Title: Data Fusion and Disease Forecasting for Dengue in Brazil

Author(s): Martinez, Kaitlyn Michelle
Manore, Carrie Anna
Del Valle, Sara Y.
Fairchild, Geoffrey
Ziemann, Amanda
Parikh, Nidhi K.
Castro, Lauren Ann
Kempfert, Katherine Candice
Romero-Alvarez, Daniel Alejandro
Generous, Eric Nicholas Alexandre

Intended for: Meeting with Potential External Collaborator

Issued: 2021-11-17

Data Fusion and Disease Forecasting for Dengue in Brazil

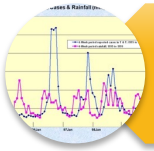
Katy Martinez

Sara Del Valle, Carrie Manore, Geoff Fairchild, Amanda Ziemann, Nick Generous, Nidhi Parikh, Katherine Kempfert, Daniel Romero Alvarez, Lauren Castro

National Nuclear Security Administration

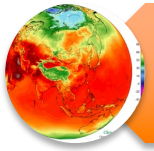
Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Data Fusion in Dengue Models for Brazil



Historical Clinical Surveillance Data

Satellite Imagery

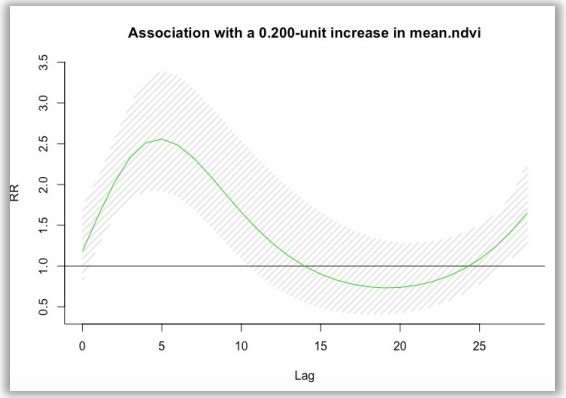


Climatological Data

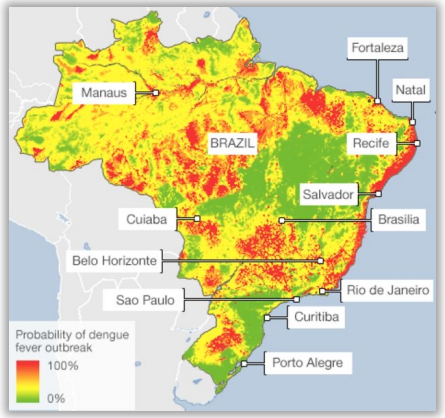
Demographic Data

Google Search Queries

+



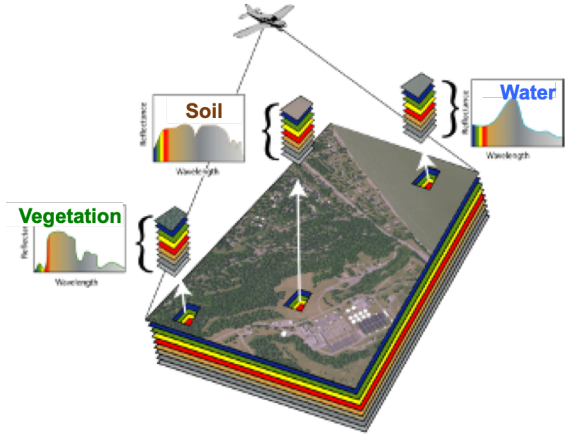
Predictive Models



Dengue Risk Map & Forecasts

Data Fusion in Dengue Models for Brazil

Satellite Imagery

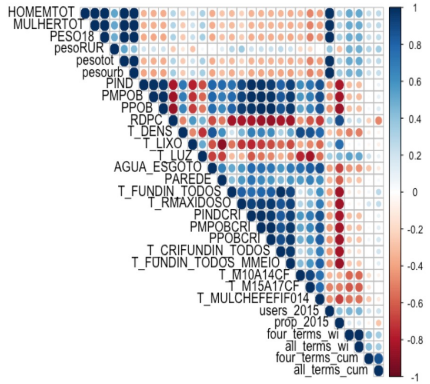


- 4 satellites (Landsat 5, 7, 8 & Sentinel 2)
- 4 indices (vegetation, water content in leaves & bodies, burn)
- 5,570 municipalities across a 7-year period, per week

Weather Stations

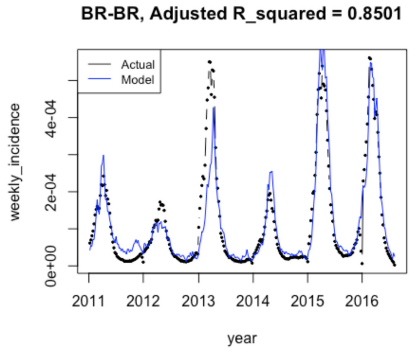
- Global Surface Summary of the Day
- Daily temperature & precipitation
- 613 weather stations covering 8.5 million km^2

Demographic



- 2010 Census data
- Identified factors correlated with dengue cases
- 200 variables — rural population, inadequate sewage, poverty, garbage collection

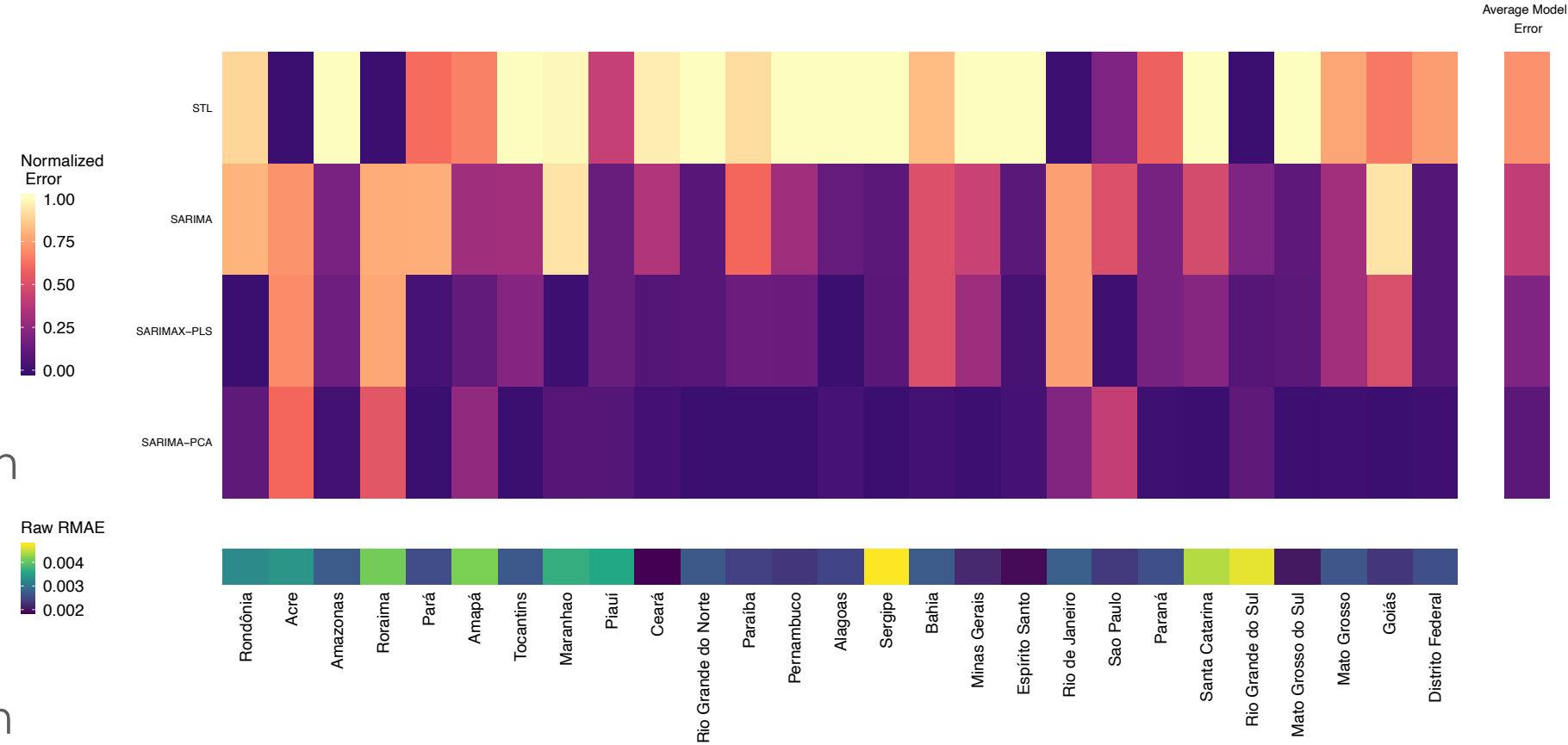
Google Queries



- Google Health Trends
- Daily time series at various spatial resolutions
- Keywords associated with dengue such as
  - Dengue
  - Dengue sintomas
  - Aedes
  - Mosquito

Now-casting Models

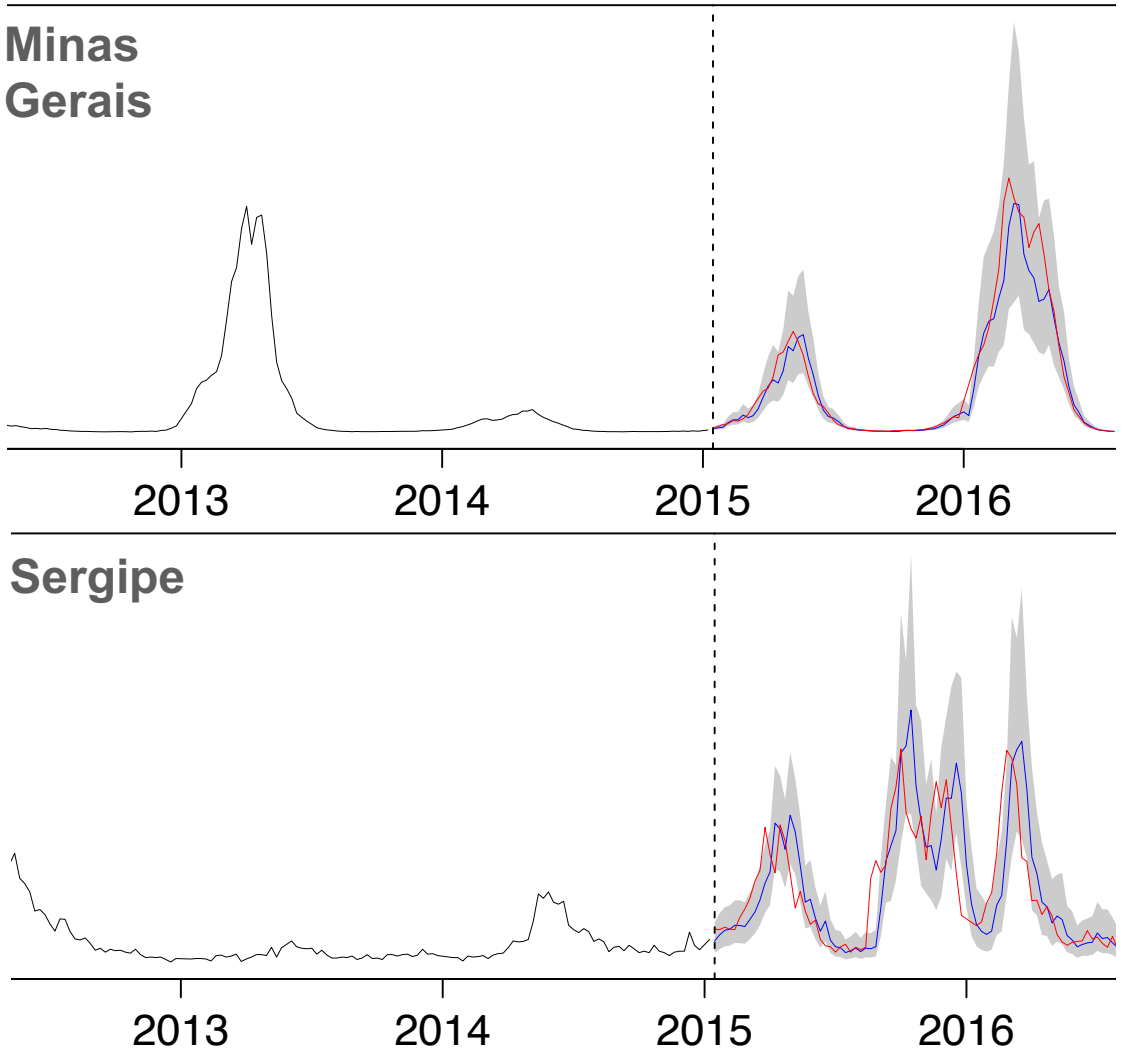
- **STL**: Multiplicative and additive seasonal trend decomposition
- **SARIMA:** Seasonal autoregressive integrated moving average
- **SARIMAX-PLS:** SARIMA with variables transformed through partial least squares
- **SARIMA-PCA:** SARIMA with variables transformed through principal component analysis (PCA)

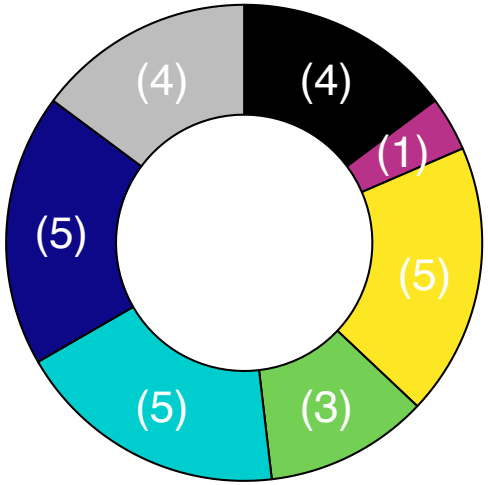


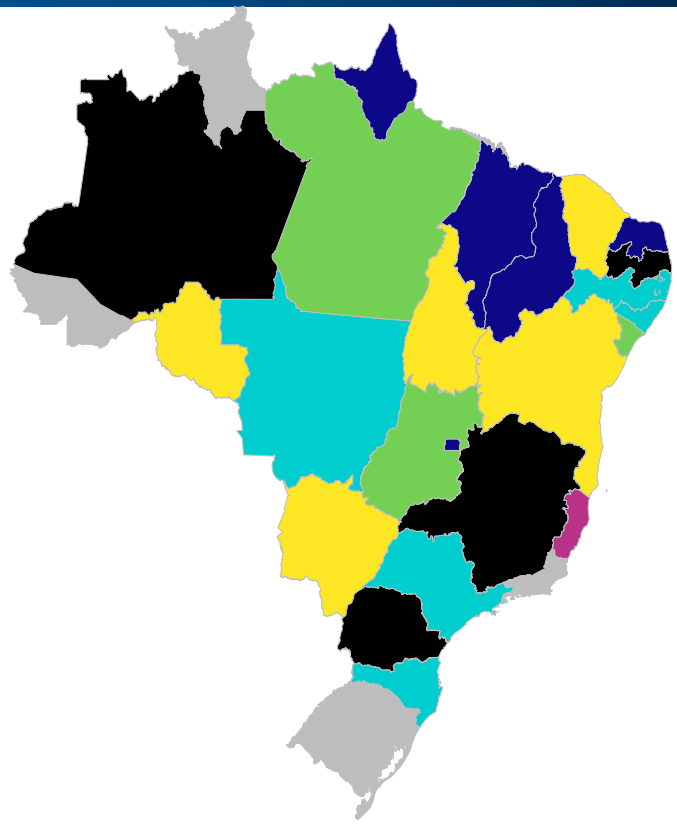
Four Models were developed for each unit of space

→ Kempfert et al., Time Series Methods and Ensemble Models to Nowcast Dengue at the State Level in Brazil. Arxiv 2020.

Heterogeneity of Data Streams Utility





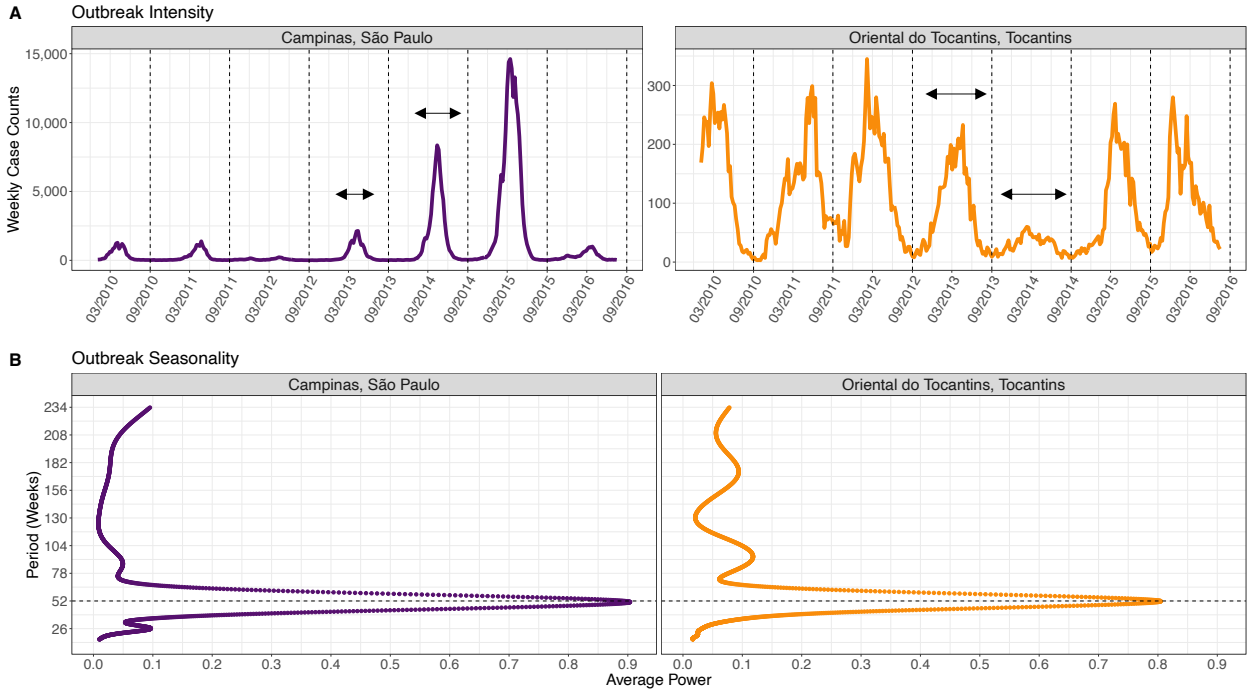


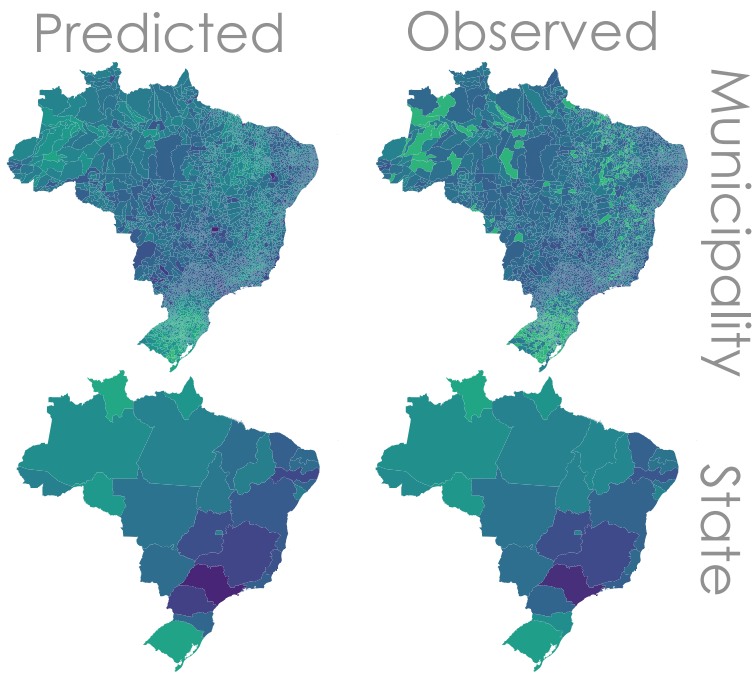
None
Satellite
Satellite + Google
Google
Google + Weather
Satellite + Weather
All

Red – Observed Dengue ••• Blue – Predicted Dengue ••• Grey – 95% Prediction Intervals

Additional Analysis and Project Directions

- Currently further exploration of dengue outbreak dynamics and predictive variables is helping to build predictive models at finer spatial scales.
- Also building integrated climate, mosquito and disease models.





→ Castro et al., Using heterogeneous data to identify signatures of dengue outbreaks at fine spatio-temporal scales across Brazil. PLOS 2021.