

## A NEW MEASURE OF CLASSIFICATION ERROR: DESIGNED FOR LANDSCAPE PATTERN INDEX

X. H. Chen <sup>a,\*</sup>, Y. Yamaguchi <sup>a</sup>, J. Chen <sup>b</sup>

<sup>a</sup> Graduate School of Environmental Studies, Nagoya University, Nagoya, 464-8601, Japan –  
chen.xuehong@j.mbox.nagoya-u.ac.jp

<sup>b</sup> State key laboratory of earth surface processes and resource ecology, Beijing Normal University, Beijing, 100875, China

Commission VIII, WG VIII/8

**KEY WORDS:** Landscape Pattern Index; Classification Error; Accuracy Assessment Index

### ABSTRACT:

Classified thematic maps based on remote sensing data are usually used to derive Landscape Pattern Index (LPI). However, the classification error can be propagated into the LPIs calculation, while it is usually ignored in the previous literatures. Correctly estimating the LPI error is vital for reliable landscape analysis; however, the widely accepted accuracy assessment method without considering spatial information is not suitable for indicating LPI error. In this paper, we developed a new accuracy assessment index by giving a certain weight to each error pixel according to its isolated degree. The result shows that the new index is a good predictor for the error of Number of Patches (NP), Total Edge (TE) and Aggregation Index (AI). The effect of sample size was also studied, showing that the correlation between new index and LPI error become higher and more stable when the sample size increases. The results suggest that the proposed new index is potential to be a practicable measure of classification error for landscape analysis as supplement of overall accuracy and Kappa coefficient.

### 1. INTRODUCTION

The quantification of spatial heterogeneity is one of the most important topics in landscape ecology (Turner, 2005). Since O'Neill, et al. (1988) firstly proposed the landscape pattern index (LPI) to quantitatively describe the landscape pattern, a large number of LPIs have been developed for categorical data. Meanwhile, an excellent software package, Fragstats (McGarigal and Marks, 2002) is also available for calculating LPIs, which stimulates a broad application of LPI in various fields, e.g. habitat conservation, urban planning, etc. (Turner, et al, 2001).

As remote sensing is an ideal tool to investigate the wide range of the earth surface with low cost, classified thematic maps based on remote sensing data were usually used for deriving LPIs (Newton, et al, 2009). Unfortunately, inevitable error in the classified result can be propagated into the LPIs calculation, while it is usually ignored in the previous literatures (Shao & Wu, 2008). Although this problem has been pointed out early (Hess, 1994), few study examined the impacts of classification errors on the calculation of LPI (Shao & Wu, 2008). Difficulty may be caused by the unstable relationship between the LPI error and commonly used accuracy assessment index (Langford et al, 2006). Traditional classification accuracy assessment index focuses on the amount of error only without considering spatial information, consequently, can not be suitable for indicating LPI error as the spatial heterogeneity is the key meaning of LPI.

Landscape analysis is hardly reliable without a correct estimation of LPI error. Therefore, there is a great need to propose a new measure of classification error, which can be a

reliable predictor for LPI error. One possible solution is to develop the object-based assessment method (e.g. Dungan, 2006; Zhan et al, 2009). Although spatial information is considered, this type of method is not practicable for remotely sensed image since it is too expensive and difficult to obtain the reference land cover classification for the entire region of interest (Stehman, 2009).

As problems mentioned above, an ideal classification error measure designed for LPI should satisfy two requirements: 1) highly correlated with LPI error; 2) convenient for sampling. Aiming for the objectives, we proposed a new accuracy assessment index, which is pixel based but integrating the spatial information through considering the adjacent pixels. A series of simulated images were used to test the performance of new index.

### 2. MEASURE OF CLASSIFICATION ERROR

#### 2.1 Traditional accuracy assessment method

The error matrix or confusion matrix (shown in table 1) is the most popular method for assessing the classification error. A series of accuracy assessment indices, such as overall accuracy, user's accuracy, producer's accuracy and kappa coefficient can be derived from the confusion matrix. These indices reflect the classification error in different aspects. Among them, overall accuracy (Eq.1) and kappa coefficient (Eq.2) indicate the classification error in map level, and have been widely used for accuracy assessment of classification quality in remote sensing community.

\* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

Table 1 Confusion matrix for accuracy assessment of classification

Classified data	Referenced data				Total
	1	2	...	n	
1	n <sub>11</sub>	n <sub>21</sub>	...	n <sub>k1</sub>	n <sub>+1</sub>
2	n <sub>12</sub>	n <sub>22</sub>	...	n <sub>k2</sub>	n <sub>+2</sub>
...	...	...	...	...	...
k	n <sub>1k</sub>	n <sub>2k</sub>	...	n <sub>k</sub>	n <sub>+k</sub>
Total	n <sub>1+</sub>	n <sub>2+</sub>	...	n <sub>k+</sub>	N

$$\text{Overall accuracy} = \frac{1}{n} \sum_{i=1}^k n_{ii} \quad (1)$$

$$\text{Kappa} = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_{i+} n_{+i}}{n^2 - \sum_{i=1}^k n_{i+} n_{+i}} \quad (2)$$

These traditional accuracy indices measure the degree to which the derived thematic map agrees with reality on the amount of error, but do not consider the spatial distribution of misclassification pixels, hindering them to be reliable indicators for LPI error.

2.2 New index to measure the classification error

LPI error is determined not only the amount of error pixels, but also and the spatial configuration of error pixels because error pixels with different spatial distribution will affect LPIs in different degree. Compared with adjacent misclassification pixels, isolated misclassification pixels more seriously affect the LPI calculation because isolated misclassification pixels enhance the fragmentation of classified map and a large group of LPIs are sensitive to the degree of fragmentation. Considering the effect of isolated misclassification pixels, we proposed a new index through giving different weights to different misclassification pixels:

$$\text{New Index} = \frac{\sum_{i=1}^m w_i}{N} \times 100\% \quad (3)$$

where *m* is the number of misclassification pixels, *N* is number of all the sampling pixels, *w<sub>i</sub>* is weight of misclassification pixel *i*, calculated as the fraction of the 8 neighbourhood pixels belonging to the different classes with the central pixel *i*. It is noticed that the weight is calculated based on the classification map instead of reference map, indicating that only point reference data is necessary for assessment. As an example shown in Fig.1, the misclassification pixel with red border in case (a) has 4 neighbour pixels belonging to the different classes with the central pixel, while the misclassification pixel in case (b) has 7 neighbor pixels belonging to the different classes with the central pixel. Therefore, the weight of the misclassification pixel in case (a) is 0.5 and that in case (a) is 0.875. Obviously the misclassification pixel in case (b) is more isolated than the one in case (a), accordingly the weight in case (b) is larger than that in case (a). When the misclassification pixel is completely isolated, the weight is equal to 1, and it is equal to 0 when the misclassification pixel is completely adjacent with the same class pixels. Through considering the spatial information of error pixels, the new accuracy assessment index is expected to predict the LIP error better than traditional indices.

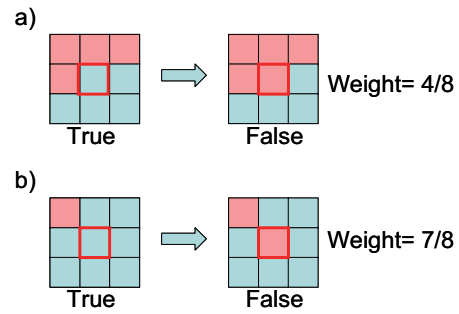


Fig. 1 Illustration of the weight calculation (a. adjacent misclassification pixel; b. isolated misclassification pixel)

3. SIMULATION STUDY

3.1 Data Simulation

A series of correct and incorrect classification maps were simulated in this study instead of using actual classification and reference maps, because it is very difficult to collect a larger number of actual classification maps from literature to perform our analysis. A map simulation software entitled as Simmap (Saura and Martinez-Millan, 2000) was used to generate a total of 75 correct base maps to represent the various landscapes without classification error (See Fig.2 for example). The size of the map is 200x200-pixel and the number of classes is three. We varied the class proportions and aggregation level. There were five sets of proportion configuration: 0.125-0.125-0.75; 0.125-0.25-0.625; 0.125-0.375-0.5; 0.25-0.25-0.5; 0.333-0.333-0.333. And three levels of aggregation (*p*), 0.1, 0.3 and 0.5, were considered. For each combination of proportion and aggregation configuration, five maps were generated.

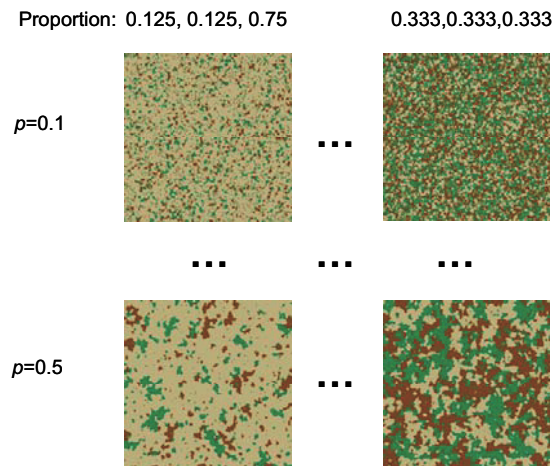


Fig. 2 Example of simulated "correct" landscapes with a range of proportion and aggregation configurations

We created incorrect maps through changing the class of certain proportion of pixels to the other two classes, and the transition probabilities of two classes are both equal to 0.5 (Fig.3). Regarding the spatial distribution of classification error, two types of error were considered: (a) randomly located misclassification; (b) double misclassification rates near patch boundaries. For each correct based map, five levels of classification error rate were added (2%, 4%, 6%, 8%, 10% for randomly located misclassification; 4%, 8%, 12%, 16%, 20% for misclassification of edge pixels). Therefore, totally 75 x 5=375 incorrect images were simulated.

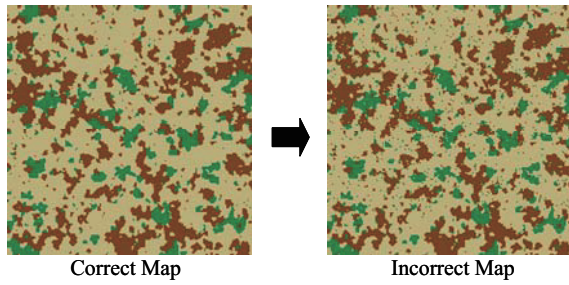


Fig. 3 Example of the simulated classification error

### 3.2 Measuring Fragmentation and LPI Error

We applied FRAGSTATS 3.0 (McGarigal et al, 2002) to calculate the LPIs. In this study, four LPI indices were calculated: (a) Number of Patches (NP); (b) Mean Patch Size (MPS); (c) Total Edge (TE); (d) Aggregation Index (AI). All of these indices were calculated on both the landscape level and class level.

For each LPI on landscape level, we defined the LPI error as follow:

$$LPI\_error = (LPI_{correct} - LPI_{incorrect}) / (LPI_{max} - LPI_{min}) \quad (4)$$

where  $LPI_{correct}$  is LPI of correct classification map,  $LPI_{incorrect}$  is LPI of incorrect classification map,  $LPI_{max}$  and  $LPI_{min}$  are the maximum and minimum value of LPIs in the series of simulated correct based maps, which normalize all the LPI error in the same scale.

### 3.3 Measuring the Classification Error

Overall accuracy, Kappa coefficient and the proposed new index were used to measure the classification error. Corresponding to the landscape level and class level of LPIs, both of these two levels of accuracy assessment indices were calculated. For the landscape level, the indices are consistent with the original definition (Eq.1-3). For the class level, some minor modification was conducted. Firstly, three-class map is combined into two-class map, one class is the objective class and the other one is the combination of the background classes. Then, the accuracy assessment indices are calculated based on the two-class map.

### 3.4 Analysis of Result

Based on the 375 simulated classification maps, we investigated the correlation between LPI error and the accuracy assessment indices on both the landscape level and class level. For the class level, we chose only one class to analyze because there are no differences among three classes.

Sample size is an important factor in assessment of classification error. The effect of sample size was also examined for the new index. For each incorrect classification map, new index was calculated based on a range of sample sizes, ranging from 10% to 90% with a step of 10%. We randomly sample 10 times for each level of sample size. Then, the average and the variation of the correlation coefficient for each sample size were prepared for analysis.

## 4. RESULT

### 4.1 Correlation Between LPI Error and Classification Accuracy Indices

Here, we used the stratified random sampling method and 10% sampling size, since 10% is close to the practicable sampling amount. Table 2 and 3 show the correlation coefficients between LPI error and the classification accuracy assessment indices respectively on landscape level and class level. The new index is better correlated with LPIs error than overall accuracy and kappa. Especially, the error of NP, TE, and AI are well correlated with new index. However, the error of MPS is not well correlated with new index, because MPS describe the shape information of the patches, which are not necessarily related with the degree of fragmentation.

Table 2 Correlation coefficient between classification accuracy assessment indices and LPI error on landscape level

Correlation Coefficient	Overall Accuracy	Kappa	New Index
Number of Patches	-0.574**	-0.659**	0.967**
Total Edge	-0.801**	-0.836**	0.952**
Mean Patch Size	0.029	-0.086	0.630**
Aggregation Index	-0.801**	-0.835**	0.952**

\*\* Correlation is significant at the 0.01 level (2-tailed)

Table 3 Correlation coefficient between classification accuracy assessment indices and LPI error on class level

Correlation Coefficient	Overall Accuracy	Kappa	New Index
Number of Patches	-0.430**	-0.759**	0.884**
Total Edge	-0.809**	-0.867**	0.932**
Mean Patch Size	0.035	-0.144**	0.445**
Aggregation Index	-0.641**	-0.701**	0.856**

\*\* Correlation is significant at the 0.01 level (2-tailed)

### 4.2 Effect of Sample Size

Fig.4 shows the correlation coefficient against sample size respectively on the landscape level. MPS was not considered because the new index can not predict the error of MPS well (Section 4.1). It can be seen that the correlation between new index and LPI error increases and the variation of correlation decrease, when the sample size increases. This result is consistent with other studies on hard and soft classification (Stehman, 1996; Chen et al., 2010).

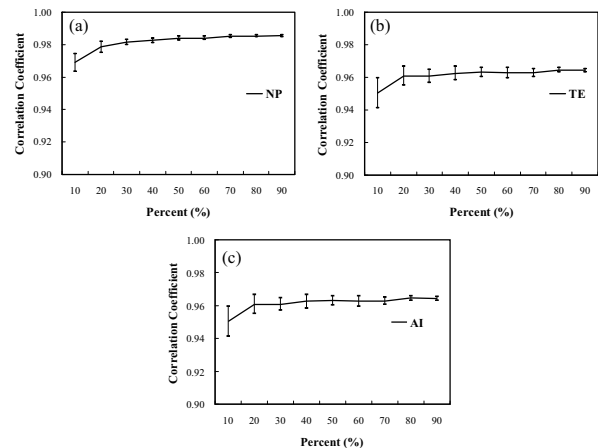


Fig. 4 Correlation between LPI error and new accuracy assess index using different sample size (a. NP, b. TE, c. AI)

## 5. DISCUSSION AND CONCLUSION

Estimating LPI error is important for landscape analysis, while traditional accuracy assessment indices can not predict LPI error well because they do not consider spatial information. In this paper, a new accuracy assessment index is proposed to indicate LPI error. Through analysis of a series of simulated images, the new index is found to be well correlated with error of NP, ED and AI, indicating that new index is a better predictor for these LPIs' error compared with traditional indices. However, different LPIs reflect different aspects of landscape pattern and it is impossible for one index to indicate error of all the LPIs. MPS are found not to be correlated well with new index in this study. However, the new index is still very meaningful as fragmentation is a very important characteristic of landscape pattern. New index is potential to be a reliable indicator of the error of LPI related with fragmentation.

One limitation of this study is that our conclusion is derived based on simulated data. The gap between simulated and actual data is obvious and difficult to be eliminated. However, the simulated data still can behalf of some representative cases. Most characteristics of landscape patterns can be captured by Simmap model (Li et al., 2004). Also, although the spatial autocorrelation is not considered, the error model can well represent the salt-and-pepper error, which usually occurs for per-pixel classification methods. On the other hand, simulation study provides sufficient data for analysis, through which statistically meaningful result can be derived. Therefore, the conclusion of this study is valid in many cases, and the proposed new index is potential to be a practicable measure of classification error for landscape analysis as supplement of overall accuracy and Kappa coefficient.

## REFERENCE

Chen J, Zhu X, Imura H et al, 2010. Consistency of accuracy assessment indices for soft classification. *ISPRS J Photogramm*, 65, pp. 156-164.

Dungan JL, 2006. Focusing on feature-based differences in map comparison. *J Geogr Syst*, 8, pp. 131-143.

Herold M, Scepan J, Clarke KC, 2002. The use of remote sensing and landscape metrics to describe structures and changes in urban land uses. *Environ Plann A*, 34, pp. 1443-1458.

Hess GR, 1994. Pattern and error in landscape ecology: a commentary. *Landscape Ecol*, 9, pp. 3-5.

Langford WT, Gergel SE, Dietterich TG et al, 2006. Map misclassification can cause large errors in landscape pattern indices: example from habitat fragmentation. *Ecosystems*, 9, pp. 474-488.

Li X, He HS, Wang X et al, 2004. Evaluating the effectiveness of neutral landscape models to represent a real landscape. *Landscape Urban Plan*, pp. 137-148.

McGarigal K, Cushman SA, Neel MC et al, 2002 FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the

University of Massachusetts, Amherst. Available at the following web site: [www.umass.edu/landeco/research/fragstats/fragstats.html](http://www.umass.edu/landeco/research/fragstats/fragstats.html).

Newton AC, Hill RA, Echeverría C et al, 2009. Remote sensing and the future of landscape ecology. *Prog Phys Geog*, 33, pp. 528-546.

O'Neill RV, Krummel JR, Gardner RH et al, 1988. Indices of landscape pattern. *Landscape Ecol*, 1, pp. 153-162

Saura S, Martinez-Millan J, 2000. Landscape patterns simulation with a modified random clusters method. *Landscape Ecol*, 15, pp. 661-678.

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

Stehman SV, 1996. Estimating the kappa coefficient and its variance under stratified random sampling. *Photogramm Eng Rem Sens*, 52, pp. 401-407.

Stehman SV, 2009. Sampling designs for accuracy assessment of land cover. *Int J Remote Sens*, 30, pp. 5243-5272.

Turner MG, 2005. Landscape Ecology: What is the state of the science? *Landscape Ecol*, 36, pp. 319-344.

Turner MG, Gardner RH, O'Neill RV, 2001. Landscape ecology in theory and practice: pattern and process. Berlin Heidelberg, New York: Springer.

Zhan Q, Molenaar M, Tempfli K et al, 2009. Quality assessment for geo-spatial objects derived from remotely sensed data. *Int J Remote Sens*, pp. 2953-2974.