLoS Negl Trop Dis, 2025. **19**(8): p. e0012501.
50. Reich, N.G., et al., *Challenges in Real-Time Prediction of Infectious Disease: A Case Study of Dengue in Thailand*. PLoS Negl Trop Dis, 2016. **10**(6): p. e0004761.
51. Bastos, L.S., et al., *A modelling approach for correcting reporting delays in disease surveillance data*. Stat Med, 2019. **38**(22): p. 4363-4377.
52. McGough, S.F., et al., *Nowcasting by Bayesian Smoothing: A flexible, generalizable model for real-time epidemic tracking*. PLoS Comput Biol, 2020. **16**(4): p. e1007735.
53. Pollett, S., et al., *Understanding dengue virus evolution to support epidemic surveillance and counter-measure development*. Infect Genet Evol, 2018. **62**: p. 279-295.
54. Yu, X. and G. Cheng, *Contribution of phylogenetics to understanding the evolution and epidemiology of dengue virus*. Animal Model Exp Med, 2022. **5**(5): p. 410-417.

55. Grubaugh, N.D., et al., *Tracking virus outbreaks in the twenty-first century*. Nat Microbiol, 2019. **4**(1): p. 10-19.
56. Hall, M.D., M.E. Woolhouse, and A. Rambaut, *Using genomics data to reconstruct transmission trees during disease outbreaks*. Rev Sci Tech, 2016. **35**(1): p. 287-96.
57. Holmes, E.C. and S.S. Twiddy, *The origin, emergence and evolutionary genetics of dengue virus*. Infect Genet Evol, 2003. **3**(1): p. 19-28.
58. Mir, D., et al., *Spatiotemporal dynamics of DENV-2 Asian-American genotype lineages in the Americas*. PLoS One, 2014. **9**(6): p. e98519.
59. Martinez, D., et al., *Emergence of Dengue Virus Serotype 2 Cosmopolitan Genotype, Colombia*. Emerg Infect Dis, 2024. **30**(1): p. 189-192.
60. Ribeiro, J.R., et al., *DENV-2 Outbreak Associated With Cosmopolitan Genotype Emergence in Western Brazilian Amazon*. Bioinform Biol Insights, 2024. **18**: p. 11779322241251581.
61. Yenamandra, S.P., et al., *Evolution, heterogeneity and global dispersal of cosmopolitan genotype of Dengue virus type 2*. Sci Rep, 2021. **11**(1): p. 13496.
62. Amorim, M.T., et al., *Emergence of a New Strain of DENV-2 in South America: Introduction of the Cosmopolitan Genotype through the Brazilian-Peruvian Border*. Trop Med Infect Dis, 2023. **8**(6).
63. Giovanetti, M., et al., *Emergence of Dengue Virus Serotype 2 Cosmopolitan Genotype, Brazil*. Emerg Infect Dis, 2022. **28**(8): p. 1725-1727.
64. Gonzalez-Elizondo, M., et al., *Shifting Dynamics of Dengue Virus Serotype 2 and Emergence of Cosmopolitan Genotype, Costa Rica, 2024*. Emerg Infect Dis, 2025. **31**(11): p. 2153-2158.
65. Kubiszeski, J.R., et al., *Detection of the Asian II genotype of dengue virus serotype 2 in humans and mosquitoes in Brazil*. Rev Soc Bras Med Trop, 2020. **53**: p. e20190439.
66. Tracey, W.D., *The taste of water*. Elife, 2019. **8**.
67. Bowatte, G., et al., *Tadpoles as dengue mosquito (Aedes aegypti) egg predators*. Biological Control, 2013. **67**(3): p. 469-474.
68. Samuel, M., et al., *Community effectiveness of indoor spraying as a dengue vector control method: A systematic review*. PLoS Negl Trop Dis, 2017. **11**(8): p. e0005837.
69. Manrique-Saide, P., et al., *Use of insecticide-treated house screens to reduce infestations of dengue virus vectors, Mexico*. Emerg Infect Dis, 2015. **21**(2): p. 308-11.
70. Minwuyelet, A., et al., *Symbiotic Wolbachia in mosquitoes and its role in reducing the transmission of mosquito-borne diseases: updates and prospects*. Front Microbiol, 2023. **14**: p. 1267832.
71. Mustafa, M.S., et al., *Wolbachia: The selfish Trojan Horse in dengue control*. Med J Armed Forces India, 2016. **72**(4): p. 373-376.
72. Utarini, A., et al., *Efficacy of Wolbachia-Infected Mosquito Deployments for the Control of Dengue*. N Engl J Med, 2021. **384**(23): p. 2177-2186.
73. Chow, J.Y., et al., *Assessing the direct and spillover protective effectiveness of Wolbachia-mediated introgression to combat dengue*. EBioMedicine, 2024. **110**: p. 105456.

74. World Mosquito Program. *Impact*. 2025; Available from: <https://www.worldmosquitoprogram.org/en/work/wolbachia-method/impact>.
75. Arham, A.F., et al., *Fogging to combat dengue: factors influencing stakeholders' attitudes in Malaysia*. BMC Public Health, 2023. **23**(1): p. 1140.
76. Kawada, H., et al., *Insecticidal and repellent activities of pyrethroids to the three major pyrethroid-resistant malaria vectors in western Kenya*. Parasites & Vectors, 2014. **7**(1): p. 208.
77. Hassan, M.R., et al., *Insecticide resistance of Dengue vectors in South East Asia: a systematic review*. Afr Health Sci, 2021. **21**(3): p. 1124-1140.
78. Gan, S.J., et al., *Dengue fever and insecticide resistance in Aedes mosquitoes in Southeast Asia: a review*. Parasit Vectors, 2021. **14**(1): p. 315.
79. Tognarelli, J., et al., *Global distribution and impact of knockdown resistance mutations in Aedes aegypti on pyrethroid resistance*. Parasit Vectors, 2025. **18**(1): p. 382.
80. Schmidt, T.L., et al., *Global, asynchronous partial sweeps at multiple insecticide resistance genes in Aedes mosquitoes*. Nat Commun, 2024. **15**(1): p. 6251.
81. Asgarian, T.S., et al., *Worldwide Status of Insecticide Resistance of Aedes aegypti and Ae. albopictus, Vectors of Arboviruses of Chikungunya, Dengue, Zika and Yellow Fever*. J Arthropod Borne Dis, 2023. **17**(1): p. 1-27.
82. Al-Amin, H.M., et al., *Insecticide resistance compromises the control of Aedes aegypti in Bangladesh*. Pest Manag Sci, 2023. **79**(8): p. 2846-2861.
83. Sa, E.L.R., et al., *Evaluation of insecticide resistance in Aedes aegypti populations connected by roads and rivers: the case of Tocantins state in Brazil*. Mem Inst Oswaldo Cruz, 2019. **114**: p. e180318.
84. Lima, E.P., et al., *Insecticide resistance in Aedes aegypti populations from Ceara, Brazil*. Parasit Vectors, 2011. **4**: p. 5.
85. Ridha, M.R., et al., *A systematic review of insecticide resistance in Aedes aegypti (Diptera: Culicidae) and implications for dengue control in Indonesia*. Vet World, 2025. **18**(3): p. 658-672.
86. Francis, S., et al., *Screening of insecticide resistance in Aedes aegypti populations collected from parishes in Eastern Jamaica*. PLoS Negl Trop Dis, 2020. **14**(7): p. e0008490.
87. Rasli, R., et al., *Insecticide resistance in dengue vectors from hotspots in Selangor, Malaysia*. PLoS Negl Trop Dis, 2021. **15**(3): p. e0009205.
88. Rasli, R., et al., *Susceptibility Status and Resistance Mechanisms in Permethrin-Selected, Laboratory Susceptible and Field-Collected Aedes aegypti from Malaysia*. Insects, 2018. **9**(2).
89. Haidy Massa, M., et al., *Insecticide Resistance Status of Aedes aegypti Adults and Larvae in Nouakchott, Mauritania*. Insects, 2025. **16**(3).
90. Sene, N.M., et al., *Insecticide resistance status and mechanisms in Aedes aegypti populations from Senegal*. PLoS Negl Trop Dis, 2021. **15**(5): p. e0009393.
91. Fernando, H.S.D., et al., *Resistance to commonly used insecticides and underlying mechanisms of resistance in Aedes aegypti (L.) from Sri Lanka*. Parasites & Vectors, 2020. **13**(1): p. 407.
92. Smith, D.L., et al., *Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens*. PLoS Pathog, 2012. **8**(4): p. e1002588.

93. Sanches, R.P. and E. Massad, *A comparative analysis of three different methods for the estimation of the basic reproduction number of dengue*. Infect Dis Model, 2016. **1**(1): p. 88-100.
94. Vicari, A., *Dengue modeling*. 2014, Pan American Health Organization / World Health Organization: Washington, DC.
95. Couret, J. and M.Q. Benedict, *A meta-analysis of the factors influencing development rate variation in Aedes aegypti (Diptera: Culicidae)*. BMC Ecol, 2014. **14**: p. 3.
96. Brady, O.J., et al., *Modelling adult Aedes aegypti and Aedes albopictus survival at different temperatures in laboratory and field settings*. Parasit Vectors, 2013. **6**: p. 351.
97. Reinhold, J.M., C.R. Lazzari, and C. Lahondere, *Effects of the Environmental Temperature on Aedes aegypti and Aedes albopictus Mosquitoes: A Review*. Insects, 2018. **9**(4).
98. Edussuriya, C., S. Deegalla, and I. Gawarammana, *An accurate mathematical model predicting number of dengue cases in tropics*. PLoS Negl Trop Dis, 2021. **15**(11): p. e0009756.
99. Yin, Q., et al., *A field-based modeling study on ecological characterization of hourly host-seeking behavior and its associated climatic variables in Aedes albopictus*. Parasit Vectors, 2019. **12**(1): p. 474.
100. Tjaden, N.B., et al., *Extrinsic Incubation Period of Dengue: Knowledge, Backlog, and Applications of Temperature Dependence*. PLoS Negl Trop Dis, 2013. **7**(6): p. e2207.
101. Chan, M. and M.A. Johansson, *The incubation periods of Dengue viruses*. PLoS One, 2012. **7**(11): p. e50972.
102. Trejo, I., et al., *Changing temperature profiles and the risk of dengue outbreaks*. PLoS Climate, 2023. **2**(2): p. e0000115.
103. Seposo, X., S. Valenzuela, and G.L. Apostol, *Socio-economic factors and its influence on the association between temperature and dengue incidence in 61 Provinces of the Philippines, 2010-2019*. PLoS Negl Trop Dis, 2023. **17**(10): p. e0011700.
104. Wu, X., et al., *Non-linear effects of mean temperature and relative humidity on dengue incidence in Guangzhou, China*. Sci Total Environ, 2018. **628-629**: p. 766-771.
105. Xu, Z., et al., *Spatiotemporal patterns and climatic drivers of severe dengue in Thailand*. Sci Total Environ, 2019. **656**: p. 889-901.
106. Campbell, K.M., et al., *The complex relationship between weather and dengue virus transmission in Thailand*. Am J Trop Med Hyg, 2013. **89**(6): p. 1066-1080.
107. Fan, J., et al., *A systematic review and meta-analysis of dengue risk with temperature change*. Int J Environ Res Public Health, 2014. **12**(1): p. 1-15.
108. Damtew, Y.T., et al., *Effects of high temperatures and heatwaves on dengue fever: a systematic review and meta-analysis*. EBioMedicine, 2023. **91**: p. 104582.
109. Abdullah, N., et al., *The association between dengue case and climate: A systematic review and meta-analysis*. One Health, 2022. **15**: p. 100452.

110. Lega, J., H.E. Brown, and R. Barrera, *Aedes aegypti* (Diptera: Culicidae) Abundance Model Improved With Relative Humidity and Precipitation-Driven Egg Hatching. J Med Entomol, 2017. **54**(5): p. 1375-1384.
111. Alto, B.W. and S.A. Juliano, *Precipitation and temperature effects on populations of Aedes albopictus* (Diptera: Culicidae): implications for range expansion. J Med Entomol, 2001. **38**(5): p. 646-56.
112. Montgomery, M.J., et al., *The effects of urbanization, temperature, and rainfall on Aedes aegypti and Aedes albopictus mosquito abundance across a broad latitudinal gradient in Central Africa*. Parasit Vectors, 2025. **18**(1): p. 135.
113. Barrera, R., et al., *Use of the CDC autocidal gravid ovitrap to control and prevent outbreaks of Aedes aegypti* (Diptera: Culicidae). J Med Entomol, 2014. **51**(1): p. 145-54.
114. Barrera, R., et al., *Sustained, area-wide control of Aedes aegypti using CDC autocidal gravid ovitraps*. Am J Trop Med Hyg, 2014. **91**(6): p. 1269-1276.
115. Lourenco-de-Oliveira, R., et al., *The invasion of urban forest by dengue vectors in Rio de Janeiro*. J Vector Ecol, 2004. **29**(1): p. 94-100.
116. Lounibos, L., *Habitat segregation among African treehole mosquitoes*. Ecological Entomology, 1981. **6**(2): p. 129-154.
117. Roiz, D., et al., *Effects of temperature and rainfall on the activity and dynamics of host-seeking Aedes albopictus females in northern Italy*. Vector Borne Zoonotic Dis, 2010. **10**(8): p. 811-6.
118. Higa, Y., et al., *Seasonal changes in oviposition activity, hatching and embryonation rates of eggs of Aedes albopictus* (Diptera: Culicidae) on three islands of the Ryukyu Archipelago, Japan Med. Entomol. Zool., 2007. **58**(1): p. 1-10.
119. Koenraadt, C.J. and L.C. Harrington, *Flushing effect of rain on container-inhabiting mosquitoes Aedes aegypti and Culex pipiens* (Diptera: Culicidae). J Med Entomol, 2008. **45**(1): p. 28-35.
120. Lamy, K., et al., *Impact of regional climate change on the mosquito vector Aedes albopictus in a tropical island environment: La Reunion*. Sci Total Environ, 2023. **875**: p. 162484.
121. Legros, M., et al., *Comparison of Two Detailed Models of Aedes aegypti Population Dynamics*. Ecosphere, 2016. **7**(10).
122. Valdez, L.D., G.J. Sibona, and C.A. Condat, *Impact of rainfall on Aedes aegypti populations*. Ecological Modelling, 2018. **385**: p. 96-105.
123. Newman, E.A., et al., *Defining the roles of local precipitation and anthropogenic water sources in driving the abundance of Aedes aegypti, an emerging disease vector in urban, arid landscapes*. Sci Rep, 2024. **14**(1): p. 2058.
124. Barrera, R., M. Amador, and A.J. MacKay, *Population dynamics of Aedes aegypti and dengue as influenced by weather and human behavior in San Juan, Puerto Rico*. PLoS Negl Trop Dis, 2011. **5**(12): p. e1378.
125. Walker, K.R., et al., *Socioeconomic and Human Behavioral Factors Associated With Aedes aegypti* (Diptera: Culicidae) Immature Habitat in Tucson, AZ. J Med Entomol, 2018. **55**(4): p. 955-963.

126. Li, Y., et al., *Effects of ambient temperature and precipitation on the risk of dengue fever: A systematic review and updated meta-analysis*. Environ Res, 2020. **191**: p. 110043.
127. Lorenz, C., et al., *Influence of climate and heatwaves on dengue transmission in Sao Paulo and Natal, Brazil*. PLoS One, 2025. **20**(10): p. e0334838.
128. Mckee, T.B., N.J. Doesken, and J.R. Kleist. *THE RELATIONSHIP OF DROUGHT FREQUENCY AND DURATION TO TIME SCALES*. 1993.
129. World Meteorological Organization, *Standardized Precipitation Index User Guide*, in *Standardized Precipitation Index User Guide*. 2012, World Meteorological Organization: Geneva.
130. Vicente-Serrano, S.M., S. Beguería, and J.I. López-Moreno, *A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index*. Journal of Climate, 2010. **23**(7): p. 1696-1718.
131. Li, C., et al., *Association Between Hydrological Conditions and Dengue Fever Incidence in Coastal Southeastern China From 2013 to 2019*. JAMA Netw Open, 2023. **6**(1): p. e2249440.
132. Cox, V.M., et al., *Spatiotemporal relationships between extreme weather events and arbovirus transmission across Brazil*. 2025, eLife Sciences Publications, Ltd.
133. Lowe, R., et al., *Combined effects of hydrometeorological hazards and urbanisation on dengue risk in Brazil: a spatiotemporal modelling study*. Lancet Planet Health, 2021. **5**(4): p. e209-e219.
134. Li, C., et al., *Projecting future risk of dengue related to hydrometeorological conditions in mainland China under climate change scenarios: a modelling study*. Lancet Planet Health, 2023. **7**(5): p. e397-e406.
135. Wang, H., Y. Pan, and Y. Chen, *Comparison of three drought indices and their evolutionary characteristics in the arid region of northwestern China*. Atmospheric Science Letters, 2017. **18**(3): p. 132-139.
136. Schmidt, C.A., et al., *Effects of desiccation stress on adult female longevity in Aedes aegypti and Ae. albopictus (Diptera: Culicidae): results of a systematic review and pooled survival analysis*. Parasit Vectors, 2018. **11**(1): p. 267.
137. Jeffrey Gutierrez, E.H., et al., *Size as a Proxy for Survival in Aedes aegypti (Diptera: Culicidae) Mosquitoes*. J Med Entomol, 2020. **57**(4): p. 1228-1238.
138. Costa, E.d.A., et al., *Impact of small variations in temperature and humidity on the reproductive activity and survival of Aedes aegypti (Diptera, Culicidae)*. Revista Brasileira de Entomologia, 2010. **54**(3): p. 488-493.
139. Abu, A.E.I., et al., *Low humidity enhances Zika virus infection and dissemination in Aedes aegypti mosquitoes*. mSphere, 2024. **9**(8): p. e0040124.
140. Manzano-Alvarez, J., et al., *Dehydration stress and Mayaro virus vector competence in Aedes aegypti*. J Virol, 2023. **97**(12): p. e0069523.
141. Canyon, D.V., R. Muller, and J.L. Hii, *Aedes aegypti disregard humidity-related conditions with adequate nutrition*. Trop Biomed, 2013. **30**(1): p. 1-8.
142. Thu, H.M., K.M. Aye, and S. Thein, *The effect of temperature and humidity on dengue virus propagation in Aedes aegypti mosquitos*. Southeast Asian J Trop Med Public Health, 1998. **29**(2): p. 280-4.
143. Mamenun, et al., *Spatiotemporal Characterization of Dengue Incidence and Its Correlation to Climate Parameters in Indonesia*. Insects, 2024. **15**(5).

144. Ramachandran, V.G., et al., *Empirical model for estimating dengue incidence using temperature, rainfall, and relative humidity: a 19-year retrospective analysis in East Delhi*. Epidemiol Health, 2016. **38**: p. e2016052.
145. Sophia, Y., et al., *Dengue dynamics, predictions, and future increase under changing monsoon climate in India*. Sci Rep, 2025. **15**(1): p. 1637.
146. Ninphanomchai, S., et al., *Predictiveness of disease risk in a global outreach tourist setting in Thailand using meteorological data and vector-borne disease incidences*. Int J Environ Res Public Health, 2014. **11**(10): p. 10694-709.
147. Sirisena, P., et al., *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, 2017. **12**(1): p. e0166806.
148. Costa-da-Silva, A.L., et al., *Landscape and meteorological variables associated with Aedes aegypti and Aedes albopictus mosquito infestation in two southeastern USA coastal cities*. bioRxiv, 2024.
149. Osalla, J., et al., *Longitudinal surveillance of Aedes aegypti (Diptera: Culicidae) in urban coastal Kenya: population dynamics, blood feeding frequency and dengue virus infection rates*. Sci Rep, 2025. **15**(1): p. 21787.
150. Reiner, R.C., Jr., et al., *Estimating the impact of city-wide Aedes aegypti population control: An observational study in Iquitos, Peru*. PLoS Negl Trop Dis, 2019. **13**(5): p. e0007255.
151. Adeleke, E.D., et al., *High Wind Speed Prevents the Establishment of the Disease Vector Mosquito Aedes albopictus in Its Climatic Niche in Europe*. Frontiers in Environmental Science, 2022. **Volume 10 - 2022**.
152. Islam, M.A., et al., *Correlation of Dengue and Meteorological Factors in Bangladesh: A Public Health Concern*. Int J Environ Res Public Health, 2023. **20**(6).
153. Sulekan, A., J. Suhaila, and N.A. Abdul Wahid, *Assessing the Effect of Climate Factors on Dengue Incidence via a Generalized Linear Model*. Open Journal of Applied Sciences, 2021. **11**(4): p. 549-563.
154. Singh, S., et al., *The Effects of Meteorological Factors on Dengue Cases in Malaysia*. Int J Environ Res Public Health, 2022. **19**(11).
155. Gomez Gomez, R.E., et al., *Association between Climate Factors and Dengue Fever in Asuncion, Paraguay: A Generalized Additive Model*. Int J Environ Res Public Health, 2022. **19**(19).
156. Yu, X., X. Wang, and S. Tang, *Assessing the influencing factors of dengue fever in Chinese mainland based on causal analysis*. Sci Rep, 2025. **15**(1): p. 15311.
157. Hossain, S., et al., *Association of climate factors with dengue incidence in Bangladesh, Dhaka City: A count regression approach*. Heliyon, 2023. **9**(5): p. e16053.
158. Rahman, M.S., M. Amrin, and M.A. Bokkor Shiddik, *Dengue Early Warning System and Outbreak Prediction Tool in Bangladesh Using Interpretable Tree-Based Machine Learning Model*. Health Sci Rep, 2025. **8**(5): p. e70726.
159. NASA Earthdata. *Normalized Difference Vegetation Index (NDVI)*. 2025 July 30, 2025; Available from: <https://www.earthdata.nasa.gov/topics/land-surface/normalized-difference-vegetation-index-ndvi>.

160. Ordóñez-Sierra, R., et al., *Spatial Risk Distribution of Dengue Based on the Ecological Niche Model of Aedes aegypti (Diptera: Culicidae) in the Central Mexican Highlands*. J Med Entomol, 2020. **57**(3): p. 728-737.
161. Andreo, V., et al., *Spatial Distribution of Aedes aegypti Oviposition Temporal Patterns and Their Relationship with Environment and Dengue Incidence*. Insects, 2021. **12**(10).
162. Tsheten, T., et al., *Spatial and temporal patterns of dengue incidence in Bhutan: a Bayesian analysis*. Emerg Microbes Infect, 2020. **9**(1): p. 1360-1371.
163. Soukavong, M., et al., *Bayesian spatio-temporal analysis of dengue transmission in Lao PDR*. Sci Rep, 2024. **14**(1): p. 21327.
164. Qi, X., et al., *The Effects of Socioeconomic and Environmental Factors on the Incidence of Dengue Fever in the Pearl River Delta, China, 2013*. PLoS Negl Trop Dis, 2015. **9**(10): p. e0004159.
165. Whiteman, A., et al., *Do socioeconomic factors drive Aedes mosquito vectors and their arboviral diseases? A systematic review of dengue, chikungunya, yellow fever, and Zika Virus*. One Health, 2020. **11**: p. 100188.
166. Holeva-Eklund, W.M., T.K. Behrens, and C.M. Hepp, *Systematic review: the impact of socioeconomic factors on Aedes aegypti mosquito distribution in the mainland United States*. Rev Environ Health, 2021. **36**(1): p. 63-75.
167. Rodrigues Mde, M., et al., *Density of Aedes aegypti and Aedes albopictus and its association with number of residents and meteorological variables in the home environment of dengue endemic area, Sao Paulo, Brazil*. Parasit Vectors, 2015. **8**: p. 115.
168. Schmidt, W.P., et al., *Population density, water supply, and the risk of dengue fever in Vietnam: cohort study and spatial analysis*. PLoS Med, 2011. **8**(8): p. e1001082.
169. Perkins, T.A., et al., *Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission*. PLoS Comput Biol, 2013. **9**(12): p. e1003327.
170. Hemme, R.R., et al., *Influence of urban landscapes on population dynamics in a short-distance migrant mosquito: evidence for the dengue vector Aedes aegypti*. PLoS Negl Trop Dis, 2010. **4**(3): p. e634.
171. Wilder-Smith, A. and D.J. Gubler, *Geographic expansion of dengue: the impact of international travel*. Med Clin North Am, 2008. **92**(6): p. 1377-90, x.
172. Stoddard, S.T., et al., *House-to-house human movement drives dengue virus transmission*. Proc Natl Acad Sci U S A, 2013. **110**(3): p. 994-9.
173. Massaro, E., D. Kondor, and C. Ratti, *Assessing the interplay between human mobility and mosquito borne diseases in urban environments*. Sci Rep, 2019. **9**(1): p. 16911.
174. Wesolowski, A., et al., *Impact of human mobility on the emergence of dengue epidemics in Pakistan*. Proc Natl Acad Sci U S A, 2015. **112**(38): p. 11887-92.
175. Kiang, M.V., et al., *Incorporating human mobility data improves forecasts of Dengue fever in Thailand*. Sci Rep, 2021. **11**(1): p. 923.
176. Zhang, Y., et al., *Modeling the relative role of human mobility, land-use and climate factors on dengue outbreak emergence in Sri Lanka*. BMC Infect Dis, 2020. **20**(1): p. 649.

177. Gibb, R., et al., *Interactions between climate change, urban infrastructure and mobility are driving dengue emergence in Vietnam*. Nat Commun, 2023. **14**(1): p. 8179.
178. Lefebvre, B., et al., *Importance of Public Transport Networks for Reconciling the Spatial Distribution of Dengue and the Association of Socio-Economic Factors with Dengue Risk in Bangkok, Thailand*. Int J Environ Res Public Health, 2022. **19**(16).
179. Bomfim, R., et al., *Predicting dengue outbreaks at neighbourhood level using human mobility in urban areas*. J R Soc Interface, 2020. **17**(171): p. 20200691.
180. Ospina-Aguirre, C., et al., *Effects of human mobility on the spread of Dengue in the region of Caldas, Colombia*. PLoS Negl Trop Dis, 2023. **17**(11): p. e0011087.
181. Schaber, K.L., et al., *Disease-driven reduction in human mobility influences human-mosquito contacts and dengue transmission dynamics*. PLoS Comput Biol, 2021. **17**(1): p. e1008627.
182. Khan, M., et al., *Interactive effects of climate change and human mobility on dengue transmission*. Ecological Modelling, 2025. **499**: p. 110924.
183. Abdul-Ghani, R., et al., *Impact of population displacement and forced movements on the transmission and outbreaks of Aedes-borne viral diseases: Dengue as a model*. Acta Trop, 2019. **197**: p. 105066.
184. Leung, X.Y., et al., *A systematic review of dengue outbreak prediction models: Current scenario and future directions*. PLoS Negl Trop Dis, 2023. **17**(2): p. e0010631.
185. Wu, S., et al., *Ensemble approaches for short-term dengue fever forecasts: A global evaluation study*. Proc Natl Acad Sci U S A, 2025. **122**(33): p. e2422335122.
186. Baharom, M., et al., *Dengue Early Warning System as Outbreak Prediction Tool: A Systematic Review*. Risk Manag Healthc Policy, 2022. **15**: p. 871-886.
187. Eisen, L., et al., *Proactive Vector control strategies and improved monitoring and evaluation practices for dengue prevention*. J Med Entomol, 2009. **46**(6): p. 1245-55.
188. Chen, X. and P. Moraga, *Assessing dengue forecasting methods: a comparative study of statistical models and machine learning techniques in Rio de Janeiro, Brazil*. Trop Med Health, 2025. **53**(1): p. 52.
189. Araujo, E.C., et al., *Leveraging probabilistic forecasts for dengue preparedness and control: the 2024 Dengue Forecasting Sprint in Brazil*. medRxiv, 2025: p. 2025.05.12.25327419.
190. Johansson, M.A., et al., *An open challenge to advance probabilistic forecasting for dengue epidemics*. Proc Natl Acad Sci U S A, 2019. **116**(48): p. 24268-24274.
191. Sebastianelli, A., et al., *A reproducible ensemble machine learning approach to forecast dengue outbreaks*. Sci Rep, 2024. **14**(1): p. 3807.
192. Earnest, A., et al., *Comparing statistical models to predict dengue fever notifications*. Comput Math Methods Med, 2012. **2012**: p. 758674.
193. Gui, H., et al., *Weather Factors Associated with Reduced Risk of Dengue Transmission in an Urbanized Tropical City*. Int J Environ Res Public Health, 2021. **19**(1).

194. Karasinghe, N., et al., *Forecasting weekly dengue incidence in Sri Lanka: Modified Autoregressive Integrated Moving Average modeling approach*. PLoS One, 2024. **19**(3): p. e0299953.
195. Luz, P.M., et al., *Time series analysis of dengue incidence in Rio de Janeiro, Brazil*. Am J Trop Med Hyg, 2008. **79**(6): p. 933-9.
196. Martinez, E.Z., E.A. Silva, and A.L. Fabbro, *A SARIMA forecasting model to predict the number of cases of dengue in Campinas, State of Sao Paulo, Brazil*. Rev Soc Bras Med Trop, 2011. **44**(4): p. 436-40.
197. Lu, X., et al., *Application of multiple linear regression model and long short-term memory with compartmental model to forecast dengue cases in Selangor, Malaysia based on climate variables*. Infect Dis Model, 2025. **10**(1): p. 240-256.
198. Hii, Y.L., et al., *Forecast of dengue incidence using temperature and rainfall*. PLoS Negl Trop Dis, 2012. **6**(11): p. e1908.
199. Sang, S., et al., *Predicting unprecedented dengue outbreak using imported cases and climatic factors in Guangzhou, 2014*. PLoS Negl Trop Dis, 2015. **9**(5): p. e0003808.
200. Baquero, O.S., L.M.R. Santana, and F. Chiaravalloti-Neto, *Dengue forecasting in Sao Paulo city with generalized additive models, artificial neural networks and seasonal autoregressive integrated moving average models*. PLoS One, 2018. **13**(4): p. e0195065.
201. Picinini Freitas, L., et al., *A statistical model for forecasting probabilistic epidemic bands for dengue cases in Brazil*. Infect Dis Model, 2025. **10**(4): p. 1479-1487.
202. Zhao, N., et al., *Machine learning and dengue forecasting: Comparing random forests and artificial neural networks for predicting dengue burden at national and sub-national scales in Colombia*. PLoS Negl Trop Dis, 2020. **14**(9): p. e0008056.
203. Nguyen, V.H., et al., *Deep learning models for forecasting dengue fever based on climate data in Vietnam*. PLoS Negl Trop Dis, 2022. **16**(6): p. e0010509.
204. Tian, N., et al., *Precision Prediction for Dengue Fever in Singapore: A Machine Learning Approach Incorporating Meteorological Data*. Trop Med Infect Dis, 2024. **9**(4).
205. Guo, P., et al., *Developing a dengue forecast model using machine learning: A case study in China*. PLoS Negl Trop Dis, 2017. **11**(10): p. e0005973.
206. Mussumeci, E. and F. Codeco Coelho, *Large-scale multivariate forecasting models for Dengue - LSTM versus random forest regression*. Spat Spatiotemporal Epidemiol, 2020. **35**: p. 100372.
207. Liu, Y., et al., *An explainable covariate compartmental model for predicting the spatio-temporal patterns of dengue in Sri Lanka*. PLoS Comput Biol, 2025. **21**(9): p. e1013540.
208. Pollett, S., et al., *Recommended reporting items for epidemic forecasting and prediction research: The EPIFORGE 2020 guidelines*. PLoS Med, 2021. **18**(10): p. e1003793.
209. Hodson, T.O., *Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not*. Geosci. Model Dev., 2022. **15**(14): p. 5481-5487.
210. Reich, N.G., et al., *Case study in evaluating time series prediction models using the relative mean absolute error*. Am Stat, 2016. **70**(3): p. 285-292.

211. Meadows, T., et al., *Epidemiological model can forecast COVID-19 outbreaks from wastewater-based surveillance in rural communities*. Water Res, 2025. **268**(Pt A): p. 122671.
212. Buckeridge, D.L., et al., *Predicting outbreak detection in public health surveillance: quantitative analysis to enable evidence-based method selection*. AMIA Annu Symp Proc, 2008. **2008**: p. 76-80.
213. Bosse, N.I., et al., *Scoring epidemiological forecasts on transformed scales*. PLoS Comput Biol, 2023. **19**(8): p. e1011393.
214. Cook, J.D., et al., *Bayesian forecasting of disease spread with little or no local data*. Sci Rep, 2023. **13**(1): p. 8137.
215. Gómez-Rubio, V., *Bayesian Inference with INLA*. 2020, Boca Raton, Florida: Chapman & Hall/CRC Press.
216. Aswi, A., et al., *Bayesian spatial and spatio-temporal approaches to modelling dengue fever: a systematic review*. Epidemiology & Infection, 2019. **147**: p. e33.
217. Li, X., F. Chadwick, and B. Swallow, *Advances in approximate Bayesian inference for models in epidemiology*. Epidemics, 2025. **53**: p. 100855.
218. Blangiardo, M., et al. *Spatial and spatio-temporal models with R-INLA*. Spatial and Spatio-temporal Epidemiology 2013 [cited 4; 33-49]. Available from: https://discovery.ucl.ac.uk/id/eprint/1415919/1/Baio_BlaCamBaiRue.pdf.
219. Moraga, P., *Spatial Statistics for Data Science: Theory and Practice with R*. Data Science Series. 2023: Chapman & Hall/CRC
220. Wang, Y., X. Chen, and F. Xue, *A Review of Bayesian Spatiotemporal Models in Spatial Epidemiology*. ISPRS International Journal of Geo-Information, 2024. **13**(3): p. 97.
221. Rue, H., S. Martino, and N. Chopin, *Approximate Bayesian Inference for Latent Gaussian models by using Integrated Nested Laplace Approximations*. Journal of the Royal Statistical Society Series B: Statistical Methodology, 2009. **71**(2): p. 319-392.
222. Rue, H., et al., *Bayesian Computing with INLA: A Review*. Annual Review of Statistics and Its Application, 2017. **4**: p. 395-421.
223. Lindgren, F. and H. Rue, *Bayesian Spatial Modelling with R-INLA*. Journal of Statistical Software, 2015. **63**(19): p. 1 - 25.
224. Rawat, S. and S. Deb, *A spatio-temporal statistical model to analyze COVID-19 spread in the USA*. J Appl Stat, 2023. **50**(11-12): p. 2310-2329.
225. van Zoest, V., et al., *Spatio-temporal predictions of COVID-19 test positivity in Uppsala County, Sweden: a comparative approach*. Sci Rep, 2022. **12**(1): p. 15176.
226. Colon-Gonzalez, F.J., et al., *Probabilistic seasonal dengue forecasting in Vietnam: A modelling study using superensembles*. PLoS Med, 2021. **18**(3): p. e1003542.
227. Lowe, R., et al., *Evaluating probabilistic dengue risk forecasts from a prototype early warning system for Brazil*. Elife, 2016. **5**.
228. Mills, C. and C.A. Donnelly, *Climate-based modelling and forecasting of dengue in three endemic departments of Peru*. PLoS Negl Trop Dis, 2024. **18**(12): p. e0012596.
229. Chou-Chen, S.W., et al., *Bayesian spatio-temporal model with INLA for dengue fever risk prediction in Costa Rica*. Environmental and Ecological Statistics, 2023. **30**(4): p. 687-713.

230. Redding, D.W., et al., *Geographical drivers and climate-linked dynamics of Lassa fever in Nigeria*. Nat Commun, 2021. **12**(1): p. 5759.
231. McCarter, M.S.J., et al., *The utility of a Bayesian predictive model to forecast neuroinvasive West Nile virus disease in the United States of America*, 2022. PLoS One, 2023. **18**(9): p. e0290873.
232. Lim, J.T., et al., *Revealing two dynamic dengue epidemic clusters in Thailand*. BMC Infect Dis, 2020. **20**(1): p. 927.
233. Thisyakorn, U., et al., *Epidemiology and costs of dengue in Thailand: A systematic literature review*. PLoS Negl Trop Dis, 2022. **16**(12): p. e0010966.
234. Phanitchat, T., et al., *Spatial and temporal patterns of dengue incidence in northeastern Thailand 2006–2016*. BMC Infectious Diseases, 2019. **19**(1): p. 743.
235. Bekoe, C., et al., *Modeling the Geographic Consequence and Pattern of Dengue Fever Transmission in Thailand*. J Res Health Sci, 2017. **17**(2): p. e00378.
236. The United Nations High Commissioner for Refugees, *RTG/MOI-UNHCR Verified Refugee Population (as of 31 December 2022)*. 2023.
237. Langkulsen, U., K. Promsakha Na Sakolnakhon, and N. James, *Climate change and dengue risk in central region of Thailand*. Int J Environ Health Res, 2020. **30**(3): p. 327-335.
238. Polwiang, S., *Estimation of dengue infection for travelers in Thailand*. Travel Med Infect Dis, 2016. **14**(4): p. 398-406.
239. Polwiang, S., *The time series seasonal patterns of dengue fever and associated weather variables in Bangkok (2003-2017)*. BMC Infect Dis, 2020. **20**(1): p. 208.
240. Lim, A., et al., *Spatial occurrence-intensity modeling of dengue incidence in southernmost provinces of Thailand*. PLoS Negl Trop Dis, 2025. **19**(7): p. e0013347.
241. Saita, S., S. Maeakhian, and T. Silawan, *Temporal Variations and Spatial Clusters of Dengue in Thailand: Longitudinal Study before and during the Coronavirus Disease (COVID-19) Pandemic*. Trop Med Infect Dis, 2022. **7**(8).
242. Limkittikul, K., J. Brett, and M. L'Azou, *Epidemiological trends of dengue disease in Thailand (2000-2011): a systematic literature review*. PLoS Negl Trop Dis, 2014. **8**(11): p. e3241.
243. Rotejanaprasert, C., et al., *Bayesian spatiotemporal modeling with sliding windows to correct reporting delays for real-time dengue surveillance in Thailand*. International Journal of Health Geographics, 2020. **19**(1): p. 4.
244. Thailand Ministry of Public Health: Department of Disease Control *Disease Distribution Dashboard - District Level* (การกระจายโรคอำเภอ). 2025.
245. Rahman, M.S., et al., *Ecological, Social, and Other Environmental Determinants of Dengue Vector Abundance in Urban and Rural Areas of Northeastern Thailand*. Int J Environ Res Public Health, 2021. **18**(11).
246. Koyadun, S., P. Butraporn, and P. Kittayapong, *Ecologic and sociodemographic risk determinants for dengue transmission in urban areas in Thailand*. Interdiscip Perspect Infect Dis, 2012. **2012**: p. 907494.
247. Polrob, W. and A. La-Up, *Nonlinear and lagged effects of climate variability on dengue incidence in an urban megacity: a distributed lag non-linear model (DLNM) based study in Bangkok, Thailand*. BMC Public Health, 2025. **25**(1): p. 4024.

248. Zafar, S., et al., *Landscape and climatic drivers of dengue fever in Lao People's Democratic Republic and Thailand: a retrospective analysis during 2002–2019*. *Landscape Ecology*, 2025. **40**(5): p. 102.
249. Chumpu, R., N. Khamsemanan, and C. Nattee, *The association between dengue incidences and provincial-level weather variables in Thailand from 2001 to 2014*. *PLoS One*, 2019. **14**(12): p. e0226945.
250. Promprou, S., M. Jaroensutasinee, and K. Jaroensutasinee. *Forecasting Dengue Haemorrhagic Fever Cases in Southern Thailand using ARIMA Models*. 2006.
251. Silawan, T., et al., *Temporal patterns and forecast of dengue infection in Northeastern Thailand*. *Southeast Asian J Trop Med Public Health*, 2008. **39**(1): p. 90-8.
252. Wongkoon, S., M. Jaroensutasinee, and K. Jaroensutasinee, *Weather factors influencing the occurrence of dengue fever in Nakhon Si Thammarat, Thailand*. *Trop Biomed*, 2013. **30**(4): p. 631-41.
253. Somboonsak, P., *Time Series Analysis of Dengue Fever Cases in Thailand Utilizing the SARIMA Model*, in *Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City*. 2020, Association for Computing Machinery: Shanghai, China. p. 439–444.
254. Bunprom, P., I. Siramaneerat, and P. Bhumkittipich, *Forecasting Dengue Fever Incidence in Thailand Using ARIMA: Implications for Public Health Planning: 10.32526/ennrj/24/20250069*. *Environment and Natural Resources Journal*, 2025: p. xx.
255. Phuthomdee, S., et al., *Dengue Forecasting Model Using SARIMA Model to Predict the Incidence of Dengue in Thailand*. *Current Applied Science and Technology*, 2018. **18**(2): p. 58-65.
256. Riley, P., et al., *SARIMA Forecasts of Dengue Incidence in Brazil, Mexico, Singapore, Sri Lanka, and Thailand: Model Performance and the Significance of Reporting Delays*. *medRxiv*, 2020.
257. Lauer, S.A., et al., *Prospective forecasts of annual dengue hemorrhagic fever incidence in Thailand, 2010-2014*. *Proc Natl Acad Sci U S A*, 2018. **115**(10): p. E2175-E2182.
258. Jain, R., et al., *Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data*. *BMC Infect Dis*, 2019. **19**(1): p. 272.
259. Colon-Gonzalez, F.J., et al., *The effects of weather and climate change on dengue*. *PLoS Negl Trop Dis*, 2013. **7**(11): p. e2503.
260. Ryan, S.J., et al., *Global expansion and redistribution of Aedes-borne virus transmission risk with climate change*. *PLoS Negl Trop Dis*, 2019. **13**(3): p. e0007213.
261. Sylvestre, E., et al., *Data-driven methods for dengue prediction and surveillance using real-world and Big Data: A systematic review*. *PLoS Negl Trop Dis*, 2022. **16**(1): p. e0010056.
262. de Lima, C.L., et al., *Temporal and Spatiotemporal Arboviruses Forecasting by Machine Learning: A Systematic Review*. *Frontiers in Public Health*, 2022. **10**.
263. World Health Organization. *WHO regional offices*. 2026; Available from: <https://www.who.int/about/who-we-are/regional-offices>.

264. Aguiar, M., et al., *How much complexity is needed to describe the fluctuations observed in dengue hemorrhagic fever incidence data?* Ecological Complexity, 2013. **16**: p. 31-40.
265. Tilburg Science Hub. *The Random Effects Model*. 2024; Available from: <https://www.tilburgsciencehub.com/topics/analyze/causal-inference/panel-data/random-effects/>.
266. World Health Organization, *Dengue: global situation, surveillance and progress – 2024 update*. Weekly Epidemiological Record, 2025. **100**(52): p. 665-678.
267. Chowell, G. and P. Skums, *Investigating and forecasting infectious disease dynamics using epidemiological and molecular surveillance data*. Phys Life Rev, 2024. **51**: p. 294-327.
268. Johansson, M.A., et al., *Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico*. Sci Rep, 2016. **6**: p. 33707.
269. Runge-Ranzinger, S., et al., *Dengue Contingency Planning: From Research to Policy and Practice*. PLoS Negl Trop Dis, 2016. **10**(9): p. e0004916.
270. Lowe, R., et al., *Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil*. Computers & Geosciences, 2011. **37**(3): p. 371-381.
271. Bosse, N., et al., *scoringutils: Utilities for Scoring and Assessing Predictions*. 2025.
272. Huang, A.T., et al., *Assessing the role of multiple mechanisms increasing the age of dengue cases in Thailand*. Proc Natl Acad Sci U S A, 2022. **119**(20): p. e2115790119.
273. Cummings, D.A., et al., *The impact of the demographic transition on dengue in Thailand: insights from a statistical analysis and mathematical modeling*. PLoS Med, 2009. **6**(9): p. e1000139.
274. Garcia-Carreras, B., et al., *Periodic synchronisation of dengue epidemics in Thailand over the last 5 decades driven by temperature and immunity*. PLoS Biol, 2022. **20**(3): p. e3001160.
275. Jeefoo, P., N.K. Tripathi, and M. Souris, *Spatio-temporal diffusion pattern and hotspot detection of dengue in Chachoengsao province, Thailand*. Int J Environ Res Public Health, 2011. **8**(1): p. 51-74.
276. Ibrahim Abdulsalam, F., et al., *Association between climate variables and dengue incidence in Nakhon Si Thammarat Province, Thailand*. Geospat Health, 2021. **16**(2).
277. Perez, L.J., et al., *Climate, inter-serotype competition and arboviral interactions shape dengue dynamics in Thailand*. Commun Biol, 2025. **8**(1): p. 601.
278. Grenfell, B.T., O.N. Bjornstad, and J. Kappey, *Travelling waves and spatial hierarchies in measles epidemics*. Nature, 2001. **414**(6865): p. 716-23.
279. Cummings, D.A., et al., *Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand*. Nature, 2004. **427**(6972): p. 344-7.
280. Ko, H.Y., et al., *Emergence and increased epidemic potential of dengue variants with the NS5(V357E) mutation after consecutive years of transmission*. iScience, 2024. **27**(11): p. 110899.

281. Hernandez-Garcia, E., et al., *Epidemiological implications of the genetic diversification of dengue virus (DENV) serotypes and genotypes in Mexico*. Infect Genet Evol, 2020. **84**: p. 104391.
282. Bennett, S.N., et al., *Epidemic dynamics revealed in dengue evolution*. Mol Biol Evol, 2010. **27**(4): p. 811-8.
283. Fontaine, A., et al., *Epidemiological significance of dengue virus genetic variation in mosquito infection dynamics*. PLoS Pathog, 2018. **14**(7): p. e1007187.
284. Aguiar, M., et al., *Are we modelling the correct dataset? Minimizing false predictions for dengue fever in Thailand*. Epidemiology and Infection, 2014. **142**(11): p. 2447-2459.
285. Division of Vector-Borne Diseases: Department of Disease Control, *Guidelines for Surveillance, Prevention, and Control of Mosquito-Borne Diseases for Public Health Officials*. 2021, Ministry of Public Health: Bangkok.
286. Hadinegoro, S.R., *The revised WHO dengue case classification: does the system need to be modified?* Paediatr Int Child Health, 2012. **32 Suppl 1**(s1): p. 33-8.
287. World Health Organization. Regional Office for South-East, A., *Comprehensive Guideline for Prevention and Control of Dengue and Dengue Haemorrhagic Fever. Revised and expanded edition*. 2011, New Delhi: WHO Regional Office for South-East Asia.
288. Bureau of Epidemiology, D.o.D.C., MoPH, Thailand. *DHF_Total: Annual Situation Summary*. 2024; Available from: <http://doe.moph.go.th/surdata/disease.php?dcontent=old&ds=262766>.
289. Department of Provincial Administration, *Official Statistics Registration Systems: Population Statistics*. Department of Provincial Administration, Ministry of Interior, Thailand.
290. Department of Medical Sciences, Ministry of Public Health, *Laboratory testing guidelines for dengue fever (Thailand)*. 2023.
291. Roesch, A.S., Harald *WaveletComp*. 2018.
292. Gauthier, G., et al., *Taking the beat of the Arctic: are lemming population cycles changing due to winter climate?* Proc Biol Sci, 2024. **291**(2016): p. 20232361.
293. StatsDirect, *LOESS Curve Fitting (Local Polynomial Regression)*. 2025.
294. Nagao, Y. and K. Koelle, *Decreases in dengue transmission may act to increase the incidence of dengue hemorrhagic fever*. Proc Natl Acad Sci U S A, 2008. **105**(6): p. 2238-43.
295. Giraud, T., *osrm: Interface Between R and the OpenStreetMap-Based Routing Service OSRM*. 2024, Comprehensive R Archive Network (CRAN).
296. GenBank. 1982; Available from: <https://www.ncbi.nlm.nih.gov/nucleotide/>.
297. Gifford, R.J. *Dengue-GLUE: Phylogenomic Analysis of Dengue Virus*. Available from: <https://github.com/giffordlabcvr/Dengue-GLUE>.
298. Aksamentov, I.R., Cornelius; Hodcroft, Emma B.; Neher, Richard A. , *Nextclade: clade assignment, mutation calling and quality control for viral genomes*. The Journal of Open Source Software, 2021.
299. Inward, R., *dengue_pipeline*. 2025.
300. Paradis, E.J., T., Kamvar, Z.N., Knaus, B., Schliep, K., Potts, A., Winter, D. *pegas: Population and Evolutionary Genetics Analysis System*. 2023; Available from: <https://cran.r-project.org/web/packages/pegas/index.html>.

301. Naimi, B. *usdm: Uncertainty Analysis for Species Distribution Models*. 2023; Available from: <https://cran.r-project.org/web/packages/usdm/index.html>.
302. Koo, C., et al., *Highly Selective Transmission Success of Dengue Virus Type 1 Lineages in a Dynamic Virus Population: An Evolutionary and Fitness Perspective*. iScience, 2018. **6**: p. 38-51.
303. Kennedy, D.A. and G. Dwyer, *Effects of multiple sources of genetic drift on pathogen variation within hosts*. PLoS Biol, 2018. **16**(3): p. e2004444.
304. Stockdale, J.E., P. Liu, and C. Colijn, *The potential of genomics for infectious disease forecasting*. Nat Microbiol, 2022. **7**(11): p. 1736-1743.
305. Hadi, Z.A., N.C. Dom, and R. Ismail, *A seasonal autoregressive integrated moving average model to predict incidence of dengue cases in Kuantan, Malaysia*. Science, Engineering and Health Studies, 2021. **15**.
306. de Lima, M.V.M. and G.Z. Laporta, *Evaluation of the Models for Forecasting Dengue in Brazil from 2000 to 2017: An Ecological Time-Series Study*. Insects, 2020. **11**(11).
307. Runge-Ranzinger, S., et al., *Dengue disease surveillance: an updated systematic literature review*. Trop Med Int Health, 2014. **19**(9): p. 1116-60.
308. Benitez-Valladares, D., et al., *Validation of the Early Warning and Response System (EWARS) for dengue outbreaks: Evidence from the national vector control program in Mexico*. PLoS Negl Trop Dis, 2021. **15**(12): p. e0009261.
309. Thayer, M.B., et al., *Dengue epidemic alert thresholds for surveillance and decision-making in Puerto Rico: development and prospective application of an early warning system using routine surveillance data*. BMJ Open, 2025. **15**(9): p. e106182.
310. Duarte, F.H.O., et al., *Leveraging graph neural networks and mobility data for COVID-19 forecasting*. 2025.
311. Ma, Z. and L. Rennert, *An epidemiological modeling framework to inform institutional-level response to infectious disease outbreaks: a Covid-19 case study*. Sci Rep, 2024. **14**(1): p. 7221.
312. Greening, B., Jr. and M.I. Meltzer, *Mathematical Modeling for Emergency Response: Using Models to Inform and Direct Response Priorities and Shape the Research Agenda*, in *Principles and Practice of Emergency Research Response*, R.A. Sorenson, E.S. Higgs, and M.P. Fallah, Editors. 2024, Springer: Cham (CH).
313. Gneiting, T. and A.E. Raftery, *Strictly proper scoring rules, prediction, and estimation*. Journal of the American Statistical Association, 2007. **102**(477): p. 359-378.
314. Meakin, S., et al., *Comparative assessment of methods for short-term forecasts of COVID-19 hospital admissions in England at the local level*. BMC Med, 2022. **20**(1): p. 86.
315. *R-INLA*. Available from: www.r-inla.org.
316. Simpson, D., et al., *Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors*. Statistical Science, 2017. **32**(1): p. 1-28.
317. INLA Team, *Model for seasonal variation*, in *INLA Documentation*.
318. Yang, W., et al., *Modified Brier score for evaluating prediction accuracy for binary outcomes*. Stat Methods Med Res, 2022. **31**(12): p. 2287-2296.

319. Tuan, D.A., *Correction: Tuan, D.A. Leveraging Climate Data for Dengue Forecasting in Ba Ria Vung Tau Province, Vietnam: An Advanced Machine Learning Approach. Trop. Med. Infect. Dis.* 2024, 9, 250. *Trop Med Infect Dis*, 2025. **10**(5).
320. Tuan, D.A., *Leveraging Climate Data for Dengue Forecasting in Ba Ria Vung Tau Province, Vietnam: An Advanced Machine Learning Approach. Trop Med Infect Dis*, 2024. **9**(10).
321. Hapuarachchi, H.C., et al., *Intra-epidemic evolutionary dynamics of a Dengue virus type 1 population reveal mutant spectra that correlate with disease transmission. Sci Rep*, 2016. **6**: p. 22592.
322. Copernicus Climate Change Service. *ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S)* 2023.
323. *Thailand administrative level 0-3 shapefiles*. 2022.
324. Singer, M., et al., *Hourly potential evapotranspiration (hPET) at 0.1degs grid resolution for the global land surface from 1981-present*. 2020.
325. Beguería, S. and S.M. Vicente-Serrano, *SPEI: Calculation of the Standardized Precipitation-Evapotranspiration Index*. 2023.
326. AppEEARS Team, *Application for Extracting and Exploring Analysis Ready Samples (AppEEARS)*, N.E.L.P.D.A.A.C.L. DAAC), Editor. 2025, USGS/Earth Resources Observation and Science (EROS) Center: Sioux Falls, South Dakota, USA.
327. Didan, K., *MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V061*. 2021, NASA EOSDIS Land Processes Distributed Active Archive Center.
328. Zeileis, A.G., G; Jeffrey, A. Ryan; Ulrich, Joshua M; Andrews, Felix, *zoo: S3 Infrastructure for Regular and Irregular Time Series (Z's Ordered Observations)*. 2025.
329. Data for Good. *Activity Space Maps*. 2025; Available from: <https://dataforgood.facebook.com/dfg/tools/activity-space-maps>.
330. Kakarla, S.G., et al., *Lag effect of climatic variables on dengue burden in India. Epidemiol Infect*, 2019. **147**: p. e170.
331. Chuang, T.W., L.F. Chaves, and P.J. Chen, *Effects of local and regional climatic fluctuations on dengue outbreaks in southern Taiwan. PLoS One*, 2017. **12**(6): p. e0178698.
332. Plos One Staff, *Correction: Effects of local and regional climatic fluctuations on dengue outbreaks in southern Taiwan. PLoS One*, 2017. **12**(7): p. e0181638.
333. Gasparini, A., B. Amrstrong, and F. Scheipl, *dlnm: Distributed Lag Non-Linear Models*. 2025.
334. Wood, S., *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*. 2025.
335. R Core Team, *R: A language and environment for statistical computing*, R.F.f.S. Computing, Editor. 2025: Vienna, Austria.
336. Kirk, D., et al., *Temperature impacts on dengue incidence are nonlinear and mediated by climatic and socioeconomic factors: A meta-analysis. PLOS Climate*, 2024. **3**(3): p. e0000152.
337. Childs, M.L., et al., *Climate warming is expanding dengue burden in the Americas and Asia. Proc Natl Acad Sci U S A*, 2025. **122**(37): p. e2512350122.

338. Gharbi, M., et al., *Time series analysis of dengue incidence in Guadeloupe, French West Indies: forecasting models using climate variables as predictors*. BMC Infect Dis, 2011. **11**: p. 166.
339. Ferdousi, F., et al., *Identification of Essential Containers for Aedes Larval Breeding to Control Dengue in Dhaka, Bangladesh*. Trop Med Health, 2015. **43**(4): p. 253-64.
340. Benedum, C.M., et al., *Statistical modeling of the effect of rainfall flushing on dengue transmission in Singapore*. PLoS Negl Trop Dis, 2018. **12**(12): p. e0006935.
341. Seidahmed, O.M. and E.A. Eltahir, *A Sequence of Flushing and Drying of Breeding Habitats of Aedes aegypti (L.) Prior to the Low Dengue Season in Singapore*. PLoS Negl Trop Dis, 2016. **10**(7): p. e0004842.
342. Lu, L., et al., *Time series analysis of dengue fever and weather in Guangzhou, China*. BMC Public Health, 2009. **9**: p. 395.
343. Sutriyawan, A., et al., *Analysis of temperature, humidity, rainfall, and wind velocity on dengue hemorrhagic fever in Bandung municipality*. Russian Journal of Infection and Immunity, 2024. **14**(1): p. 155-162.
344. Cheng, Y., et al., *Integrating meteorological data and hybrid intelligent models for dengue fever prediction*. BMC Public Health, 2025. **25**(1): p. 1516.
345. Huang, C.C., et al., *Spatial Clustering of Dengue Fever Incidence and Its Association with Surrounding Greenness*. Int J Environ Res Public Health, 2018. **15**(9).
346. Cunha, M., et al., *Disentangling associations between vegetation greenness and dengue in a Latin American city: Findings and challenges*. Landsc Urban Plan, 2021. **216**: p. None.
347. Araujo, R.V., et al., *Sao Paulo urban heat islands have a higher incidence of dengue than other urban areas*. Braz J Infect Dis, 2015. **19**(2): p. 146-55.
348. Pham, N.T., C.T. Nguyen, and M.R.B. Pineda-Cortel, *Time-series modelling of dengue incidence in the Mekong Delta region of Viet Nam using remote sensing data*. Western Pac Surveill Response J, 2020. **11**(1): p. 13-21.
349. Sarma, D.K., et al., *An assessment of remotely sensed environmental variables on Dengue epidemiology in Central India*. PLoS Negl Trop Dis, 2022. **16**(10): p. e0010859.
350. Buczak, A.L., et al., *A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data*. BMC Med Inform Decis Mak, 2012. **12**: p. 124.
351. Chen, X. and P. Moraga, *Dengue forecasting and outbreak detection in Brazil using LSTM: integrating human mobility and climate factors*. Infect Dis Model, 2026. **11**(1): p. 338-354.
352. Greenland, S. and J.M. Robins, *Identifiability, exchangeability, and epidemiological confounding*. Int J Epidemiol, 1986. **15**(3): p. 413-9.
353. Jager, K.J., et al., *Confounding: what it is and how to deal with it*. Kidney Int, 2008. **73**(3): p. 256-60.
354. Harrell, F.E., Jr., K.L. Lee, and D.B. Mark, *Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors*. Stat Med, 1996. **15**(4): p. 361-87.

355. VanderWeele, T.J. and M.J. Knol, *A Tutorial on Interaction*. Epidemiologic Methods, 2014. **3**(1): p. 33-72.
356. Rothman, K.J., S. Greenland, and A.M. Walker, *Concepts of interaction*. Am J Epidemiol, 1980. **112**(4): p. 467-70.
357. Wong, M.C., et al., *A climate model for predicting the abundance of Aedes mosquitoes in Hong Kong*. Meteorological Applications, 2011. **18**: p. 105-110.
358. Kamiya, T., et al., *Temperature-dependent variation in the extrinsic incubation period elevates the risk of vector-borne disease emergence*. Epidemics, 2019. **30**: p. 100382.
359. Rahman, M.S., et al., *Mapping the spatial distribution of the dengue vector Aedes aegypti and predicting its abundance in northeastern Thailand using machine-learning approach*. One Health, 2021. **13**: p. 100358.
360. Rose, N.H., et al., *Climate and Urbanization Drive Mosquito Preference for Humans*. Curr Biol, 2020. **30**(18): p. 3570-3579 e6.
361. Wilke, A.B.B., et al., *Proliferation of Aedes aegypti in urban environments mediated by the availability of key aquatic habitats*. Sci Rep, 2020. **10**(1): p. 12925.
362. Martin, M.E., et al., *Satellite observation to assess dengue risk due to Aedes aegypti and Aedes albopictus in a subtropical city of Argentina*. Med Vet Entomol, 2023. **37**(1): p. 27-36.
363. Lorenz, C., et al., *Predicting Aedes aegypti infestation using landscape and thermal features*. Sci Rep, 2020. **10**(1): p. 21688.
364. Salje, H., et al., *Dengue diversity across spatial and temporal scales: Local structure and the effect of host population size*. Science, 2017. **355**(6331): p. 1302-1306.
365. Susilawaty, A., et al., *Climate factors and dengue fever occurrence in Makassar during period of 2011-2017*. Gac Sanit, 2021. **35 Suppl 2**: p. S408-S412.
366. Cheong, Y.L., et al., *Assessing weather effects on dengue disease in Malaysia*. Int J Environ Res Public Health, 2013. **10**(12): p. 6319-34.
367. Depradine, C. and E. Lovell, *Climatological variables and the incidence of Dengue fever in Barbados*. International Journal of Environmental Health Research, 2004. **14**(6): p. 429-441.
368. Marcantonio, M., T. Reyes, and C.M. Barker, *Quantifying Aedes aegypti dispersal in space and time: a modeling approach*. Ecosphere, 2019. **10**(12): p. e02977.
369. Muir, L.E. and B.H. Kay, *Aedes aegypti survival and dispersal estimated by mark-release-recapture in northern Australia*. Am J Trop Med Hyg, 1998. **58**(3): p. 277-82.
370. Doeurk, B., et al., *Impact of temperature on survival, development and longevity of Aedes aegypti and Aedes albopictus (Diptera: Culicidae) in Phnom Penh, Cambodia*. Parasit Vectors, 2025. **18**(1): p. 362.
371. Chen, Y., et al., *Measuring the effects of COVID-19-related disruption on dengue transmission in southeast Asia and Latin America: a statistical modelling study*. Lancet Infect Dis, 2022. **22**(5): p. 657-667.
372. Araujo, J.L.B., et al., *The impact of COVID-19 mobility restrictions on dengue transmission in urban areas*. PLoS Negl Trop Dis, 2024. **18**(11): p. e0012644.

373. Liyanage, P., J. Rocklov, and H.A. Tissera, *The impact of COVID-19 lockdown on dengue transmission in Sri Lanka; A natural experiment for understanding the influence of human mobility*. PLoS Negl Trop Dis, 2021. **15**(6): p. e0009420.
374. Cavany, S.M., et al., *Pandemic-associated mobility restrictions could cause increases in dengue virus transmission*. PLoS Negl Trop Dis, 2021. **15**(8): p. e0009603.
375. Horstick, O., et al., *Dengue vector-control services: how do they work? A systematic literature review and country case studies*. Trans R Soc Trop Med Hyg, 2010. **104**(6): p. 379-86.
376. Tsuda, H. and T. Tebakari, *A macroscopic analysis of the demographic impacts of flood inundation in Thailand (2005–2019)*. Progress in Earth and Planetary Science, 2023. **10**(1): p. 36.
377. Ten Bosch, Q.A., et al., *The Role of Serotype Interactions and Seasonality in Dengue Model Selection and Control: Insights from a Pattern Matching Approach*. PLoS Negl Trop Dis, 2016. **10**(5): p. e0004680.
378. MacCormack-Gelles, B., et al., *Epidemiological characteristics and determinants of dengue transmission during epidemic and non-epidemic years in Fortaleza, Brazil: 2011-2015*. PLoS Negl Trop Dis, 2018. **12**(12): p. e0006990.
379. Naish, S., et al., *Climate change and dengue: a critical and systematic review of quantitative modelling approaches*. BMC Infect Dis, 2014. **14**: p. 167.
380. Khongwichit, S., et al., *Molecular surveillance of arboviruses circulation and co-infection during a large chikungunya virus outbreak in Thailand, October 2018 to February 2020*. Sci Rep, 2022. **12**(1): p. 22323.
381. Vogels, C.B.F., et al., *Arbovirus coinfection and co-transmission: A neglected public health concern?* PLoS Biol, 2019. **17**(1): p. e3000130.
382. Sanchez Tejada, G., et al., *Early warning and response system for dengue outbreaks: Moving from research to operational implementation in Mexico*. PLOS Glob Public Health, 2023. **3**(9): p. e0001691.
383. Al Mobin, M., *Forecasting dengue in Bangladesh using meteorological variables with a novel feature selection approach*. Sci Rep, 2024. **14**(1): p. 32073.
384. Snyder, J. and G. Maglasang, *Dengue in Cebu City, Philippines: A Pilot Study of Predictive Models and Visualizations for Public Health*. Am J Trop Med Hyg, 2024. **110**(1): p. 179-187.
385. Morbey, R.A., et al., *Machine learning forecasts for seasonal epidemic peaks: Lessons learnt from an atypical respiratory syncytial virus season*. PLoS One, 2023. **18**(9): p. e0291932.
386. Chen, Y., et al., *The utility of LASSO-based models for real time forecasts of endemic infectious diseases: A cross country comparison*. J Biomed Inform, 2018. **81**: p. 16-30.
387. Chanprasopchai, P., P. Pongsumpun, and I.M. Tang, *Effect of Rainfall for the Dynamical Transmission Model of the Dengue Disease in Thailand*. Comput Math Methods Med, 2017. **2017**: p. 2541862.
388. Kliengchuay, W., et al., *Variation of vegetation cover and the relationship with land surface temperature across Thailand (2007 to 2022)*. Sci Rep, 2025. **15**(1): p. 27823.

389. Chaiwino, W., et al., *Analysis of long-term rainfall trend and extreme in upper northern Thailand*. Sci Rep, 2025. **15**(1): p. 33380.
390. Leenawarat, D., et al., *Influences of tropical monsoon and El Niño Southern Oscillations on surface chlorophyll-a variability in the Gulf of Thailand*. Frontiers in Climate, 2022. **Volume 4 - 2022**.
391. Wearing, H.J. and P. Rohani, *Ecological and immunological determinants of dengue epidemics*. Proc Natl Acad Sci U S A, 2006. **103**(31): p. 11802-7.
392. Reich, N.G., et al., *Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity*. J R Soc Interface, 2013. **10**(86): p. 20130414.
393. Namwong, T., et al., *Development of Surveillance System and 3-3-1 Measure Monitoring for Dengue Control Using Application on Google Drive, Yasothon Province, Thailand*. Journal of Health Science of Thailand, 2019. **28**(3): p. 402-410.
394. Rakprasit, J., et al., *Healthcare use for communicable diseases among migrant workers in comparison with Thai workers*. Ind Health, 2017. **55**(1): p. 67-75.
395. English, K. *Thailand's Population Drops Below 66 Million as Births Hit 75-Year Low*. Khaosod English, 2025.
396. Barber, S. and R. Sciortino, *Thailand Migration Report 2024*. 2024, UN Network on Migration in Thailand: Bangkok, Thailand.
397. National Statistical Office of Thailand, *Thailand Population and Housing Census 2010*. 2010, National Statistical Office, Ministry of Information and Communication Technology: Bangkok, Thailand.
398. da Costa, V.G., A.C. Marques-Silva, and M.L. Moreli, *A meta-analysis of the diagnostic accuracy of two commercial NS1 antigen ELISA tests for early dengue virus detection*. PLoS One, 2014. **9**(4): p. e94655.
399. Shan, X., et al., *Evaluation of the diagnostic accuracy of nonstructural protein 1 Ag-based tests for dengue virus in Asian population: a meta-analysis*. BMC Infect Dis, 2015. **15**: p. 360.
400. Beesley, L.J., D. Osthus, and S.Y. Del Valle, *Addressing delayed case reporting in infectious disease forecast modeling*. PLoS Comput Biol, 2022. **18**(6): p. e1010115.
401. Rotejanaprasert, C., D. Areechokchai, and R.J. Maude, *Two-step spatiotemporal anomaly detection corrected for lag reporting time with application to real-time dengue surveillance in Thailand*. BMC Med Res Methodol, 2024. **24**(1): p. 10.
402. Liu, B., M.F. Hossain, and S. Hossain, *A comparative evaluation of multiple machine learning approaches for forecasting dengue outbreaks in Bangladesh*. Sci Rep, 2025. **15**(1): p. 35931.
403. Hewage, P., et al., *Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station*. Soft Computing, 2020. **24**(21): p. 16453-16482.
404. Ribeiro, L.M., E.J.d.S. Luz, and J.C. Gertrudes. *How Effective Are Time Series Foundation Models for Epidemiological Data Analysis?* in *2025 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. 2025.
405. Draidi Areed, W., et al., *A district-level ensemble model to enhance dengue prediction and control for the Mekong Delta Region of Vietnam*. PLoS Negl Trop Dis, 2025. **19**(9): p. e0013571.

406. Cheng, J., et al., *Extreme weather conditions and dengue outbreak in Guangdong, China: Spatial heterogeneity based on climate variability*. Environ Res, 2021. **196**: p. 110900.
407. Siriyasatien, P., et al., *Analysis of significant factors for dengue fever incidence prediction*. BMC Bioinformatics, 2016. **17**: p. 166.
408. Kesorn, K., et al., *Morbidity Rate Prediction of Dengue Hemorrhagic Fever (DHF) Using the Support Vector Machine and the Aedes aegypti Infection Rate in Similar Climates and Geographical Areas*. PLoS One, 2015. **10**(5): p. e0125049.
409. Fustec, B., et al., *Complex relationships between Aedes vectors, socio-economics and dengue transmission-Lessons learned from a case-control study in northeastern Thailand*. PLoS Negl Trop Dis, 2020. **14**(10): p. e0008703.
410. Rotejanaprasert, C., et al., *Perspectives and challenges in developing and implementing integrated dengue surveillance tools and technology in Thailand: a qualitative study*. PLoS Negl Trop Dis, 2024. **18**(8): p. e0012387.
411. Overgaard, H.J., et al., *Correction to: Assessing dengue transmission risk and a vector control intervention using entomological and immunological indices in Thailand: study protocol for a cluster-randomized controlled trial*. Trials, 2018. **19**(1): p. 703.
412. Overgaard, H.J., et al., *Assessing dengue transmission risk and a vector control intervention using entomological and immunological indices in Thailand: study protocol for a cluster-randomized controlled trial*. Trials, 2018. **19**(1): p. 122.
413. Kucharski, A.J., et al., *Using paired serology and surveillance data to quantify dengue transmission and control during a large outbreak in Fiji*. Elife, 2018. **7**.
414. Jing, Q.L., et al., *Imported cases and minimum temperature drive dengue transmission in Guangzhou, China: evidence from ARIMAX model*. Epidemiol Infect, 2018. **146**(10): p. 1226-1235.
415. Peng, H.J., et al., *A local outbreak of dengue caused by an imported case in Dongguan China*. BMC Public Health, 2012. **12**: p. 83.
416. Chungue, E., J.P. Boutin, and J. Roux, *Dengue surveillance in French Polynesia: an attempt to use the excess number of laboratory requests for confirmation of dengue diagnosis as an indicator of dengue activity*. Eur J Epidemiol, 1991. **7**(6): p. 616-20.
417. Peterson, A.T., et al., *Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases*. Trans R Soc Trop Med Hyg, 2005. **99**(9): p. 647-55.
418. Wu, P.C., et al., *Weather as an effective predictor for occurrence of dengue fever in Taiwan*. Acta Trop, 2007. **103**(1): p. 50-7.
419. Halide, H. and P. Ridd, *A predictive model for Dengue Hemorrhagic Fever epidemics*. Int J Environ Health Res, 2008. **18**(4): p. 253-65.
420. Tipayamongkhogul, M., et al., *Effects of the El Nino-southern oscillation on dengue epidemics in Thailand, 1996-2005*. BMC Public Health, 2009. **9**: p. 422.
421. Wongkoon, S., et al., *Predicting DHF Incidence in Northern Thailand using Time Series Analysis Technique*. International Journal of Biological and Medical Sciences, 2009. **4**(3): p. 117-121.

422. Aburas, H.M., B.G. Cetiner, and M. Sari, *Dengue confirmed-cases prediction: A neural network model*. Expert Systems with Applications, 2010. **37**(6): p. 4256-4260.
423. Hu, W., et al., *Dengue fever and El Nino/Southern Oscillation in Queensland, Australia: a time series predictive model*. Occup Environ Med, 2010. **67**(5): p. 307-11.
424. Althouse, B.M., Y.Y. Ng, and D.A. Cummings, *Prediction of dengue incidence using search query surveillance*. PLoS Negl Trop Dis, 2011. **5**(8): p. e1258.
425. Chan, E.H., et al., *Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance*. PLoS Negl Trop Dis, 2011. **5**(5): p. e1206.
426. Gerardi, D.O. and L.H.A. Monteiro, *System Identification and Prediction of Dengue Fever Incidence in Rio de Janeiro*. Mathematical Problems in Engineering, 2011. **2011**(1): p. 720304.
427. Martinez, E.Z. and E.A. Silva, *Predicting the number of cases of dengue infection in Ribeirao Preto, Sao Paulo State, Brazil, using a SARIMA model*. Cad Saude Publica, 2011. **27**(9): p. 1809-18.
428. Yu, H.-L., et al., *A spatio-temporal climate-based model of early dengue fever warning in southern Taiwan*. Stochastic Environmental Research and Risk Assessment, 2011. **25**(4): p. 485-494.
429. Capistran, M.A., J.A. Christen, and J.X. Velasco-Hernandez, *Towards uncertainty quantification and inference in the stochastic SIR epidemic model*. Math Biosci, 2012. **240**(2): p. 250-9.
430. Karim, M.N., et al., *Climatic factors influencing dengue cases in Dhaka city: a model for dengue prediction*. Indian J Med Res, 2012. **136**(1): p. 32-9.
431. Lowe, R., et al., *The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil*. Stat Med, 2012. **32**(5): p. 864-83.
432. Wongkoon, S., M. Jaroensutasinee, and K. Jaroensutasinee, *Assessing the temporal modelling for prediction of dengue infection in northern and north-eastern, Thailand*. Trop Biomed, 2012. **29**(3): p. 339-48.
433. Wongkoon, S., M. Jaroensutasinee, and K. Jaroensutasinee, *Development of temporal modeling for prediction of dengue infection in northeastern Thailand*. Asian Pac J Trop Med, 2012. **5**(3): p. 249-52.
434. Dom, N.C., et al., *Generating temporal model using climate variables for the prediction of dengue cases in Subang Jaya, Malaysia*. Asian Pacific Journal of Tropical Disease, 2013. **3**(5): p. 352-361.
435. Sitepu, M.S., et al., *Temporal patterns and a disease forecasting model of dengue hemorrhagic fever in Jakarta based on 10 years of surveillance data*. Southeast Asian J Trop Med Public Health, 2013. **44**(2): p. 206-17.
436. Buczak, A.L., et al., *Prediction of high incidence of dengue in the Philippines*. PLoS Negl Trop Dis, 2014. **8**(4): p. e2771.
437. Eastin, M.D., et al., *Intra- and interseasonal autoregressive prediction of dengue outbreaks using local weather and regional climate for a tropical environment in Colombia*. Am J Trop Med Hyg, 2014. **91**(3): p. 598-610.

438. Mohsin, M.F.M., A. Abu Bakar, and A.R. Hamdan, *Outbreak detection model based on danger theory*. Appl Soft Comput, 2014. **24**: p. 612-622.
439. Ongluk, P., P. Siriyasatien, and K. Kesorn, *New Key Factors Discovery to Enhance Dengue Fever Forecasting Model*. Advanced Materials Research, 2014. **931-932**: p. 1457-1461.
440. Sang, S., et al., *Predicting local dengue transmission in Guangzhou, China, through the influence of imported cases, mosquito density and climate variability*. PLoS One, 2014. **9**(7): p. e102755.
441. Torres, C., et al., *Fuzzy model identification of dengue epidemic in Colombia based on multiresolution analysis*. Artif Intell Med, 2014. **60**(1): p. 41-51.
442. Yu, H.L., et al., *An online spatiotemporal prediction model for dengue fever epidemic in Kaohsiung (Taiwan)*. Biom J, 2014. **56**(3): p. 428-40.
443. Chan, T.C., T.H. Hu, and J.S. Hwang, *Daily forecast of dengue fever incidents for urban villages in a city*. Int J Health Geogr, 2015. **14**: p. 9.
444. Pepin, K.M., et al., *Utility of mosquito surveillance data for spatial prioritization of vector control against dengue viruses in three Brazilian cities*. Parasit Vectors, 2015. **8**: p. 98.
445. Phung, D., et al., *Identification of the prediction model for dengue incidence in Can Tho city, a Mekong Delta area in Vietnam*. Acta Trop, 2015. **141**(Pt A): p. 88-96.
446. Tung, C.Y., T.C. Chou, and J.W. Lin, *Using prediction markets of market scoring rule to forecast infectious diseases: a case study in Taiwan*. BMC Public Health, 2015. **15**: p. 766.
447. Ali, M.A., et al., *ID-Viewer: a visual analytics architecture for infectious diseases surveillance and response management in Pakistan*. Public Health, 2016. **134**: p. 72-85.
448. Li, E.Y., C.Y. Tung, and S.H. Chang, *The wisdom of crowds in action: Forecasting epidemic diseases with a web-based prediction market system*. Int J Med Inform, 2016. **92**: p. 35-43.
449. Mekpariyup, J. and K. Saithanu, *A new approach to detect epidemic of DHF by combining ARIMA model and adjusted Tukey's control chart with interpretation rules*. Interv Med Appl Sci, 2016. **8**(3): p. 118-120.
450. Ramadona, A.L., et al., *Prediction of Dengue Outbreaks Based on Disease Surveillance and Meteorological Data*. PLoS One, 2016. **11**(3): p. e0152688.
451. Shi, Y., et al., *Three-Month Real-Time Dengue Forecast Models: An Early Warning System for Outbreak Alerts and Policy Decision Support in Singapore*. Environ Health Perspect, 2016. **124**(9): p. 1369-75.
452. Wongkoon, S., M. Jaroensutasinee, and K. Jaroensutasinee, *Spatio-temporal climate-based model of dengue infection in Southern, Thailand*. Trop Biomed, 2016. **33**(1): p. 55-70.
453. Yamana, T.K., S. Kandula, and J. Shaman, *Superensemble forecasts of dengue outbreaks*. J R Soc Interface, 2016. **13**(123).
454. Yu, H.-L., C.-H. Lee, and L.-C. Chien, *A spatiotemporal dengue fever early warning model accounting for nonlinear associations with hydrological factors: a Bayesian maximum entropy approach*. Stochastic Environmental Research and Risk Assessment, 2016. **30**(8): p. 2127-2141.

455. Zhang, Y., et al., *Developing a Time Series Predictive Model for Dengue in Zhongshan, China Based on Weather and Guangzhou Dengue Surveillance Data*. PLoS Negl Trop Dis, 2016. **10**(2): p. e0004473.
456. de Almeida Marques-Toledo, C.A., et al., *Dengue prediction by the web: Tweets are a useful tool for estimating and forecasting Dengue at country and city level*. PLoS Negl Trop Dis, 2017. **11**(7): p. e0005729.
457. Dom, N.C., et al., *Integration of GIS-Based Model with Epidemiological Data as a Tool for Dengue Surveillance*. EnvironmentAsia, 2017. **10**(2): p. 135-146.
458. Lee, J.S., et al., *Early warning signal for dengue outbreaks and identification of high risk areas for dengue fever in Colombia using climate and non-climate datasets*. BMC Infect Dis, 2017. **17**(1): p. 480.
459. Li, C., et al., *Modeling and projection of dengue fever cases in Guangzhou based on variation of weather factors*. Science of The Total Environment, 2017. **605-606**: p. 867-873.
460. Lowe, R., et al., *Climate services for health: predicting the evolution of the 2016 dengue season in Machala, Ecuador*. The Lancet Planetary Health, 2017. **1**(4): p. e142-e151.
461. Mahdiana, D., et al., *A Model for Forecasting the Number of Cases and Distribution Pattern of Dengue Hemorrhagic Fever in Indonesia*. International Journal of Advanced Computer Science and Applications, 2017. **8**(11).
462. Martinez-Bello, D.A., A. Lopez-Quilez, and A. Torres-Prieto, *Bayesian dynamic modeling of time series of dengue disease case counts*. PLoS Negl Trop Dis, 2017. **11**(7): p. e0005696.
463. Ray, E.L., et al., *Infectious disease prediction with kernel conditional density estimation*. Stat Med, 2017. **36**(30): p. 4908-4929.
464. Sun, W., L. Xue, and X. Xie, *Spatial-temporal distribution of dengue and climate characteristics for two clusters in Sri Lanka from 2012 to 2016*. Scientific Reports, 2017. **7**(1): p. 12884.
465. Xu, L., et al., *Climate variation drives dengue dynamics*. Proc Natl Acad Sci U S A, 2017. **114**(1): p. 113-118.
466. Yang, S., et al., *Advances in using Internet searches to track dengue*. PLoS Computational Biology, 2017. **13**(7): p. e1005607.
467. Agarwal, N., et al., *Data mining techniques for predicting dengue outbreak in geospatial domain using weather parameters for New Delhi, India*. Current Science, 2018. **114**(11): p. 2281-2291.
468. Buczak, A.L., et al., *Ensemble method for dengue prediction*. PLOS ONE, 2018. **13**(1): p. e0189988.
469. Carvajal, T.M., et al., *Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines*. BMC Infectious Diseases, 2018. **18**(1): p. 183.
470. Chen, Y., et al., *Neighbourhood level real-time forecasting of dengue cases in tropical urban Singapore*. BMC Medicine, 2018. **16**(1): p. 129.
471. Cortes, F., et al., *Time series analysis of dengue surveillance data in two Brazilian cities*. Acta Tropica, 2018. **182**: p. 190-197.

472. Hussain-Alkhateeb, L., et al., *Early warning and response system (EWARS) for dengue outbreaks: Recent advancements towards widespread applications in critical settings*. PLOS ONE, 2018. **13**(5): p. e0196811.
473. Johnson, L.R., et al., *Phenomenological forecasting of disease incidence using heteroskedastic Gaussian processes: A dengue case study*. The Annals of Applied Statistics, 2018. **12**(1): p. 27-66.
474. Koh, Y.-M., et al., *A model comparison algorithm for increased forecast accuracy of dengue fever incidence in Singapore and the auxiliary role of total precipitation information*. International Journal of Environmental Health Research, 2018. **28**(5): p. 535-552.
475. Sánchez-González, G., et al., *Prediction of dengue outbreaks in Mexico based on entomological, meteorological and demographic data*. PLOS ONE, 2018. **13**(8): p. e0196047.
476. Tosepu, R., et al., *Climatic Factors Influencing Dengue Hemorrhagic Fever in Kolaka District, Indonesia*. Environment and Natural Resources Journal, 2018. **16**(2): p. 1-10.
477. Withanage, G.P., et al., *A forecasting model for dengue incidence in the District of Gampaha, Sri Lanka*. Parasites & Vectors, 2018. **11**(1): p. 262.
478. Attanayake, A.M.C.H., S.S.N. Perera, and U.P. Liyanage, *COMBINING FORECASTS OF ARIMA AND EXPONENTIAL SMOOTHING MODELS*. Advances and applications in statistics, 2019. **59**: p. 199-208.
479. Bett, B., et al., *Spatiotemporal analysis of historical records (2001–2012) on dengue fever in Vietnam and development of a statistical model for forecasting risk*. PLOS ONE, 2019. **14**(11): p. e0224353.
480. Chakraborty, T., S. Chattopadhyay, and I. Ghosh, *Forecasting dengue epidemics using a hybrid methodology*. Physica A: Statistical Mechanics and its Applications, 2019. **527**: p. 121266.
481. Gabriel, A.F.B., A.P. Alencar, and S.G.E.K. Miraglia, *Dengue outbreaks: unpredictable incidence time series*. Epidemiology and Infection, 2019. **147**: p. e116.
482. Guo, P., et al., *An ensemble forecast model of dengue in Guangzhou, China using climate and social media surveillance data*. Science of The Total Environment, 2019. **647**: p. 752-762.
483. Jayaraj, V.J., et al., *Developing a dengue prediction model based on climate in Tawau, Malaysia*. Acta Tropica, 2019. **197**: p. 105055.
484. Li, R., et al., *Climate-driven variation in mosquito density predicts the spatiotemporal dynamics of dengue*. Proc Natl Acad Sci U S A, 2019. **116**(9): p. 3624-3629.
485. Liu, D., et al., *A dengue fever predicting model based on Baidu search index data and climate data in South China*. PLOS ONE, 2019. **14**(12): p. e0226841.
486. Marques-Toledo, C.A., et al., *Probability of dengue transmission and propagation in a non-endemic temperate area: conceptual model and decision risk levels for early alert, prevention and control*. Parasit Vectors, 2019. **12**(1): p. 38.
487. Mincham, G., et al., *Development of a mechanistic dengue simulation model for Guangzhou*. Epidemiol Infect, 2019. **147**: p. e125.

488. Nayak, S.D.P. and K.A. Narayan, *Prediction of dengue outbreaks in Kerala state using disease surveillance and meteorological data*. International Journal Of Community Medicine And Public Health, 2019. **6**(10): p. 4392-4400.
489. Pineda-Cortel, M.R.B., B.M. Clemente, and P.T.T. Nga, *Modeling and predicting dengue fever cases in key regions of the Philippines using remote sensing data*. Asian Pacific Journal of Tropical Medicine, 2019. **12**(2): p. 60-66.
490. Rahayu, I.T., Nurhasanah, and R. Adriat, *Prediction of Dengue Hemorrhagic Fever Cases Based on Weather Parameters Using Back Propagation Neural Networks (Case Study in Pontianak City)*. Jurnal Pendidikan Fisika Indonesia, 2019. **15**(2): p. 114-121.
491. Rangarajan, P., S.K. Mody, and M. Marathe, *Forecasting dengue and influenza incidences using a sparse representation of Google trends, electronic health records, and time series data*. PLoS Comput Biol, 2019. **15**(11): p. e1007518.
492. Rypdal, M. and G. Sugihara, *Inter-outbreak stability reflects the size of the susceptible pool and forecasts magnitudes of seasonal epidemics*. Nat Commun, 2019. **10**(1): p. 2374.
493. Stolerman, L.M., P.D. Maia, and J.N. Kutz, *Forecasting dengue fever in Brazil: An assessment of climate conditions*. PLOS ONE, 2019. **14**(8): p. e0220106.
494. Zhu, B., et al., *Prediction model for dengue fever based on interactive effects between multiple meteorological factors in Guangdong, China (2008-2016)*. PLoS One, 2019. **14**(12): p. e0225811.
495. Appice, A., et al., *A Multi-Stage Machine Learning Approach to Predict Dengue Incidence: A Case Study in Mexico*. IEEE Access, 2020. **8**: p. 52713-52725.
496. Bal, S. and S. Sodoudi, *Modeling and prediction of dengue occurrences in Kolkata, India, based on climate factors*. Int J Biometeorol, 2020. **64**(8): p. 1379-1391.
497. Benedum, C.M., et al., *Weekly dengue forecasts in Iquitos, Peru; San Juan, Puerto Rico; and Singapore*. PLoS Negl Trop Dis, 2020. **14**(10): p. e0008710.
498. Brett, T.S. and P. Rohani, *Dynamical footprints enable detection of disease emergence*. PLoS Biol, 2020. **18**(5): p. e3000697.
499. Chakraborty, A. and V. Chandru, *A Robust and Non-parametric Model for Prediction of Dengue Incidence*. Journal of the Indian Institute of Science, 2020. **100**(4): p. 893-899.
500. Chen, P., et al., *Early dengue outbreak detection modeling based on dengue incidences in Singapore during 2012 to 2017*. Stat Med, 2020. **39**(15): p. 2101-2114.
501. Cheng, Y.C., et al., *Real-time dengue forecast for outbreak alerts in Southern Taiwan*. PLoS Negl Trop Dis, 2020. **14**(7): p. e0008434.
502. Liu, K., et al., *Enhancing fine-grained intra-urban dengue forecasting by integrating spatial interactions of human movements between urban regions*. PLoS Negl Trop Dis, 2020. **14**(12): p. e0008924.
503. Majhi, J., et al., *Dynamics of dengue outbreaks in gangetic West Bengal: A trend and time series analysis*. J Family Med Prim Care, 2020. **9**(11): p. 5622-5628.
504. Medina, J.R.C., et al., *Time Series Forecasting Using a SARMA Model Fitted on Dengue Cases (2010–2017): A Potential Tool for an Early Warning System in Quezon City, Philippines*. Southeast Asian Journal of Tropical Medicine and Public Health, 2020. **51**(Supplement 1): p. 31-46.

505. Mustaffa, Z., et al., *Predictive Analysis of Dengue Outbreak Based on an Improved Salp Swarm Algorithm*. Cybern. Inf. Technol., 2020. **20**(4): p. 156–169.
506. Puengpreeda, A., S. Yhusumrarn, and S. Sirikulvadhana, *Weekly Forecasting Model for Dengue Hemorrhagic Fever Outbreak in Thailand*. Engineering Journal, 2020. **24**(3): p. 71-87.
507. Stoner, O. and T. Economou, *Multivariate hierarchical frameworks for modeling delayed reporting in count data*. Biometrics, 2020. **76**(3): p. 789-798.
508. Tao, L.J., et al., *Correction to 'Explicit characterization of human population connectivity reveals long run persistence of interregional dengue shocks'*. J R Soc Interface, 2020. **17**(169): p. 20200620.
509. Tao, L.J., et al., *Explicit characterization of human population connectivity reveals long run persistence of interregional dengue shocks*. J R Soc Interface, 2020. **17**(168): p. 20200340.
510. Tsai, P.-J., *Prediction of Dengue Fever Outbreak Based on Case Household Locations in Southern Taiwan*. Open Journal of Preventive Medicine, 2020. **10**(7): p. 175-193.
511. Wagner, C.E., et al., *Climatological, virological and sociological drivers of current and projected dengue fever outbreak dynamics in Sri Lanka*. J R Soc Interface, 2020. **17**(167): p. 20200075.
512. Xu, J., et al., *Forecast of Dengue Cases in 20 Chinese Cities Based on the Deep Learning Method*. Int J Environ Res Public Health, 2020. **17**(2).
513. Benkimoun, S., et al., *Dynamic mapping of dengue basic reproduction number*. Results in Physics, 2021. **29**: p. 104687.
514. Caldwell, J.M., et al., *Climate predicts geographic and temporal variation in mosquito-borne disease dynamics on two continents*. Nat Commun, 2021. **12**(1): p. 1233.
515. Cheng, J., et al., *Extreme weather events and dengue outbreaks in Guangzhou, China: a time-series quasi-binomial distributed lag non-linear model*. Int J Biometeorol, 2021. **65**(7): p. 1033-1042.
516. Cheng, X., et al., *Impact of interventions on the incidence of natural focal diseases during the outbreak of COVID-19 in Jiangsu Province, China*. Parasit Vectors, 2021. **14**(1): p. 483.
517. Ferdousi, T., L.W. Cohnstaedt, and C.M. Scoglio, *A Windowed Correlation-Based Feature Selection Method to Improve Time Series Prediction of Dengue Fever Cases*. IEEE Access, 2021. **9**: p. 141210-141222.
518. Heidrich, P. and T. Götz, *Simulation and prediction of dengue outbreaks based on an SIR model with a time-dependent transmission rate including meteorological data. An example for Colombo and Jakarta*. International Journal of Biomathematics, 2022. **15**(02): p. 2150073.
519. Liu, K., et al., *Facilitating fine-grained intra-urban dengue forecasting by integrating urban environments measured from street-view images*. Infect Dis Poverty, 2021. **10**(1): p. 40.
520. McGough, S.F., et al., *A dynamic, ensemble learning approach to forecast dengue fever epidemic years in Brazil using weather and population susceptibility cycles*. J R Soc Interface, 2021. **18**(179): p. 20201006.

521. Nuraini, N., et al., *Climate-based dengue model in Semarang, Indonesia: Predictions and descriptive analysis*. Infect Dis Model, 2021. **6**: p. 598-611.
522. Ray, A., T. Chakraborty, and D. Ghosh, *Optimized ensemble deep learning framework for scalable forecasting of dynamics containing extreme events*. Chaos, 2021. **31**(11): p. 111105.
523. Sriklin, T., S. Kajornkasirat, and S. Puttinaovarat, *Dengue Transmission Mapping with Weather-Based Predictive Model in Three Southernmost Provinces of Thailand*. Sustainability, 2021. **13**(12): p. 6754.
524. Xavier, L.L., et al., *Analysis of climate factors and dengue incidence in the metropolitan region of Rio de Janeiro, Brazil*. PLoS One, 2021. **16**(5): p. e0251403.
525. Yi, C., L.W. Cohnstaedt, and C.M. Scoglio, *SEIR-SEI-EnKF: A New Model for Estimating and Forecasting Dengue Outbreak Dynamics*. IEEE Access, 2021. **9**: p. 156758-156767.
526. Bracher, J. and L. Held, *Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction*. International Journal of Forecasting, 2022. **38**(3): p. 1221-1233.
527. Deb, S. and S. Deb, *An Ensemble Method for Early Prediction of Dengue Outbreak*. Journal of the Royal Statistical Society Series A: Statistics in Society, 2022. **185**(1): p. 84-101.
528. Heidrich, P., et al., *Prediction of dengue cases based on human mobility and seasonality—An example for the city of Jakarta*. Mathematical Methods in the Applied Sciences, 2021. **44**(17): p. 13633-13658.
529. Ismail, S., et al., *The practicality of Malaysia dengue outbreak forecasting model as an early warning system*. Infect Dis Model, 2022. **7**(3): p. 510-525.
530. Kakarla, S.G., et al., *Weather integrated multiple machine learning models for prediction of dengue prevalence in India*. Int J Biometeorol, 2022. **67**(2): p. 285-297.
531. Koplewitz, G., et al., *Predicting dengue incidence leveraging internet-based data sources. A case study in 20 cities in Brazil*. PLoS Negl Trop Dis, 2022. **16**(1): p. e0010071.
532. Li, Z., *Forecasting Weekly Dengue Cases by Integrating Google Earth Engine-Based Risk Predictor Generation and Google Colab-Based Deep Learning Modeling in Fortaleza and the Federal District, Brazil*. Int J Environ Res Public Health, 2022. **19**(20).
533. Li, Z., et al., *Improving Dengue Forecasts by Using Geospatial Big Data Analysis in Google Earth Engine and the Historical Dengue Information-Aided Long Short Term Memory Modeling*. Biology (Basel), 2022. **11**(2).
534. Li, L., et al., *Dengue Risk Forecast with Mosquito Vector: A Multicomponent Fusion Approach Based on Spatiotemporal Analysis*. Comput Math Methods Med, 2022. **2022**: p. 2515432.
535. Ligue, K.D. and K.J. Ligue, *Deep learning approach to forecasting Dengue cases in Davao City using Long Short-term Memory (LSTM)*. Philippine Journal of Science, 2022. **151**(3): p. 797-812.
536. Mahajan, A., et al., *A Novel Stacking-Based Deterministic Ensemble Model for Infectious Disease Prediction*. Mathematics, 2022. **10**(10): p. 1714.

537. Mukhaiyar, U., et al., *Forecasting of dengue hemorrhages fever cases with autoregression distributed lag model using Poisson regression approach*. Biostatistics & Epidemiology, 2022. **7**(1): p. e2064636.
538. Ong, J., et al., *Fine-scale estimation of effective reproduction numbers for dengue surveillance*. PLoS Comput Biol, 2022. **18**(1): p. e1009791.
539. Othman, M., et al., *Model Forecasting Development for Dengue Fever Incidence in Surabaya City Using Time Series Analysis*. Processes, 2022. **10**(11): p. 2454.
540. Roster, K., C. Connaughton, and F.A. Rodrigues, *Machine-Learning-Based Forecasting of Dengue Fever in Brazilian Cities Using Epidemiologic and Meteorological Variables*. American Journal of Epidemiology, 2022. **191**(10): p. 1803-1812.
541. Sanchez-Gendríz, I., et al., *Data-driven computational intelligence applied to dengue outbreak forecasting: a case study at the scale of the city of Natal, RN-Brazil*. Sci Rep, 2022. **12**(1): p. 6550.
542. Sathish, V., S. Mukhopadhyay, and R. Tiwari, *Autoregressive and moving average models for zero-inflated count time series*. Statistica Neerlandica, 2022. **76**(2): p. 190-218.
543. Souza, C., et al., *Predicting dengue outbreaks in Brazil with manifold learning on climate data*. Expert Systems with Applications, 2022. **192**: p. 116324.
544. Surendran, S.N., et al., *Reduced dengue incidence during the COVID-19 movement restrictions in Sri Lanka from March 2020 to April 2021*. BMC Public Health, 2022. **22**(1): p. 388.
545. Khan, M., et al., *Dengue transmission under future climate and human population changes in mainland China*. Applied Mathematical Modelling, 2023. **114**: p. 785-798.