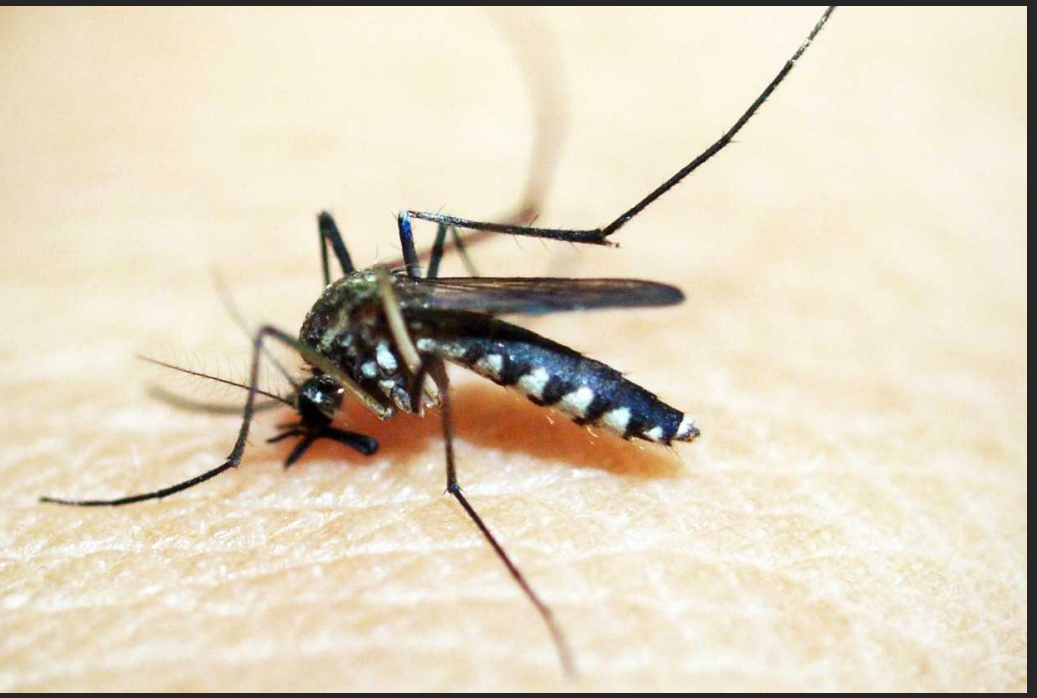
Project 4

SG DENGUE CASES PREDICTION

DSIF9 Team 5

Song Kai, Hong Xiang, Zhu Ye (Juliana)



Contents

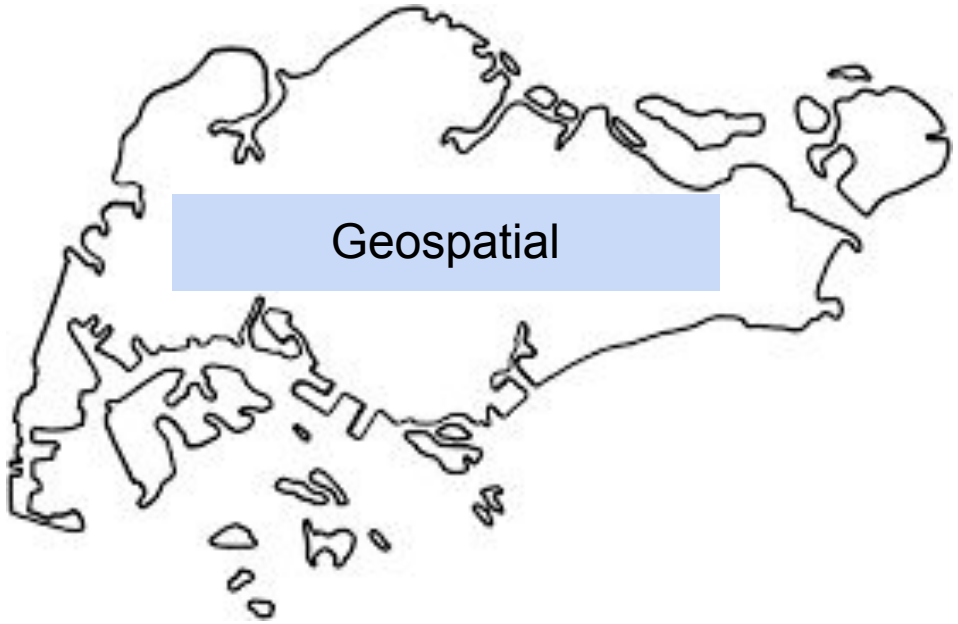
Background

- Dengue fever is a viral infection that spreads from Aedes mosquitoes to people and is commonly found in tropical countries including Singapore.

- Why is dengue fever a significant health concern?
  - Rate of transmission
  - Lethal outcomes of infection

Problem Statement

- Predict the number of dengue cases in Singapore on a temporal and spatial level, to enable targeted intervention and better resource allocation.



Temporal

Dengue Cases

Weather

- Rainfall
- Wind speed
- Temperature

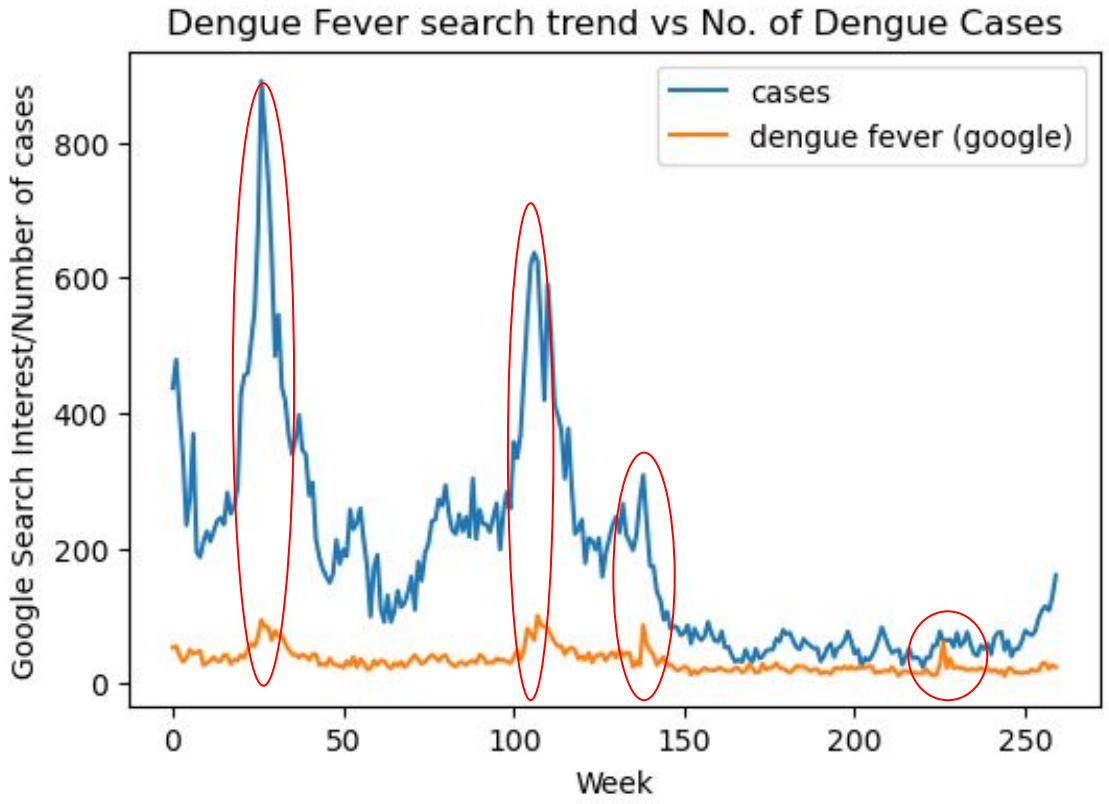
Trend words

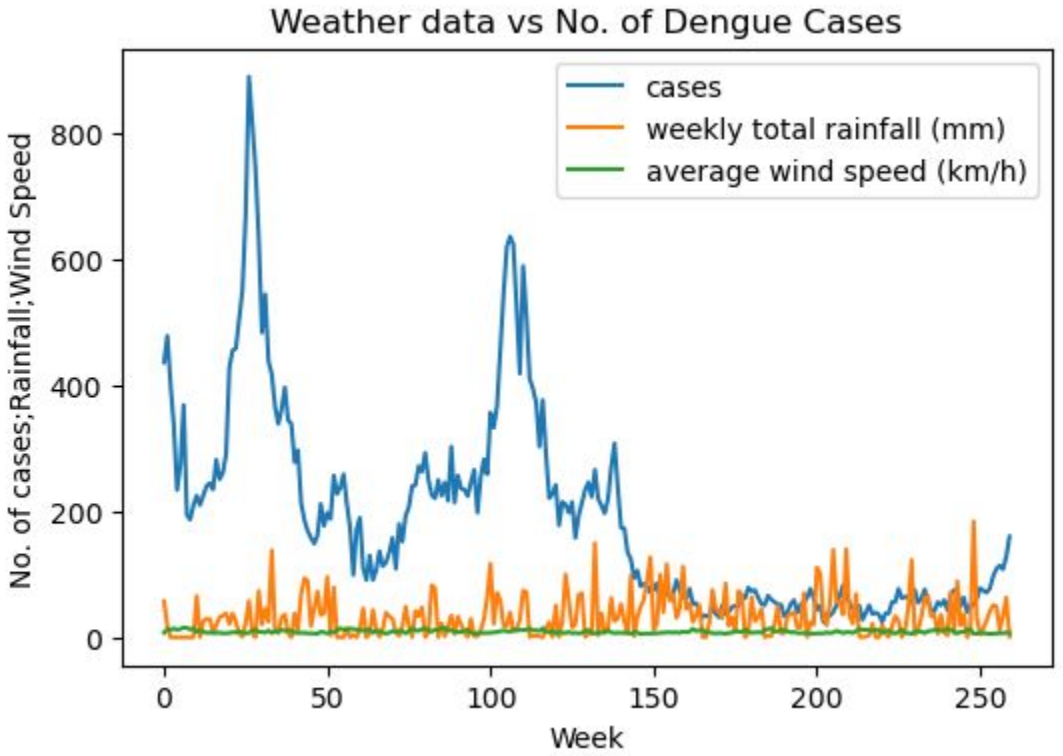
- Dengue Fever
- Headache
- Nausea
- Vomiting
- Muscle Ache
- Rashes
- Mosquito Bite

Temporal EDA

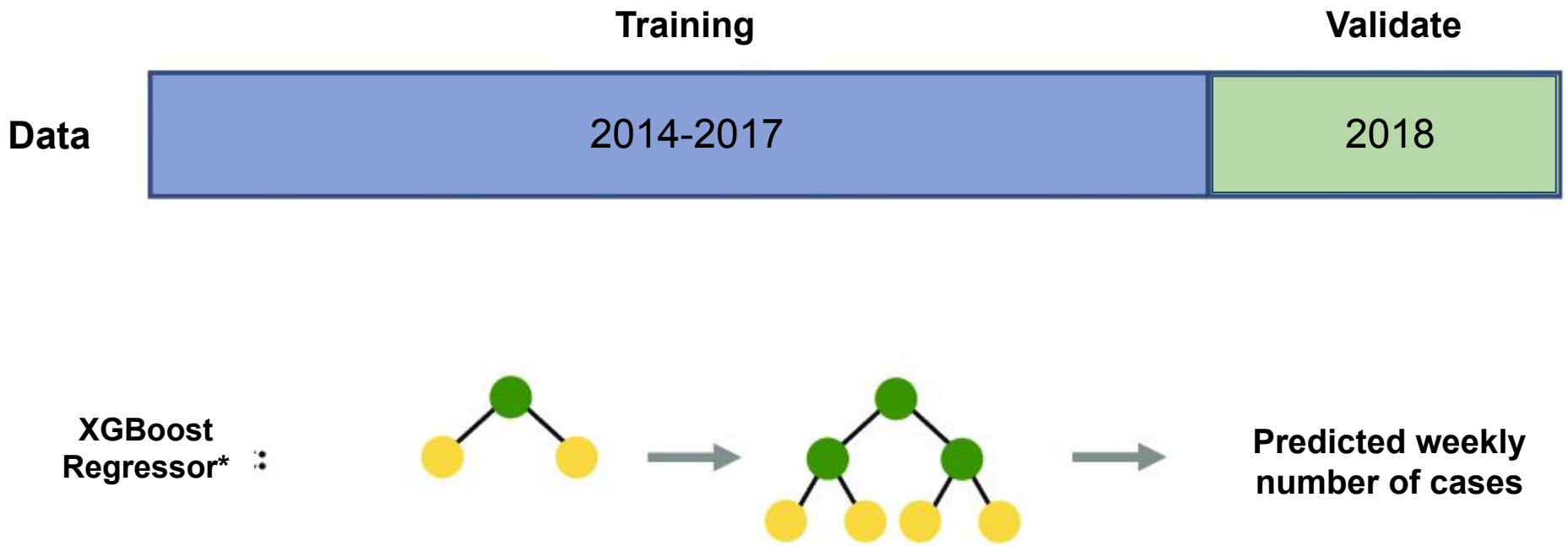
Dataset (Weekly data for 2014-2018)

- Number of Dengue Cases
- Google Search Trends: Dengue Fever, Headache, Nausea, Vomiting, Muscle Ache, Rashes, Mosquito Bite
- Weather Data: Rainfall, Temperature, Wind Speed



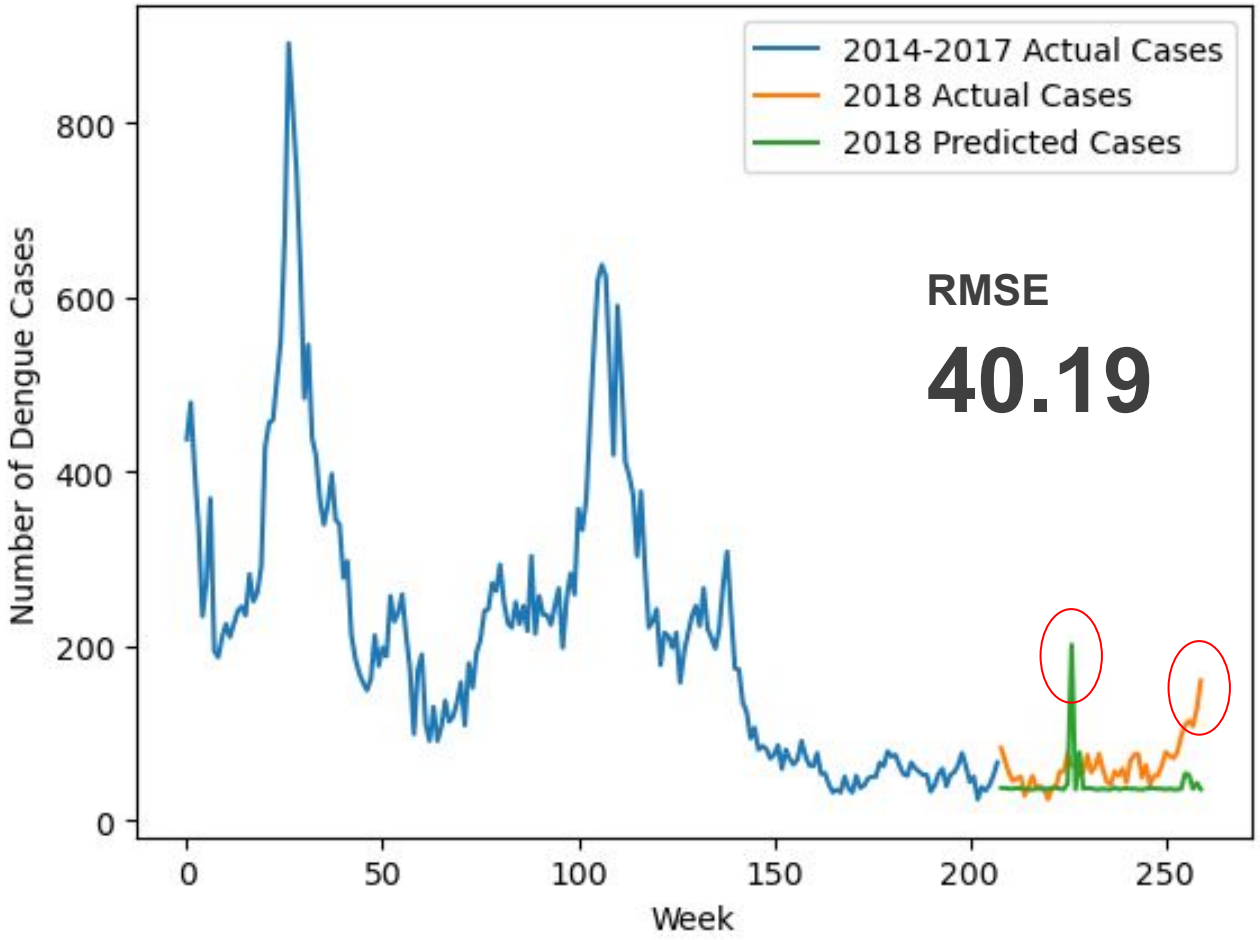


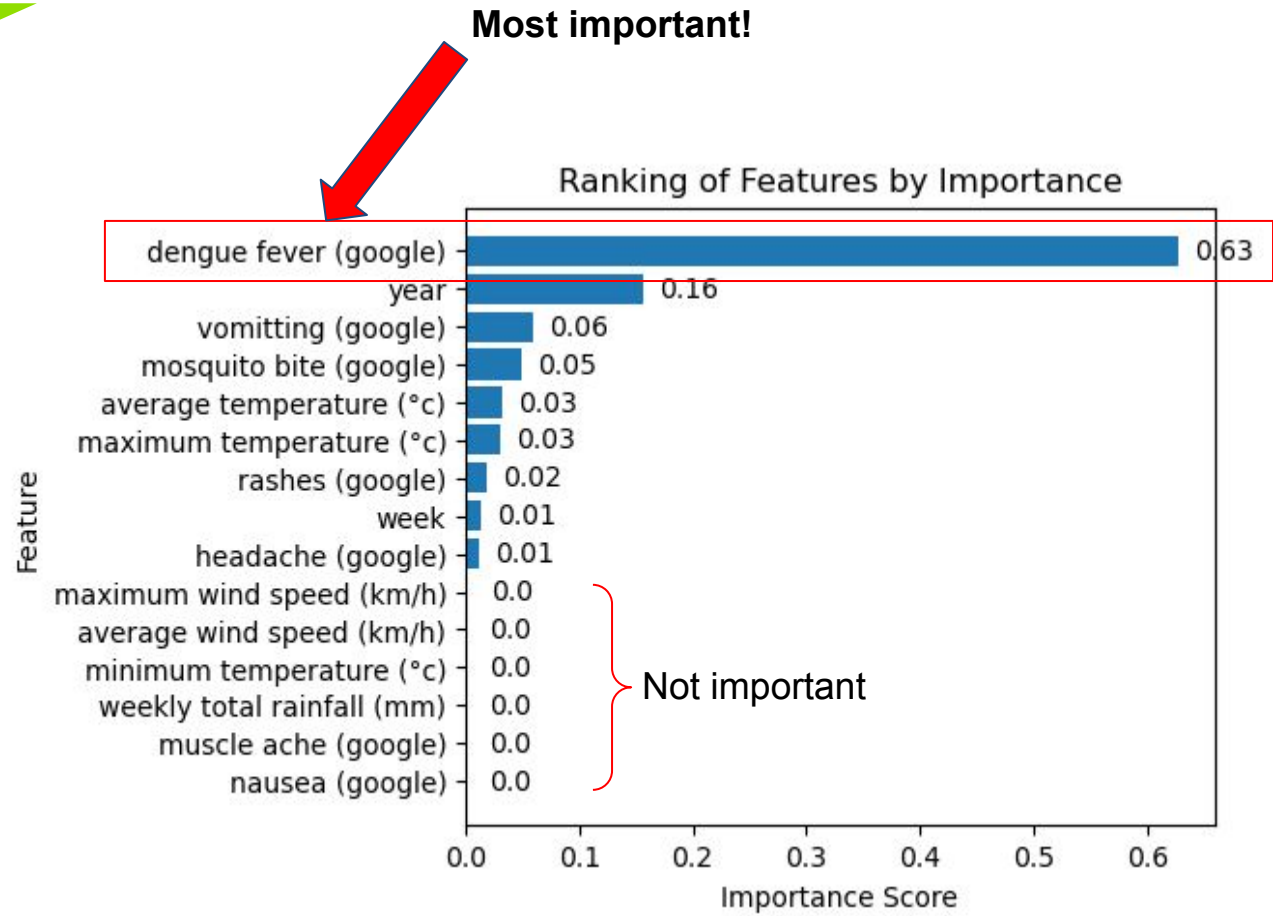
Temporal Prediction Method



* XGBoost Regressor was selected as it performed best compared against LGBM/RF/LRM.

Temporal Prediction Results



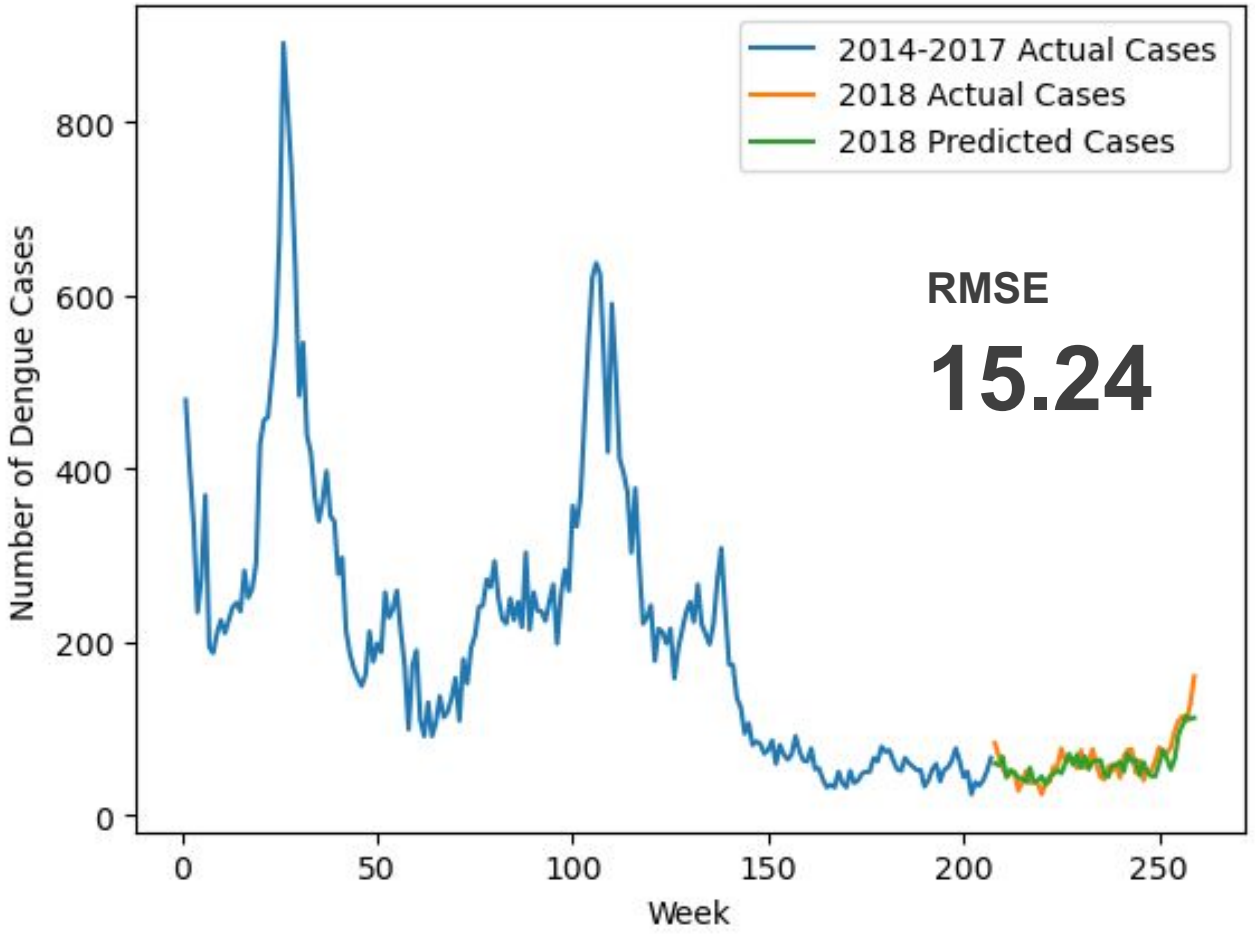


RMSE of Baseline Models

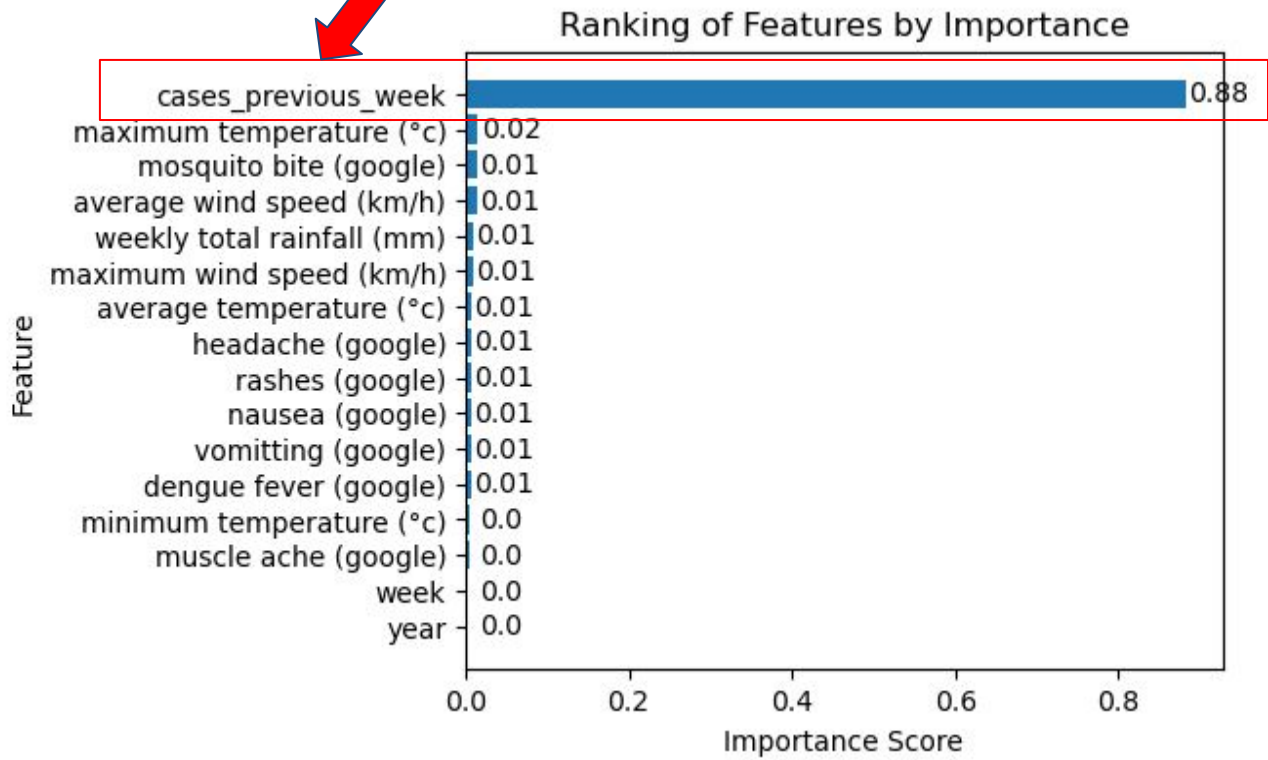
Median of previous year: 27.78

Mean of previous year: 27.32

Temporal Prediction Results (Lag Features)



Most important!

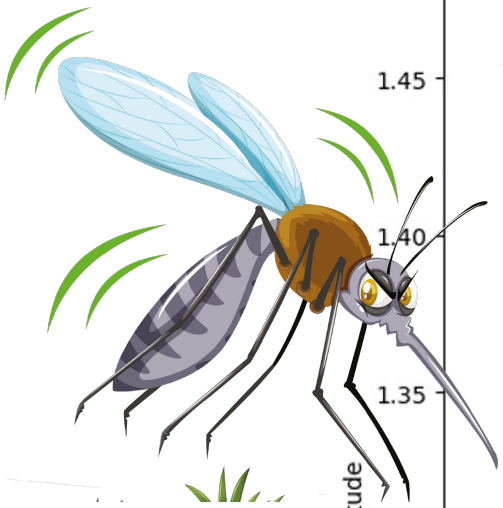


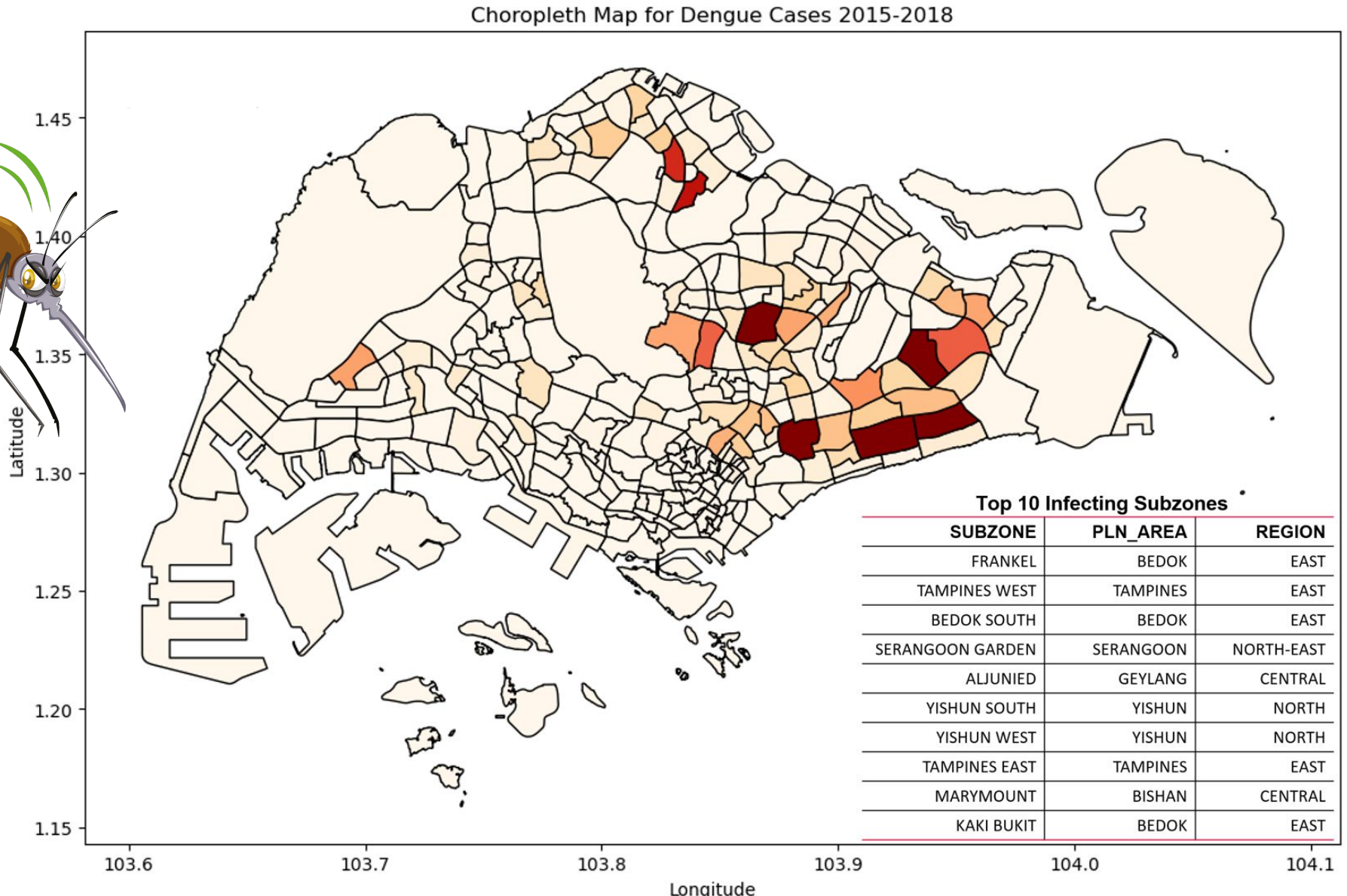
RMSE of Baseline Models

Median of previous year: 27.78

Mean of previous year: 27.32

Spatial EDA





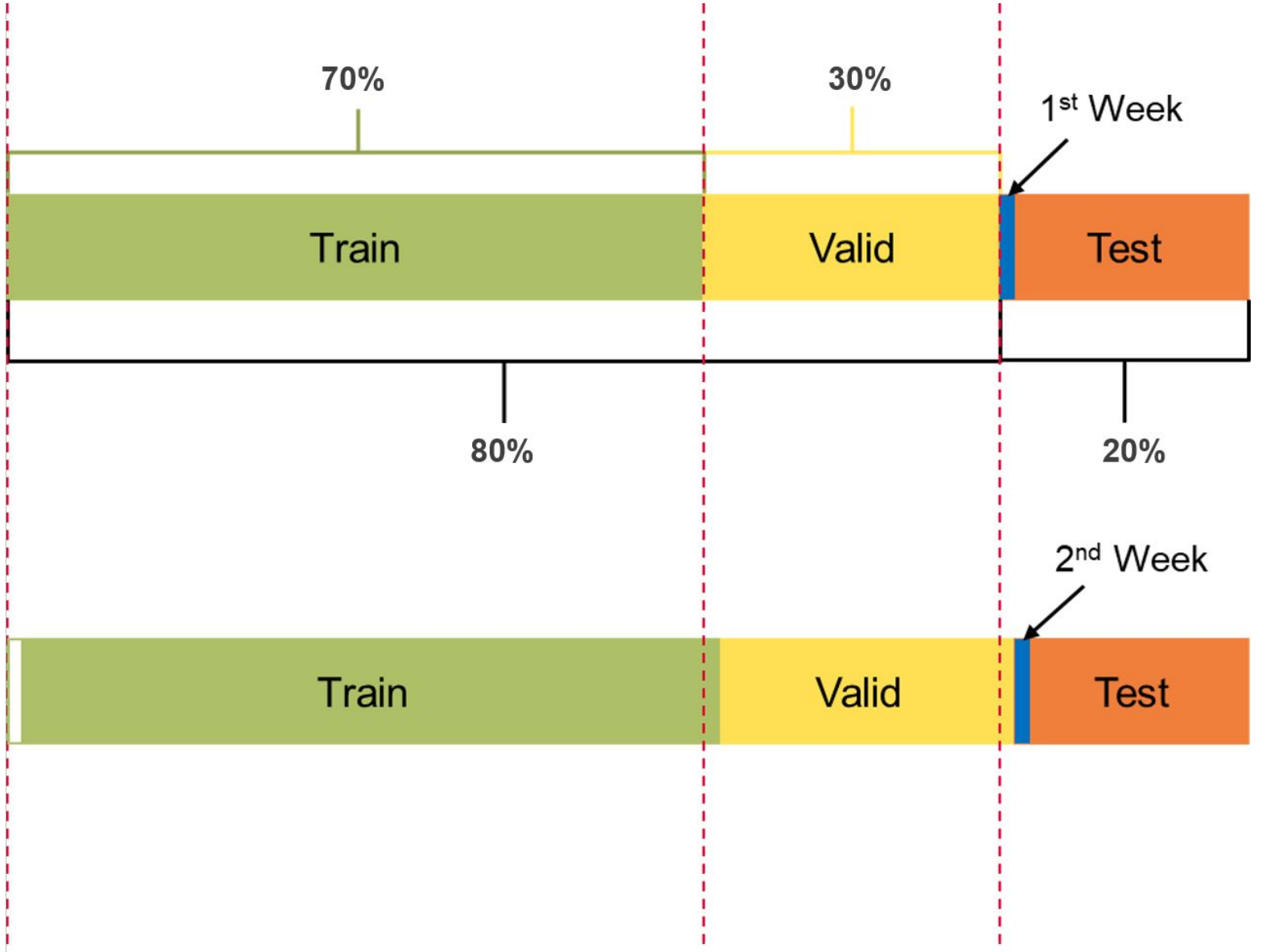
Top 10 Infecting Subzones

SUBZONE	PLN_AREA	REGION
FRANKEL	BEDOK	EAST
TAMPINES WEST	TAMPINES	EAST
BEDOK SOUTH	BEDOK	EAST
SERANGOON GARDEN	SERANGOON	NORTH-EAST
ALJUNIED	GEYLANG	CENTRAL
YISHUN SOUTH	YISHUN	NORTH
YISHUN WEST	YISHUN	NORTH
TAMPINES EAST	TAMPINES	EAST
MARYMOUNT	BISHAN	CENTRAL
KAKI BUKIT	BEDOK	EAST

Spatial Prediction Methods

WALK FORWARD

Prediction on Weekly Basis

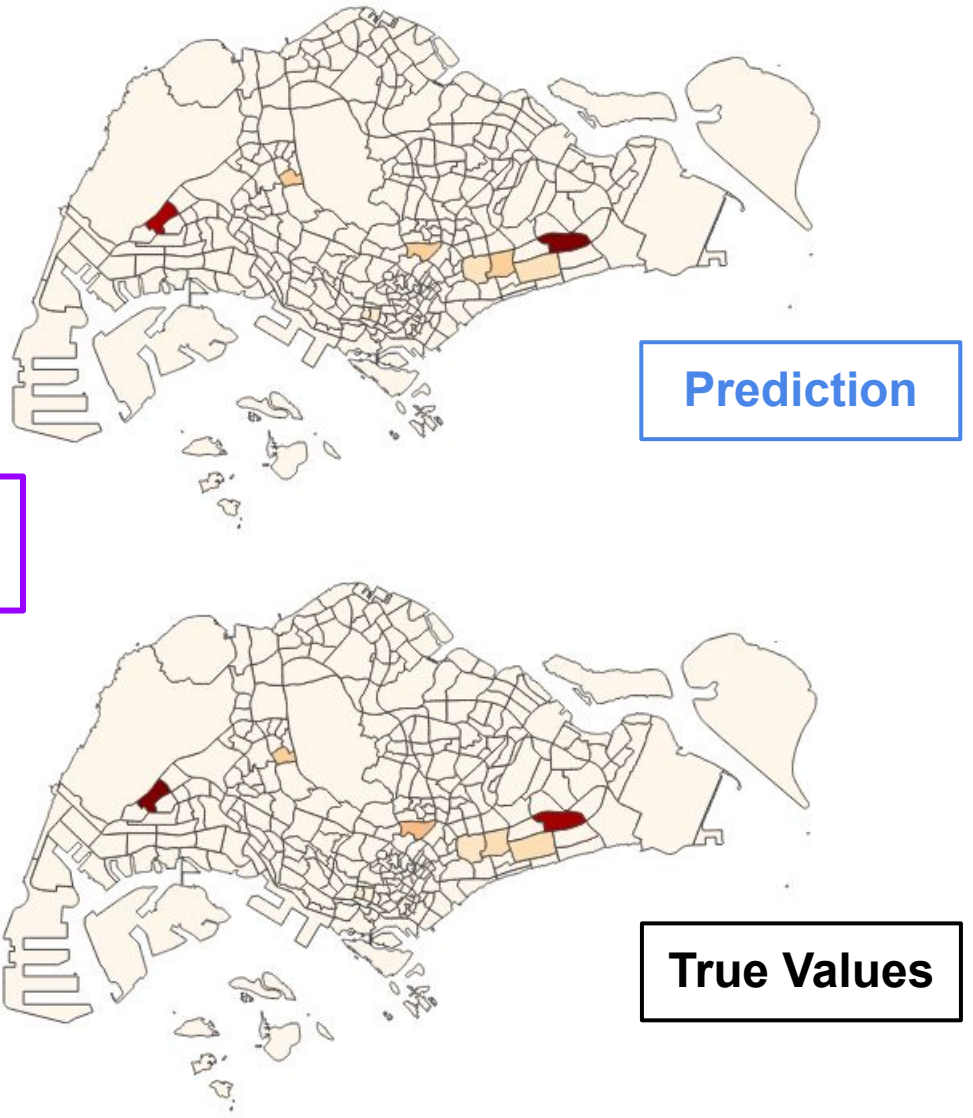


Spatial Prediction Results

First Week of Test Set Prediction vs. True Values

Number Of Cases (Pred)	Number Of Cases (True)	SUBZONE	PLN_AREA	REGION
2.000000	1.0	HENDERSON HILL	BUKIT MERAH	CENTRAL
2.000000	2.0	REDHILL	BUKIT MERAH	CENTRAL
2.767476	3.0	ALJUNIED	GEYLANG	CENTRAL
3.000000	4.0	FRANKEL	BEDOK	EAST
5.072390	4.0	GEYLANG EAST	GEYLANG	CENTRAL
16.442773	22.0	YUNNAN	JURONG WEST	WEST
18.367027	20.0	BEDOK NORTH	BEDOK	EAST
5.006327	8.0	BALESTIER	NOVENA	CENTRAL
5.263494	6.0	DAIRY FARM	BUKIT PANJANG	WEST

Valid set	Avg. SMAPE
Prediction	33.2%
Base (lag1)	35.2%



Cost-Benefit Analysis

Item	Amount	
	2014	2018
Number of dengue cases (low risk)	18306	3257
Number of dengue cases (high risk)	20	26
Cost of dengue treatment	$50 (consultation), $150 (A&E)	
Estimated medical costs	$918,300	$166,750
Number of Gravitraps	1500	50,000
Cost of Gravitrap	$20	

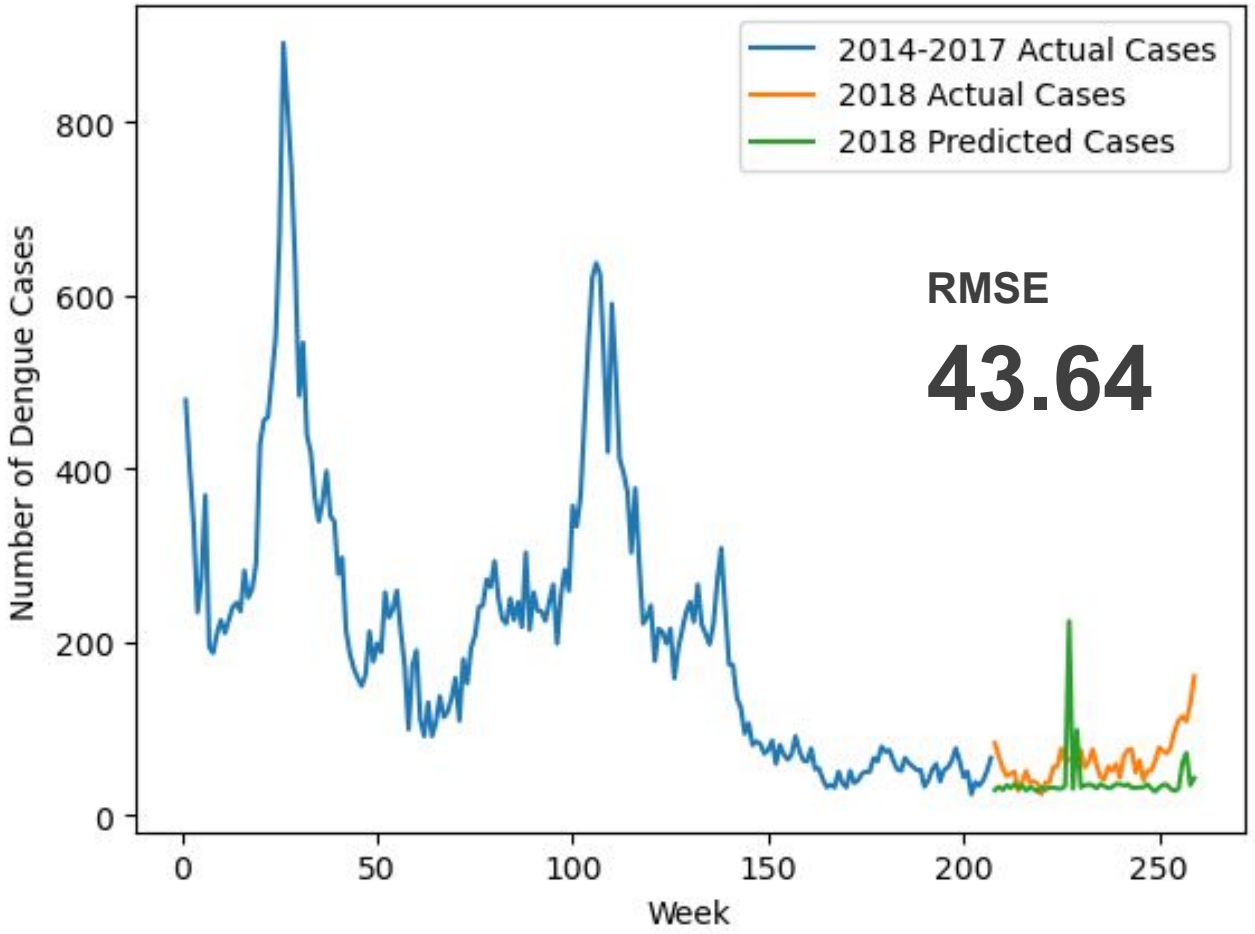
ROI for every $1 spent on Gravitrap: ~$15.50

Recommendations

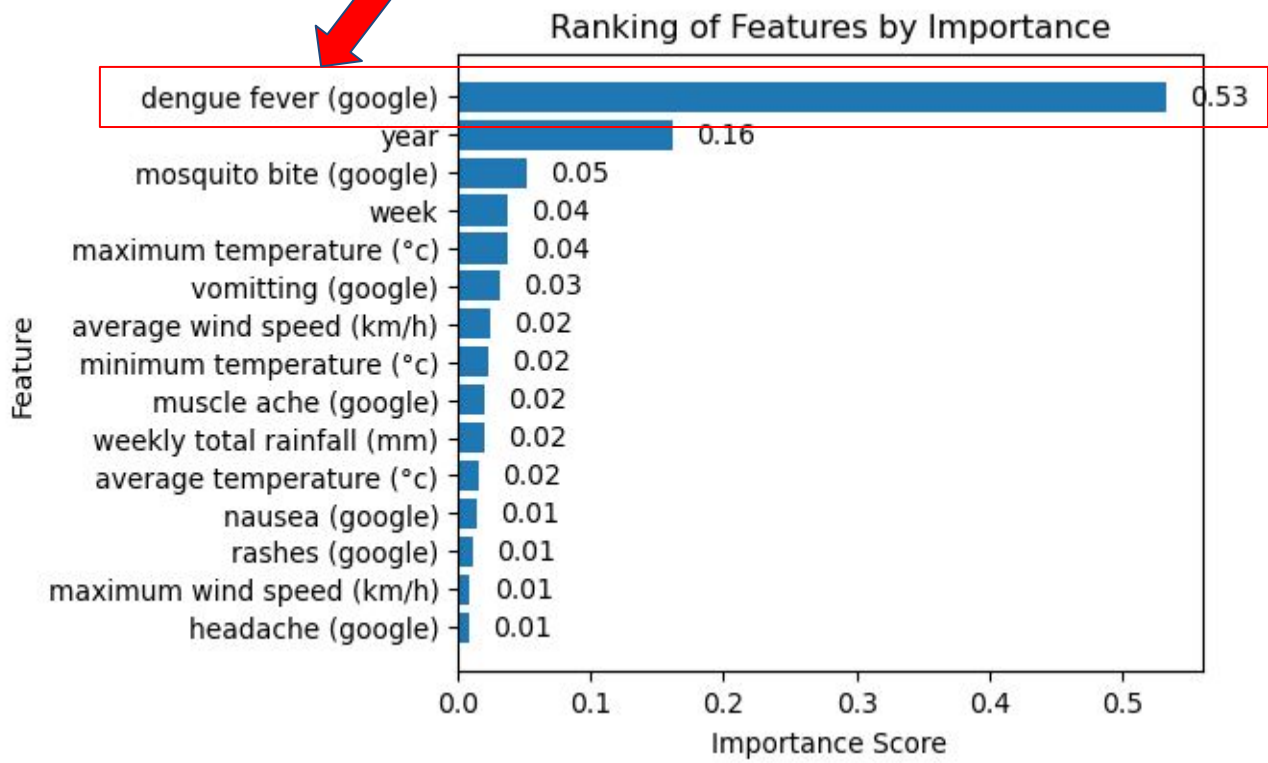
- Short term: Optimization of Gravitraps distribution to assess impact on dengue cases
- Long term: Assess feasibility of closing NEA lab for Project Wolbachia

Thank You

Temporal Prediction Results (Lag Features)



Most important!



RMSE of Baseline Models

Median of previous year: 27.78

Mean of previous year: 27.32