Using Space-for-Time to Understand the Empirical Dynamics of Tropical Rainforests

Hannah O'Grady[1,2], Les Kaufman[3], Ethan Deyle[3]

[1] Mount Holyoke College, [2] Boston University Bioinformatics BRITE REU Program, Summer 2022, [3] Boston University Department of Biology

BOSTON UNIVERSITY

MHC

Abstract

- In this project we used Empirical Dynamic Modeling (EDM), an equations-free machine-learning approach, to study the growth of tropical rainforests[2,3] at a sub-plot level.
- We tested methods which use different spatial patterns, such as elevation, cardinal direction, and growth in adjacent areas, to inform how the data from neighboring regions was used to predict change in aboveground biomass in a central area.
- We also included the change in elevation of a region as a variable to test whether the change in elevation itself, or the relationships it creates, impact the prediction of growth.
- We then compared these methods to traditional, temporal EDM[1] as well as randomly categorized spatial areas.
- We also compared how each method performed in areas with high (>7.97m) and low (<7.97m) changes in elevation.

Objectives

- Explore different methods of using spatial patterns to extend an existing machine-learning approach which is based on long-term time data.
- Determine if we can use these spatial patterns in combination with or instead of time-series data to explore a vital ecological system where change happens slowly and long-term data are rare.

Methods

Grid Data

Continuous spatial data was summarized into discrete areas.

Create Control Models

Predict growth using traditional temporal EDM with independent spatial replicates and use EDM with random spatial embeddings.

Create Test Models

Use EDM with spatial sequences based on elevation, cardinal direction, and relative growth in neighboring cells.

Compare different models

Determine if certain methods result in better predictions and which variables are best for predicting in each method.

Compare on Different Slopes

Test whether different methods perform better on steep or less steep areas,

Results

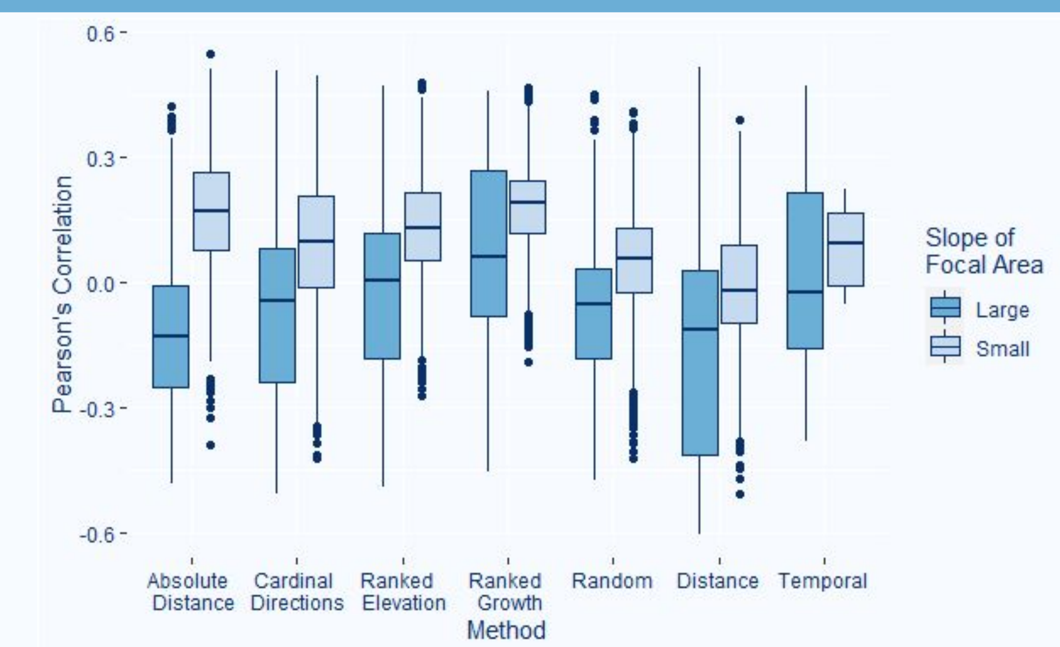

Figure 1. - Correlation between predictions and test data of different methods of EDM on areas with high change in elevation (>7.97m) and areas with low change in elevation (<7.97m) .

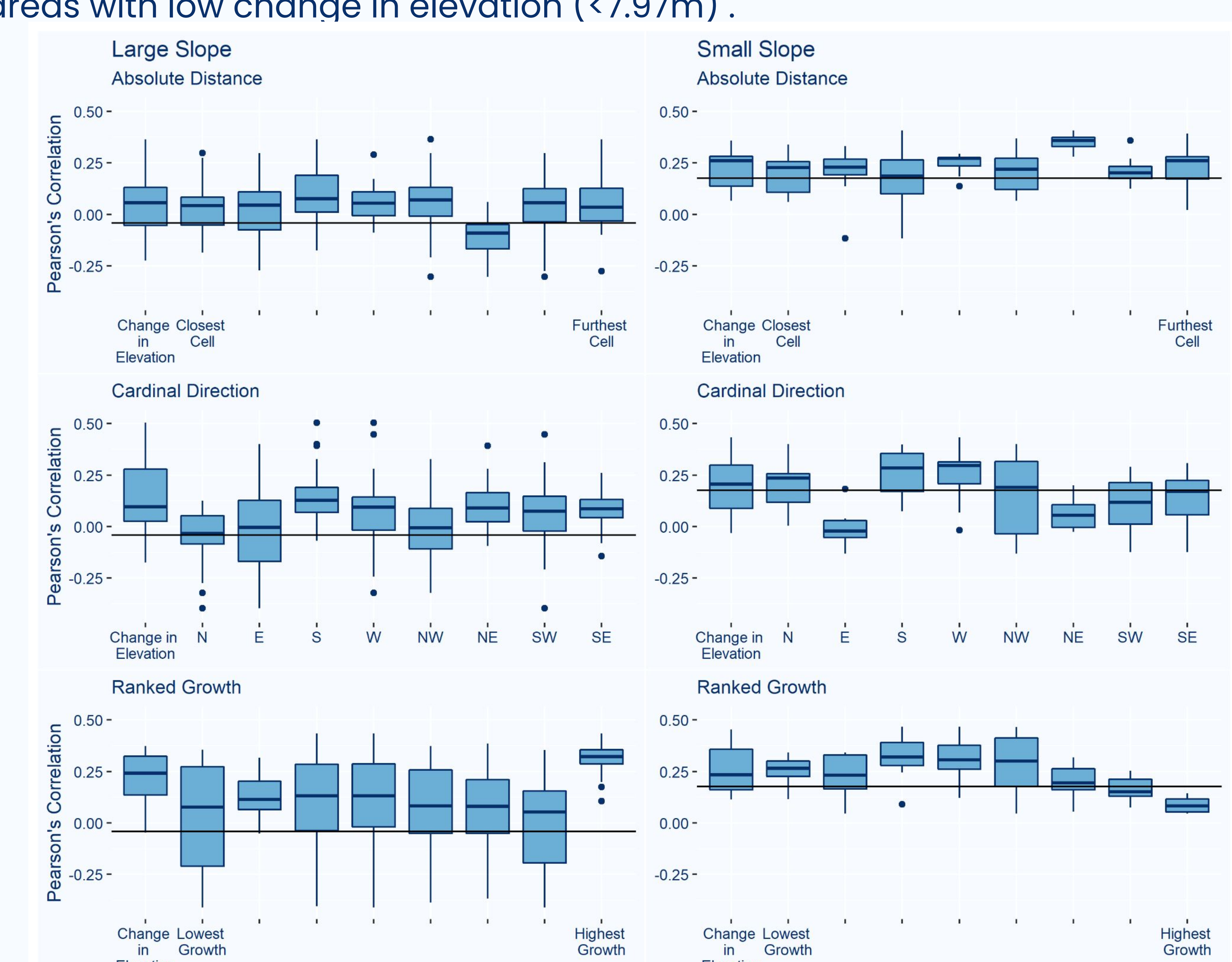

Figure 2. - The correlation between predictions and test data in aboveground biomass in areas with large versus small changes in elevation when using variables created from different categorization methods. The predictive accuracy of EDM without using spatial data is represented in each slope category by the solid lines.

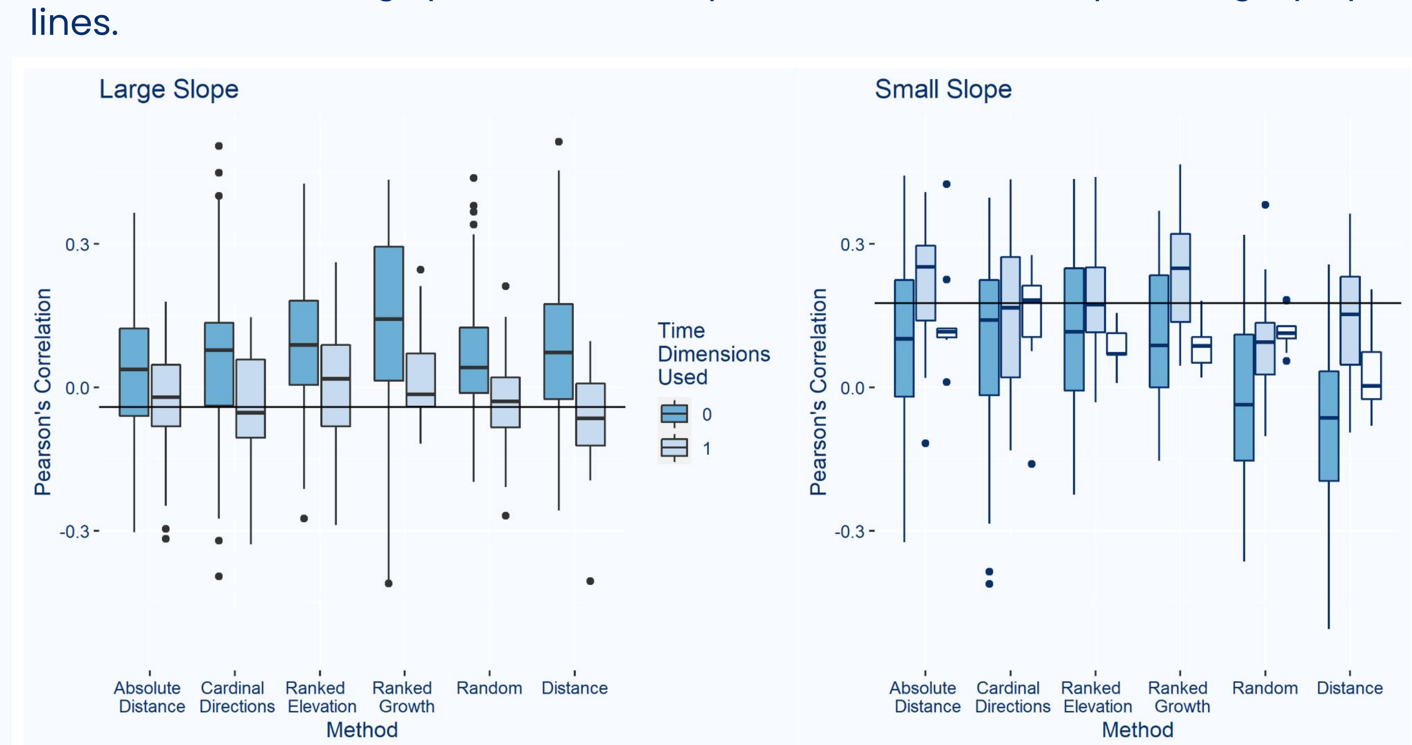

Figure 3. - Change in the correlation between predictions and test data of each method when different amounts of temporal dimensions are combined with the spatial categories of that method. The predictive accuracy of EDM without using spatial data is represented in each slope category by the solid lines.

Conclusions

- We were able to achieve comparable predictive accuracy to traditional EDM using fewer timepoints by utilizing spatial data.
- Average predictive accuracy for most methods was greater for areas with smaller change in elevation than those with larger change in elevation. However, categorizing spatial neighbors based on their growth performed well on both slope categories.
- The amount of temporal data needed to create the best models differed in areas with small and large changes in elevation.
- The variables which best predict change in aboveground biomass depend not only on the method that is used, but also the change in elevation of the region the method is applied to.

Future Work

- Apply this methodology to different ecosystems.
- Develop methods based on other spatial relationships such as light availability.
- Apply existing methods for detecting causality using EDM to the spatial patterns explored in these models rather than the temporal patterns they are traditionally applied to.

References

1. Clark, Adam Thomas, *et al.* "Spatial Convergent Cross Mapping to Detect Causal Relationships from Short Time Series." *Ecology*, vol. 96, no. 5, 2015, pp. 1174–1181., https://doi.org/10.1890/14-1479.1.
2. Condit, Richard, *et al.* (2019), Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years, Dryad, Dataset, https://doi.org/10.15146/5xcp-0d46
3. Davies SJ, *et al.*. 2021ForestGEO: understanding forest diversity and dynamics through a global observatory network. *Biol. Conserv.* 253, 108907. (doi:10.1016/j.biocon.2020.108907)

Acknowledgements

- This work was funded, in part, by NSF grant DBI-1949968, awarded to the Boston University Bioinformatics BRITE REU program
- Data for this research comes from the Smithsonian Tropical Research Institute site at Barro Colorado Island.