Dengue Cases Prediction

Irvan Indra Wahab

Table of Contents

01

Introduction

Overview, problem statement, and objectives

02

Data Preparation

Data collection, cleaning, and preprocessing

03

EDA

Data exploration and visualization

04

Modeling

Time series modeling and performance evaluation

05

Recommendation

Cost-Benefit analysis and conclusion

01

Introduction

Overview, problem statement and objectives

World Health Organization

DENGUE and severe dengue

Mosquito-borne VIRAL infection leading to:

DENGUE
Flu-like illness

SEVERE DENGUE
Life-threatening complications

Estimated
390 MILLION
dengue virus infections / year

KILLS one person every 12 minutes

Leading cause of death in children

Fastest spreading, epidemic-prone infectious disease

Threatens 4 BILLION people in 128 countries

Causes **BILLION US $ in economic loss**

Dengue in Singapore

Singapore an ideal environment for mosquito breeding

- High temperatures, rainfall, and humidity
- Densely populated city with three million of residents live in HDB flats

Crucial to predict dengue cases and implement dengue prevention program

Allow MOH/NEA and the public to take necessary preventive measures

02

Data Preparation

Data collection, cleaning, and preprocessing

Data Sources

Dengue Cases

Temperature

Rainfall

Google Trends

Data Cleaning

Filtering

Filtering diseases to only include 'Dengue Fever' or 'Dengue Haemorrhagic Fever'

Aggregating

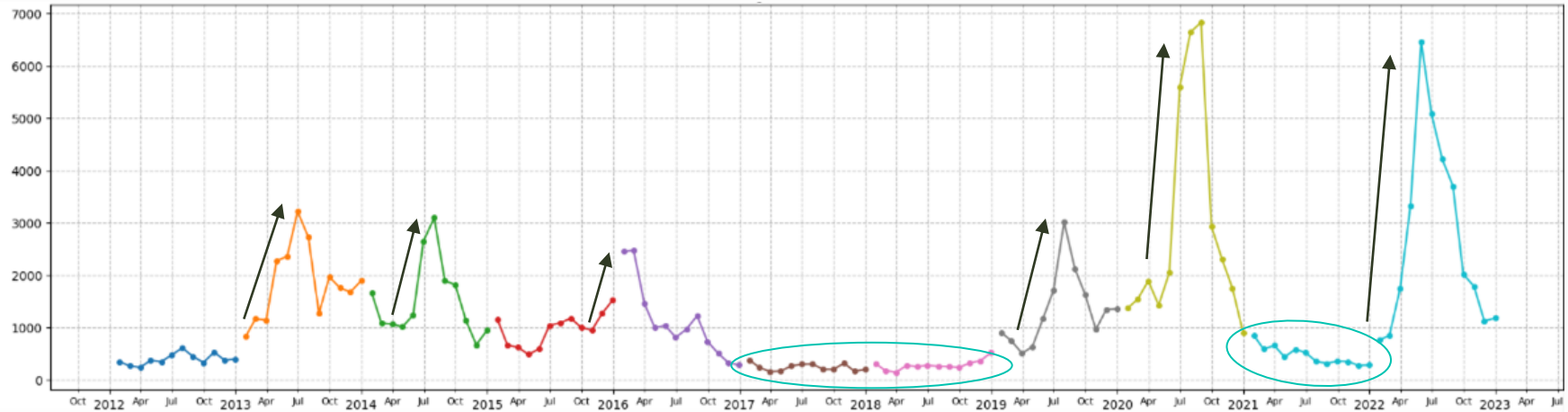
Aggregating the weekly time series data into a monthly time series format.

03 EDA

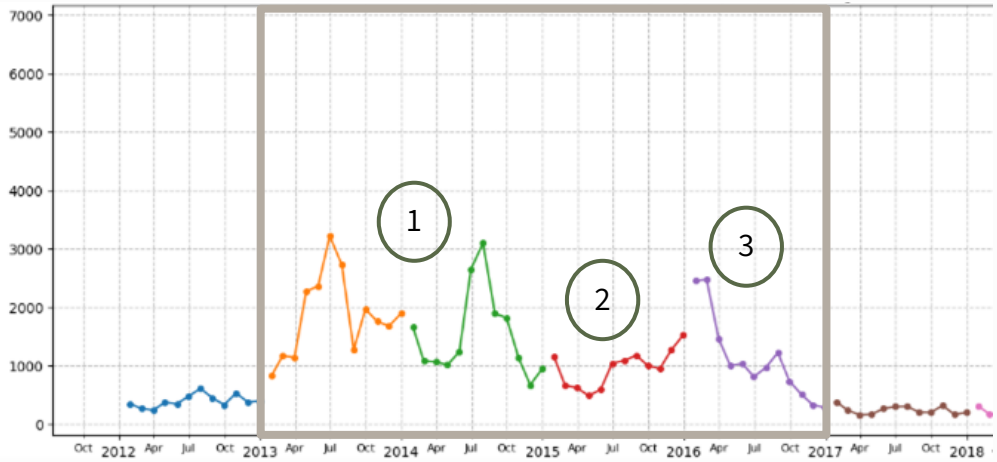
Dengue Trend, Weather Conditions, Maps, and Google Trends

Dengue Cases Across Years

- Monthly Dengue Trend in Singapore from 2012 to 2022
- Multiple spikes in cases are accompanied by some years with lower incidence
- Each year is depicted by a different color in the graph.



Serotype Switch: 2013 - 2016



1) A large epidemic stretching over two years in 2013 - 2014 associated with a switch from serotype 2 (DENV-2) to serotype 1 (DENV-1)

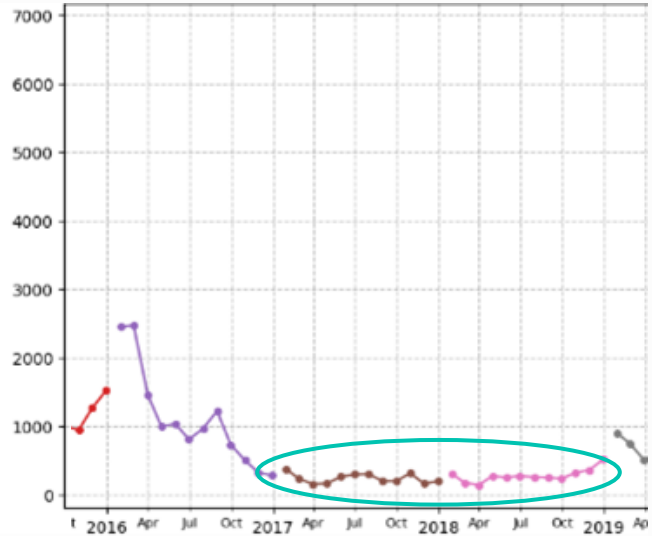
1) Drastic drop in dengue cases from the gravitraps deployment, ramping up of inspections, and population immunity

1) Outbreak following a switch from DENV-1 to DENV-2 but quickly brought under control

Controlled Dengue: 2017 - 2018

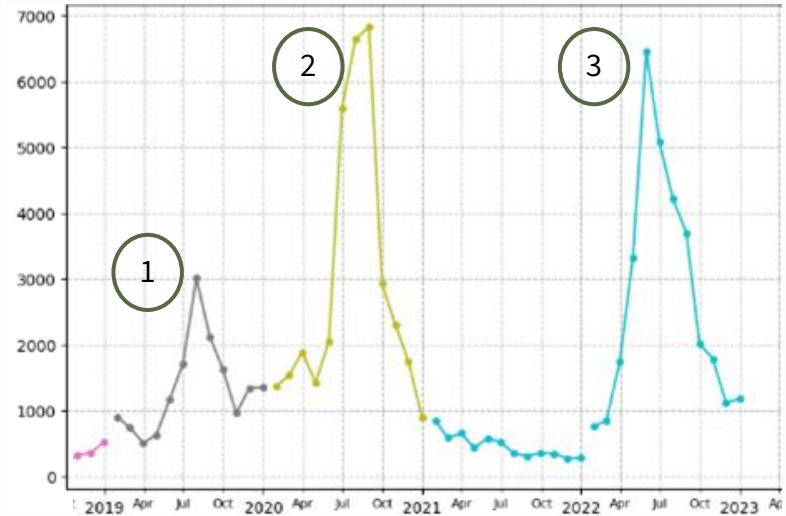
The lowest figure in the past several years with no uptick of cases observed, attributed to:

- Immunity built after a high number of cases from 2013-2014 and early 2016
- Community and NEA step up inspection
- Implementation of network of Gravitraps



Covid-19: 2019 - 2022

1) Uptick due to the switch to DENV-3 serotype and the warmer months



3) Resurgence due to:
- Waning immunity
- Rise of DENV-3
- Higher mosquito population due to increased construction

2) Resurgence of epidemic dengue during Covid-19 due to
- The rise of DENV-3 and DENV-4 serotypes
- The work-from-home arrangement increased probability of contact with mosquitos

Weather Conditions

Weather

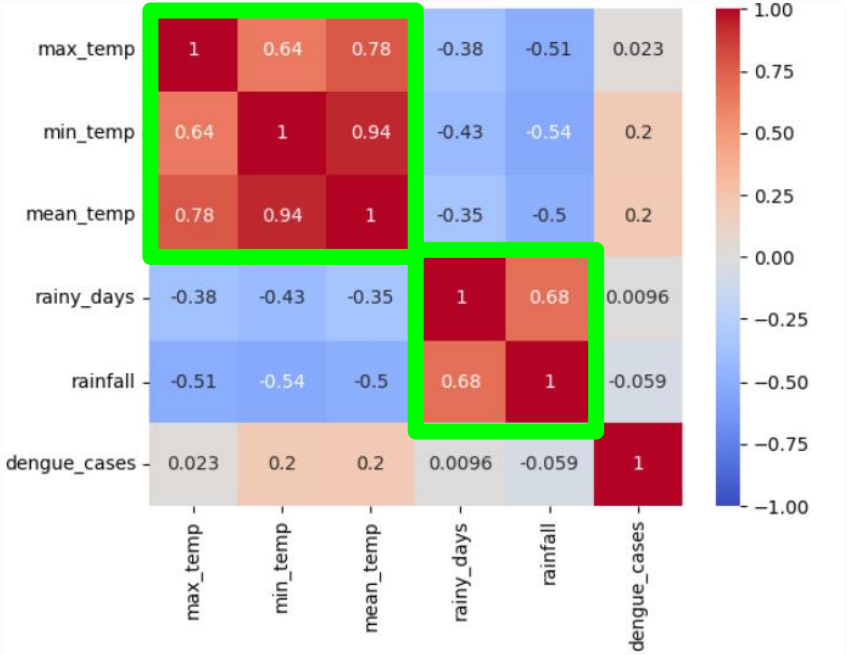
The weather conditions across Singapore

Temperature

Max, min, and mean temperature

Rainfall

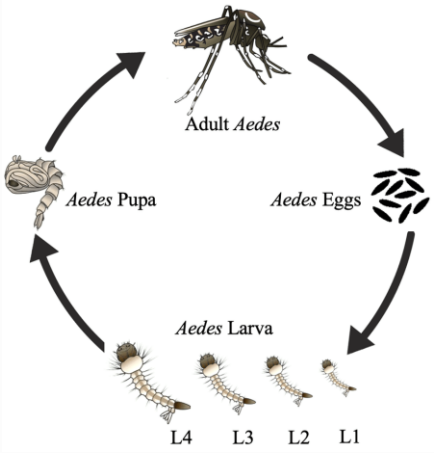
Total number of rainy days and total rainfall

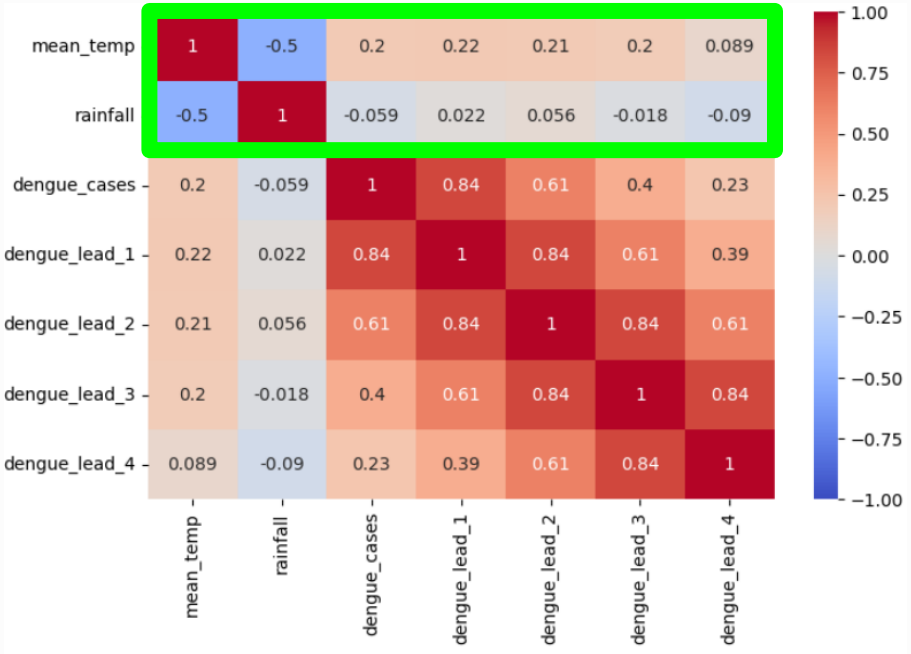


Variables used are **Mean Temperature** and **Total Rainfall** to avoid multicollinearity

Weather vs Dengue

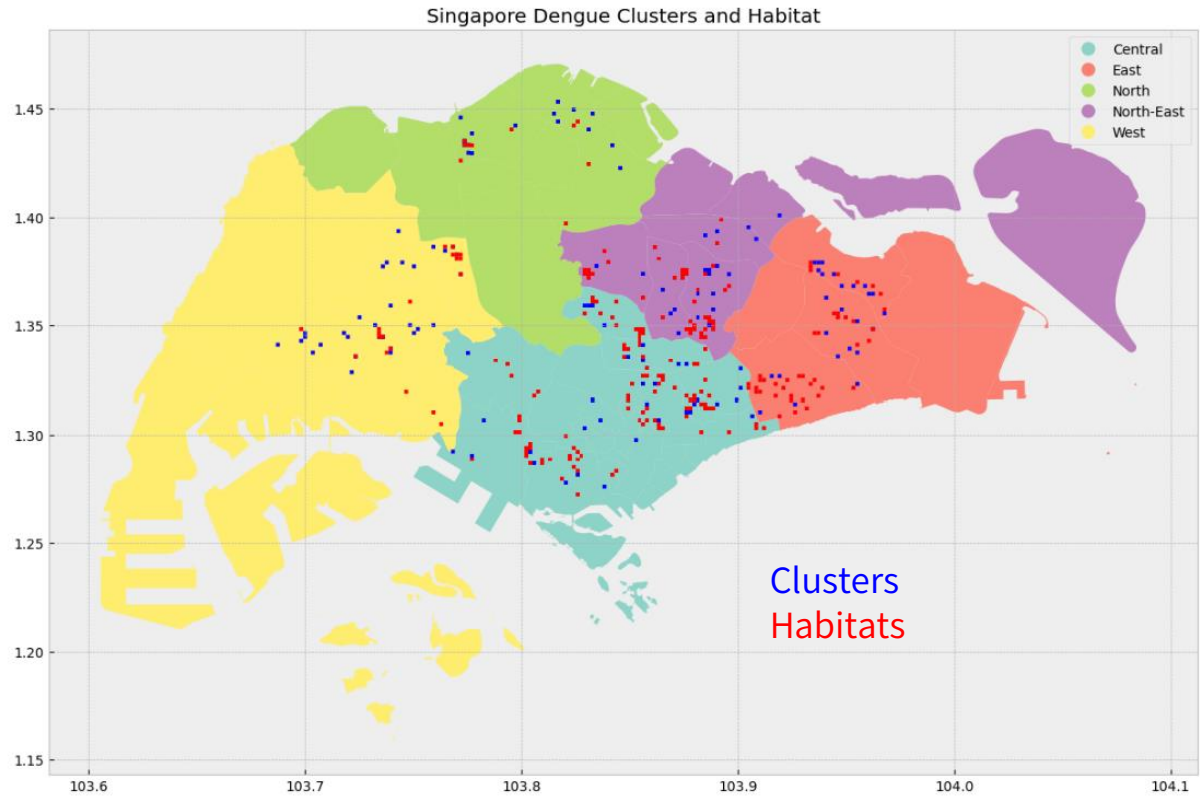
Research suggests that mean temperature and rainfall precede dengue cases by 3–4 months.





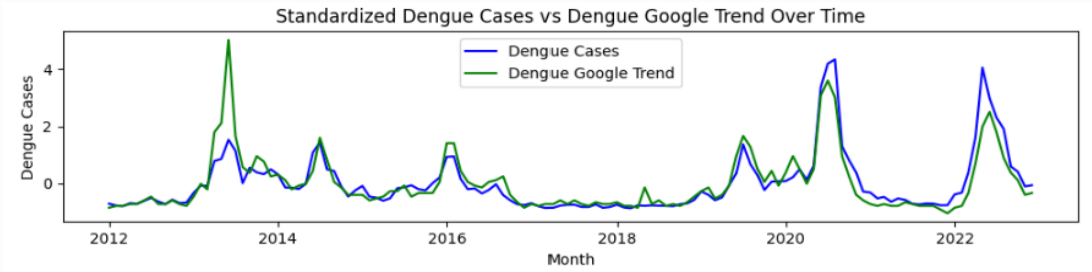
Weak correlation between weather conditions to dengue cases and lead duration

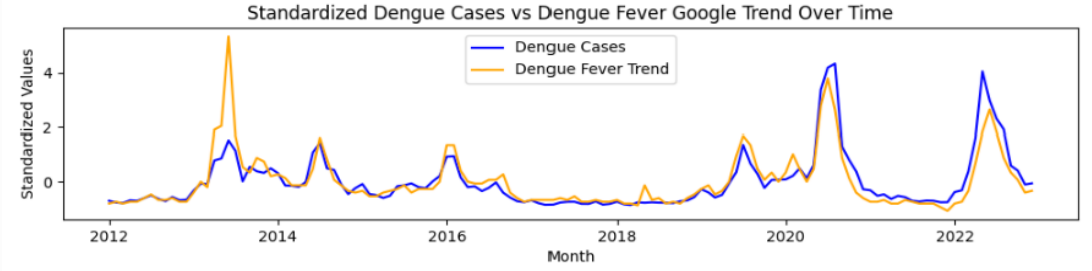
Dengue Clusters and Habitat

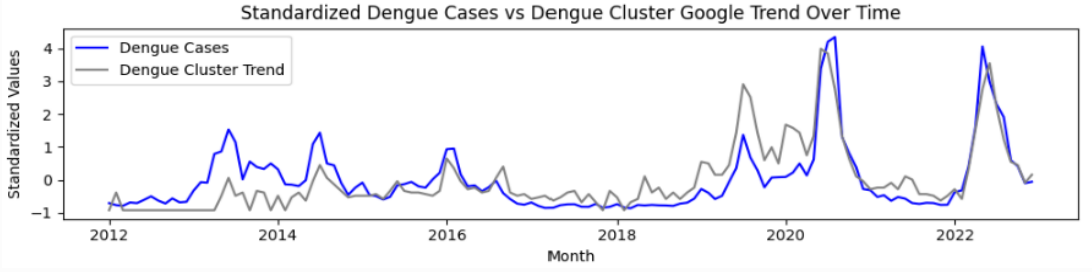


- GeoJSON data from data.gov.sg
- Clusters and habitats in close proximity
- There are clusters that are not near habitats

Dengue Google Trends







- Keywords searched:
  - Dengue
  - Dengue Fever
  - Dengue Cluster

- The trends show that the dengue related google search exhibit similar trends with the dengue cases, all points being standardized

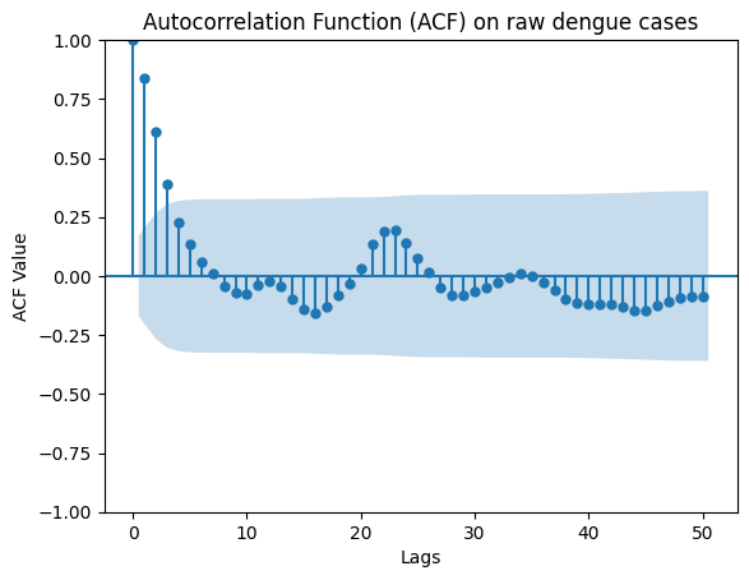
Dengue Google Trend

- Strong correlation between dengue cases and google trends
- Applied lead 1 month of dengue cases and still observe strong correlation
- This shows that Google Trends can potentially be used for an early warning system before dengue outbreak

04 Modelling

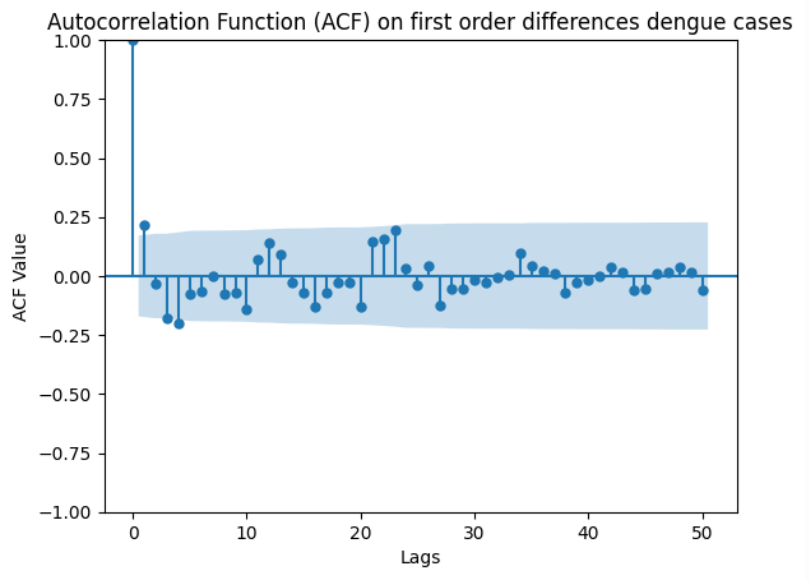
Time Series modelling and evaluation

Stationarity



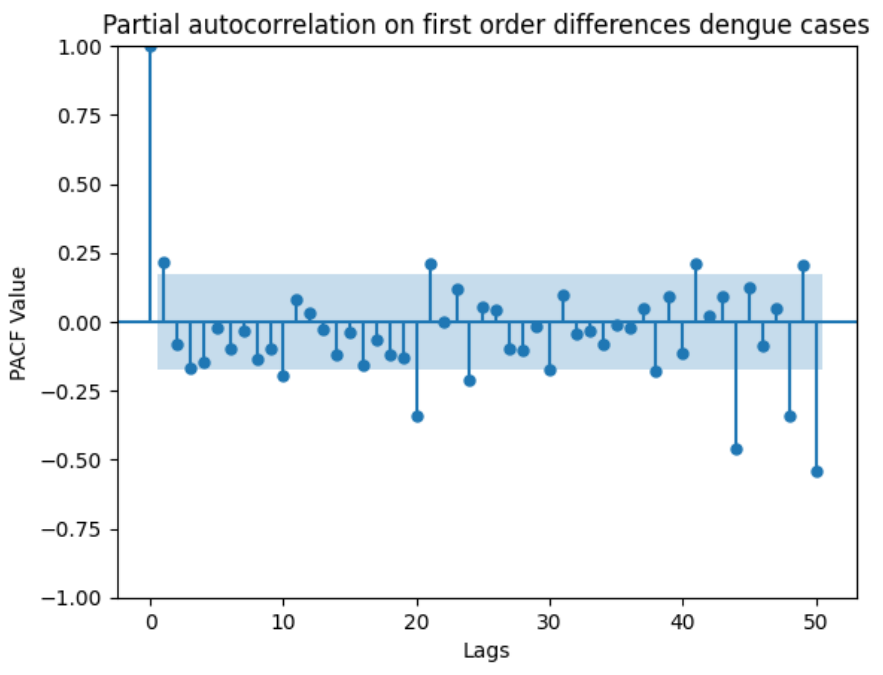
Slow decay suggests that the series is non-stationary, hence differencing is needed.

No.of cases in one month is highly dependent on no.of cases in previous months



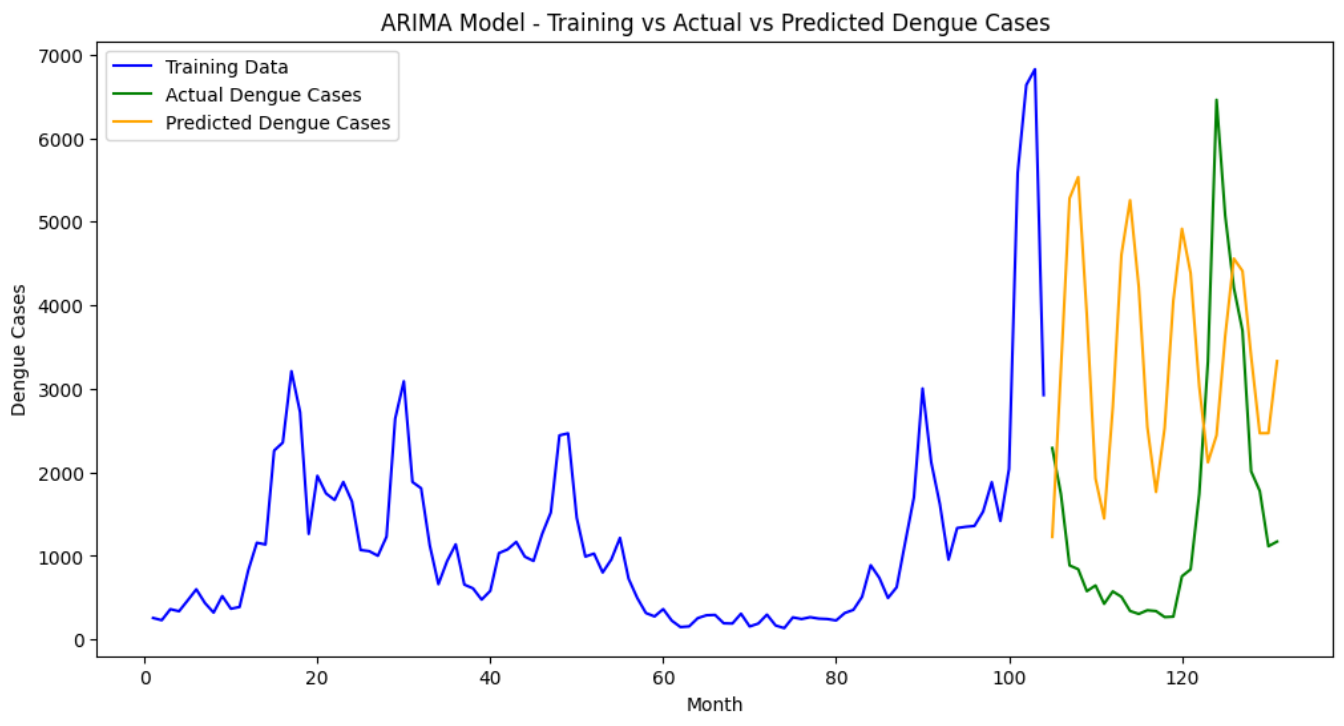
Series is now stationary, which should show a rapid decline in autocorrelation after a few lags.

Stationarity



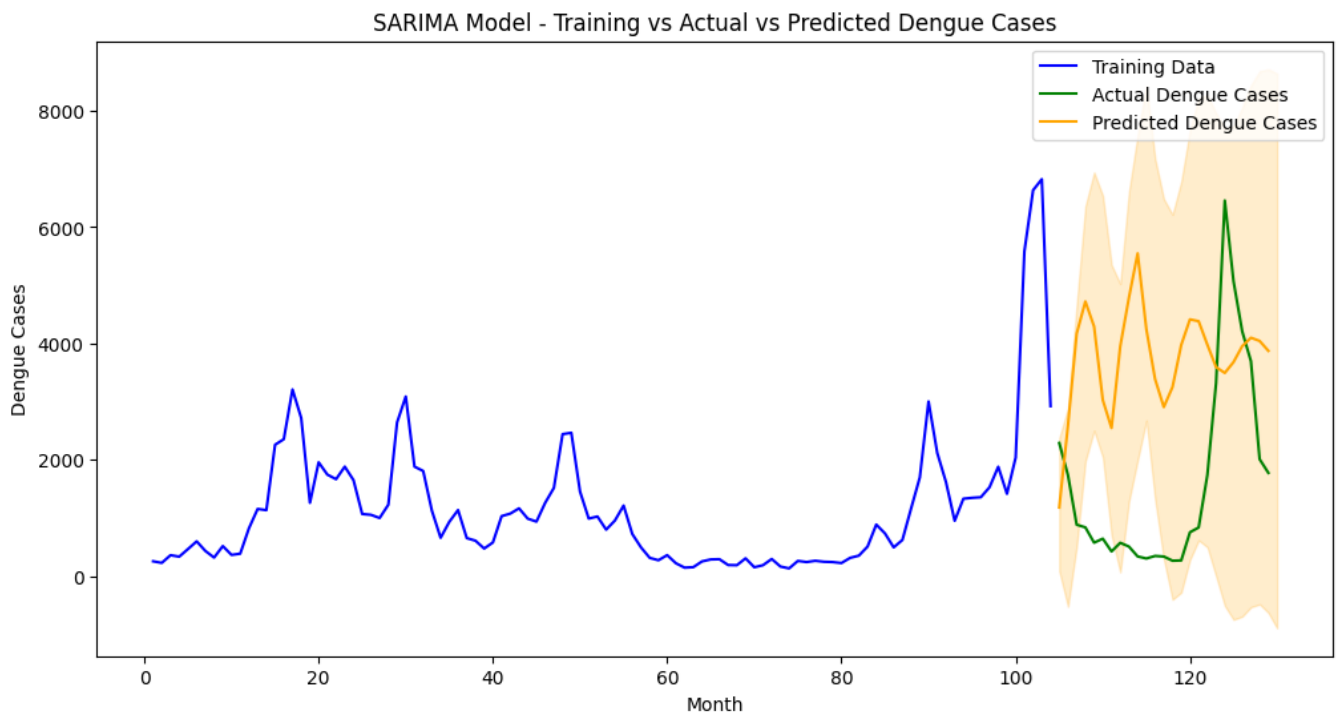
- PACF measures the direct effect of a lagged value on the current value, excluding the influence of intermediate values.
- Significant spikes indicate strong partial correlations at those lags.

ARIMA (2,1,3)



Test RMSE: 2,764

SARIMA (4, 1, 0)

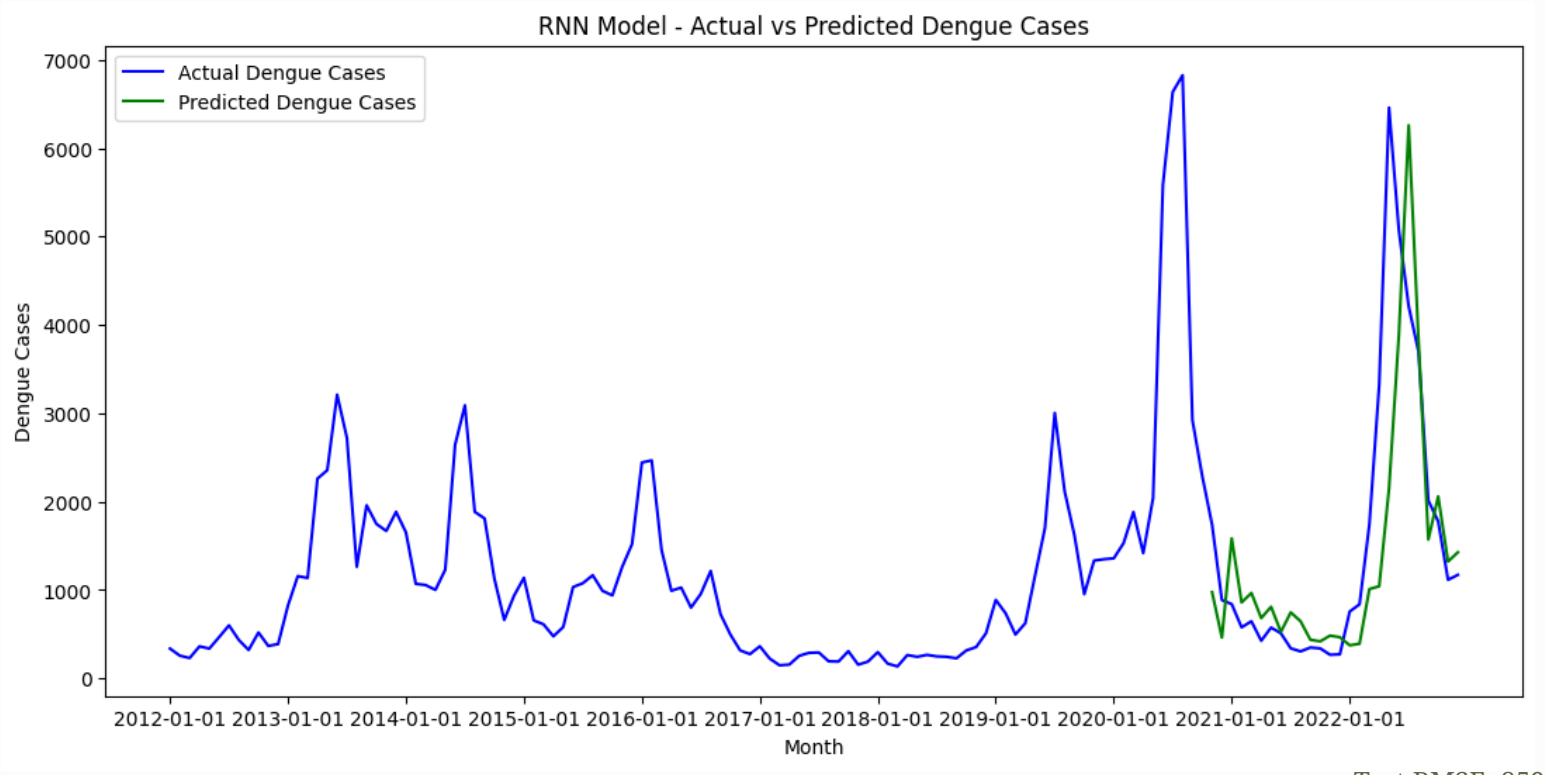


Test RMSE: 2,919

Data preprocessing

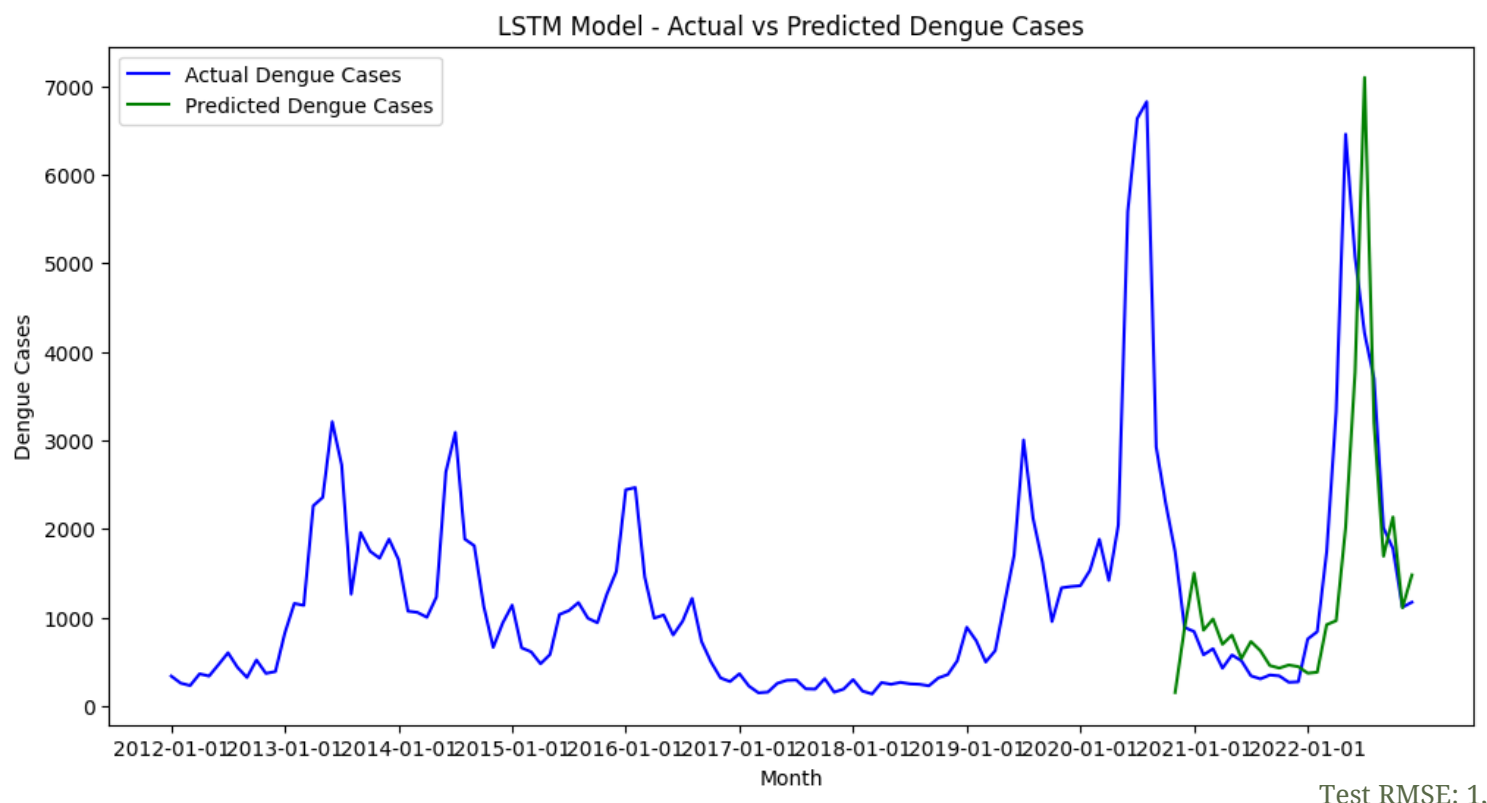
- Min Max scaler to standardise data
- Varied lookback period - indicates how many previous data points will be used to predict the current data point.
- Train test split of 80/20 ratio
- Dengue cases aggregated by month

RNN - single feature



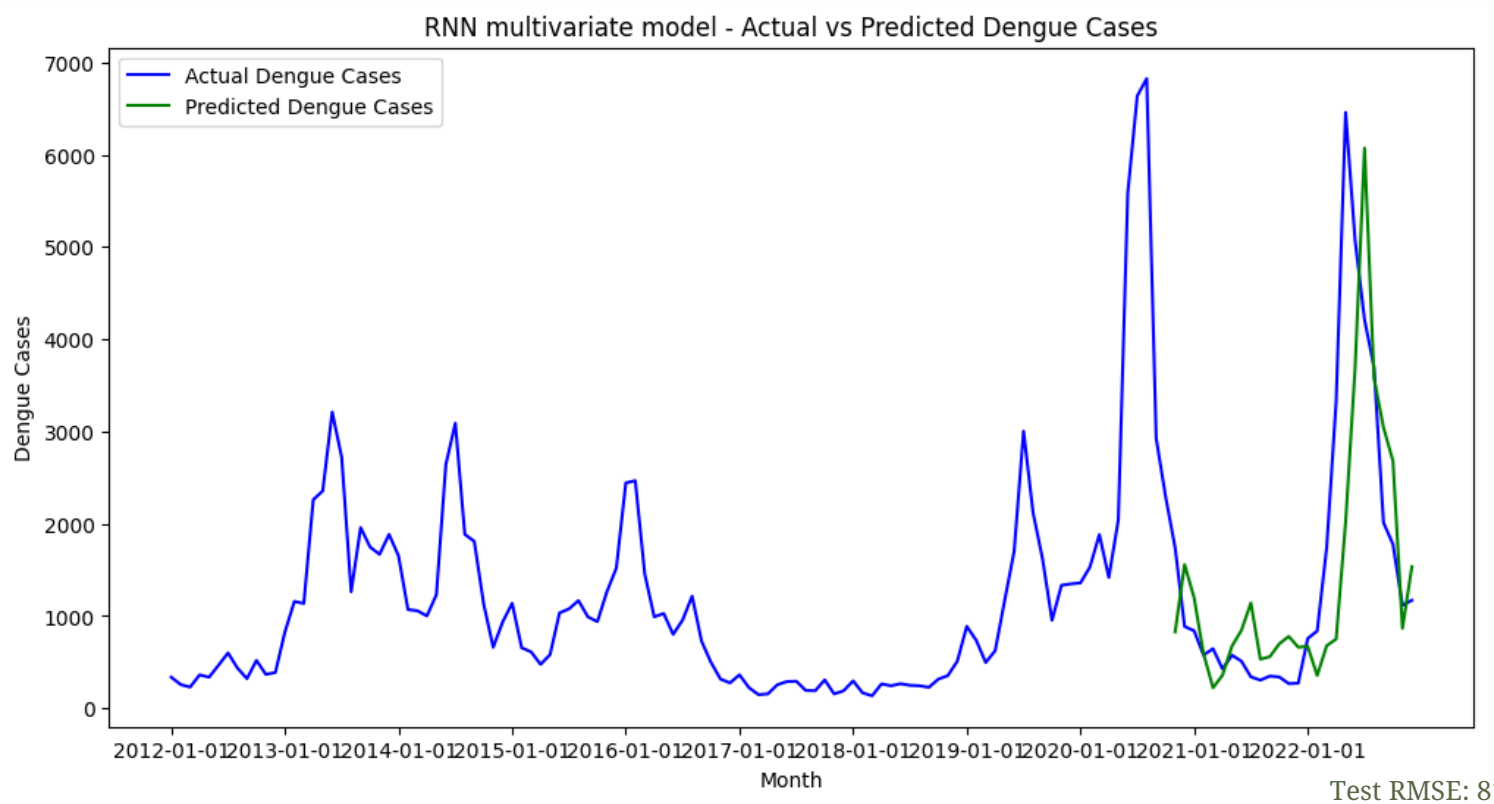
Test RMSE: 859

LSTM - single feature



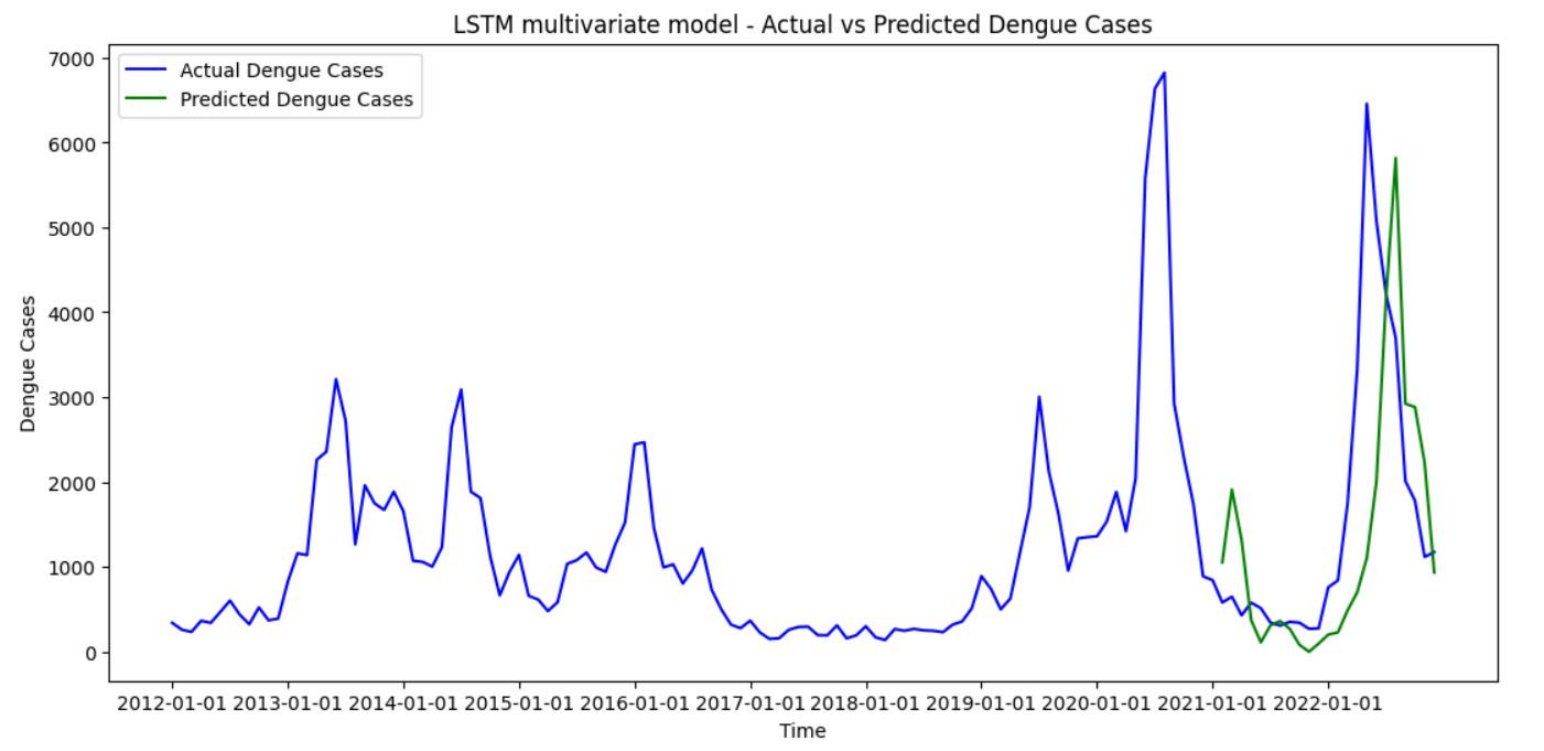
Test RMSE: 1,002

RNN - Multivariate time series forecasting



Test RMSE: 817

LSTM - Multivariate time series forecasting



Test RMSE: 1,215

Model Evaluation

- Log transformation to scale the data over Min Max scaler or standard scaler for the following reasons:
  - A log transformation reduces the impact of outliers.
  - Overall RMSE scores performed slightly worse for the test set, but lesser overfitting
- Using multiple features for Neural Networks led to a worsening for RMSE scores
- In the single feature LSTM vs RNN, the LSTM model did better than RNN because during backpropagation, RNNs suffer from the vanishing gradient problem, where the gradient can become very small and the weights of the neurons are not effectively updated

05

Recommendation

Cost-Benefit Analysis and Conclusion

PROJECT WOLBACHIA
SINGAPORE

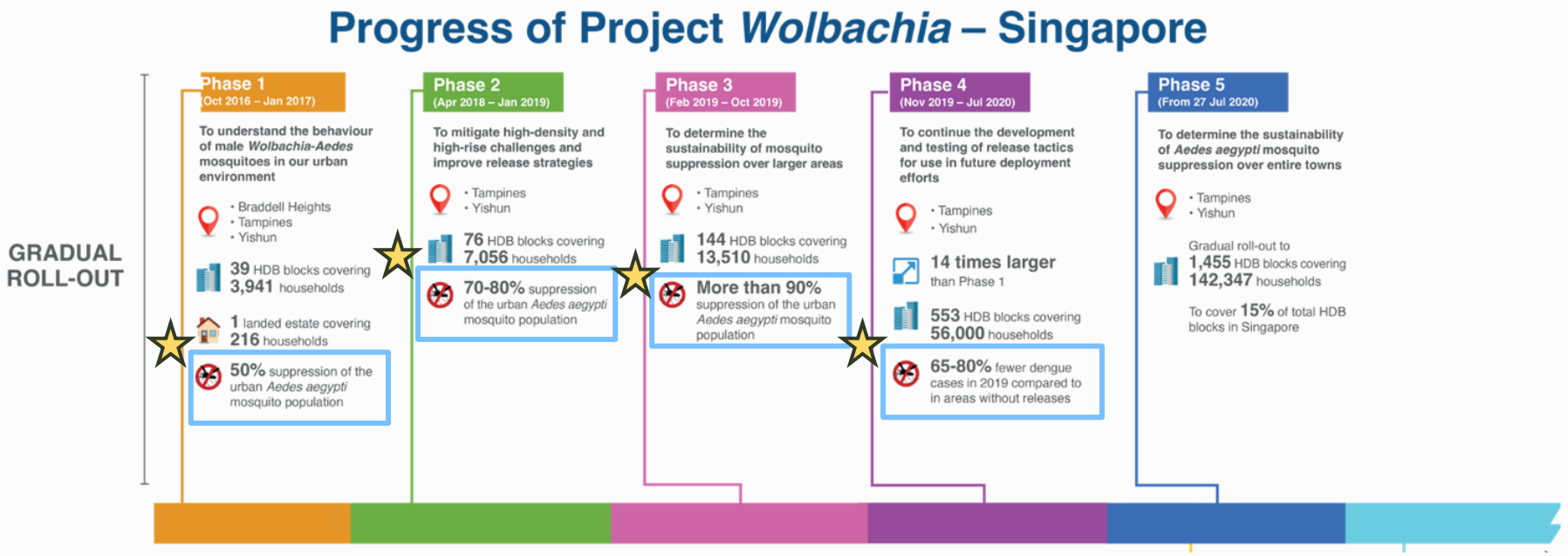
How does the suppression technology work?

- Release of non-biting male Wolbachia-carrying mosquitoes to mate with urban female mosquitoes
- When they mate, the eggs derived from these matings do not hatch, thereby, reducing the population of Aedes mosquitoes in the field

What are some common facts?

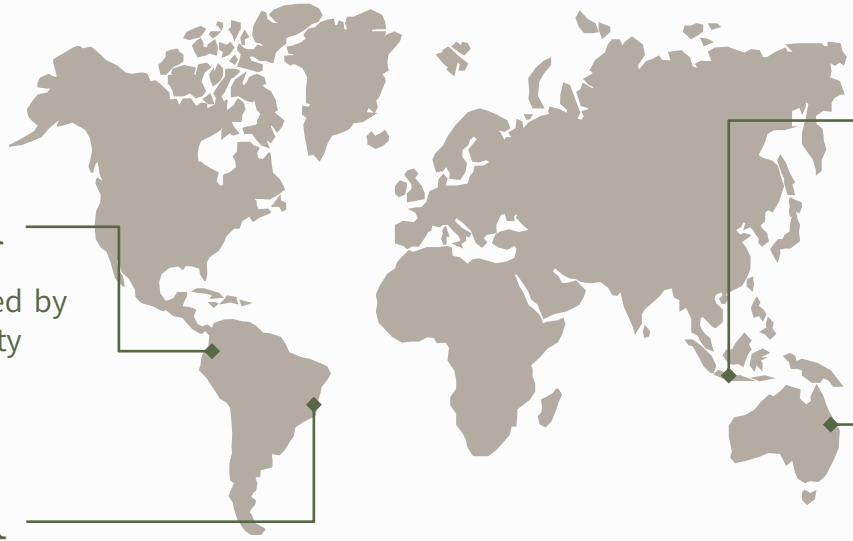
- Wolbachia is a safe, naturally occurring bacterium
- Wolbachia DOES NOT harm humans or animals
- Male mosquitoes do not bite or transmit disease
- Research suggests implementation cost for the whole of Singapore is **SGD 40 million**

Source: NEA



Average dengue cases reduction of 70%

Wolbachia Global



Colombia

Dengue incidence reduced by 50% in the applied city

Brazil

Dengue is significantly lower in Wolbachia-treated neighbourhoods of Rio and Niteroi

Indonesia

In a cluster randomised trial, dengue was 77% lower

Australia

Dengue has been effectively eliminated as a public health concern in Far North Queensland

Economic Impact

Calculating the cost per dengue case

Direct

Cost incurred from consultation and hospitalization

Consultation

Blood test, regular consultation, medication

Hospitalization

Average hospital fee for 3 days and doctor attendance

Indirect

Reduction of work productivity and time required to provide caretaking

Productivity

Typically dengue fever lasts for up to 1 week, resulting in workers unable to work

Caretaking

Symptomatic children and elderly requires caretaking

Total Economic Impact

	Direct Cost		Indirect Cost	
	Consultation	Hospitalization	Productivity	Caretaking
Percentage	0.83	0.17	0.80	0.20
Fee (SGD)	200	7,000	1,250*	320**
Total (SGD)	166	1190	1,000	64
Grand Total (SGD)	**1,356**		**1,064**	

*7 days out of work due to dengue fever
**2 days out of work due to caretaking of children and elderly

1 day out of work equivalent to the median daily annual income, derived from median gross monthly income from MOM data

Recommendations

- Project Wolbachia implementation
- Google Trend as early warning system before dengue outbreak
- Cluster does not always form in habitat - be vigilant

Thank You!