LA-UR-18-27438
Approved for public release; distribution is unlimited.

Title: Forecasting Dengue in Brazil with Time Series Modeling in Parallel.

Author(s): Gafur, Jamil
Kempfert, Katherine Candice

Intended for: Parallel Computing Summer School end of summer lighting talk

Issued: 2018-08-06

Forecasting Dengue in Brazil with Time Series Modeling in Parallel

Jamil Gafur

Katherine C. Kempfert

Motivation

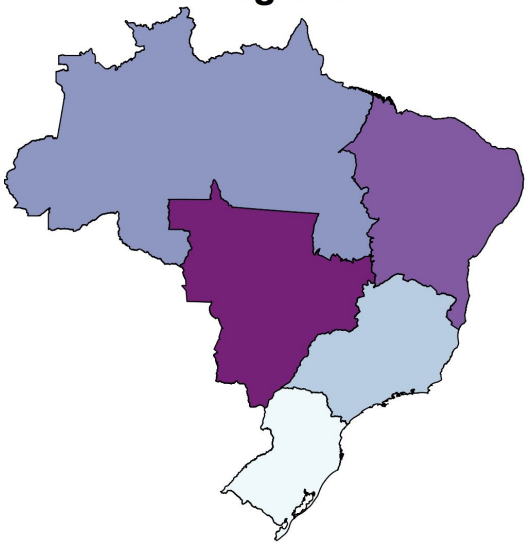
- Dengue is a mosquito-borne virus that causes 390 million infections per year, straining public health and the economy (Bhatt et al., 2013)
- Early detection can significantly mitigate outbreaks
- We develop a large-scale, parallel workflow to predict the spread of dengue in real-time

Spatial Levels of Brazil

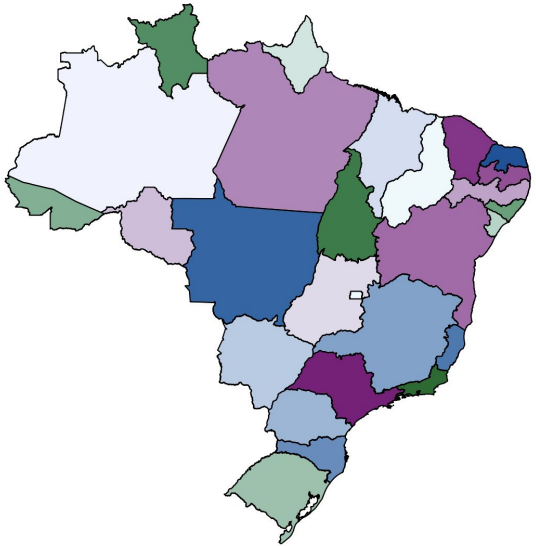
For each level, we built *k* independent time series models, one per unit of space.

This is parallel!

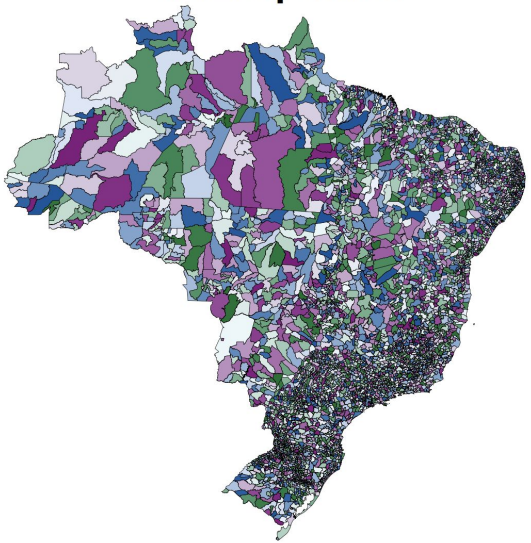
Regions



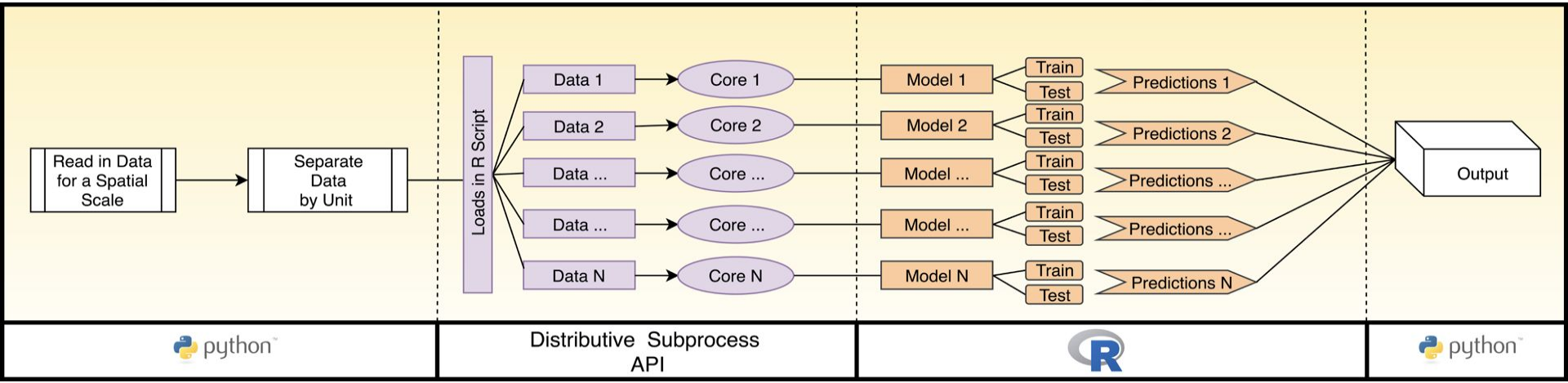
States



Municipalities

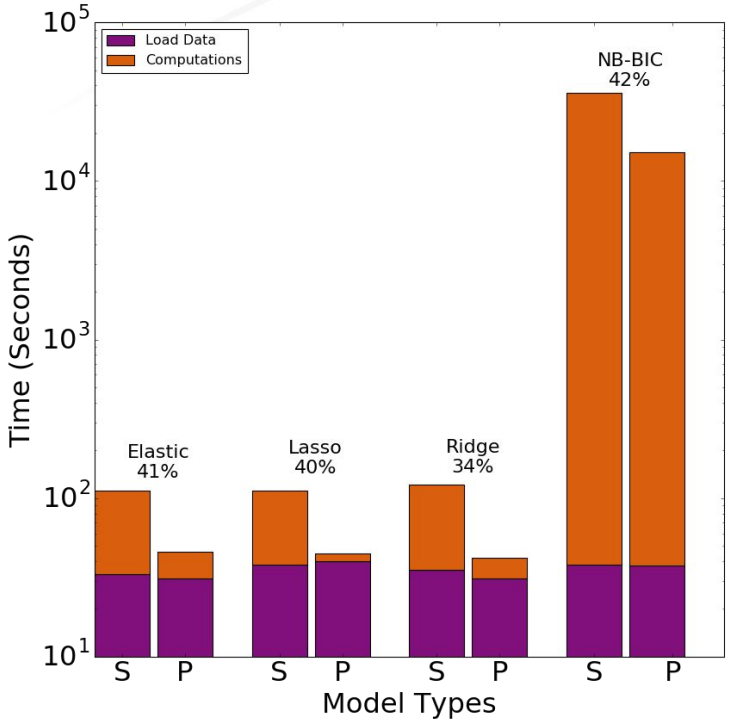


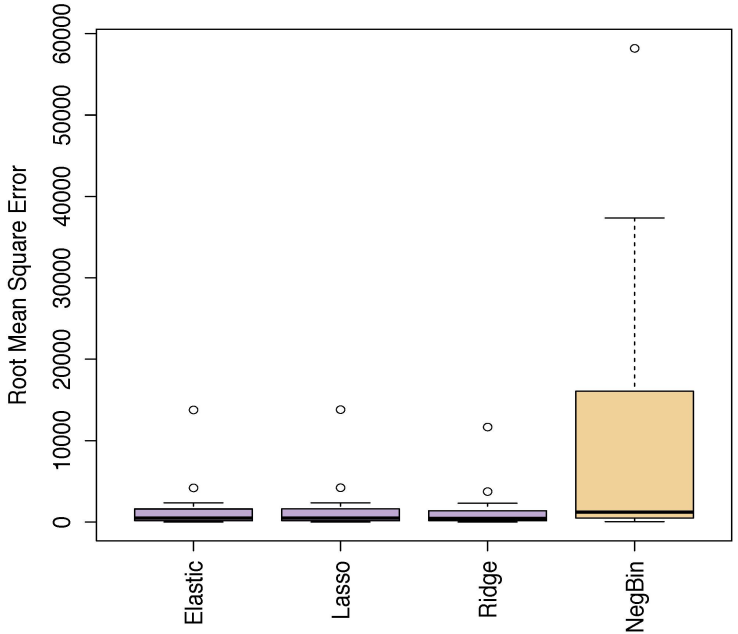
Multi-Language Parallel Workflow



Regularized regression models (lasso, ridge, and elastic net) and the **negative binomial** generalized linear model (GLM) are used (Friedman et al., 2009), (Wood et al., 2015).

Time Series Modeling Speed Up Using API



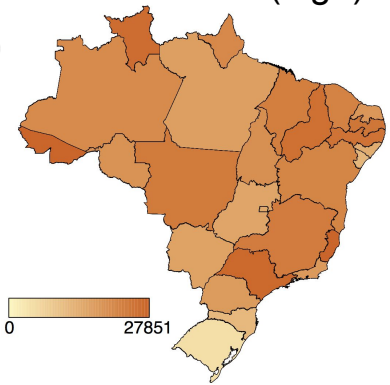


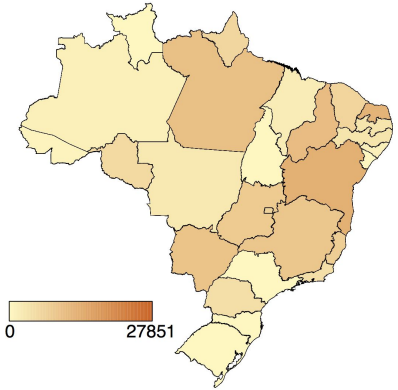
Predictions at State Level

Week 319 (high)

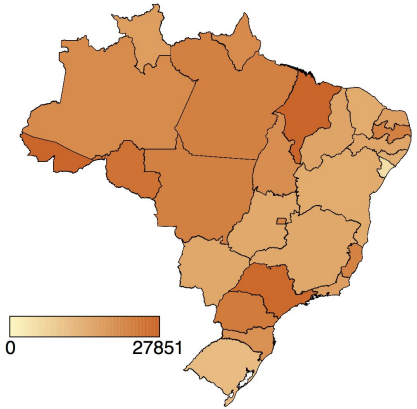
Week 344 (low)

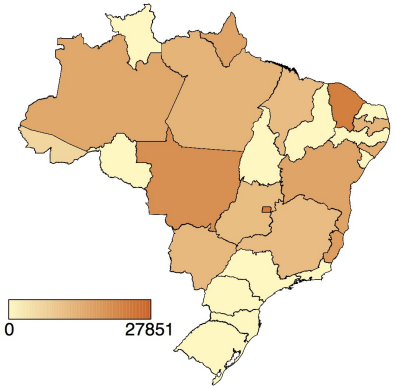
True Dengue





Predicted Dengue





Forward

- We achieve high predictive accuracy for the states while decreasing runtime by up to 42%
  - Repeat analysis for all spatial levels and make comparisons
  - Enable easier user implementation of API while decreasing runtime

Acknowledgments

- We would like to thank our mentors Carrie Manore and Geoffrey Fairchild and the Parallel Computing co-leads Hai Ah Nam, Bob Robey, and Kris Garrett.
- Computations conducted on Darwin cluster at Los Alamos National Laboratory. Support for this work was provided by U.S. Department of Energy at Los Alamos National Laboratory supported by Contract No. DE-AC52-06NA25396

References

- Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., ... & Myers, M. F. (2013). The global distribution and burden of dengue. *Nature*, 496(7446), 504.
- Friedman, J., Hastie, T., & Tibshirani, R. (2009). glmnet: Lasso and elastic-net regularized generalized linear models. *R package version*, 1(4).
- Wood, S., & Wood, M. S. (2015). Package `mgcv' *R package version*, 1, 29.