Impacts of error on the predicted pattern of change in a post-classification change analysis

Amy C. Burnicki
Department of Geography
University of Wisconsin-Madison
Madison, WI, USA
burnicki@wisc.edu

Abstract—**One of the most common uses of time-series classified imagery is the monitoring of changes in land-cover composition and structure over time. A common approach to map changes in land-cover is post-classification change detection. The resulting accuracy of a post-classified map of change depends directly on the patterns of error associated with the time-series land-cover maps. This research examined the impact of classification error on the spatial pattern of change observed in a change map. A series of land-cover-change maps were produced using a simulation approach that controlled the: 1. amount and pattern of change occurring between the time-1 and time-2 classified maps; and 2. amount and pattern of classification error associated with the time-1 and time-2 classified maps. Both error-free and error-perturbed maps of change were produced and compared using landscape pattern indices. Results showed an increase in fragmentation within the land-cover-change classes (e.g., increase in number of land-cover-change patches, total edge, shape complexity) under all error conditions. Fragmentation was greatest when the spatial autocorrelation of the change class increased. Regardless of the pattern of change considered, errors associated with classified maps significantly altered the pattern of change simulated in the post-classification change analysis.**

Keywords: land-cover change; simualtion; error propagation; landscape metrics

I. INTRODUCTION

Improving our understanding of the uncertainty associated with a map of land-cover change is needed given the importance placed on quantifying the causes and impacts of a changing landscape. Researchers in the field of land change science have emphasized the value of linking observed spatial patterns of land-cover change to the underlying driving forces of the change in order to explain the dynamics of the land-change system (e.g., Nagendra, Munroe and Southworth, 2004). The accuracy of this linkage and our ability to understand the underlying processes that give rise to observed changes in landscape structure depends on an accurate quantification of the spatial pattern of the observed land-cover changes.

A common approach to quantify landscape structure, as observed in categorical land-cover/-use maps, is the application of landscape pattern indices. Numerous studies have used a variety of landscape pattern indices to track changes in landscape composition and structure across time-series classified maps. Typically, indices are calculated for each land-cover class within the classified maps and changes in index values are compared over time. There has also been a growing number of studies that calculate landscape pattern indices for change trajectory classes produced through post-classification comparison; i.e., the landscape pattern metrics are directly calculated on classes of land-cover change (e.g., Nagendra, Southworth and Tucker, 2003; Zhou, Li and Kurban, 2008). The most common approach to map changes in land-cover is post-classification change-detection, where two or more classified land-cover maps acquired for the same area over time are overlaid and changes in the assigned land-cover class are recorded. Certainly, the accuracy of a map of change produced using post-classification comparison depends on the magnitude and structure of the classification errors associated with each land-cover map in the time-series. How these classification errors propagate through a landscape pattern analysis has been a subject of investigation by several researchers (Wickham, O'Neill, Riitters, Wade and Jones, 1997; Brown, Duh and Drzyzga, 2000; Langford, Gergel, Dietterich and Cohen, 2006; Linke *et al.*, 2009), and an increased understanding of the sensitivity of landscape pattern indices to classification errors has been advocated (Shao and Wu, 2008).

Work by Wickham, O'Neill, Riitters, Wade and Jones (1997) and Langford, Gergel, Dietterich and Cohen (2006) demonstrated the benefits of using a simulation-based approach to investigate the impacts of classification error on landscape pattern indices. Both studies simulated a series of error-perturbed classified maps to explore the effect of varying magnitudes and spatial patterns of error on landscape pattern indices. The work by Langford, Gergel, Dietterich and Cohen (2006) was particularly successful in demonstrating the impacts of classification error on pattern indices over a range of landscapes by also controlling the underlying structure of the landscape in the simulation. Their work clearly illustrated that errors in pattern indices vary with respect to the relative proportions of class types and the level of spatial autocorrelation in the simulated landscape.

The research objective of this work was to continue the investigation of the impact of classification error on landscape pattern indices by examining how classification error propagates through a post-classification change analysis and impacts the spatial structure of observed land-cover changes. A limitation of many of the aforementioned studies is their examination of a single thematic map; i.e., they do not consider the impact of error on landscape change. The purpose of this work was to directly assess the interaction between the pattern of change and pattern of error occurring within a time-series of classified maps and determine whether the impact of classification error on observed land-cover changes varies with respect to the underlying structure of the landscape and the spatial characteristics of the classification errors.

Because classified maps display different patterns of error, and due to the great variety in the patterns of change observed in post-classification change analyses, several patterns of error and patterns of change were investigated. By examining a range of change patterns and a series of error patterns, a more thorough and comprehensive assessment of the impact of classification error on the observed spatial structure of the change map resulted.

II. Research Design

This work builds on previous research efforts to increase our understanding of the interactions between patterns of error and patterns of change in post-classification change analysis (Burnicki, Brown and Goovaerts, 2007). A simulated modeling approach was created in which the pattern of change occurring between the time-1 and time-2 land-cover maps and the patterns of error associated with the time-1 and time-2 land-cover maps were controlled during simulation. The result of a single model run was the creation of an error-free map of change and a series of error-perturbed change maps. By comparing these maps, it was possible to quantify the impact of classification error on the overall accuracy and spatial pattern of the resulting map of change. In the initial application of this simulation modeling approach, the error-free and error-perturbed maps of change were compared using summary statistics – the overall accuracy of the map of change and user's accuracy for the change classes, i.e., statistics based on the change-classification error matrix. A limitation of this analysis was that the accuracy of the spatial patterns of change observed in the error-perturbed maps of change was not assessed.

In this research study, I examined the impact of classification error on the spatial pattern of land-cover change observed in maps of change by comparing simulated error-free and error-perturbed maps of change using landscape pattern indices. The evaluation of the change maps based on pattern indices enabled the direct assessment of whether the structure of the predicted land-cover changes was significantly altered by the presence of classification error. Specifically, this research investigated the following research questions: 1. How do the values of the landscape pattern indices calculated to compare the error-free and error-perturbed maps of change differ with respect to the spatial pattern of error simulated for the time-series classified maps?; 2. How do the values of the landscape pattern indices for the error-free and error-perturbed maps of change differ when the temporal dependence between classification errors within the time-series increases?; and 3. Do the observed changes in landscape pattern indices under varying spatial patterns of error depend on the pattern of change simulated to produce the subsequent time-2 map?

III. Methods

A. *Simulating Maps of Change*

The simulation approach used in this study to produce the series of error-free and error-perturbed maps of change is described in detail in Burnicki, Brown and Goovaerts (2007). A brief overview is presented here, as it pertains to the simulated maps created and used in this analysis.

The simulated time-1 and time-2 classified maps were composed of two land-cover classes occurring in equal proportions (forest and not forest, each occupying 50% of the landscape). This resulted in four possible land-cover-change classes (remained forest, change to forest, change to not forest and remained not forest), in which each land-cover-change class occurred in approximately equal proportions. Although this research used simulated time-series classified maps, the parameters specified for all simulated surfaces, minus the error surfaces, were based on land-cover and land-cover change maps created from Landsat imagery acquired for Washtenaw County, Michigan.

The simulation approach is divided into two main tasks: A. creation of the error-free map of change; and B. creation of a series of error-perturbed maps of change (Fig. 1). The creation of the error-free map of change required the simulation of two map surfaces – the initial time-1 map and the change surface used to produce the subsequent time-2 map. The simulated initial time-1 map was held constant in this study; i.e., all model runs were based on the same time-1 map.

Unlike the time-1 map, twelve different patterns of change were simulated in this analysis to fully explore the impact of classification error across a variety of landscape structures. The simulated maps of change exhibited one of two change patterns – land-cover changes either occurred randomly across the mapped surface or were correlated to the time-1 land-cover-class boundaries. Change maps were further subdivided based on the amount of simulated change or the strength of the correlation, and then by the degree of spatial autocorrelation associated with the change classes. For change maps exhibiting random change, two amounts of

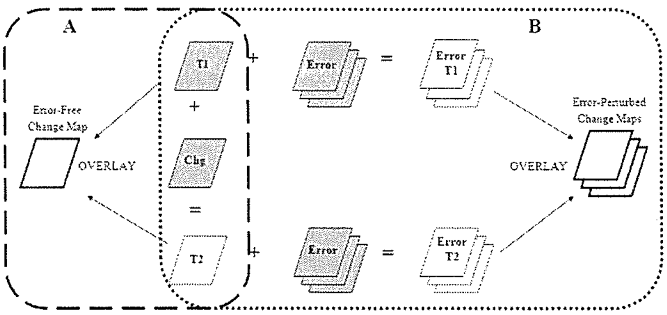

Figure 1. Overall design of the simulated modeling framework, divided into its two component parts: creation of an error-free map of change (A) and creation of a series of error-perturbed maps of change. Surfaces in grey were simulated.

total change between the time-1 and time-2 surfaces were defined: a low percentage of overall change (~12% total change) and a high percentage of overall change (~22% total change). For change maps in which changes were correlated to class boundaries, two correlation values were defined: a weak correlation of change to land-cover-class boundaries (0.3) and a strong correlation (0.7). Note: All correlated change maps exhibited a low percentage of overall change. Finally, all change maps were replicated by varying the degree of spatial autocorrelation associated with the change classes. A low, medium and high level of spatial autocorrelation was simulated for each change pattern.

The error-perturbed maps of change were created by simulating a series of error surfaces for both the time-1 and time-2 maps (Fig. 1). Unlike previous surfaces, the parameters defined to simulate the error surfaces were not based on real-world data. Instead an overall misclassification rate of 25% was defined for all error surfaces, resulting in error-perturbed time-1 and time-2 maps with overall accuracies of 75%. Additionally, the degree of spatial autocorrelation in error for all surfaces was held constant and defined to fall between the low and medium level of spatial autocorrelation used to produce the change surfaces. This decision was made to specifically address the impact of increasing spatial autocorrelation in change on landscape pattern metrics when the degree of spatial autocorrelation in error was held constant.

Five patterns of error were simulated to investigate the impact of varying patterns and increasing temporal dependence in classification error on the observed spatial structure of land-cover change. Two patterns of error were simulated for the time-1 error surfaces – the first pattern allowed errors to randomly occur across the mapped surface, while the second had errors weakly correlated (0.3) to land-cover-class boundaries. For the first pattern of error (random time-1 error), the error surfaces for the time-2 maps displayed increasing levels of temporal correlation to their time-1 counterparts, ranging from 0 (or no correlation) to 0.2 to 0.4. For the second pattern of error (correlated time-1 error), the companion error surfaces at time-2 displayed two levels of temporal correlation – 0.2 and 0.4. Thirty versions of each error pattern (i.e., realizations) were produced to account for random variability in the simulation process.

In summary, the simulation model was run sixty times, accounting for the twelve patterns of change and five patterns of error, with each run resulting in the creation of thirty error-perturbed maps of change and a single error-free change map.

B. *Evaluating Change Fragmentation*

To compare the spatial structure of land-cover change observed in the error-perturbed suite of change maps to their companion error-free map of change, a series of landscape pattern indices were calculated using FRAGSTATS version 3 (McGarigal, Cushman, Neel and Ene, 2002). The following landscape indices were calculated at the class-level (i.e., values were summarized across all patches that belonged to the same land-cover-change class): number of patches (NP), total amount of edge (TE), area of each patch summarized by determining the average and standard deviation (AREA_MN and AREA_SD), complexity of patch shape summarized by determining the average and standard deviation (SHAPE_MN and SHAPE_SD), distance to the nearest neighboring patch of the same land-cover-class type summarized by determining the average and standard deviation (ENN_MN and ENN_SD), and a measure of the aggregation or 'clumpiness' of the landscape (CLUMPY; see McGarigal, Cushman, Neel and Ene, 2002 for a full description of indices).

The aforementioned indices were selected to quantify both landscape composition and configuration; i.e., the relative proportions of each land-cover-class type and the spatial distribution of land-cover-class patches. Their selection was based on the belief that they would be most successful in capturing the hypothesized increased fragmentation in the changed landscape resulting from the inclusion of classification error. Patches were defined as contiguous pixels of the same land-cover-class type based on orthogonal neighbors, in which the four pixels sharing a common boundary were considered for patch membership. It is anticipated that the error-perturbed maps of change would display increased values for all landscape pattern metrics except ENN_MN.

IV. Results

For the purposes of this paper, only results for a single pattern of error are presented – errors occurring randomly and independently between the time-1 and time-2 error surfaces. Similar trends were observed across all patterns of error with regards to the relationship between the values of the landscape pattern indices and patterns of change. However, there were differences between the patterns of error with respect to the magnitude of index values and their relative increases or decreases across the patterns of change. Additionally, only the results for a single land-cover-change class (change to forest) are discussed because both land-cover-change classes had extremely similar index values and changes across the various patterns of change.

The introduction of classification error notably increased the degree of fragmentation in the error-perturbed maps of change (Fig. 2), impacting both the persistent and transitioning land-cover-change classes. Table 1 confirmed the significance of this increase and illustrates NP and ENN_MN, which represent indices that experienced large changes in value when comparing the error-free and error-perturbed change maps. For all patterns of change, the number of patches increased and the average distance to the nearest neighboring patch of the same land-cover-change type decreased when classification errors were incorporated into the time-series. In other words, both indices captured the tendency for error to fracture patches of land-cover-change, resulting in an increased number of patches of the same type occurring in close proximity. The increase in the number of patches was particularly dramatic, with increases of greater than 600 new patches for some error-perturbed maps of change. Relative to the number of patches in the error-free change map, this represented increases of 40% more patches in the landscape.

Examining differences in NP and ENN_MN across the patterns of change, a few notable trends emerge. First, increasing the level of spatial autocorrelation in change resulted in an increase in the number of change patches. In other words, the change maps with highest levels of clustering in change experienced the greatest fragmentation. Second, the differences between index values for the error-free and error-perturbed maps of change were smallest for the

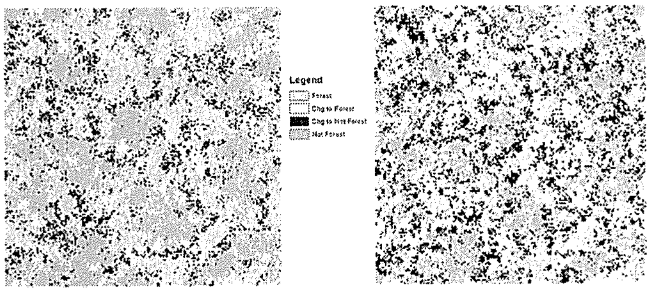

Figure 2. Comparison of an error-free and error-perturbed map of change. This example illustrates the pattern of change exhibiting a high percentage of total change randomly occurring across the mapped extent and the pattern of error exhibiting random, independent time-1 and time-2 errors.

maps exhibiting the highest amount of overall change – i.e., greater similarity in the structure of the changed landscape resulted when more of the landscape experienced a change in land-cover. In addition, there was a greater similarity between the error-free and error-perturbed change maps when changes were increasingly correlated to classification boundaries (0.7 versus 0.3).

V. Discussion And Conclusions

The comparison of the error-free and error-perturbed maps of change illustrated significant changes in value for the selected landscape pattern indices calculated for each land-cover-change class, with dramatic increases in the number of land-cover-change patches and a decrease in the distances between like patches. Thus, the propagation of classification error in a post-classification change analysis has the potential to greatly increase the degree of fragmentation observed in a map of change. The sensitivity of NP to the addition of classification error follows results of previous investigations

TABLE I. DIFFERENCE BETWEEN THE AVERAGE LANDSCAPE INDEX VALUE ACROSS THE 30 ERROR-PERTURBED MAPS OF CHANGE MINUS THE INDEX VALUE FOR THE ERROR-FREE MAP OF CHANGE, SUMMARIZED FOR ALL PATTERNS OF CHANGE AND THE PATTERN OF ERROR EXHIBITING RANDOM, INDEPENDENT TIME-1 AND TIME-2 ERRORS.

Random Change

	Low % of Overall Chg			High % of Overall Chg		
	Low SpAuto	Med SpAuto	High SpAuto	Low SpAuto	Med SpAuto	High SpAuto
NP	509	584	643	146	285	330
ENN_MN	-0.39	-0.37	-0.37	-0.18	-0.16	-0.10

Correlated Change

	Weak Correlation (Low % Chg)			Strong Correlation (Low % Chg)		
	Low SpAuto	Med SpAuto	High SpAuto	Low SpAuto	Med SpAuto	High SpAuto
NP	452	587	646	431	566	619
ENN_MN	-0.31	-0.31	-0.23	-0.25	-0.23	-0.21

(Brown, Duh and Drzyzga, 2000; Linke *et al.*, 2009). The increased level of fragmentation under increasing spatial autocorrelation in change may suggest that larger land-cover change patches have a greater potential to fracture. Examining the change patterns with the lowest level of spatial autocorrelation, the error-perturbed maps of change had both the lowest number of new land-cover-change patches and the largest change in shape complexity. This suggests the structure of the land-cover-change patches changed as opposed to an increase in the number of patches occurring in the landscape.

The increase in the degree of fragmentation was also related to the total amount of change occurring within the time-series and the underlying pattern of change. Thus, predicting the impact of error on landscape pattern indices is made more challenging by the interaction between the patterns of error and change occurring in the time-series. The limited results presented in this analysis suggest that landscape indices will have increased accuracy when a higher proportion of the map undergoes change and as changes are increasingly tied to land-cover class boundaries. However, given the relatively small percentage of a landscape experiencing a change in land-cover in typical land-cover change analyses, measurements of landscape structure, as calculated using landscape pattern indices, must be treated with caution.

References

Brown, D.G., Duh, J.-D. and Drzyzga. (2000). Estimating error in an analysis of forest fragmentation change using North American Landscape Characteristization (NALC) data. *Remote Sensing of Environment*. 71, 106-117.

Burnicki, A.C., Brown, D.G. and Goovaerts, P. (2007). Simulating error propagation in land-cover change analysis: the implications of temporal dependence. *Computers, Environment and Urban Systems*. 31, 282-302.

Langford, W.T., Gergel, S.E., Dietterich, T.G. and Cohen, W. (2006). Map misclassification can cause large errors in landscape pattern indices: examples from habitat fragmentation. *Ecosystems*. 9, 474-488.

Linke, J., McDermid, G.J., Pape, A.D., McLane, A.J., Laskin, D.N., Hall-Beyer, M. and Franklin, S.E. (2009). The influence of patch-delineation mismatches on multi-temporal landscape pattern analysis. *Landscape Ecology*. 24, 157-170.

McGarigal, K., Cushman, S.E., Neel, M.C. and Ene, E. (2002). *FRAGSTATS: spatial pattern analysis program for categorical maps.* Computer software program produced at the University of Massachusetts, Amherst. Available online: www.umass.edu/landeco/research/fragstats/fragstats.html.

Nagendra, H., Southworth, J. and Tucker, C. (2003). Accessibility as a determinant of landscape transformation in western Honduras: linking pattern and process. *Landscape Ecology*. 18, 141-158.

Nagendra, H., Munroe, D.K. and Southworth, J. (2004). From pattern to process: landscape fragmentation and the anlysis of land use/land cover change. *Agriculture, Ecosystems and Environment*. 101, 111-115.

Shao, G. and Wu, J. (2008). On the accuracy of landscape pattern analysis using remote sensing data. *Landscape Ecology*. 23, 505-511.

Wickham, J.D., O'Neill, R.V., Riitters, H.K., Wade, T.G. and Jones, K.B. (1997). Sensitivity of selected landscape metrics to land-cover misclassification in land-cover composition. *Photogrammetric Engineering and Remote Sensing*. 63(4), 397-402.

Zhou, Q., Li, B. and Kurban, A. (2008). Spatial pattern analysis of land cover change trajectories in Tarim Basin, northwest China. *International Journal of Remote Sensing*. 29 (19), 5495-5509.