FORECASTING DENGUE EPIDEMICS

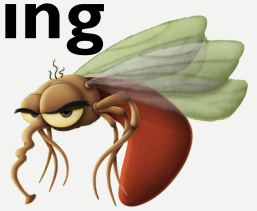
CHRISTA TAYLOR

2018

SIGNIFICANCE OF THE PROBLEM

- Dengue virus is spread by mosquitos resulting in dengue fever
- Each year approximately 50 to 528 million people are infected with the virus, 500,000 are severe life-threatening and 10 to 20 thousand people die from dengue hemorrhagic fever or dengue shock syndrome*
- Dengue affects people from all over the world; No vaccine

Accurately predicting dengue outbreaks assists in focusing efforts in prevention and save lives

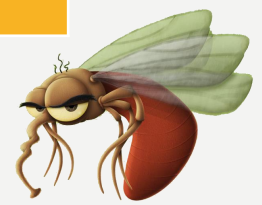


*Whitehorn & Farrar, 2010

BACKGROUND

- Dengue data and evaluation via drivendata.org
- Two locations: Iquitos, Peru and San Juan, Puerto Rico
- Predict the total cases of Dengue by year and week of the year in San Juan and Iquitos

	San Jose, Puerto Rico			Iquitos, Peru			
	Observ	Start Date	End Date	Observ	Start Date	End Date	Total
Training	936	4/30/1990	7/1/2000	520	7/1/2000	6/25/2010	1456
Test	260	4/29/2008	4/30/2013	156	7/2/2010	6/25/2013	416
Total	**1196**	**7/2/2010**		**676**			**1872**

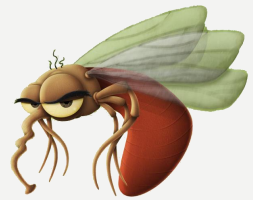


TYPES OF DATA

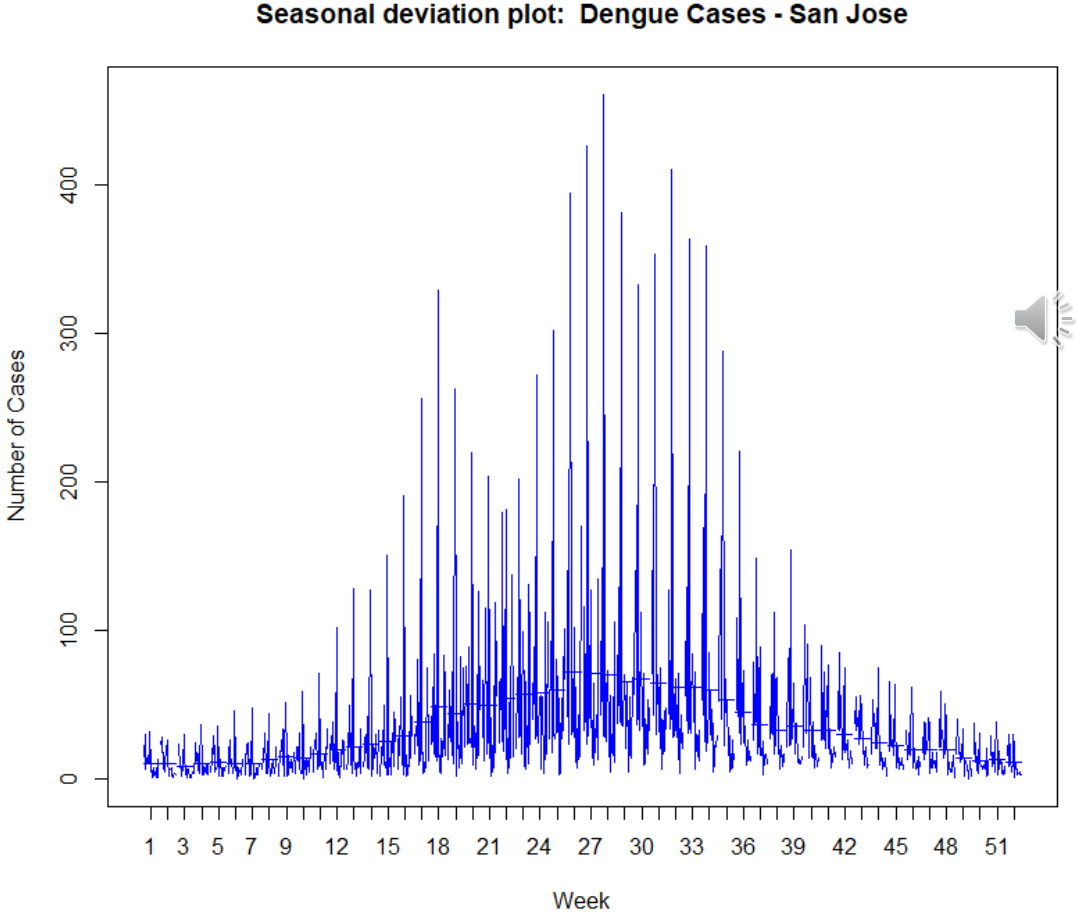
- City: San Juan and Iquitos
- Dates – Year, Week of the Year
- Geographic Coordinates
- Total Cases of Dengue

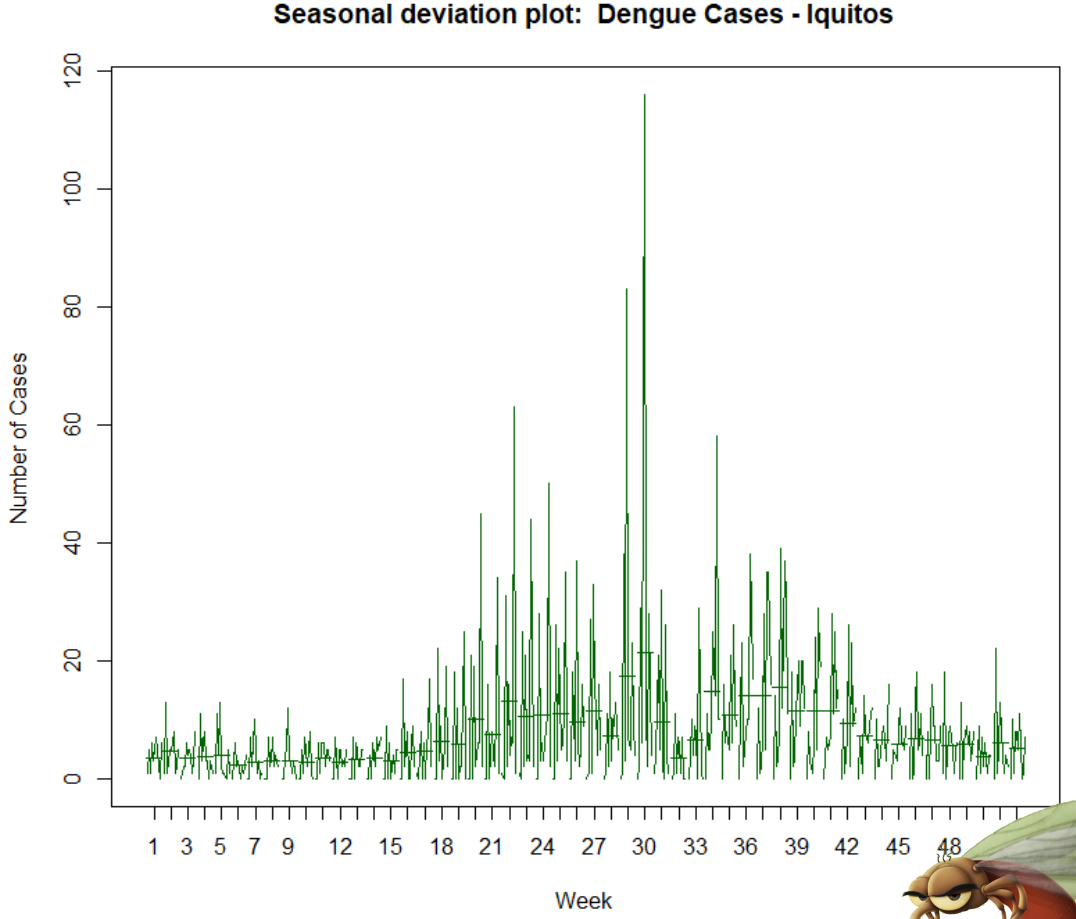
Meteorology:

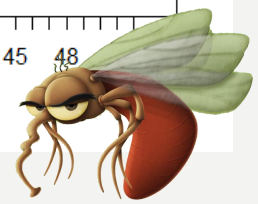
- Precipitation
- Air Temperature – Max., Min., Avg.
- Dew Point
- Relative Humidity
- Diurnal Temperature



SEASONAL DEVIATION

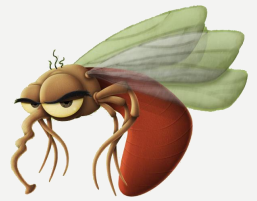






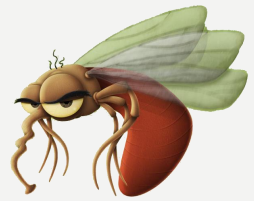
MODEL EVALUATION

- Forecasts evaluated for each target using two metrics:
- Point forecasts using Relative Mean Absolute Error (MAE) to assess performance relative to a seasonal autoregressive model and to other forecasts
- MAE and logarithmic score calculated across all seasons and forecasted weeks of the season along specific seasons to identify potential differences in model strengths



INTERESTING SIDETRACKS

- Missing data: Equivalent Temperatures Kelvin = Celsius temperature plus 273.15°
- "Aegypti' optimal temperature of 22° degrees to 32°
- Mosquito Grow Index $= \frac{Station\ Max\ Temp + Station\ Min\ Temp}{2} - 20°$
- Base temperature (20°) may not be ideal but was used as a starting point based on with the 16 degrees was considered less optimal



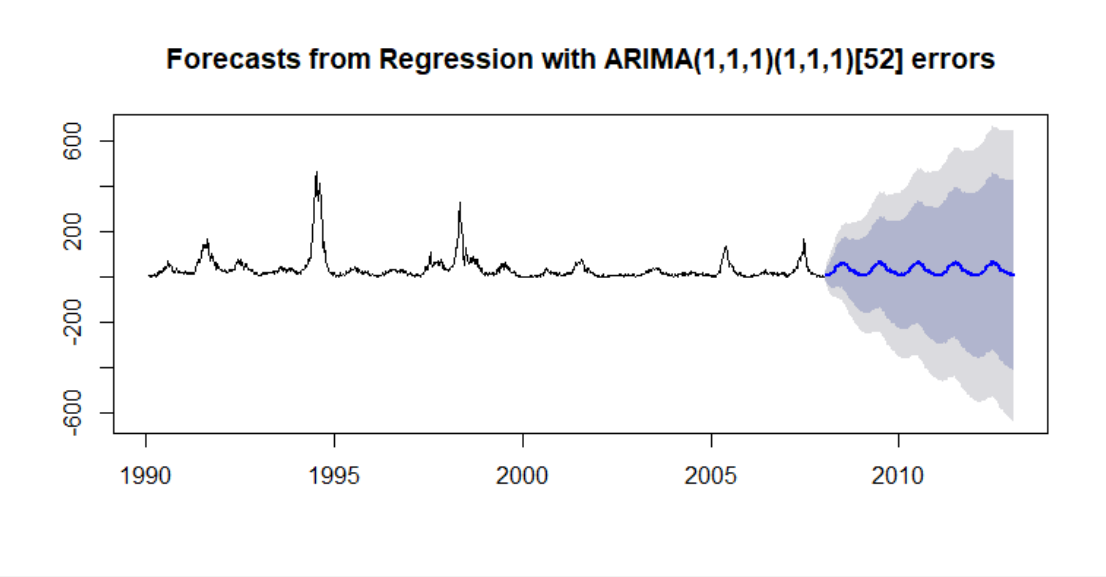
LITERATURE

- *Climate Variability, Social and Environmental Factors, and Ross River Virus Transmission: Research Development and Future Research Needs:* Poisson, logistic, ARIMA and SARIMA
- *Time series* analysis *of malaria in Afghanistan: using ARIMA models to predict future trends in incidence:* ARIMA models
- *Weather variables and the El Niño Southern Oscillation may drive the epidemics of dengue in Guangdong Province - GAM, Random forest and Wavelet*
- *Spatial and Temporal Dynamics of Dengue Fever in Peru: 1994-2006:* Stepwise regression
- *Identification of the prediction model for dengue incidence in Can Tho city, a Mekong Delta area in Vietnam:* SARIMA w/ a Box and Jenkins and Poisson

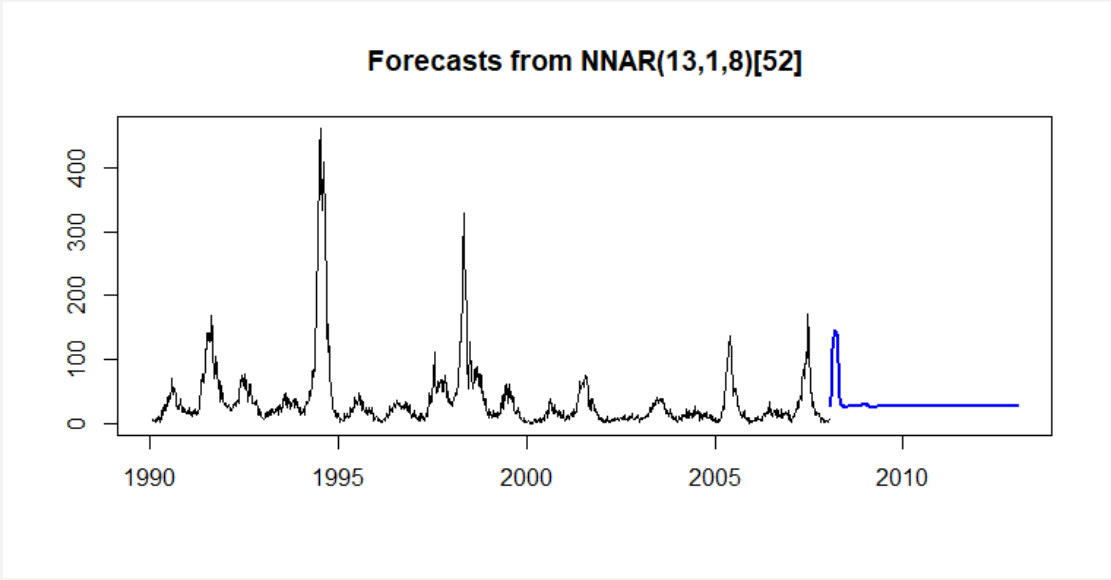
MODELS

- Seasonal ARIMA: SJ(1,1,1) IQ-(0,1,1)(1,1,1)
- ARIMA w/ Regression Terms: SJ- (1,1,1), IQ-(3,1,0)(1,1,1)
- Exponential Smoothing – Holt Winters
- Neural Networks: h=(10, 13, 15, 18, 20, 25, 30)

San Jose



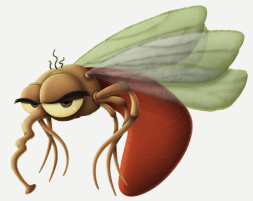
Iquotos



RESULTS

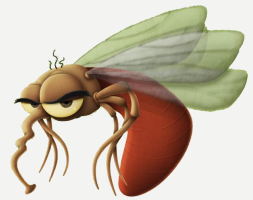
Snippet of results:

Model Reference	Score	
Arima w/ Reg Terms	25.4038	Best
Neural Network	25.7236	
Arima w/ created var	25.9111	
SARIMA	26.0361	Worst



LIMITATIONS & FUTURE ITERATIONS

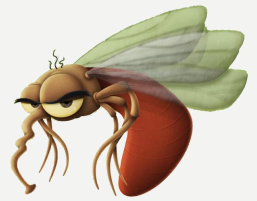
- Inability to utilize mapping coordinates
- Meteorology background - humidity and temperature interaction
- Living population and turnover of mosquitos
- Precipitation, temperature, evaporation variable creation
- Still water considerations
- Mosquito lifecycle and the interplay of weather conditions



LEARNING

Interesting from numerous perspectives.:

1. "Out of the Box" models were very effective at optimizing the various seasonal components
2. Geographic coordinates was a new challenge
3. Understanding variables and related knowledge is beneficial i.e. interaction between meteorology and mosquito lifecycle
4. Second modelling contest – very additive




MOSQUITO
BITE
Dengue