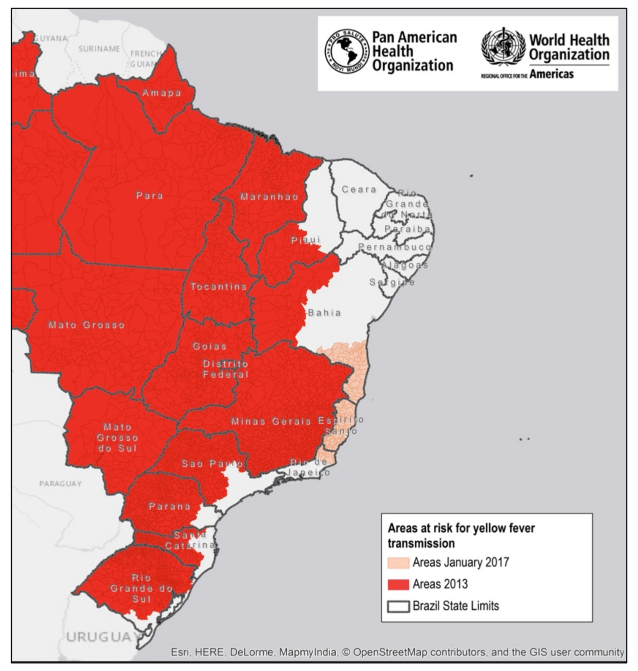
PREDICTING SPATIO-TEMPORAL DYNAMICS OF YELLOW FEVER IN BRAZIL

Michelle V. Evans[1,2], RajReni B. Kaul[1], Courtney C. Murdock[1,2,3], John M. Drake[1,2]

[1]Odum School of Ecology, University of Georgia, Athens, GA [2]Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA [3]Department of Infectious Disease, University of Georgia, Athens, GA

YELLOW FEVER IN BRAZIL



- Yellow fever is a mosquito-borne flavivirus, causing acute viral hemorrhagic disease
- Endemic to Brazil, and has been spreading from Amazon basin towards southeastern region since the early 2000s[1]
- Brazil outbreak of 2016/2017 had more cases than previous fifteen years combined
- Many cases were outside of previously defined area of risk (see map at left)
- We need a better understanding of the spatiotemporal risk of yellow fever in order to predict outbreaks

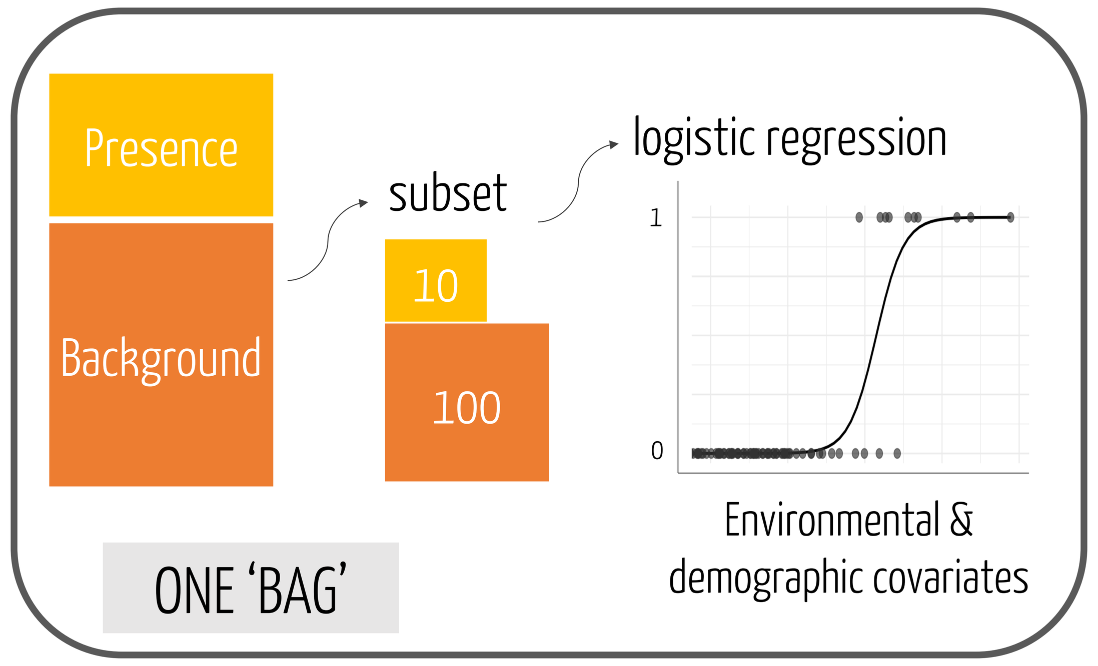
Previous risk maps[2]...

- Assume a single homogenous process across the whole country
- Are not temporally explicit, relying on data averaged across years
- Use few environmental covariates, and no socio-demographic covariates

How does risk of yellow fever spillover in Brazil shift across space and time?

BAGGED LOGISTIC REGRESSION

We used a statistical method called bagging to create an ensemble of multiple less-informative models trained on small subsets of the dataset which, when combined, are more informative than each part and ensures models trained on sparse datasets are not overfit[3].



Full model has 500 individual bags
Used a 70/30 training/testing split

MODELING DIFFERENT ECOLOGICAL CONTEXTS

- Created initial country-wide, or national, model, modeling single process across all of Brazil
- Split into two regional models based on the species richness of non-human primates, which serve as zoonotic reservoirs for the virus

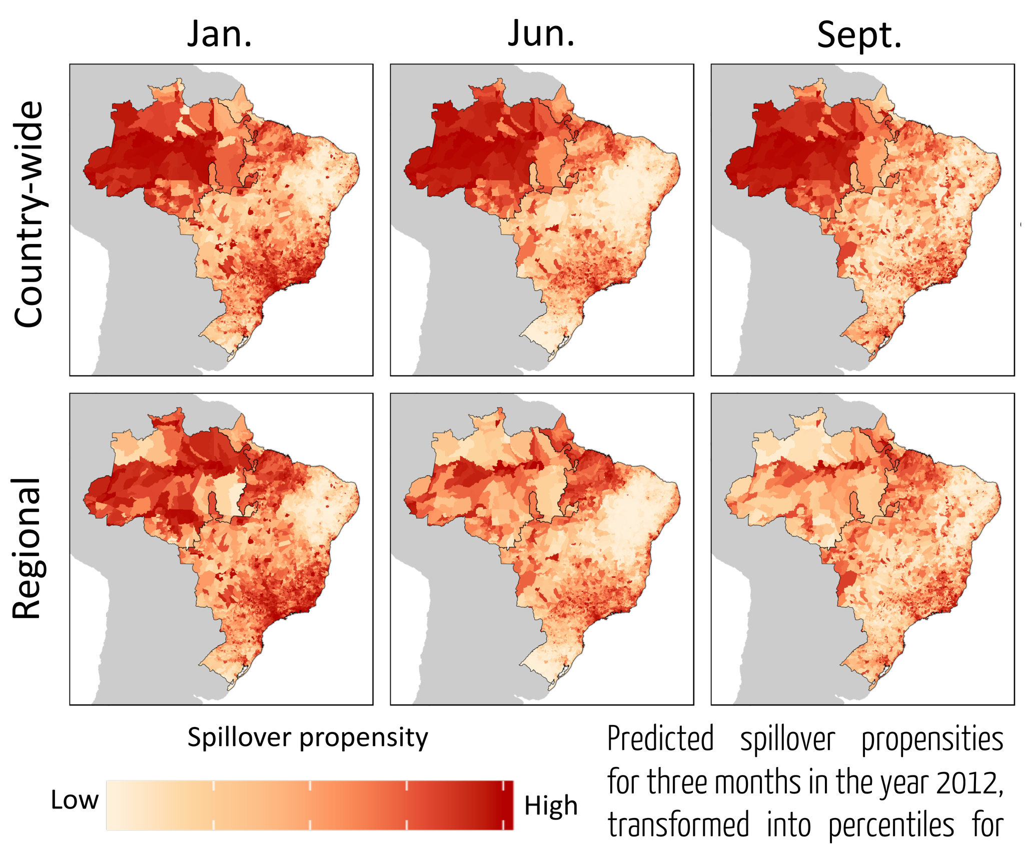
  Low reservoir richness (LRR): <6 species
  High reservoir richness (HRR): >5 species

SPILLOVER HOTSPOTS IN S.E. BRAZIL

The NW region is predicted to have a constant high risk of spillover by the national model, with more seasonal dynamics in the regional model

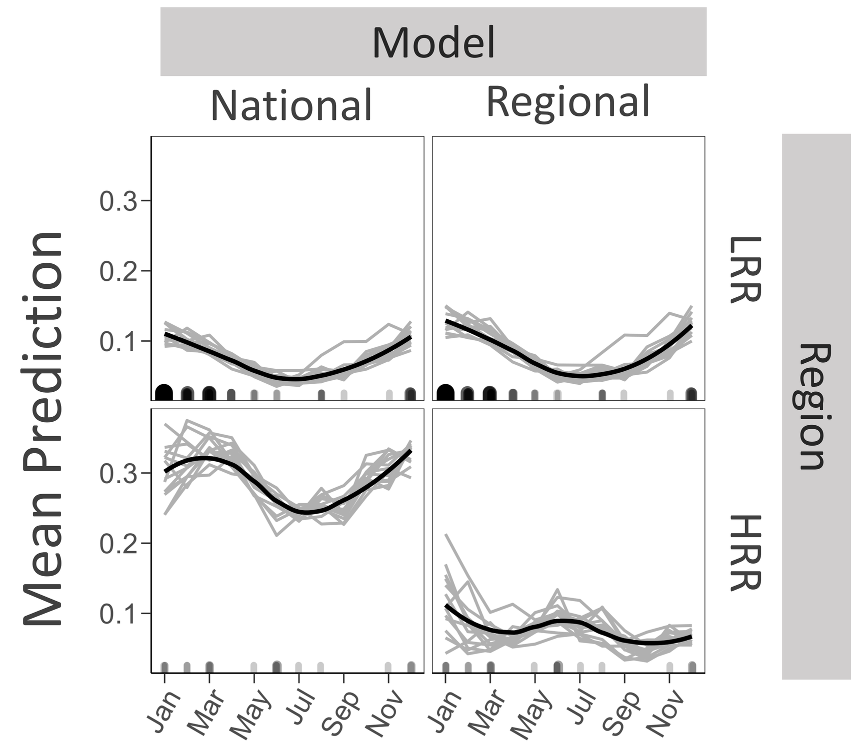
Both models highlight coastal urban areas as likely areas for potential spillover of yellow fever

Overall risk is highest during the first months of the year, during the rainy season



Predicted spillover propensities for three months in the year 2012, transformed into percentiles for comparison across models

SEASONALITY OF RISK DIFFERED ACROSS MODELS



Mean raw predictions across each region per month for the two models. Individual years are plotted in gray with the overall mean in black. Marginal rug along x-axis represents frequency of yellow fever spillover in that region by month.

The seasonality of the national model did not differ across regions

HRR was constantly high in national model, but alternated between high and low risk in regional model, better matching observed spillover frequency

There was little difference in LRR seasonality across models

DYNAMIC AND EXTREME COVARIATES

	Spatial resolution	Temporal resolution	Extreme?
Yellow fever incidence	Municipality	Monthly	-
Population density	Municipality	Yearly	-
Land surface temperature	0.05˚	Monthly	Yes
NDVI	1 km	Monthly	Yes
Rainfall	0.25˚	Monthly	Yes
Fire density	1 km	Monthly	Yes
Reservoir species richness	Municipality	Static	-
Agricultural/reservoir overlap	1 km	Yearly	-

We created extreme variables by scaling original covariates to the maximum value for that calendar month and municipality in order to distinguish between seasonal trends and climactic extremes.

MODEL PERFORMANCE

Model performance was high across all models, as measured by the Area Under the Receiver Operator Curve (AUC), a metric of the model's ability to correctly sort positive from negative observations.

	No. Presence	No. Background	Training AUC	Testing AUC
National	116	867244	0.84	0.78
Regional - LRR	95	834505	0.85	0.78
Regional - HRR	21	32739	0.88	0.85

CONCLUSIONS

Our model accurately characterizes the risk of yellow fever currently seen in the southeastern region of Brazil, raising concern for the risk of establishment of an urban cycle of the disease

Explicitly incorporating temporal and spatial heterogeneity reveals a difference in seasonality depending on ecological context (NHP richness)

Bagged logistic regressions should be applied to other systems with sparse datasets, such as emerging infectious diseases

Acknowledgements: We would like to thank Annakate Schatz and Spencer Hall for their help in collecting data and optimizing the model.

MVEVANS@UGA.EDU
MV_EVANS
MVEVANS89

References

1. Cavalcante, K. R. L. J., Tauil, P. L., Cavalcante, K. R. L. J. & Tauil, P. L. Risk of re-emergence of urban yellow fever in Brazil. *Epidemiol. E Serviços Saúde* **26,** 617–620 (2017).
2. Hamrick, P. N. *et al.* Geographic patterns and environmental factors associated with human yellow fever presence in the Americas. *PLOS Neglected Tropical Diseases* **11,** e0005897 (2017).
3. Breiman, L. Bagging predictors. *Mach Learn* **24**: 123 (1996).