

Evaluating dengue forecasting models to predict Zika and Chikungunya in Brazil.

Flávio Codeço Coelho¹✉, Elisa Mussumeci¹, Marcelo Orgler¹, Linneu Holanda¹
¹Fundação Getulio Vargas, Brazil; Correspondence to: fccoelho@fgv.br

Poster DOI: ADD here from figshare
github.com/AlertaDengue — @fccoelho

The Mosquito *Aedes aegypti* is a vector for multiple viruses around the globe. Its distribution is restricted to tropical and subtropical climates, which reflects its sensitivity to temperature, humidity and other weather constraints. The modulation of its life cycle by climate, shapes the seasonality of the diseases it transmits. In Brazil, *Aedes aegypti* has been mainly associated with the transmission of dengue, making it a marked seasonal disease. In recent years, *A. aegypti* has also been notably responsible for epidemics of the Zika and Chikungunya virus. In this paper we explore the performance of dengue forecast models trained on the longer available incidence timeseries to predict the weekly incidence of Zika and Chikungunya as well. We will use a LSTM (long short term memory) recursive neural network model, which we have shown, in a previous work, to yield accurate forecasts for weekly dengue incidence. We will also compare it to a Random Quantile Forest model. Climate variables such as temperature, humidity, and atmospheric pressure are also used as predictors. A spatial component built from the incidence at neighboring cities is also included. We present results of the forecast of total incidence of arboviral disease as well as of each disease separately and discuss the relative performances of the model for each of these tasks.

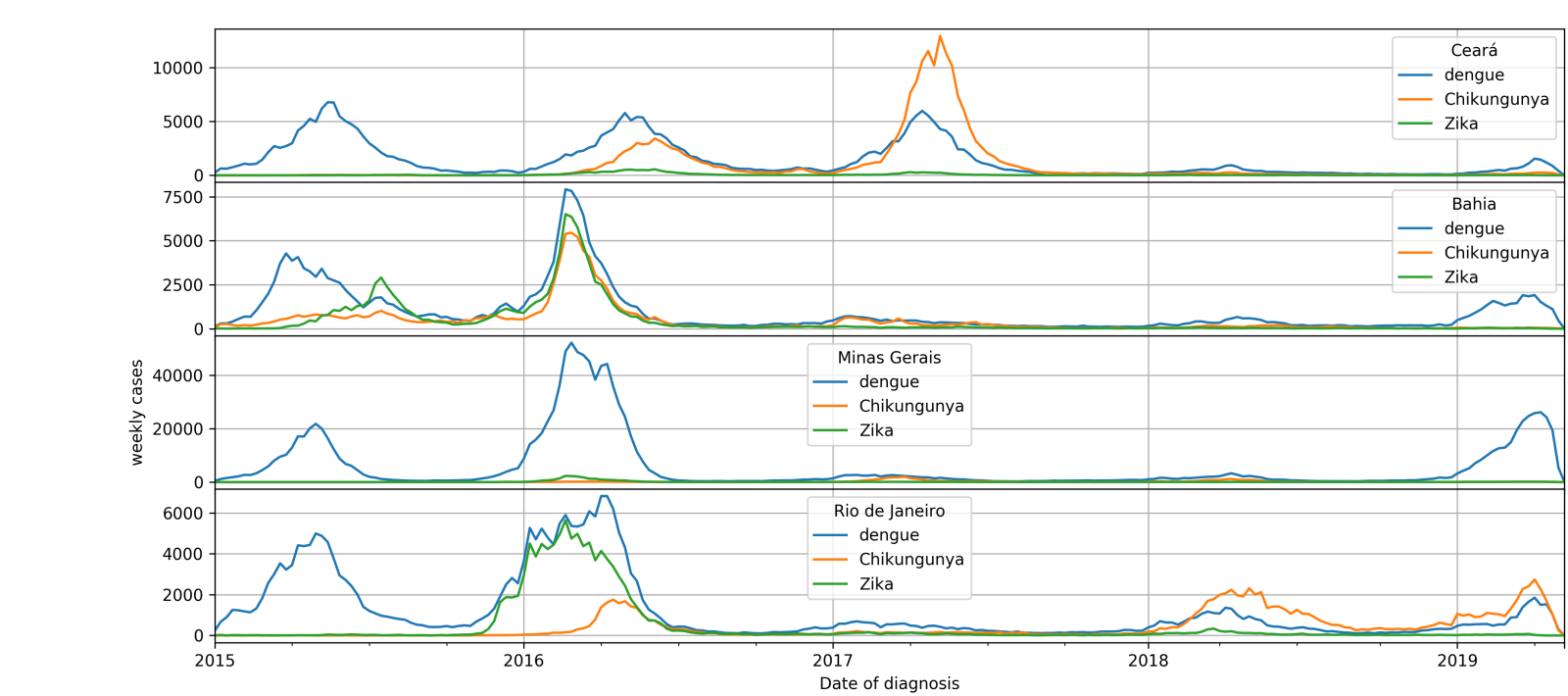


Figure 1: Weekly incidence for the state of Ceará, Brazil. The three arboviroses incidences are correlated. The data used in this work comes from the Infodengue project, which monitors these arboviroses in Brazil. In the figure 1 we can see the incidence series for states in Brazil which had significant outbreaks of the 3 arboviroses since 2016 when Chikungunya and Zika arrived in Brazil.

Three diseases, one vector.

Given their similar transmission mechanisms, these diseases present similar seasonal patterns. Being seasonal diseases with serious consequences, being able to forecast their incidence is very important for public health authorities. Among these, Dengue is the one with the longest historic records. Therefore its worth investigating how usefull our knowledge about dengue dynamics is to predict other arboviroses, in this case Zika and Chikungunya. Here we will revisit a forecast model developed by the authors for dengue, and evaluate its performance in predicting the other two arboviroses.

Dengue forecast models

LSTM

A LSTM model is a recurrent deep neural network model developed to handle predictions of timeseries. We used a LSTM model with 3 LSTM units followed by 3 dropout layers. The model was trained for 300 epochs using a mean-log squared-error (MLSE) loss function and a Nesterov Adam optimizer[2]. A look back of 4 weeks and a forecasting window of 4 weeks were used.



Figure 2: Multi-colour channel registration with NanoJ-Core. **a)** Composite image of multi-colour beads, prior to (left) and after (right) channel registration using NanoJ-Core. Scale bars: 25 μm , insets: 0.5 μm . **b)** Vectorial representation of the shift between the two channels (left, displacement vector length 50 times larger for representation purposes), horizontal (middle) and vertical (right) shift maps obtained and applied to the data shown in a). Scale bars: 25 μm .

Random Quantile Forest (RQF)

Random Forest models calculate an ensemble of regression trees from random subsets of data. RQF model are an extension to regular random forests in which the full conditional distribution of Y given $X = x$ is calculated. As a result, it is a non-parametric, consistent and accurate way to determine conditional quantiles from high-dimensional predictors[1].

Results

Both models were fitted to available data yielding...



Figure 3: Live-cell super-resolution microscopy with NanoJ-SRRF. **a)** Comparison of widefield and SRRF reconstruction from UtrCH-GFP actin labelling. Scale bar: 5 μm . **b)** Time-course of the inset shown in a), obtained at 33.3 Hz and displayed every 30 s. Scale bar: 1 μm . **c)** Colour-coded time course. Scale bar: 1 μm .

Discussion



Figure 4: Quality and resolution assessment with NanoJ-SQUIRREL. **a)** A super-resolution rendering and acquired widefield image of fixed Alexa647 labelled microtubules. **b)** Left: SQUIRREL error map highlighting discrepancies between the super-resolution and diffraction-limited images in (a). Right: Magnified insets at indicated positions on error map. **c)** Left: SQUIRREL resolution map of the super-resolution image in (a). Right: Magnified insets for indicated resolution blocks. Whole image scale bars = 5 μm , inset scale bars = 1 μm .



Conclusion

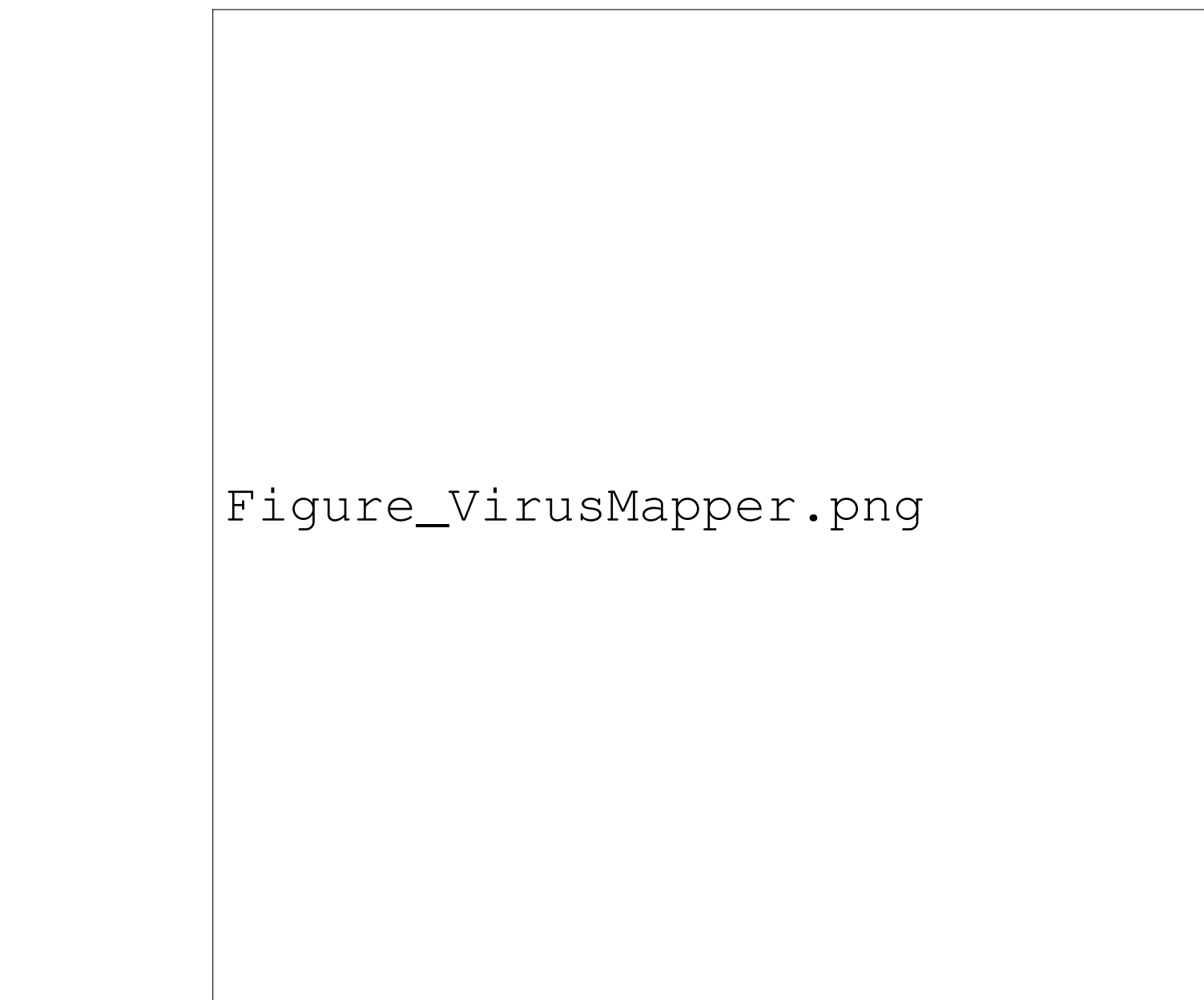


Figure 5: Quantitative SPA-based modelling. Multicomponent model of the Vaccinia virus by imaging in super-resolution hundreds of fluorescently labelled viruses and modelling their structure through VirusMapper



Figure 6: Schematics of the NanoJ-Fluidics system. **a)** 3D side view of a single syringe pump. **b)** 2D top view of a syringe pump array (representing 4 pumps out of 128 maximum) and a fluid extraction peristaltic pump, both controlled by an Arduino UNO. **c)** Example of possible workflows

References

[1] Nicolai Meinshausen. Quantile regression forests. *Journal of Machine Learning Research*, 7(Jun):983–999, 2006.

[2] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton. On the importance of initialization and momentum in deep learning. In *International conference on machine learning*, pages 1139–1147, 2013.

Funded by

