The Total Operating Characteristic requires improvements to its use and software.

Tanner Honnef

DECEMBER 2023

A THESIS

Submitted to the faculty of Clark University,
Worcester, Massachusetts,
in partial fulfillment of
the requirements for
the degree of Bachelor of Geography
in the Department of Geography

THESIS COMMITTEE

Robert Gilmore Pontius Jr, Ph.D.
Chief Instructor

Christopher Williams, Ph.D.
Committee Member

ABSTRACT

The Total Operating Characteristic (TOC) is an improvement on the quantitative method called the Relative Operating Characteristic (ROC), both of which show the association between actual presence and an index of presence for a binary variable. TOC shows all the information that the ROC shows while the TOC shows additional information to give the total information in the contingency table for each threshold. TOC is helpful when applied properly according to best practices. I analyze ten articles that cited the TOC. Results show that that all ten articles neglect to apply at least one of seven best practices. Then I illustrate best practices by applying the TOC to analyze the gain of temporary crops in Western Bahia, Brazil. The illustration shows how to compare two TOC curves. The first index uses exclusively distance from change of temporary crop as a driver variable while the second index integrates distance to change of temporary crop, elevation, and slope in the Multi-Layer Perceptron neural net. TOC curves for the two indices are plotted in the same parallelogram. The first index derived from exclusively distance to change has a higher AUC than the index from MLP because of a complication in the algorithm to rank the many observations. A more effective algorithm would have likely computed the AUC for MLP to be larger than the AUC for distance to change. My thesis concludes by recommending further development of software to generate and to compare TOC curves.

ACKNOWLEDGEMENTS

Professor Pontius, Professor Williams, and Thomas Bilintoh provided advice for this work. Clark Labs facilitated this work by creating the GIS software TerrSet®. The United States National Aeronautics and Space Administration (NASA) supported this work via the grant *Irrigation as climate change adaptation in the Cerrado biome of Brazil evaluated with new quantitative methods, socio-economic analysis, and scenario models* Award # 80NSSC23K0508. MapBiomas provided data for free.

Table of Contents

LIST OF FIGURES

page

LIST OF TABLES

1 Introduction

1.1 Background

The Total Operating Characteristic can measure the association between change of a particular categorical variable such as the gain or loss of forest and a corresponding ranked variable such as distance to forest edge or elevation. TOC was designed to examine the change of a land cover category during a given time interval (Pontius Jr & Si, 2014). TOC is summarized by the Area Under the Curve (AUC) which is a metric commonly used by scientists which does not show as much detail as TOC. AUC is one number while TOC shows the total information from a contingency table for each threshold (Pontius Jr & Si, 2014).

1.2 Research Objective

The Total Operating Characteristic (TOC) is designed to show diagnostic ability for multiple thresholds (Pontius Jr & Si, 2014). This manuscript provides an example of the use of TOC to analyze temporary crop gross gain using distance to change in Western Bahia, Brazil in comparison to another model which uses distance to change, elevation, and slope. It also provides examples from the literature of how TOC should and should not be used based on a set of best practices.

1.3 Literature Review

Some literature incorrectly uses the AUC of the TOC to claim that a model is good. Chen et al., 2020 claims that the results of the model in the analysis are "relatively reasonable" because it yields a high AUC. Another incorrect use of TOC is not masking pixels that are not eligible for change. The TOC curve in Naghibi et al., 2016 goes straight up the left bound of the parallelogram. This likely indicates that pixels not eligible for change were not masked when the TOC curve was generated. Ranagalage et al., 2019 is an example of a work that models land use and land cover change but does not use TOC to assess the goodness of fit of the model. Instead, the publication uses Kappa as compared to simpler techniques of assessing goodness of fit (Pontius and Millones, 2011).

1.4 Study area

Figure 1 shows the study area used for illustration, which is in Bahia, Brazil. Much of the land in Western Bahia has experienced land change because of the gain of temporary crops such as soybeans, cotton, and corn (Pousa et al., 2019). The data is masked to the nine western most municipalities in Bahia where much of the landscape is temporary crops. Figure 2 shows the municipalities to which the datum is masked.

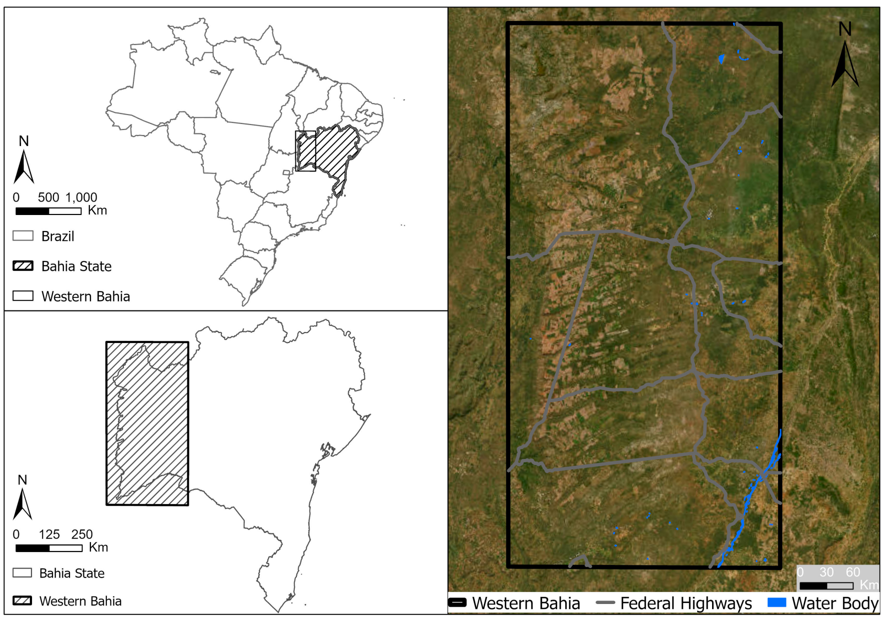


Figure 1. Maps of the study area in Bahia, Brazil (Pontius et al., 2023).

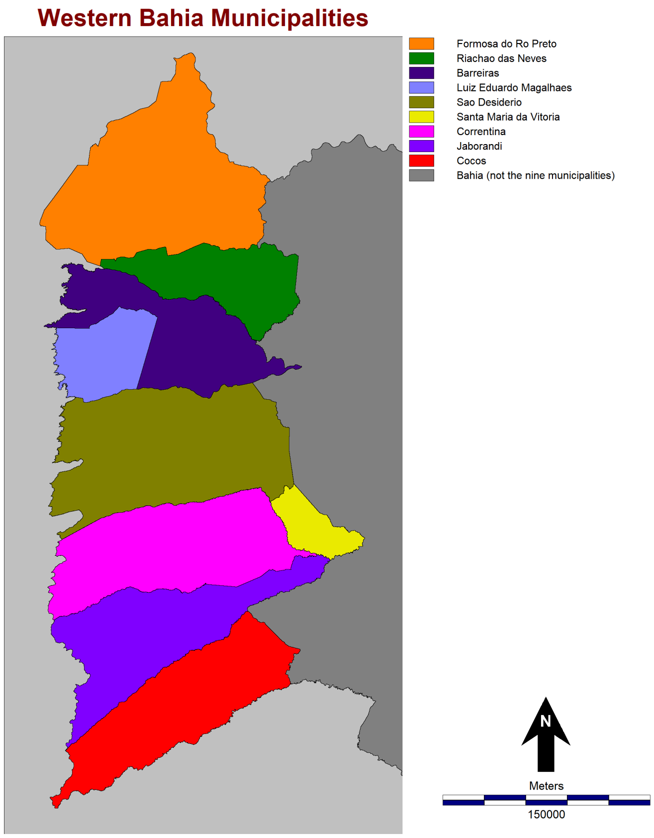
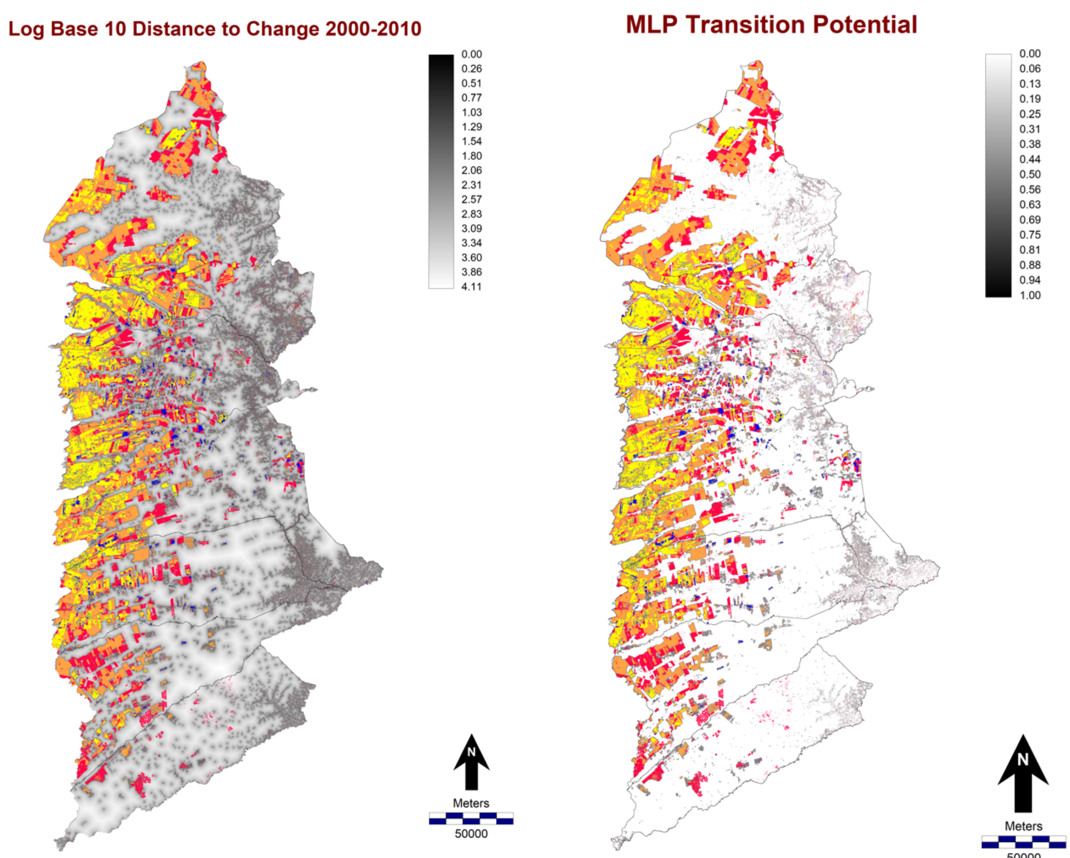


Figure 2. Municipalities within the study area of Bahia, Brazil

2 Methods

2.1 Data

The 30-meter resolution data shows presence and absence of temporary crops. The data used in the analysis has been provided by MapBiomas (Souza et al., 2020). The three time points are 2000, 2010, and 2020. Each pixel shows either presence or absence of temporary crops. Temporary crops are crops that have a growing cycle of less than one year. Figure 3 shows how the landscape changes between each of the time points. Yellow represents the persistence of temporary crops from 2000-2010, orange represents the gain of temporary crops from 2000-2010, blue represents the loss of temporary crops from 2000-2010, and red represents the gain of temporary crops from 2010-2020. In figure 3a, the darkest pixels are closest to change from 2000-2010 while the lightest pixels are furthest from change during the same period.



(a) **(b)**

Figure 3. Map of log base 10 transformed distance from change from 2000 to 2010 (a). Map of MLP transition potential values for distance from change, elevation, and slope (b). Yellow is persistence of temporary crops from 2000-2010. Orange is gain of temporary crops from 2000 to 2010. Blue is loss of temporary crops from 2000 to 2010. Red is gain of temporary crops from 2010 to 2020.

2.2 Example of Appropriate TOC Use in Land Change Science

Two runs were used to demonstrate how TOC should be used. These runs focus on the goodness of fit of the validation interval. Both runs use a calibration interval of 2000 to 2010 and a validation interval of 2010 to 2020. The model predicts temporary crops' gross gain, not its loss.

2.2.1 TOC using distance to change

The index raster for the TOC curve is the log base 10 distance from change during the calibration interval. The boolean raster is the land that changed from absence to presence during the validation interval. The mask raster is pixels representing absence of temporary crops at the end of the calibration interval. A TOC curve above the uniform line will show that gains of presence have a positive correlation with pixels that changed during the calibration interval. A steep slope will show a stronger correlation between distance to change and gains of presence while gradual slopes will show a weaker correlation between distance to change and presence.

Figure 3a represents the distance from change from 2000 to 2010 transformed by log base 10. In figure 3, red pixels represent those that changed from absence to presence from 2010-2020. The pixels suitable for change in the analysis are those which are red, blue, or are on a scale from black to white.

2.2.2 TOC using MLP and multiple variables

The second run is done in the Land Change Modeler in TerrSet with Multi-Layer Perceptron (MLP). The driver variables are distance from change from 2000 to 2010, elevation, and slope. Figure 4 shows the driver variables used in the analysis. TOC curves for both runs are generated and overlayed on the same parallelogram.

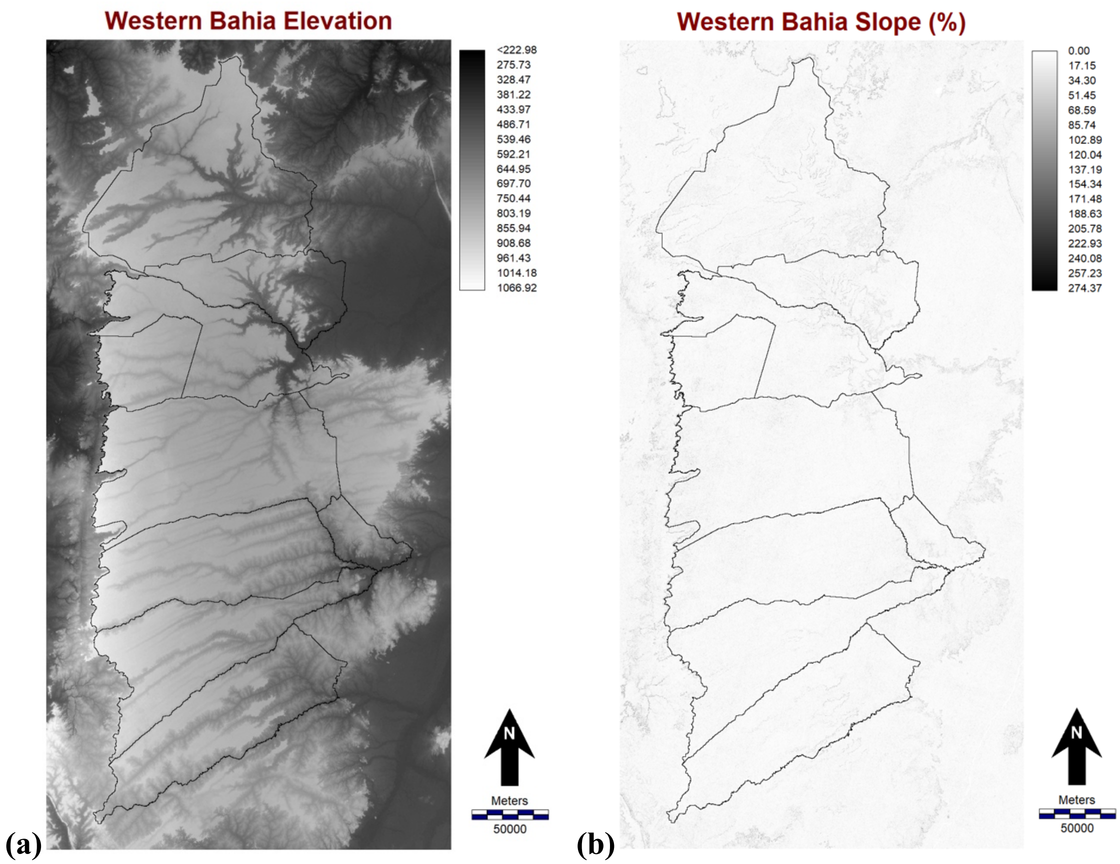


Figure 4. Map of elevation driver used in Multi-Layer Perceptron run (a). Map of slope driver used in Multi-Layer Perceptron run (b). The black outline represents the area to which the data are masked.

2.3 Literature Review

The literature review analyzes ten publications that cited Pontius Jr & Si, 2014. Each of the publications were assessed on whether they follow the best practices of the Total Operating Characteristic. Table 1 shows the criteria by which each publication was evaluated.

Table 1. The criteria by which each piece of literature is evaluated

ID	Description
1	Assesses data quality in an appropriate manner
2	Shows the maps in a clear manner
3	Masks pixels that are not candidates for the particular type of change
4	Avoids using AUC to claim model performance such as acceptable, fair, good, excellent
5	Shows the TOC curve as opposed to the ROC curves
6	Interprets the shape of the TOC curve to the left of the point of true quantity versus to the right of the point of true quantity
7	Compares a baseline suitability map to an alternate suitability map

3 Results

3.1 Example of Appropriate TOC Use in Land Change Science

Figure 5 has two TOC curves. The blue TOC curve represents the goodness of fit of the validation on the single variable of distance to change during the calibration interval. The orange TOC curve represents the goodness of fit of the validation on the three variables which are distance to change during the calibration, elevation, and slope using the Multi-Layer Perceptron. The steep segments near the lower left corner of the TOC parallelogram represent gain of temporary crops near existing temporary crops while the less steep segments in the top right of the parallelogram represent less gain of presence occurring further from previous crop change. The AUC of the blue TOC curve is 0.766 and the AUC of the orange TOC curve is 0.755.

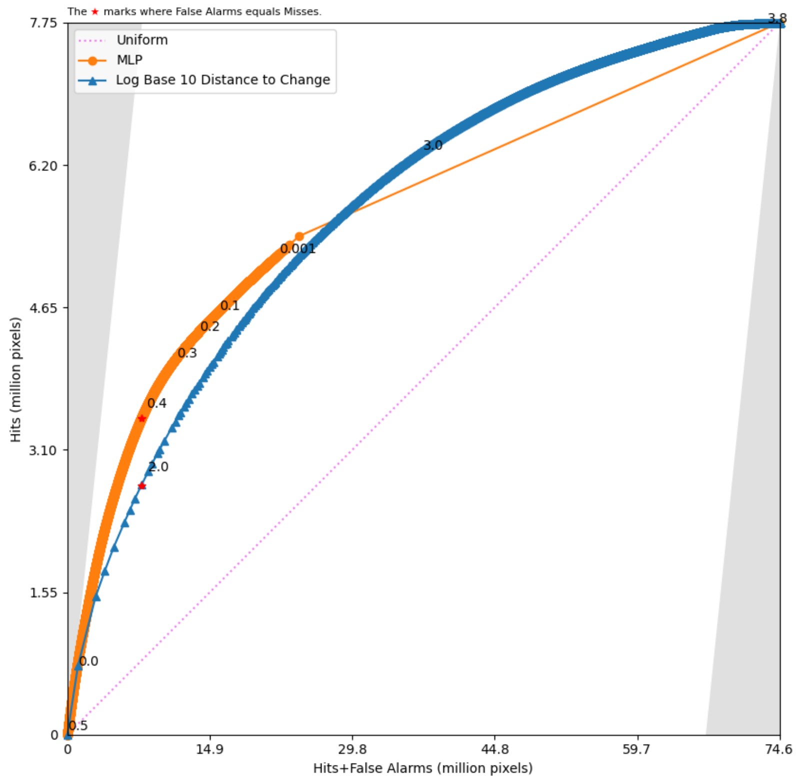


Figure 5. TOC curve showing goodness of fit of the validation interval. Blue represents the model with a single variable of distance to change from 2000 to 2010. Orange represents the model using the Multi-Layer Perceptron run with distance to change from 2000 to 2010, elevation, and slope as input variables. The red star represents the true quantity of change.

3.2 Literature Review

Out of the ten selected publications, none of the sources used all of the best practices outlined in this manuscript. Table 2 outlines the criteria each of the sources uses in their analysis. Cushman et al. (2020) uses all the best practices but does not mention that it analyzes the segments of the TOC curve that are to the left of the true quantity. Other sources such as

Chen et al. (2020) and Estoque & Murayama (2016) do not follow hardly any of the best practices outlined in this manuscript.

None of the publications explicitly analyze the TOC curve to the left of the true quantity. All the sources either analyze the segments throughout the whole curve or do not mention specific segments in their discussions so it cannot be concluded that they are analyzing segments to the left of the true quantity.

Table 2. The criteria each publication uses when modeling with TOC. ✓ means a publication uses that criterion. X means the publication does not use that criterion. ? means the publication does not say whether it uses the criterion. Refer to table 1 for the Criteria ID.

Publication	**Criteria ID**							
	1	2	3	4	5	6	7	Total ✓
Chen et al. (2020)	X	X	X	X	X	X	X	**0**
Estoque & Murayama (2016)	X	✓	?	X	X	?	X	**1**
Zhuang et al. (2022)	✓	✓	?	X	X	X	X	**2**
Chakraborti et al. (2018)	X	✓	X	X	✓	X	X	**2**
Amato et al. (2018)	X	✓	?	?	✓	?	X	**2**
Naghibi et al. (2016)	X	✓	X	✓	✓	X	X	**3**
Kamusoko & Gamba (2015)	✓	✓	✓	X	✓	X	?	**4**
Deng & Quan (2022)	✓	✓	X	✓	✓	?	X	**4**
Shojaei et al. (2022)	X	✓	✓	✓	✓	X	✓	**5**
Cushman et al. (2017)	✓	✓	✓	✓	✓	?	✓	**6**
Total ✓	**4**	**9**	**3**	**4**	**7**	**0**	**2**	

4 Discussion

TOC is a method that reveals more information than ROC. TOC outputs the size of the extent, the size of the presence of the boolean dependent variable, the thresholds of the index variable, and a contingency table at each threshold. In contrast, ROC shows two ratios at each threshold (Pontius Jr & Si, 2014).

In figure 3, there are many grey pixels on both maps in the southeast portion of the map. These are either a result of small, likely unplanned agricultural activity or noise. Many of these temporary crop pixels are a single pixel or a small cluster that is surrounded by absence. When using a model that only factors distance to change, each single pixel of gain of presence surrounded by absence generates 4 pixels with the second highest possible transition potential value. The highest transition potential value is on pixels that were presence at the beginning of the calibration interval and absence at the end. This leads to allocation of pixels in areas which are less likely to become developed. In this example, the western portion of the map is dominated by rapidly expanding large scale agriculture, but many of the high transition potential values can be found in the southeastern and northeastern parts of the map where there is limited large scale agriculture. Small clusters of pixels which represent noise or small-scale agriculture cause these high transition potential values to be generated in areas where change is less likely. One way to resolve this is to pass over the data with a spatial filter that eliminates single pixels of presence or small groups of presence pixels before performing analysis.

The greyscale pixels in figure 3 represent the log transformed distance to change of temporary crops within the calibration interval in the study area of Western Bahia, Brazil. The data is log transformed for visual purposes only. This does not affect the shape of the TOC curve. The distance to change map is the index raster used in the first model run. With the distance to change being the sole driver variables of the first model, the model predicts that pixels closest to change at the beginning of the calibration interval will transition first while pixels furthest from change will transition last. The boolean raster input shows the pixels that transitioned from absence to presence during the validation interval. As a result, this model predicts the gain of presence, not the loss of presence.

Figure 5 shows the resulting TOC curves overlayed on the same parallelogram. The curves are above the uniform line, which indicates that the models predict more accurately than random. The points near the origin show pixels of presence gain, which are near pixels that changed during the calibration interval. The points in the top right of the parallelogram show areas where pixels furthest from change during the calibration interval are gaining presence. Steep slopes in segments near the origin represent higher intensities of pixels per threshold being converted from absence to presence. Flatter slopes in the upper right corner represent smaller intensities of pixels per threshold being converted from presence to absence.

The Multi-Layer Perceptron assigns values of pixels in the study area that are white values extremely close to 0. Many of the white pixels are a value times 10^{-13} while the values closer to pixels that changed are between 0.2 and 0.5. Many of these values very close to 0 are unique which means they will have their own threshold when the TOC curve generator is set to unique values. A TOC curve with a threshold for each unique value was not able to be generated because of the computational power required to generate it. Instead, the MLP curve in Figure 5 is set to have a threshold for every 0.0001 units in the transition potential map. This allowed run time to be reduced, but this also yielded a long segment in the top right corner of the parallelogram. Around 67% of the pixels that are eligible for change in this analysis fall in that segment. If MLP ranked transition potential values differently or the TOC Curve Generator had another way of displaying pixels that are all extremely close to 0, the computational power required to perform this analysis could be reduced, and results could be more meaningful.

The AUC of the distance to change curve is higher than that of the MLP curve. The distance to change curve is run so that each distance value has a unique threshold while the MLP curve has a threshold for every 0.0001 units of the transition potential. In areas where there are thresholds on both curves, the MLP curve is above the distance to change curve. The distance to change goes above the MLP curve in the top right where the MLP curve does not have thresholds. It is likely that if it was possible to run the TOC Curve Generator to create results for unique values on the MLP curve, the AUC of the MLP curve would exceed the AUC of the distance to change curve.

None of the sources reviewed followed all the best practices of TOC outlined in this manuscript. As outlined as one of the best practices when using TOC in this manuscript, assessing data quality is crucial before analyzing the results of the model. In the model that assesses distance to change, there are many lone pixels of temporary crop which are likely a result of data error. If scientists are unable to understand the quality of their data, their interpretation of the results is likely to be misleading. If data quality were not assessed in this study, it would be expected that gain of temporary crop would be in the eastern portion of the study area where the highest transition potential values are located. In reality, most of the gain of temporary crop occurs in the western portion of the study area.

5 Conclusions

This manuscript gives a clear outline on how to generate a TOC curve based on a single factor and multiple factors. It addresses issues with the generation of TOC curves and with data quality both of which are relevant to the case study. This work provides a set of best practices for using TOC and assesses literature on their use of those best practices. None of the literature analyzed in this manuscript follows all the best practices that have been outlined for creating and analyzing TOC. If the best practices outlined in this manuscript are used more frequently, then the literature concerning TOC will become clearer. Future research concerning the topic should develop a more efficient algorithm to compute TOC curves in a way that does not require as much computational power to generate many unique threshold values.

BIBLIOGRAPHY

Amato, F., Tonini, M., Murgante, B., & Kanevski, M. (2018). Fuzzy definition of Rural Urban Interface: An application based on land use change scenarios in Portugal. *Environmental Modelling & Software*, *104*, 171–187. https://doi.org/10.1016/j.envsoft.2018.03.016

Bilintoh, T. (2023). Methods to Analyze Land Transitions during Multiple Time Intervals*.* Unpublished manuscript.

Chakraborti, S., Das, D. N., Mondal, B., Shafizadeh-Moghadam, H., & Feng, Y. (2018). A neural network and landscape metrics to propose a flexible urban growth boundary: A case study. *Ecological Indicators*, *93*, 952–965. https://doi.org/10.1016/j.ecolind.2018.05.036

Chen, S., Feng, Y., Ye, Z., Tong, X., Wang, R., Zhai, S., Gao, C., Lei, Z., & Jin, Y. (2020). A cellular automata approach of urban sprawl simulation with Bayesian spatially-varying transformation rules. *GIScience & Remote Sensing*, *57*(7), 924–942. https://doi.org/10.1080/15481603.2020.1829376

Cushman, S. A., Macdonald, E. A., Landguth, E. L., Malhi, Y., & Macdonald, D. W. (2017). Multiple-scale prediction of forest loss risk across Borneo. *Landscape Ecology*, *32*(8), 1581–1598. https://doi.org/10.1007/s10980-017-0520-0

Deng, Z., & Quan, B. (2022). Intensity Characteristics and Multi-Scenario Projection of Land Use and Land Cover Change in Hengyang, China. *International Journal of Environmental Research and Public Health*, *19*(14), 8491. https://doi.org/10.3390/ijerph19148491

Estoque, R. C., & Murayama, Y. (2016). Quantifying landscape pattern and ecosystem service value changes in four rapidly urbanizing hill stations of Southeast Asia. *Landscape Ecology*, *31*(7), 1481–1507. https://doi.org/10.1007/s10980-016-0341-6

Kamusoko, C., & Gamba, J. (2015). Simulating Urban Growth Using a Random Forest-Cellular Automata (RF-CA) Model. *ISPRS International Journal of Geo-Information*, *4*(2), 447–470. https://doi.org/10.3390/ijgi4020447

Liu, Z., & Pontius Jr, R. G. (2021). The Total Operating Characteristic from Stratified Random Sampling with an Application to Flood Mapping. *Remote Sensing*, *13*(19), 3922. https://doi.org/10.3390/rs13193922

Naghibi, F., & Delavar, M. (2016). Discovery of Transition Rules for Cellular Automata Using Artificial Bee Colony and Particle Swarm Optimization Algorithms in Urban Growth Modeling. *ISPRS International Journal of Geo-Information*, *5*(12), 241. https://doi.org/10.3390/ijgi5120241

MapBiomas Project - Collection 8 of the Annual Land Use Land Cover Maps of Brazil, accessed on 11/16/23 through the link: https://mapbiomas.org/

Pontius Jr, R. G. (2022). Metrics that make a difference. Springer Nature Switzerland AG: Cham, Switzerland.

Pontius Jr, R. G., & Si, K. (2014). The total operating characteristic to measure diagnostic ability for multiple thresholds. *International Journal of Geographical Information Science*, *28*(3), 570–583. https://doi.org/10.1080/13658816.2013.862623

Pontius, R. G., & Millones, M. (2011). Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing, 32(15), 4407–4429. https://doi.org/10.1080/01431161.2011.552923

Pontius Jr, Robert Gilmore, Thomas Bilintoh, Gustavo de L. T. Oliveira, Julia Z. Shimbo. (2023). Trajectories of losses and gains of soybean cultivation during multiple time intervals in western Bahia, Brazil. Space Week Nordeste. Fortaleza, Brazil.

Pousa, R., Costa, M. H., Pimenta, F. M., Fontes, V. C., Brito, V. F. A. D., & Castro, M. (2019). Climate Change and Intense Irrigation Growth in Western Bahia, Brazil: The Urgent Need for Hydroclimatic Monitoring. *Water*, *11*(5), 933. https://doi.org/10.3390/w11050933

Shojaei, H., Nadi, S., Shafizadeh-Moghadam, H., Tayyebi, A., & Van Genderen, J. (2022). An efficient built-up land expansion model using a modified U-Net. *International Journal of Digital Earth*, *15*(1), 148–163. https://doi.org/10.1080/17538947.2021.2017035

Souza, C. M., Z. Shimbo, J., Rosa, M. R., Parente, L. L., A. Alencar, A., Rudorff, B. F. T., Hasenack, H., Matsumoto, M., G. Ferreira, L., Souza-Filho, P. W. M., De Oliveira, S. W., Rocha, W. F., Fonseca, A. V., Marques, C. B., Diniz, C. G., Costa, D., Monteiro, D., Rosa, E. R., Vélez-Martin, E., … Azevedo, T. (2020). Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sensing*, *12*(17), 2735. https://doi.org/10.3390/rs12172735

Zhuang, H., Chen, G., Yan, Y., Li, B., Zeng, L., Ou, J., Liu, K., & Liu, X. (2022). Simulation of urban land expansion in China at 30 m resolution through 2050 under shared socioeconomic pathways. *GIScience & Remote Sensing*, *59*(1), 1301–1320. https://doi.org/10.1080/15481603.2022.2110197