An Explicit Index for Assessing the Accuracy of Cover-Class Areas

Guofan Shao, Wenchun We, Gang Wu, Xinhua Zhou, and Jianguo Wu

Abstract
We present a new index, called Relative Errors of Area (REA), for assessing the accuracy of cover-class areal percentage (%LAND) that is extracted from thematic maps after classifying remotely sensed data. We demonstrate how to derive REA from an error matrix and its relationship with user's and producer's accuracy. We compare the REA index with other accuracy indices in a hypothetical and two real case studies. The accuracy of cover-class areal estimates is highly correlated with the REA index, but not with other classification accuracy indices such as the overall classification accuracy. In general, users should beware of using thematic maps with low REA values. Moreover, the estimates of cover-class area can be revised by using REA if cell values of the major diagonal in an error matrix are available.

Introduction
Cover-class areal percentage, or %LAND, commonly derived from thematic maps that are produced by classifying remotely sensed data, is broadly used to measure natural resources, evaluate ecological conditions, and quantify land-use and land-cover changes in space and time (O'Neill *et al.*, 1988; McGarigal and Marks, 1994; Frohn, 1998). Unfortunately, all thematic maps contain classification errors. Thus, statistical analyses with %LAND inevitably contain errors. Furthermore, the uncertainties of %LAND are not readily estimable with the existing classification accuracy measures (Hess, 1994; *Shao et al.*, 2001; Wu and Shao, 2002). Without knowing the uncertainties of landscape metrics, it is difficult to compare different landscapes, to detect landscape change over time, or to relate detected landscape pattern to ecological processes (Hess, 1994). Readers could be even wrongly impressed by the values of landscape metrics if an explicit accuracy estimate is unknown. Unless landscape indices can be explicitly assessed, quantitative analysis with landscape indices is still a controversial issue.

Studies have addressed the issues relevant to the problem of the accuracy or uncertainty in using landscape indices. The studied issues include correlations among landscape indices (Riitters *et al.*, 1995; Cain *et al.*, 1997), scale factors of landscape mapping with remote sensing data (Woodcock and Strahler, 1987; Moody and Woodcock, 1995), scale effects on landscape analysis (Wickham and Riitters, 1995; Jelinski and Wu, 1996; Qi and Wu, 1996; Wu *et al.*, 1997; Fortin, 1999), sensitivity of landscape indices to land-cover misclassification and composition (Moody and Woodcock, 1995; Wickham *et al.*, 1997; Shao *et al.*, 2001), uncertainties of diversity indices derived from error matrices (Hess and Bay, 1997), and error propagation with environmental modeling (Jelinski and Wu, 1996; Heuvelink, 1998). However, little success has been achieved regarding the effect of classification errors on the accuracy of landscape indices.

With the advent of more advanced digital satellite remote sensing techniques, the necessity of performing accuracy assessments of image classification has received renewed interest (Congalton, 1991). Classification accuracy can be assessed with various measures or statistics derived from an error matrix (Congalton and Green, 1999). These statistics include overall accuracy, producer's accuracy, user's accuracy, Kappa statistic (including weighted Kappa statistic), and classification success indices (Koukoulas and Blackburn, 2001). On the one hand, each accuracy measure is appropriate for certain, but not for all, applications. On the other hand, the overall accuracy or Kappa statistic is more commonly used in practice than user's and producer's accuracy. In this paper, we compare these accuracy statistics with a new index, called Relative Errors of Area (REA), in estimating the uncertainty of %LAND, which is defined as the total area of a cover class divided by the total landscape area multiplied by 100 (McGarigal and Marks, 1994; http://flash.lakeheadu.ca/~rrempel/patch/; last accessed 13 January 2003).

An Illustration of Classification Errors with a Contrived Data Set
In a hypothetical situation (Figure 1), two classifications have the same overall accuracy of 90 percent. For the first classification, the producer's accuracy is 79 percent for class 1 and 100 percent for class 2, and user's accuracy is 100 percent for class 1 and 83 percent for class 2; for the second classification, the producer's accuracies of class 1 and 2 are 100 percent and 80 percent, respectively, and user's accuracy is 83 percent for class 1 and 100 percent for class 2. Although the overall accuracy is relatively high or satisfactory in practice, which is higher than the minimum level of 85 percent defined by the U.S. Geological Survey (Anderson *et al.*, 1976), the area estimations based on the two classifications are obviously different. While the

G. Shao is with the Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110015, China, and the Department of Forestry and Natural Resources, Purdue University, 195 Marsteller Street, West Lafayette, IN 47907–2033 (shao@purdue.edu).

W. We is with the Department of Forestry and Natural Resources, Purdue University, 1159 Forestry Building, West Lafayette, IN 47907–2033.

G. Wu is with the Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

X. Zhou is with the School of Natural Resource Sciences, University of Nebraska, Lincoln, NE 68583.

J. Wu is with the Department of Plant Biology, Arizona State University, Tempe, AZ 85287–1601.

Photogrammetric Engineering & Remote Sensing
Vol. 69, No. 8, August 2003, pp. 907–913.

0099-1112/03/6908–907$3.00/0

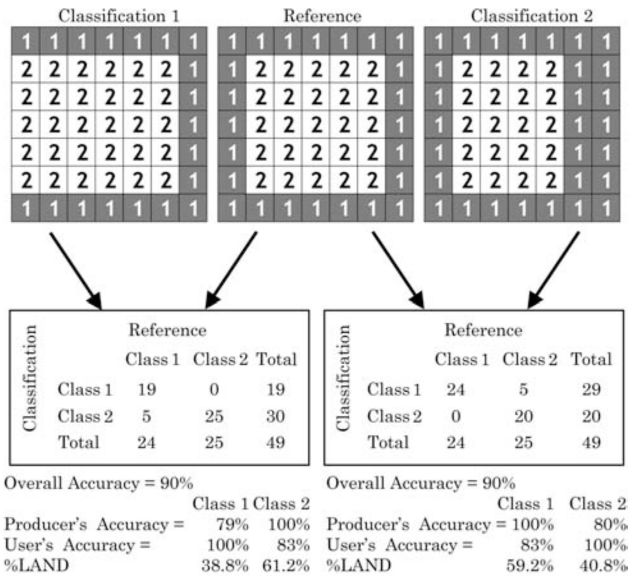

Figure 1. A hypothetical illustration to demonstrate that land-use changes can be wrongly detected even if overall classification accuracy is relatively high.

value of %LAND is about 50 percent for class 1 or 2 in the reference map, %LAND ranges from about 40 percent to 60 percent for each class in both classifications. The errors of %LAND are ±10 for each class in both classifications. The relative difference in area estimations between the two classifications is (29−19)/24 = 41 percent for class 1 and is (30−20)/25 = 40 percent for class 2.

This example has practical implications for areal estimations or determining changes in %LAND over time or across space. Unless the overall classification accuracy is 100 percent, which is impossible in reality, %LAND estimates based on remote sensing data may have sizable errors and, in turn, result in misinterpretations in further statistical or quantitative analysis.

The Derivation of Relative Errors of Area (REA)

To deal with the problem of classification errors associated with thematic maps based primarily on remote sensing data, we propose a new index that can be used to assess the accuracy of landscape area. This index provides information on the uncertainty in derived cover-class area or %LAND estimates.

If a thematic map contains n classes, its accuracy can be assessed with the following error matrix (Congalton, 1991; Congalton and Green, 1999):

$$\begin{array}{c|cccc|c} & & \text{Reference} & & & \text{Row Total} \\ \hline \text{Classification} & f_{11} & f_{12} & \cdots & f_{1n} & f_{1+} \\ & f_{21} & f_{22} & \cdots & f_{2n} & f_{2+} \\ & \vdots & \vdots & \vdots & \vdots & \vdots \\ & f_{n1} & f_{n2} & \cdots & f_{nn} & f_{n+} \\ \hline \text{Column Total} & f_{+1} & f_{+2} & \cdots & f_{+n} & N \end{array} \tag{1}$$

where N is the total number of sampling points, f_{ij} (i and j = 1, 2, ..., n) is the joint frequency of observations assigned to class i by classification and to class j by reference data, f_{i+} is the total frequency of class i as derived from the classification, and f_{+j} is the total frequency of class j as derived from the reference data.

For a given cover-class k ($1 \le k \le n$), the reference value of %LAND (LR_k) is computed as

$$LR_k = \frac{f_{+k}}{N} = \frac{\sum_{i=1}^{n} f_{ik}}{N} = \frac{\sum_{\substack{i=1 \\ i \ne k}}^{n} f_{ik} + f_{kk}}{N}. \tag{2}$$

The classification value of %LAND (LC_k) is derived as

$$LC_k = \frac{f_{k+}}{N} = \frac{\sum_{j=1}^{n} f_{kj}}{N} = \frac{\sum_{\substack{j=1 \\ j \ne k}}^{n} f_{kj} + f_{kk}}{N}. \tag{3}$$

Thus, the difference between LC_k and LR_k is

$$LC_k - LR_k = \frac{f_{k+} - f_{+k}}{N} = \frac{\sum_{j=1}^{n} f_{kj} - \sum_{i=1}^{n} f_{ik}}{N} = \frac{\sum_{\substack{j=1 \\ j \ne k}}^{n} f_{kj} - \sum_{\substack{i=1 \\ i \ne k}}^{n} f_{ik}}{N}. \tag{4}$$

If $LC_k - LR_k = 0$, there are two possibilities: classification errors are zero, or commission errors (CE) and omission errors (OE) are the same for cover-class k. The first possibility is normally untrue in reality. In many situations, the second possibility is also untrue. If $CE_k > OE_k$ and $LC_k - LR_k > 0$, the value of %LAND of class k is overestimated; if $CE_k < OE_k$ and $LC_k - LR_k < 0$, the value of %LAND of class k is underestimated. Therefore, the components of CE_k and OE_k in Equation 4 determine the accuracy of %LAND for cover-class k.

Mathematically, CE_k is given as

$$CE_k = \sum_{\substack{j=1 \\ j \ne k}}^{n} f_{kj} \tag{5}$$

and OE_k is expressed as

$$OE_k = \sum_{\substack{i=1 \\ i \ne k}}^{n} f_{ik}. \tag{6}$$

The balance between CE_k and OE_k indicates the absolute errors of area estimate for cover-class k. The relative errors of area (REA) are then defined as

$$REA_k = \frac{\sum_{\substack{j=1 \\ j \ne k}}^{n} f_{kj} - \sum_{\substack{i=1 \\ i \ne k}}^{n} f_{ik}}{f_{kk}} \times 100 \tag{7}$$

where f_{kk} is an element of the kth row and kth column in an error matrix. It represents the frequency of sample points that are correctly classified.

According to Congalton and Green (1999), User's Accuracy of class k (UA_k) can be expressed as

$$UA_k = \frac{f_{kk}}{f_{k+}} = \frac{f_{kk}}{\sum_{j=1}^{n} f_{kj}} = \frac{f_{kk}}{f_{kk} + \sum_{\substack{j=1 \\ j \ne k}}^{n} f_{kj}} \tag{8}$$

and Producer's Accuracy of class k (PA_k) can be expressed as

$$PA_k = \frac{f_{kk}}{f_{+k}} = \frac{f_{kk}}{\sum_{i=1}^{n} f_{ik}} = \frac{f_{kk}}{f_{kk} + \sum_{\substack{i=1 \\ i \ne k}}^{n} f_{ik}}. \tag{9}$$

By substituting Equations 8 and 9 into Equation 7, it is easily derived that

$$\mathrm{REA}_k = \left(\frac{1}{UA_k} - \frac{1}{PA_k}\right) \times 100. \tag{10}$$

Thus, REA can be obtained using information on the error matrix with Equation 7 or the user and producer's accuracy with Equation 10. If $\mathrm{REA}_k = 0$, $\%\mathrm{LAND}_k$ has no errors; $\mathrm{REA}_k > 0$, $\%\mathrm{LAND}_k$ is overestimated; if $\mathrm{REA}_k < 0$, $\%\mathrm{LAND}_k$ is underestimated. In the hypothetical example above, the first classification underestimates class 1 ($\mathrm{REA}_1 = -26.6\%$) and overestimates class 2 ($\mathrm{REA}_2 = 20.5\%$), and the second classification overestimates class 1 ($\mathrm{REA}_1 = 20.5\%$) and underestimates class 2 ($\mathrm{REA}_2 = -25.0\%$).

To calibrate areal estimates with REA_k, we introduce a constant K_k for class k as follows:

$$\%\mathrm{LAND}_{k,c} = \%\mathrm{LAND}_k + K_k \times \mathrm{REA}_k \tag{11}$$

where $\%\mathrm{LAND}_{k,c}$ = calibrated $\%\mathrm{LAND}_k$.

Assuming that LR_k (see Equation 2) is the unbiased estimate of $\%\mathrm{LAND}_{k,c}$ and LC_k (see Equation 3) is the unbiased estimate of $\%\mathrm{LAND}_k$, K_k can be derived as

$$K_k = \frac{\dfrac{-\sum\limits_{\substack{j=1\\ j\neq k}}^{n} f_{kj} + \sum\limits_{\substack{i=1\\ i\neq k}}^{n} f_{ik}}{N}}{\dfrac{\sum\limits_{\substack{j=1\\ j\neq k}}^{n} f_{kj} - \sum\limits_{\substack{i=1\\ i\neq k}}^{n} f_{ik}}{f_{kk}}} = -\frac{f_{kk}}{N}. \tag{12}$$

K_k can be computed as long as cell values of the major diagonal in an error matrix are available. In the hypothetical example above, $K_1 = -19/49 = -0.39$ and $K_2 = -25/49 = -0.51$ for classification 1 and $K_1 = -24/49 = -0.49$ and $K_2 = -25/49 = -0.41$ for classification 2. Both classifications result in the same values of calibrated %LAND: calibrated $\%\mathrm{LAND}_1 = 49$ percent ($38.8\% + 0.39 \times 26.6\%$ or $59.2\% - 0.49 \times 20.5$) and calibrated $\%\mathrm{LAND}_2 = 51$ percent ($61.2\% - 0.51 \times 20.5\%$ or $40.8\% + 0.41 \times 25.0\%$).

Case Studies

Case 1: Mapping a Two-Class Landscape

Two sub-scenes of Landsat Thematic Mapper (TM) data, path 116 and row 31 (128° and 42° N), were acquired on 12 September 1987 and 04 September 1993. The study site is covered with typical old-growth, broadleaved-coniferous mixed forest (Barnes *et al.*, 1993). It is one of the four vegetation zones on the north slope of Changbai Mountain lying between China and North Korea. The goal of the classification was to map the landscape into two classes: forested and clear-cut areas.

The 1987 data were rectified into a 30-m resolution image in the UTM coordinate system by referencing to 1:50,000-scale topographic maps. The 1993 data were rectified against the 1987 data, and the co-registration model RMS (root-mean-square) error was less than 0.5 pixels based on ten control points. A composite data set was made by stacking the 1993 and 1987 images. We refer to the 1993 imagery as Uni-Temporal Data (UTD) and the stacked imagery as Bi-Temporal Data (BTD) in this paper. Three approaches were used to determine the number of spectral classes in each class, and are called Image Sample (IS), Field Sample (FS), and Hybrid Sample (HS) (Wu and Shao 2002). For each spectral class, a training sample contained two fields and each field was sized around 50 pixels. Three classifiers — Minimum Distance (MD) that takes only the first-order statistics, the class mean, Quadratic Likelihood (QL) that includes both the first- and second-order statistics, and Extraction and Classification of Homogeneous Objects (ECHO) — were used for supervised classifications. MD and QL are referred to as spectral classifiers because they consider only spectral information, whereas ECHO is a spectral-spatial classifier or sample-based classifier (Kettig and Landgrebe, 1976). After initial classifications, spectral classes were grouped into two information classes: forest and clear-cut. The classification experiment resulted in 18 thematic maps (2 data × 3 training samples × 3 classifiers). A manually digitized thematic map from 1:20,000-scale infrared color aerial photographs covering an area of 7 by 12 km was used as reference data for accuracy assessment. The reference map was re-rectified and re-sampled to match the resolution of the Landsat TM data used for this study. A set of 1,000 simple-random samples (pixels) was located on the reference map, and the values for all the sample locations were used to build error matrix tables.

The overall accuracy for the 18 thematic maps ranged from 82.6 to 93.2 percent whereas their KHAT (Congalton and Green, 1999) values ranged from 0.63 to 0.86 (Figure 2). The range of producer's accuracy is between 82 and 98 percent for forest and 64 and 95 percent for clear-cut; the range of user's accuracy is between 79 and 96 percent for forest and 80 and 96 percent for clear-cut. The %LAND is estimated to be between 49.7 and 70.4 percent for forest and 29.6 and 50.3 percent for clear-cut.

REA ranges from −17 to +23 percent for forest and from −46 to +20 percent for clear-cut. No simple relationships are found between %LAND and overall accuracy. The estimations of %LAND are related to producer's accuracy or user's accuracy but almost perfect relationships are found between REA and %LAND of forest (Figure 3). At the 95 percent confidence level, the slope of the regression is 0.52 between REA (x) and %LAND (y), which is about the same as the mean of K values ($\bar{K} = 0.53$) of class forest among the 18 maps. When Equation 11 is used to perform calibrations, the range of %LAND of forest is reduced from 49.7 to 70.4 percent to 55.5 to 59.1 percent (Figure 3).

Case 2: Mapping a Multiple-Class Landscape

A Landsat Thematic Mapper (TM) image of path 22 and row 32, acquired on 05 October 1992, was used to map land-use and land-cover classes of an area of Tippecanoe County, Indiana. Based on the classification system of the U.S. Geological Survey (Anderson *et al.*, 1976), four Level-I land-use and land-cover classes of urban, agricultural land, forest, and water were defined from the image (Shao *et al.*, 2001).

Color infrared aerial photographs acquired in 1993 at a scale of 1:10,000 for the same area were scanned and rectified to the UTM projection coordinates with a 1-m resolution. The aerial images were used as background images while displaying the TM image. The comparison between the aerial and TM images on the computer helped assign class names to training areas and reference samples. The classification was carried out by 23 students (18 graduates and five undergraduates) from a remote sensing class. The students could choose any of the following techniques: supervised classification, unsupervised classification, original TM bands, PCA (Principal Components Analysis) transformation, or Tasseled Cap transformation. Following a 3- by 3-pixel majority filtering to classified images, the accuracy of the 23 four-class thematic maps was evaluated. To assure that all the maps share the same standard of classification accuracy, the instructor provided all the students with a single set of 250 reference samples, which was determined

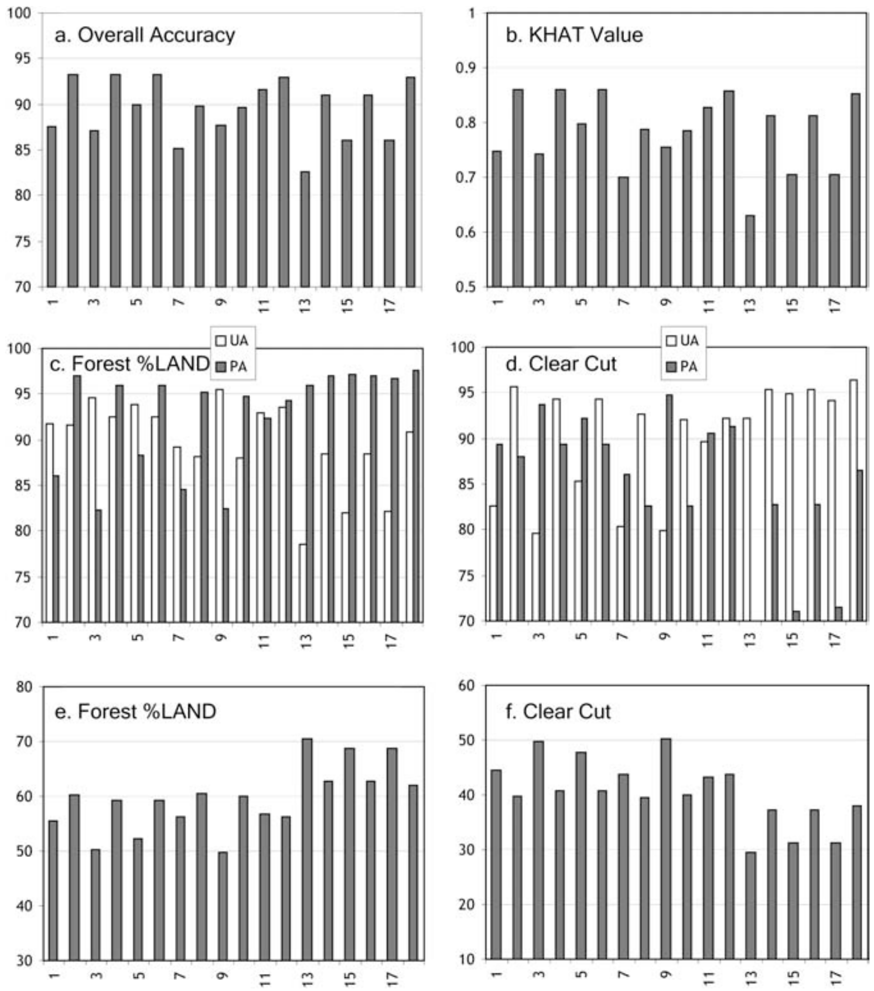

Figure 2. The overall accuracy (a), Kappa statistic (b), user's accuracy (c&d), producer's accuracy (c&d), and %LAND (e&f) of an image data classification with two classes ($N = 18$).

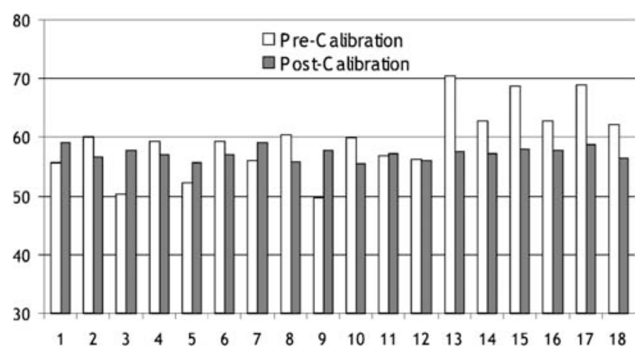

Figure 3. A comparison of forest %LAND values before and after calibrations.

with the stratified random sampling technique provided by Erdas Imagine software (http://gis.leica-geosystems.com/ Products; last accessed 13 January 2003). The sample unit was a pixel. The student could repeat classifications with various classification techniques until they reached a satisfactory overall accuracy. This resembled closely the common practice in image classification.

Each map was different in terms of classification accuracy and area estimation (Table 1). The overall accuracy of the 23 thematic maps ranged from 78 to 89 percent. The range of producer's accuracy is 32 to 92 percent for urban, 73 to 97 percent for agriculture, 47 to 92 percent for forest, and 78 to 100 percent for water; the range of user's accuracy is 62 to 96 percent for urban, 72 to 93 percent for agriculture, 62 to 100 percent for forest, and 84 to 100 percent for water. The %LAND is estimated to be 5.3 to 24.0 percent for urban, 49.7 to 78.3 percent for agriculture, 13.1 to 31.5 percent for forest, and 0.8 to 1.4 percent for water.

The REA ranges from −200 to +52 percent for urban, from −30 to +37 percent for agriculture, from −100 percent to +43 percent for forest, and from −23 to +18 percent for water. The closest relationship is found between REA and %LAND of each land-use and land-cover class among six classification accuracy indices (Table 2). When Equation 11 is used to perform calibrations, the range of %LAND is reduced to 11.2 to 17.4 percent for urban, 61.5 to

Table 1. The Minimum, Maximum, and Mean Values of Classification Accuracy and %LAND of Individual Classes

	OA	PA Urban	PA Agri-culture	PA Forest	PA Water	UA Urban	UA Agri-culture	UA Forest	UA Water	%LAND Urban	%LAND Agri-culture	%LAND Forest	%LAND Water
Min.	77.6	32.0	72.8	48	78.0	62.2	71.8	61.5	84.4	5.3	49.8	13.1	0.8
Max.	89.2	92.0	97.6	92	100	95.7	93.3	100	100	24.0	78.3	31.5	1.4
Mean	83.3	70.1	89.3	74.2	96.0	79.0	85.1	85.4	89.8	12.8	64.8	21.2	1.1

Notes: OA = Overall Accuracy, PA = Producer's Accuracy, UA = User's Accuracy

Table 2. A Summary of R^2 Values Following a Linear Regression Analysis between Accuracy Indices and %LAND of Individual Classes (N = 23)

Accuracy Index	Urban	Agriculture	Forest	Water
Overall Accuracy	0.07	0.00	0.00	0.00
Producer's Accuracy	0.74	0.80	0.68	0.62
User's Accuracy	0.74	0.74	0.72	0.39
Relative Errors of Area (*REA*)	0.86	0.94	0.87	0.75
Classification Success Index*	0.03	0.00	0.00	0.00
Individual Classification* Success Index	0.04	0.02	0.02	0.00

*From Koukoulas and Blackburn (2001).

65.4 percent for agriculture, 21.2 to 25.6 percent for forest, and 0.9 to 1.3 percent for water.

Discussion and Concluding Remarks

Among dozens of landscape indices, %LAND is relatively simple but essential in almost all applications. Unlike some spatially explicit indices, such as edge density, connectivity, and shape index, the errors of %LAND are supposed to be directly related to classification accuracy. A number of studies have been focused on statistical relationships between class areal estimates and classification accuracy (e.g., Bauer *et al.*, 1978; Card, 1982, Hay, 1988; Czaplewski, 1992, Dymond, 1992; Woodcock, 1996). Compared with the other works, REA is relatively simple and explicit because it can be directly computed based on user's and producer's accuracy of individual cover classes.

The implications of REA are obviously different from those of the overall classification accuracy because the former is related to the difference between user's and producer's accuracy values and the latter is approximately the average of user's and producer's accuracy for individual cover classes. This explains why REA has close relationships with %LAND but not the overall classification accuracy. However, the overall accuracy is still useful for revealing the potential uncertainty of %LAND. Figure 4a indicates that the variation in estimated %LAND decreases if overall accuracy increases. Only if overall accuracy is 100 percent, will %LAND have no variations at all and, therefore, is perfectly accurate. If thematic maps have relatively low overall classification accuracy, the differences between user's and producer's accuracy could be high. In this case, REA is particularly meaningful. If user's and producer's accuracy is unavailable, maps with higher values of the overall classification accuracy are still potentially preferable to those with lower values of the overall classification accuracy. This is because a small difference in overall accuracy may result in a big difference in user's and producer's accuracy, which, in turn, results in a big difference in cover-class areal estimates.

The relationship between REA and %LAND theoretically follows a linear function. In reality, R^2 is less than 1. This is because the source, an error matrix, from which REA is derived, contains sampling errors (Congalton, 1988). In this paper, Case 1 has 1,000 simple-random samples for a two-class map, whereas Case 2 has 250 points for a four-class map. The larger sample size results in stronger relationships between REA and %LAND than does the smaller sample size. To assure unbiased representatives of an error matrix, sample points should be located with the simple random sampling scheme. If stratified sampling methods are used, the producer's accuracy should be computed by weighting the cell proportions by the proportion of each land cover on the map (Stehman and Czaplewski, 1998). In this case, the constant K in Equation 11 should be computed by using the weighted value of cell f_{kk}.

Anderson *et al.* (1976) proposed the minimum level of classification accuracy was at least 85 percent. Based on what we have learned from the hypothetical and real examples, 85 percent of overall accuracy is not high enough for assuring accurate estimations of cover-class area. At the continental or global scale, the practical level of classification accuracy is even lower than 85 percent (Scepan, 1999; Vogelmann *et al.*, 2001). The effects of the low classification accuracy on downstream applications are still unknown. In many situations, users have no choice but to simply believe that classification accuracy measures, particularly overall accuracy, provide sufficient needed information. Such interpretation is dangerous because it may help produce misleading results or conclusions when a thematic map's low classification accuracy is used. One of advantages of REA is that it combines the user's and producer's accuracy into one index and, therefore, is readily interpretable. In this case, classification accuracy can be translated into an explicit measure of map quality for map users. If an error matrix is available, REA can be used to revise areal estimates of individual cover classes. This promotes a change from "merely referring to" to "actually using" the information of classification accuracy.

Acknowledgments

The authors would like to thank their home institutions for their support for conducting this study. Remote sensing data and ground data were provided by the Changbai Mountain Forest Ecosystem Research Station at the Chinese Academy of Sciences and the Laboratory for Application of Remote Sensing (LARS) at Purdue University. The thematic maps for the second case study were provided by 23 students in FNR 558 of 1999. Anonymous reviewers provided useful comments for revising the manuscript. The final stage of ground checking was supported by the "100 people" project from the Chinese Academy of Sciences.

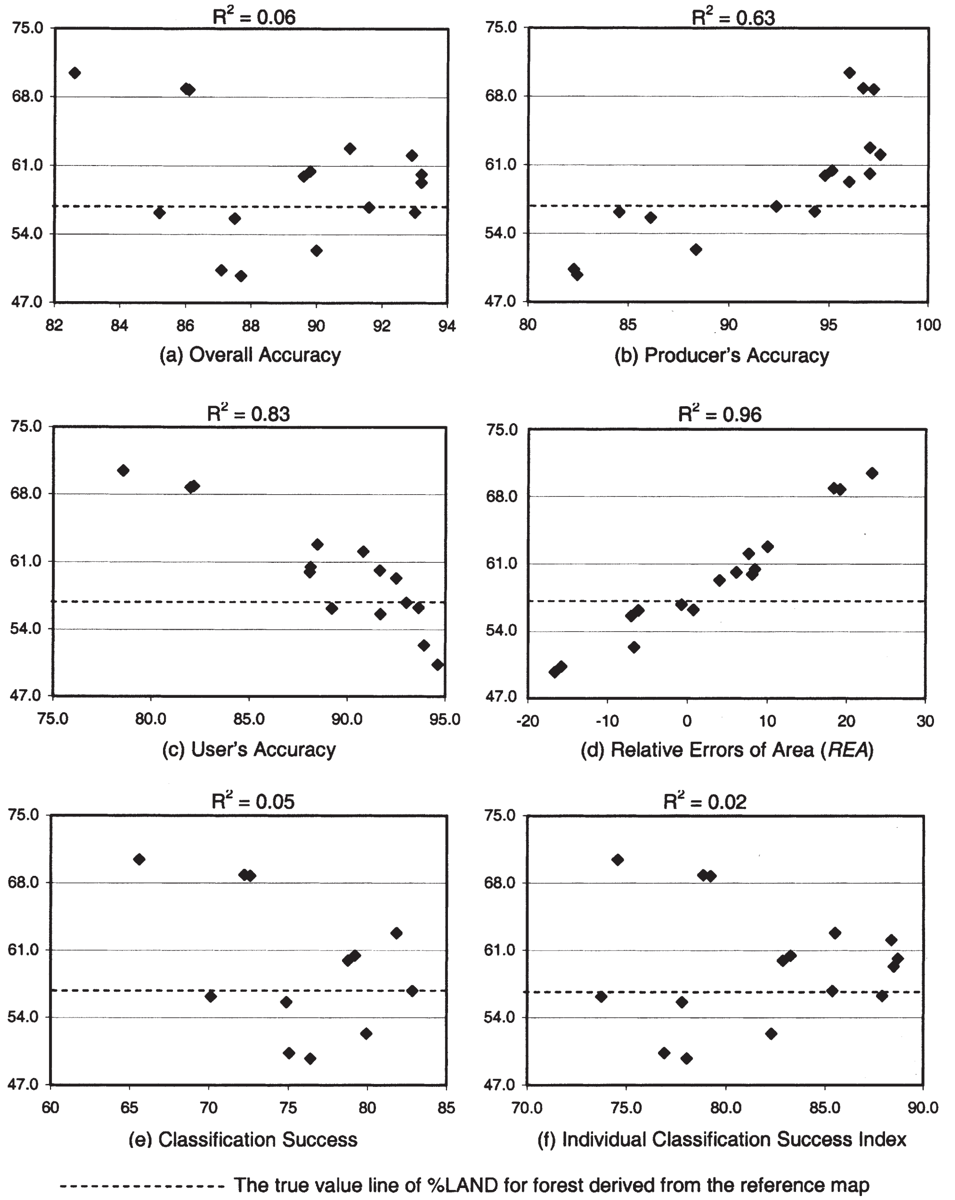

Figure 4. The relationships of %LAND values (*Y* axis) of forest with various accuracy indices.

References

Anderson, J.R., E.E. Hardy, J.T. Toach, and R.E. Witmer, 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*, U.S. Geological Survey Professional Paper 964, U.S. Government Printing Office, Washington, D.C., 28 pp.

Barnes, B.V., Z. Xu, and S. Zhao, 1993. Forest ecosystems in an old-growth pine-mixed hardwood forest of the Changbai Shan Preserve in northeastern China, *Canadian Journal of Forest Research*, 22(2):144–160.

Bauer, M.E., M.M. Hixson, B.J. Davis, and J.B. Etheridgs, 1978. Area estimation of crops by digital analysis of Landsat data, *Photogrammetric Engineering & Remote Sensing*, 44(8):1033–1043.

Cain, D.H., K. Riitters, and K. Orvis, 1997. A multi-scale analysis of landscape statistics, *Landscape Ecology*, 12(4):199–212.

Card, D.H., 1982. Using map categorical marginal frequencies to improve estimates of thematic map accuracy, *Photogrammetric Engineering & Remote Sensing*, 48(3):431–439.

Congalton, R.G., 1988. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data, *Photogrammetric Engineering & Remote Sensing*, 54(5):593–600.

———, 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37(1):35–46.

Congalton, R.G., and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, New York, N.Y., 160 p.

Czaplewski, R.L., 1992. Misclassification bias in areal estimates. *Photogrammetric Engineering & Remote Sensing*, 58(2):189–192.

Dymond, J.R., 1992. How accurately do image classifier estimate area? *International Journal of Remote Sensing*, 13(9):1735–1342.

Frohn, R.C., 1998. *Remote Sensing for Landscape Ecology: New Metric Indicators for Monitoring, Modeling, and Assessment of Ecosystems*, Boca Raton, Lewis Publications, Florida, 99 p.

Hess, G., 1994. Pattern and error in landscape ecology: A commentary, *Landscape Ecology*, 9(1):3–5.

Hess, G.R., and J.M. Bay, 1997. Generating confidence intervals for composition-based landscape indexes, *Landscape Ecology*, 12(5):309–320.

Heuvelink, G.B.M., 1998. *Error Propagation in Environmental Modelling with GIS*, Taylor & Francis Inc., Bristol, United Kingdom, 144 p.

Jelinski, D.E., and J. Wu, 1996. The modifiable areal unit problem and implications for landscape ecology, *Landscape Ecology*, 11(3):129–140.

Kettig, R.L., and D.A. Landgrebe, 1976. Classification of multispectral image data by extraction and classification of homogeneous objects, *IEEE Transactions on Geoscience Electronics*, GE-14:19–26.

Koukoulas, S., and G.A. Blackburn, 2001. Introducing new indices for accuracy evaluation of classified images representing semi-natural woodland environments, *Photogrammetric Engineering & Remote Sensing*, 67(4):499–510.

McGarigal, K., and B.J. Marks, 1994. *FRAGSTATS: Spatial Patterns Analysis Program for Quantifying Landscape Structure*, unpublished software, Department of Forest Science, Oregon State University, Corvallis, Oregon.

Moody, A., and C.E. Woodcock, 1995. The influence of scale and the spatial characteristics of landscapes on landcover mapping using remote sensing, *Landscape Ecology*, 10(6):363–379.

O'Neill, R.V., J.R. Krummel, R.H. Gardner, G. Sugihara, B. Jackson, D.L. DeAngelis, B.T. Milne, M.G. Turner, B. Zygnut, S.W. Christensen, V.H. Dale, and R.L. Graham, 1988. Indices of landscape pattern, *Landscape Ecology*, 1(2):152–162.

Qi, Y., and J. Wu, 1996. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices, *Landscape Ecology*, 11(1):39–49.

Riitters, K.H., R.V. O'Neill, C.T. Hunsaker, J.D. Wickham, D.H. Yankee, S.P. Timmins, K.B. Jones, and B.L. Jackson, 1995. A factor analysis of landscape pattern and structure metrics, *Landscape Ecology*, 10(1):23–39.

Shao., G., D. Liu, and G. Zhao, 2001. Relationships of image classification accuracy and variation of landscape statistics, *Canadian Journal of Remote Sensing*, 27(1):33–43.

Scepan, J., 1999. Thematic validation of high-resolution global land-cover data sets, *Photogrammetric Engineering & Remote Sensing*, 65(9):1051–1060.

Stehman, S.V., and R.L. Czaplewski, 1998. Design and analysis for thematic map accuracy assessment: Fundamental principles, *Remote Sensing of Environment*, 64(3):331–344.

Vogelmann, J.E., D. Helder, R. Morfitt, M.J. Choate, J.W. Merchant, and H. Bulley, 2001. Effects of Landsat 5 thematic mapper and Landsat 7 enhanced thematic mapper plus radiometric and geometric calibrations and corrections on landscape characterization, *Remote Sensing of Environment*, 78(1–2):55–70.

Wickham, J.D., and K.H. Riitters, 1995. Sensitivity of landscape metrics to pixel size, *International Journal of Remote Sensing*, 16(18):3585–3594.

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

Wookcock, C.E., 1996. On the roles and goals for map accuracy assessment: A remote sensing perspective, *In Proceedings of 2nd International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Fort Collins, Colorado (Gen. Tech. Rep. RM-GTR-277, USDA Forest Service Rocky Mountain Forest and Range Experiment Station), pp. 535–540.

Woodcock, C.E., and A.H. Strahler, 1987. The factor of scale in remote sensing, *Remote Sensing of Environment*, 21(3):311–332.

Wu, J., W. Gao, and P.T. Tueller, 1997. Effects of changing spatial scale on the results of statistical analysis with landscape data: A case study, *Geographic Information Sciences*, 3(1):20–41.

Wu, W., and G. Shao, 2002. Optimal combinations of data, classifiers, and sampling methods for accurate characterizations of deforestation, *Canadian Journal of Remote Sensing*, 28(4):601–609.

(Received 13 November 2001; accepted 24 October 2002; revised 04 November 2002)