Quantifying the Time Series Pattern of a Binary Variable: Land Change Across 36 Years in Brazil

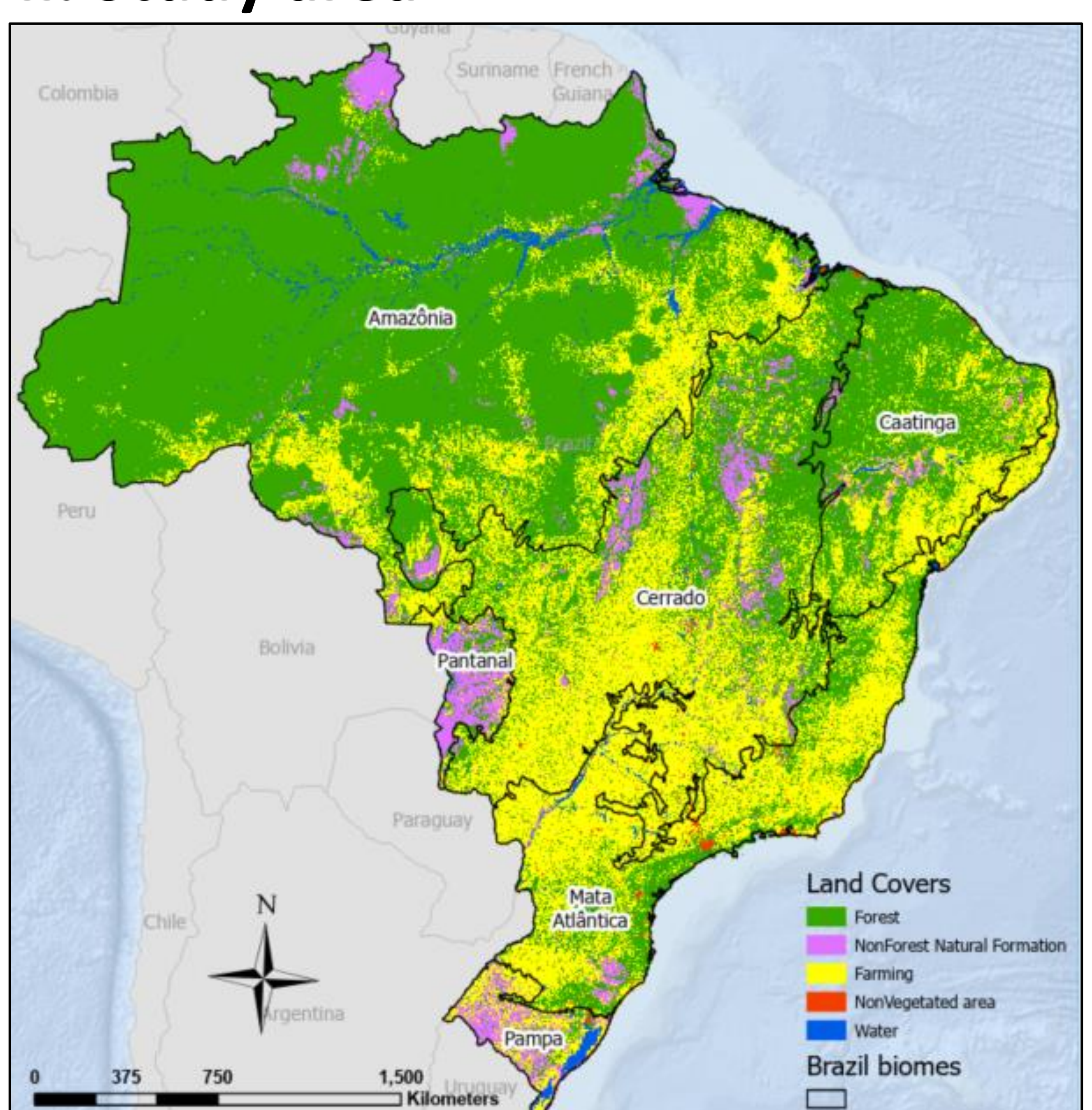
Claire Wang(clwang@clarku.edu)
Graduate School of Geography, Clark University, MapBiomas

I. Abstract

How would scientists recognize particular land change patterns or errors in a map that is part of a time series? Analyzing binary status of land categories can address the question. Binary variables in time series represent the presence or absence of a class at each time point. There are many possible ways to analyze such a time series for land cover classes. Temporal composition and configuration of the presence versus absence of a class during a time series can reveal important characteristics of land change but the profession lacks sufficient methods to characterize such patterns, which my research addresses. Temporal composition means the number of times a category is present while temporal configuration concerns the sequential arrangement of the presence of the class during a time series. I developed concepts and created a computer program to characterize composition and configuration in ways that have implications for data quality assessment and land change science. The data come from MapBiomas' collection 6.0, annual land cover data of 36 years.

The methods (1) differentiate temporal change patterns across spatial scales, (2) give insights into data quality, and (3) improve understanding of land change drivers. My methods are more sophisticated than common error detections because my methods analyze time series in an integrated manner to highlight noteworthy land changes. The methods can be widely applied to other time series of categorical variables via modification of the criteria according to a project's goals.

II. Study area



Figures 1. Land covers in a five-category classification in Brazil, 2020.

Objective: To define and to visualize noteworthy temporal patterns of land change in Brazil.

Noteworthy data are anomalous land changes or errors in a time series. My computer code allows users to define noteworthy data in a time series according to three criteria.

III. Criteria for filtering noteworthy data

a)Longest presence < 3 time points

b)Shortest presence < its theoretical upper limit / 3

c)Number of changes < 5

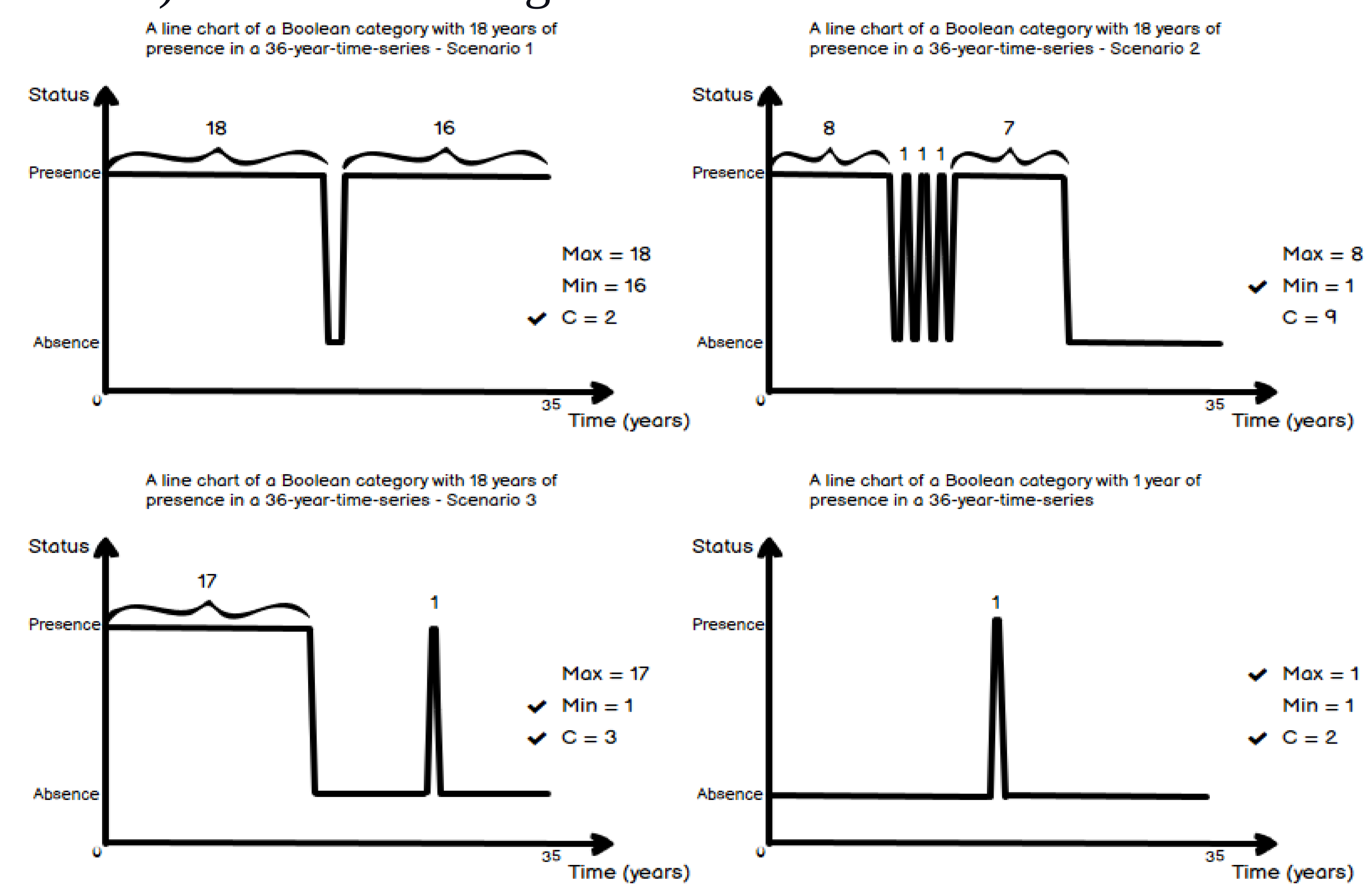


Figure 2. Line charts showing changes of a land category in Boolean status in 36 years. Charts in the second row meet two of the three criteria for noteworthy data.

IV. Results

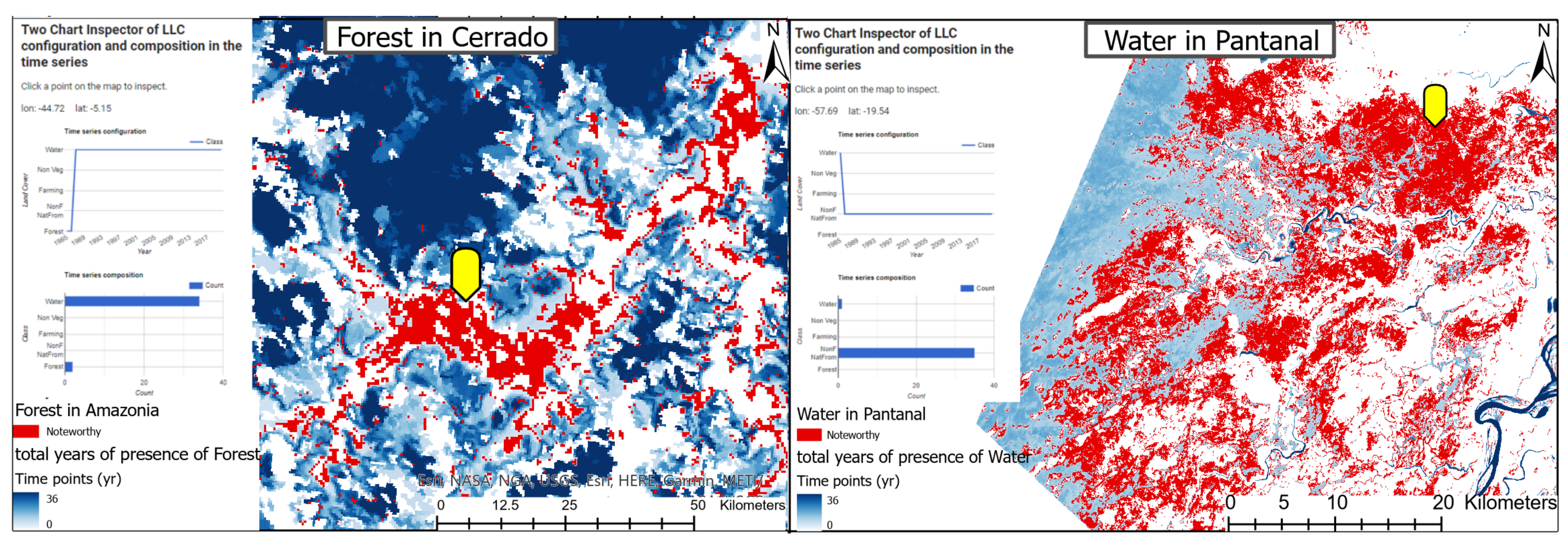


Figure 3. Large areas of forest and water lost at the beginning of the time series.

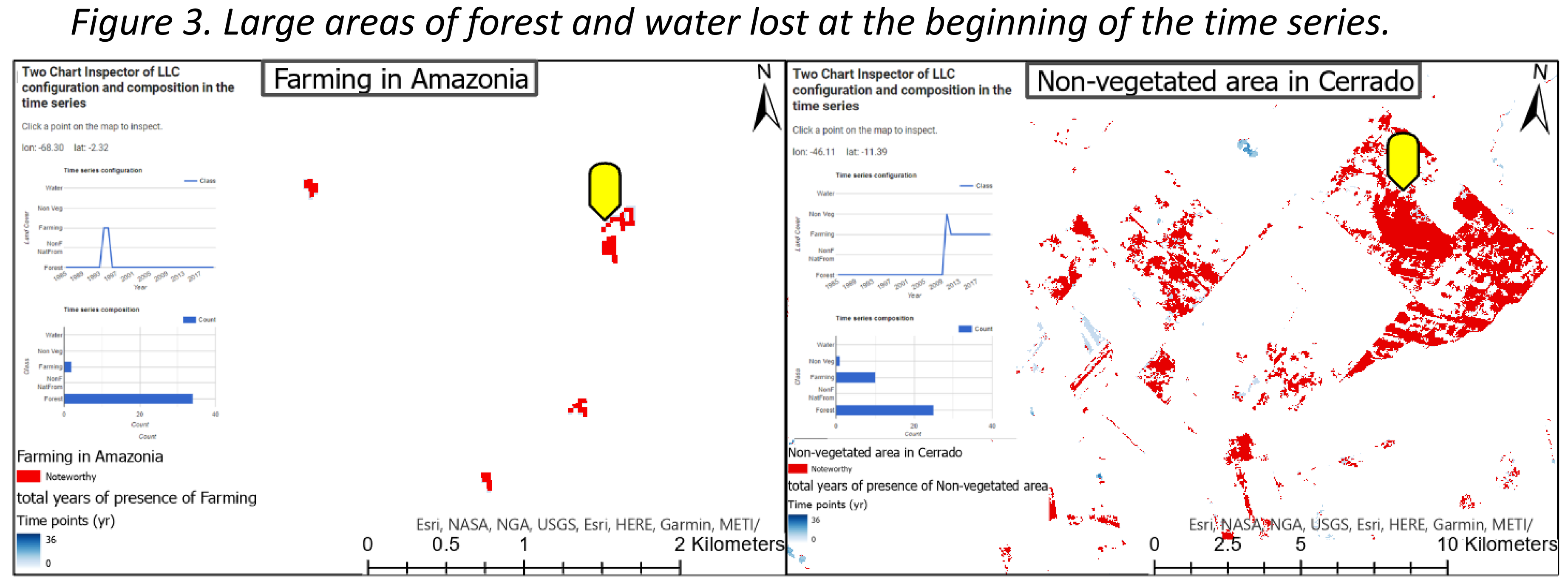


Figure 4. Anomalous but possible real changes in farming and non-vegetated area.

The maps above show common noteworthy land change patterns. Forest shows loss of large patches with regular shapes, such as rivers, at the beginning of the time series. Water also exhibits a similar pattern due to the lack of a temporal filter. These patterns are not likely to be real and can mistakenly overestimate forest loss.

The methods also capture anomalous but real data. Farming has short and abrupt presences due to a policy change around 1995. For non-vegetated area, the noteworthy data are possibly pre-farming clearance or dry years. The anomalous but real changes improve researchers' understanding of land change history.

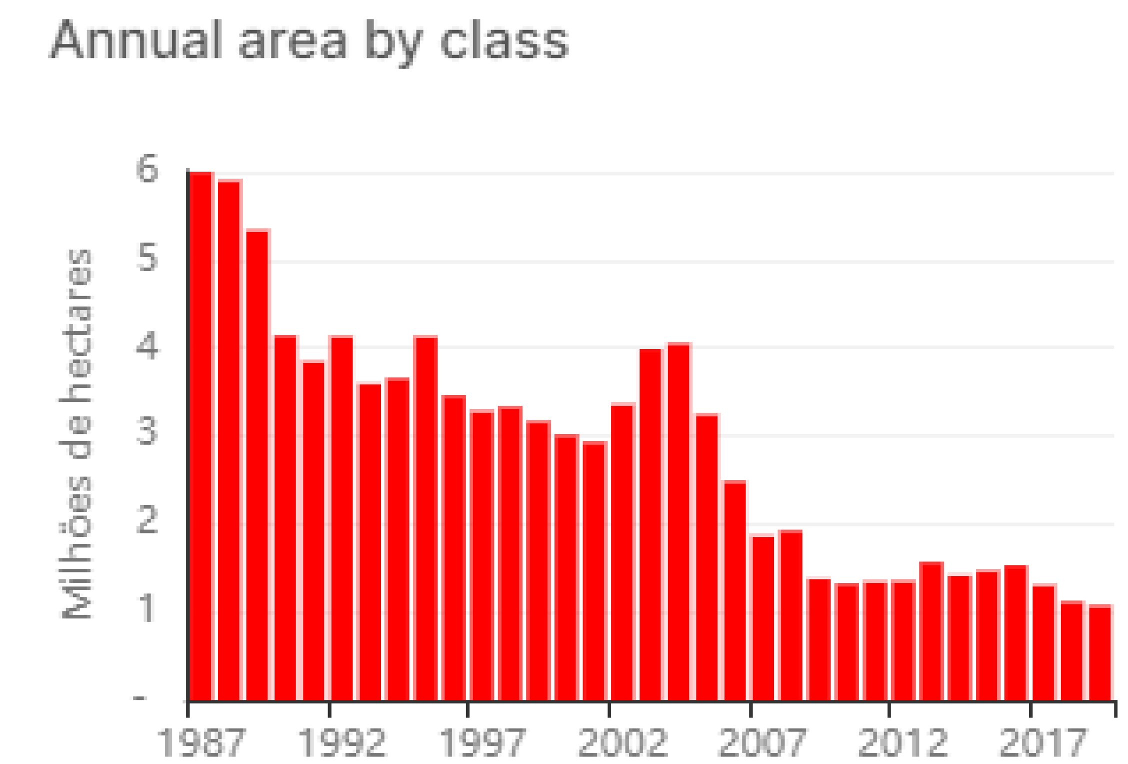


Figure 5. A bar plot from MapBiomas with time series on the x-axis and area of deforestation in million hectares on the y-axis. The plot shows more deforestation at the beginning of the time series, possibly due to the accidental overestimation of forest loss.

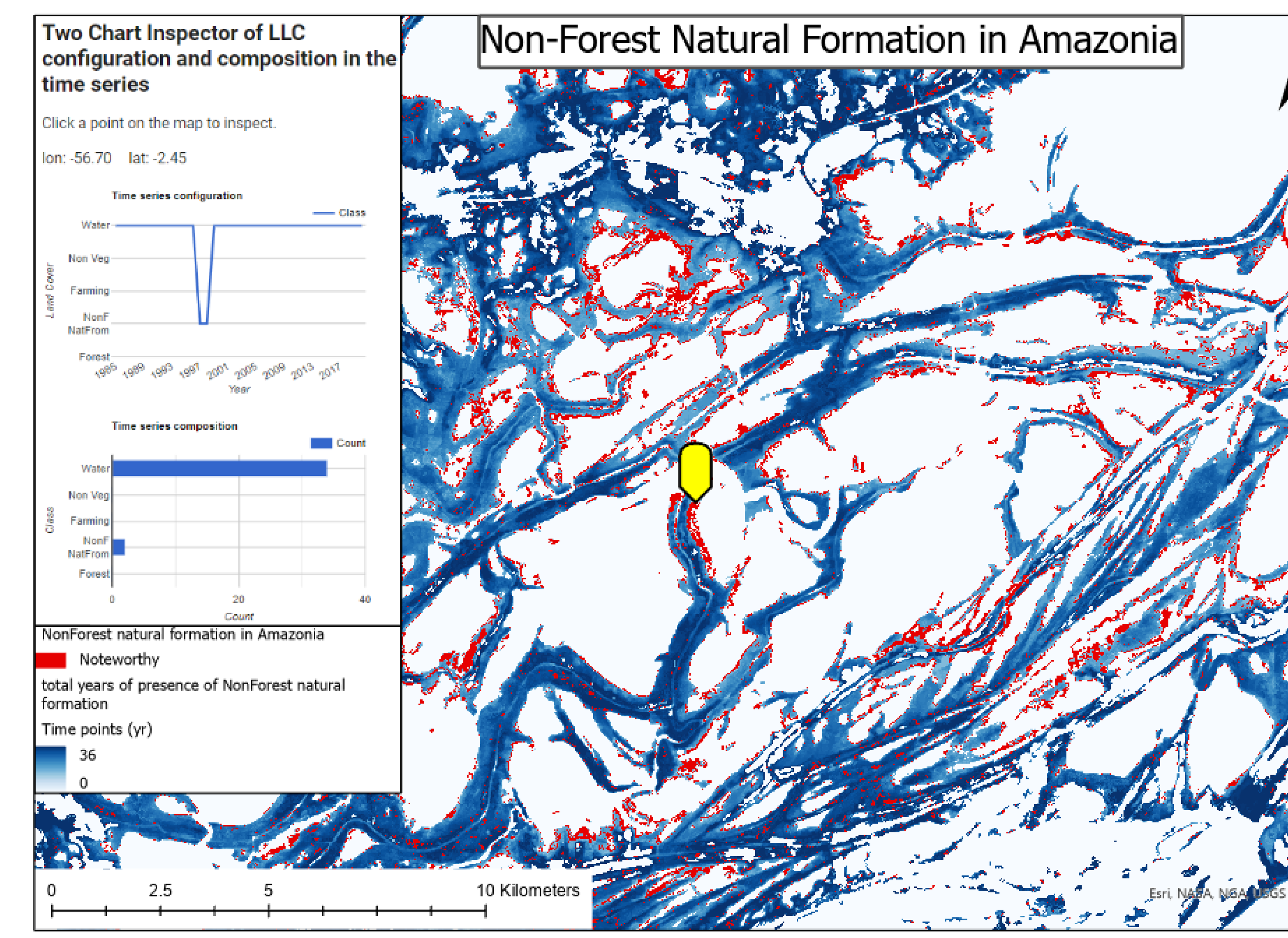


Figure 6. Noteworthy data often appear on the edge of land cover patches.

V. Conclusions

This research quantifies temporal patterns for categorical variables involved in land change. I developed novel metrics to analyze time series configurations and compositions. The methods use distribution and clustering patterns to detect noteworthy data in a single land cover in a time series. The findings are as follows:

1. Common temporal filters fail to recognize errors at the beginning of the time series, which my methods address.
2. The methods can capture anomalous but real land change. Researchers can modify the criteria to capture errors only after having better understanding of the land change history in Brazil.
3. There are overall more noteworthy data on the edges of land cover patches where multiple land covers may interact.

Acknowledgements

The Humanize Institute supported this work via a grant entitled "Methods and software to understand and build the MapBiomas data". Land cover data derive from MapBiomas (www.mapbiomas.org). Scientists at MapBiomas provided helpful feedback. Any questions or concerns regarding the above research can be directed to [clwang@clarku.edu.](clwang@clarku.edu)