Landscape Ecology **12**: 309–320, 1997.

Generating confidence intervals for composition-based landscape indexes

George R. Hess[1*] & Jeff M. Bay[2**]
[1]*Forestry Department, Box 8002, North Carolina State University, Raleigh NC 27695-8002, U.S.A.*
[2]*Statistics Department, Box 7632, North Carolina State University, Raleigh NC 27695-7632, U.S.A.*

Keywords: landscape indexes, landscape diversity, accuracy assessment, uncertainty, classification error, error matrixes, bootstrapping

Abstract

Many landscape indexes with ecological relevance have been proposed, including diversity indexes, dominance, fractal dimension, and patch size distribution. Classified land cover data in a geographic information system (GIS) are frequently used to calculate these indexes. However, a lack of methods for quantifying uncertainty in these measures makes it difficult to test hypothesized relations among landscape indexes and ecological processes. One source of uncertainty in landscape indexes is classification error in land cover data, which can be reported in the form of an error matrix. Some researchers have used error matrices to adjust extent estimates derived from classified land cover data. Because landscape diversity indexes depend only on landscape composition – the extent of each cover in a landscape – adjusted extent estimates may be used to calculate diversity indexes. We used a bootstrap procedure to extend this approach and generate confidence intervals for diversity indexes. Bootstrapping is a technique that allows one to estimate sample variability by resampling from the empirical probability distribution defined by a single sample. Using the empirical distribution defined by an error matrix, we generated a bootstrap sample of error matrixes. The sample of error matrixes was used to generate a sample of adjusted diversity indexes from which estimated confidence intervals for the diversity indexes were calculated. We also note that present methods for accuracy assessment are not sufficient for quantifying the uncertainty in landscape indexes that are sensitive to the size, shape, and spatial arrangement of patches. More information about the spatial structure of error is needed to calculate uncertainty for these indexes. Alternative approaches should be considered, including combining traditional accuracy assessments with other probability data generated during the classification procedure.

Introduction

Many landscape indexes with ecological relevance have been proposed in the literature (overviews may be found in O'Neill *et al.* 1988a; Turner 1989; Turner and Gardner 1991; see Riitters *et al.* 1995 for an extensive list). These include diversity indexes (Magurran 1988; Turner 1989), edge indexes (Patton 1975; Bowen and Burgess 1981), fractal dimension (Mandelbrot 1977; Burrough 1981; Krummel *et al.* 1987; Milne 1988, 1991), contiguity (LaGro 1991), contagion (O'Neill *et al.* 1988a; Li and Reynolds 1993), dominance (O'Neill *et al.* 1988a), and patch size distribution (Bowen and Burgess 1981). These indexes may reflect the ability of organisms to inhabit and traverse a landscape (O'Neill *et al.* 1988b; Johnson *et al.* 1992; Flather *et al.* 1992; With 1994), the potential for materials or disturbances to move from one part of the landscape to another (Turn-

* G.R. Hess: Phone: 919.515.7437; Fax: 919.515.8149; E-mail: grhess@ncsu.edu; ** J.M. Bay: Phone: 919.515.3311; Fax: 919.515.3593; E-mail: jeff_bay@ncsu.edu

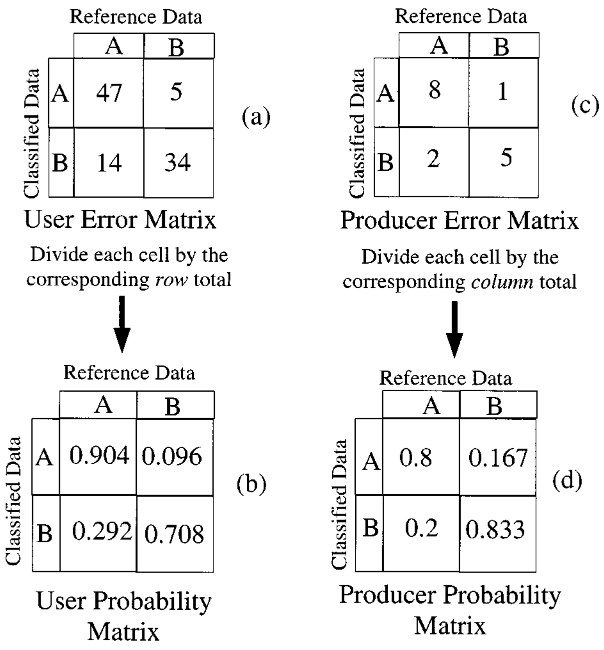

Fig. 1. Error matrixes describe the accuracy of each classification category in various ways.

er *et al.* 1989; Barrett *et al.* 1990), or the types of processes that are shaping the landscape (Krummel *et al.* 1987; Rex and Malanson 1992; Zipperer 1993). However, these hypothesized relations have been subject to very limited testing (Wiens 1992).

Landscape indexes are usually calculated using classified land cover data in a geographic information system (GIS). These data represent the final product of a complicated set of processes and analyses that introduce error at many points (Walsh *et al.* 1987; Lunetta *et al.* 1991; Congalton and Green 1993; Janssen and van der Wel 1994). Yet, methods for quantifying the uncertainty in landscape indexes derived from these data are all but absent (Hess 1994). Without measures of uncertainty it is difficult to evaluate correlations between landscape indexes and ecological processes, to detect differences among landscapes, or to detect changes in a landscape over time. Wickham and Norton (1994) used a bootstrap procedure to generate confidence intervals for some landscape indexes, but their method does not consider error in the land cover data. This paper describes a bootstrap procedure for quantifying the uncertainty in composition-based landscape indexes that is introduced by error in land cover data. Composition-based indexes depend only on the extent of each cover in a landscape. We take the perspective of a data user who has been given land cover data, a user error matrix, and the task of generating landscape indexes using these data.

Classification error is often reported as an error matrix that describes the accuracy of each classification category, as well as the nature of the confusion among categories (Figure 1) (Story and Congalton 1986; Congalton 1988a, 1991). An error matrix is constructed by comparing the classified data to reference data (*e.g.*, ground truth, aerial photography) for a sample of pixels. Error matrix columns usually represent the reference data, which are assumed to be correct, and rows represent the classified data from the remotely sensed scene. A *user error matrix* is created by selecting the sample pixels from the classified data. Each cell in a user error matrix represents the number of times that a pixel which was classified into the category specified by the classified data (row) was actually in the category specified by the reference data (column) (Figure 1a). A *producer error matrix* is created by selecting the sample pixels from the reference data. Each cell in a producer error matrix represents the number of times that a pixel which was actually in the category specified by the reference data (column) was classified into the category specified by the classified data (row) (Figure 1c). In both cases, the main diagonal represents correctly classified pixels.

If the count in each user error matrix cell is divided by its marginal row total, proportions in the resulting matrix represent the probability that a pixel is actually in the column's reference data category given that it has been classified into the row's classified data category (Figure 1b). We call this adjusted matrix the *user probability matrix*, **U**, after Prisley and Smith (1987). The elements of the main diagonal of the user probability matrix are called the user's (Story and Congalton 1986) or consumer's (Aronoff 1982) accuracy. If the count in each producer error matrix cell is divided by its marginal column total, proportions in the resulting matrix represent the probability that a pixel is classified into the row's classified data category given that it is actually in the column's reference data category (Figure 1d). This is called the *producer probability matrix* (Story and Congalton 1986).

Methods

Landscape indexes that depend only on landscape composition – the extent of each cover in the landscape – can be adjusted using the conditional probabilities of misclassification calculated from an error matrix (Tenenbein, 1972; Prisley and Smith 1987; Card 1982; Hay 1988; Buckland and Elston 1994). These indexes include various measures of diversity, richness, evenness, and dominance (Appendix A). If the pixel counts from the scene being analyzed are represented in a column vector, **V**, then an adjusted pixel count, **W**, can be obtained by

$$W^T = V^T U \tag{1}$$

where **U** is the user probability matrix. This calculation is equivalent to Fisher's (1991) algorithm for "controlled perturbation of polygonal maps", with the error matrix supplying the information used to control map perturbations. Fisher's algorithm reduces to Equation (1) when a user probability matrix is used to adjust a large number of pixels. The adjusted pixel count is the maximum likelihood estimate under the following assumptions (Tenenbein, 1972):

- each pixel is eligible for selection,
- sample pixels for the error matrix are selected randomly,
- each pixel is classified independently, and
- there is no spatial autocorrelation of error among pixels.

An adjusted landscape composition measure is obtained by using the adjusted pixel counts.

One must understand the sampling protocol used to create the error matrix when making adjustments (Buckland and Elston 1994). The adjusted pixel counts will be maximum likelihood estimates when the observations in the error matrix are selected in an unconstrained manner, or if the number of observations in each row or column are fixed. If sampling is unconstrained or the number of observations for each row (classified data category) are fixed, then the conditional probabilities used to make adjustments can be estimated directly from the user probability matrix as described above. If the number of observations for each column (reference data category) are fixed, then only the probabilities of a pixel being in the classified category given that it is actually in the reference category can be estimated directly. However, Bayes' Rule can be used to estimate the conditional probabilities that we use to make the adjustments in our methods (Green *et al.*, 1993).

We have extended this approach to generate confidence intervals for landscape indexes that depend only on landscape composition. A user error matrix is constructed from a sample of *m* pixels from a classified scene. A different sample would likely yield a different user error matrix, different probabilities of misclassification, and a different adjusted landscape index. One could obtain an estimate of the effect of this sampling variability on a landscape index by constructing a large number of error matrixes and calculating an adjusted landscape index using each error matrix. However, the expense of constructing an error matrix makes such an approach infeasible. Bootstrapping is a technique that allows one to estimate sample variability by resampling from the empirical probability distribution defined by a single sample (Efron and Tibshirani 1993). By bootstrapping the user error matrix and calculating an adjusted landscape index using each bootstrapped error matrix, one can estimate the effect of sampling variability on the adjusted index.

Given an original sample of size *m*, a bootstrap sample is a random sample of size *m*, with replacement, from the empirical probability distribution of the original sample. The empirical probability distribution, defined by placing probability mass $1/m$ on each of the *m* observations in the original sample, is used as a surrogate for the true underlying, but unknown, probability distribution. In our procedure, the error matrix is used as the empirical probability distribution representing the true, but unknown, classification error rates. The variability among the bootstrapped error matrixes estimates the variability introduced by sampling during error matrix construction. A bootstrap sample

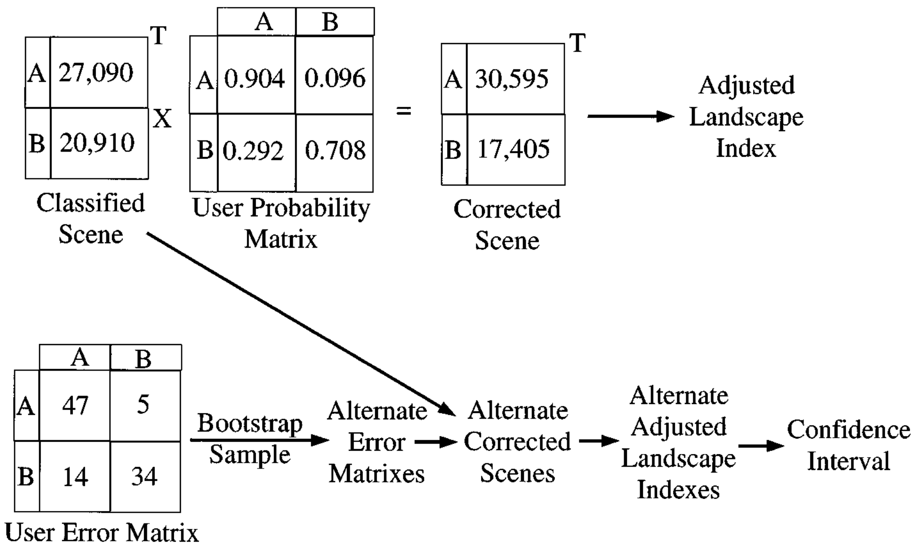

Fig. 2. Overview of procedure for generating confidence for landscape pattern measures that depend only on landscape composition.

of the error matrix can be drawn in a number of ways. One method is to rewrite the error matrix as a two-column list of reference and classified data, with one entry for each of the m observations in the error matrix; the list should have m entries. Next, generate m random numbers between one and m and, for each random number, select the corresponding observation in the list. The randomly selected observations define a bootstrapped error matrix. The bootstrapped error matrix is used to adjust the landscape index (Equation 1).

The bootstrapped error matrix should be selected following the same protocol used in creating the original error matrix. If the observations for the original error matrix were drawn randomly from the scene, then the observations for a bootstrapped error matrix should be drawn randomly from the original error matrix. If the original error matrix is drawn in a constrained manner, the bootstrap samples must be constrained in the same manner. For example, if one cover type is rare in a scene, observations for the original error matrix may be drawn after stratifying the scene by data class and predetermining the number of observations from each class. This procedure guarantees that examples of rare land cover types are included in the error matrix. In such a case, a bootstrapped error matrix should be constrained to have the same number of observations in each row as the original error matrix. The observations for each row in the bootstrapped error matrix are drawn from the corresponding row of the original error matrix. The code we have developed (Appendix B) can handle a random sample and a sample that is stratified by classified data (row).

The pixel counts from the classified scene are adjusted using the original user probability matrix and the adjusted pixel counts are used to generate adjusted landscape indexes (Figure 2). A large number, B, of bootstrap samples is generated by resampling from the empirical probability distribution defined by the original error matrix. For each bootstrap sample the bootstrapped error matrix is used to compute an adjusted value for the landscape index, resulting in a sample of B adjusted values of the index. The variability of the bootstrap sample of adjusted landscape indexes estimates the variability introduced by sampling error in error matrix production. (One could also use the error matrix bootstrap samples to estimate variability for entries in the error matrix.) Efron and Tibshirani (1993) recommend a minimum of B=1,000 for generating confidence intervals. An empirical, two-tailed 1-α confidence interval for the landscape index can be calculated by sorting the bootstrapped indexes in ascending order and using the $\alpha/2$ and (1-$\alpha/2$) percentiles. For example, if B=1,000, the 95% confidence interval would be defined by the 25th and 976th entries in

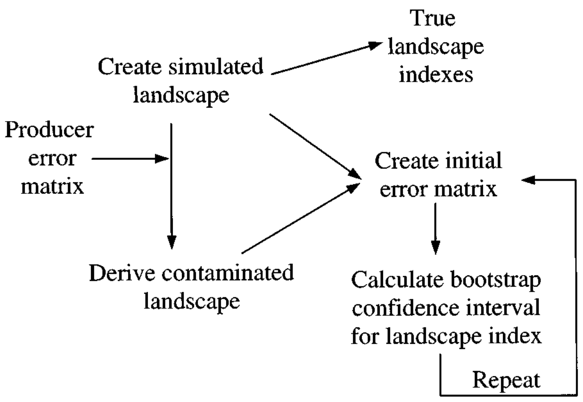

Fig. 3. Overview of simulation procedure we used to test our method.

the sorted list. We used Efron and Tibshirani's (1993) bias-corrected and accelerated (BCa) method to determine the 95% confidence intervals for each composition-based index we tested. This method adjusts for possible bias in an estimator and accounts for the possible change in the standard deviation of an estimator as the true value changes. When an estimator is unbiased and its standard deviation does not depend on the true value it is estimating, then the BCa method will, on average, give the same confidence limits as the percentile method described above.

We tested our method using two landscape diversity indexes: Shannon diversity and Simpson dominance (Figure 3, Appendix A). Two other common indexes, Shannon evenness and Simpson diversity (Appendix A), are simple transformations of Shannon diversity and Simpson dominance, respectively. Adjusted values and confidence intervals for these indexes can be obtained by transforming the adjusted values and confidence intervals of the Shannon diversity and Simpson dominance indexes (see Appendix A for a cautionary note). Starting with a two-category, 48,000 pixel scene of known composition (Table 1) we contaminated the scene according to predefined producer error matrixes of high, medium, and low accuracy (Table 2), creating corrupted scenes (Table 1). These corrupted scenes represent the data an end user would obtain, probably from a GIS coverage. Next, we drew a simple random sample, without replacement, of 100 paired observations from the original and contaminated landscapes to construct an initial error matrix. This represents the user error matrix one might receive with the GIS data (Table 3). We used the initial error matrix to adjust the landscape index for the corrupted landscape (Table 4). Then, we generated a bootstrap sample of 1,000 error matrices from the initial error matrix, each containing 100 observations. Each bootstrapped error matrix was used to produce an adjusted index, resulting in a bootstrap sample of indexes. We used this bootstrap sample to calculate 95% confidence intervals for each landscape diversity index (Table 4). We repeated the entire bootstrap procedure 2,000 times for each index and for each producer accuracy (Table 5).

Results

For each landscape diversity index the adjusted index was an improved estimate of the true index, and the 95% confidence interval contained the true landscape diversity index (Table 4). Results from repeating the procedure 2,000 times showed that the 95% confidence interval contained the true index approximately 95% of the time (Table 5). This confirms that our method calculates valid confidence intervals for products with a wide range of producer accuracy.

A critical question for the end user, who has

Table 1. Proportion of pixels in true scene compared to scenes contaminated according to high, medium, and low accuracy producer error matrixes. The true pixel counts are shown in parentheses with the true proportions. The unadjusted pattern measures in Table 4 were calculated from these data.

Cover	True scene (pixel count)	High accuracy	Medium accuracy	Low accuracy
A	0.625 (30,000)	0.659	0.564	0.438
B	0.375 (18,000)	0.341	0.436	0.562

Table 2. High, medium, and low producer probability matrixes used to contaminate the true scene in our simulated two category landscape.

A. High accuracy

Truth → Data	A	B
A	1	0.091
B	0	0.909

B. Medium accuracy

Truth → Data	A	B
A	0.8	0.167
B	0.2	0.833

C. Low accuracy

Truth → Data	A	B
A	0.2	0.833
B	0.8	0.167

only a scene and an error matrix, is "How many bootstrap error matrixes should I draw?" In other words, what value of B is large enough? Based on improvements in the coefficient of variation of the bootstrapped statistic, Efron and Tibshirani (1993) recommend a minimum of B=1000 for generating confidence intervals. To confirm this, we ran simulations in which we varied the number of bootstrap samples for a single initial error matrix from B = (50, 100, 200, 400, 600, 800 . . . 1,600). For each value of B, we ran 100 simulations and plotted the coefficient of variation of the upper and lower 95% confidence intervals for the Simpson dominance index (Figure 4). Little precision was gained after 800 bootstrap samples.

We also applied our method to two scenes consisting of three classes. In one scene the data were unbalanced in the sense that one class covered an area that was an order of magnitude smaller than the area covered by each of the other two (*i.e.*, 1,500 pixels compared to 10,000 and 30,000 pixels). The second scene was balanced: areas for the three classes were all of the same order of magnitude (*i.e.*, 3,000; 1,000; 1,000). For the unbalanced scene there was slight undercoverage. Results from repeating the procedure 1,000 times

Table 3. Sample user error and probability matrixes generated by comparing 100 pixels selected randomly, without replacement, in the contaminated scene to the true scene. We used these user probability matrixes to generate the adjusted values and bootstrap confidence intervals in Table 4.

A. High accuracy

Truth → Data	A pixel count (user probability)	B pixel count (user probability)
A	63 (0.969)	2 (0.031)
B	0 (0)	35 (1.0)

B. Medium accuracy

Truth → Data	A pixel count (user probability)	B pixel count (user probability)
A	47 (0.904)	5 (0.096)
B	14 (0.292)	34 (0.708)

C. Low accuracy

Truth → Data	A pixel count (user probability)	B pixel count (user probability)
A	10 (0.270)	27 (0.730)
B	61 (0.968)	2 (0.032)

showed that the 95% confidence interval contained the true value 93% of the time, regardless of whether we used stratified or simple random sampling. The coverage probabilities were very close to 95% for the balanced scene.

Discussion

The relative merits of diversity indexes and their interpretation have been expounded and debated in the ecological literature for some time (*e.g.*, Hurlbert 1977; Pielou 1975, 1977; Magurran 1988). Changes in landscape diversity indexes are hypothesized to reflect changes in the level of human impacts, species diversity, habitat use by wildlife, and the nutrient content and productivity of aquatic systems (Romme and Knight 1982,

Table 4. Results of our method applied to contaminated scenes of high, medium, and low accuracy. The 95% confidence intervals were obtained from a bootstrap sample of size B=1,000.

A. High producer accuracy.

Diversity Measure	True value	Unadjusted	Adjusted	95% CI
Shannon diversity	0.662	0.642	0.654	(0.641 , 0.676)
Simpson dominance	0.531	0.551	0.538	(0.506 , 0.551)

B. Medium producer accuracy.

Diversity Measure	True value	Unadjusted	Adjusted	95% CI
Shannon diversity	0.662	0.685	0.655	(0.604 , 0.684)
Simpson dominance	0.531	0.508	0.538	(0.508 , 0.585)

C. Low producer accuracy.

Diversity Measure	True value	Unadjusted	Adjusted	95% CI
Shannon diversity	0.662	0.686	0.639	(0.579 , 0.674)
Simpson dominance	0.531	0.508	0.522	(0.517 , 0.607)

Table 5. Portion (%) of 2,000 bootstrap-generated 95% confidence intervals that contained the true value for two landscape pattern measures and three levels of producer accuracy.

Producer Accuracy	Shannon diversity	Simpson dominance
High	94.7	95.0
Medium	94.3	94.1
Low	94.5	93.4

Turner and Ruscher 1987, Turner 1989). Turner *et al.* (1995) note that the ability to quantify and monitor landscape pattern exceeds the ability to interpret its ecological effects. They also note the need to determine what constitutes a significant statistical and ecological change in landscape pattern and to relate these changes to ecologically relevant responses. Lack of information about the uncertainty in landscape indexes limits our ability

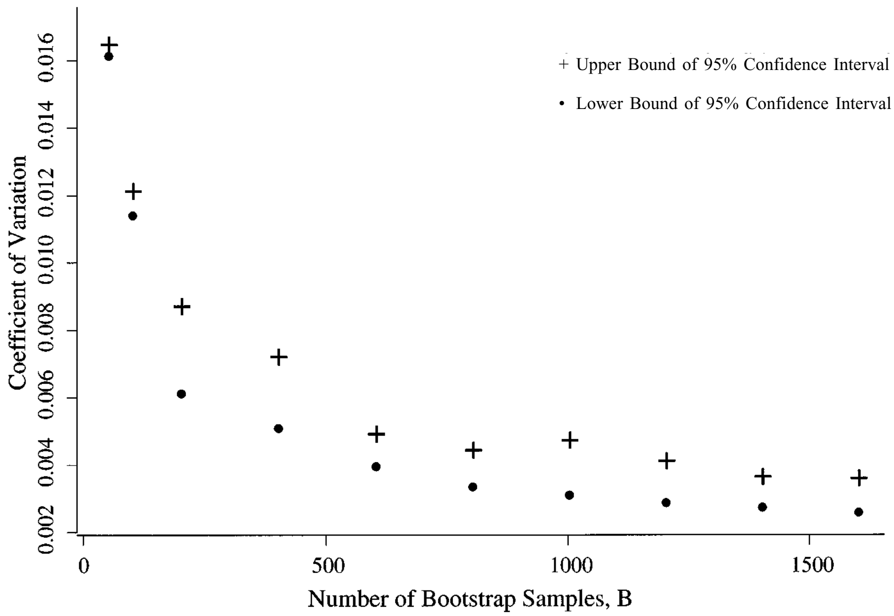

Fig. 4. The coefficient of variation of 100 upper and lower bounds of 95% confidence intervals generated by our method for the Simpson dominance of the medium accuracy scene plotted against the size of the bootstrap sample.

to identify statistically significant changes and, consequently, our ability to draw statistically valid conclusions about their correlation with ecologically relevant responses. Statistical differences, although not always ecologically meaningful, may indicate the presence of underlying processes that will lead to ecological differences over time.

We have provided a method for assessing the variability associated with landscape indexes that depend only on the relative proportions of each cover. With this information it is possible to compare statistically landscape diversity indexes from different scenes or to compare the same scene at different points in time. For example, to compare diversity indexes for two different geographic regions that have a single error matrix, one could compute the difference between the indexes using the original error matrix. Our method can be used to compute the confidence interval of the difference (*i.e.,* compute the difference for each bootstrap error matrix and then determine the confidence interval for the true difference). If the confidence interval excludes zero, the difference is statistically significant at the level corresponding to the confidence interval (*e.g.,* 0.05 for a 95% confidence interval). To compare indexes for two geographical regions that each have their own error matrix, one could compute a confidence interval for each geographical region and then check to see whether the confidence intervals overlap. If the confidence intervals do not overlap, then one could conclude that the indexes are significantly different at the $1\text{-}(1\text{-}\alpha)^2$ level, where $(1\text{-}\alpha)$ is the level of the confidence intervals. These tests could also be used to compare the same region at different time periods.

Our procedure estimates the variability of the sampling method used to construct the original error matrix. It does not correct for a biased sampling method. If the original error matrix is a biased estimate of the classification error rate, the adjusted pixel counts and landscape index will be biased estimates, and the confidence interval may not contain the true values with probability $1\text{-}\alpha$. Violations of the maximum likelihood assumptions may lead to a biased error matrix and, therefore, biased estimates of landscape indexes. The estimates should have small bias if the assumptions hold approximately. For serious violations, the adjusted landscape indexes are likely to be poor estimates of the true values.

Spatial autocorrelation may be a source of bias in error matrix construction. Error among pixels can be spatially autocorrelated, and the autocorrelation structure varies with landscape type and classification scheme (Congalton 1988b). Small patches of pixels may be difficult to classify and perhaps be misclassified together. Congalton (1988b) demonstrated that even large areas can be misclassified in some landscapes. Classification error rates may be over- or underestimates of the true error rate, depending upon the degree of spatial autocorrelation and error matrix construction techniques. Another potential problem is that pixels used for accuracy assessment are often selected from homogenous groups of pixels rather than entirely at random. This violates the assumption that each pixel in the classified data is eligible for selection (*i.e.*, pixels in heterogeneous groups are not eligible). If pixels in heterogeneous groups are more prone to misclassification, the result would be a biased estimate of the classification error rate. The user who receives GIS data and an error matrix should be aware of these potential problems, but can do little to correct them.

Our approach to calculating uncertainty in composition-based indexes is based on the idea of recreating alternative versions of the “true” landscape composition by using the error matrix to correct errors in the classified data. The method works because composition-based indexes depend only on the number of pixels in each classification category, and because the error matrix provides information about the average uncertainty in pixel classification. Most landscape indexes depend not only on landscape composition, but also on landscape physiognomy – the size, shape, and spatial arrangement of patches (see Riitters *et al.* 1995). Physiognomy-based indexes are sensitive not only to the number of pixels in each classification category, but also to their spatial arrangement. Consequently, these indexes are sensitive not only to errors in the number of pixels, but also to the exact location of those errors within a scene. For example, edge and contagion indexes are sensitive to changes in the number of pixels in one class adjacent to pixels in another class.

Extending our approach to physiognomy-based indexes requires the recreation of alternative versions of the "true" landscape pattern and structure, as well as composition. In attempting to recreate a "true" landscape, we found ourselves asking many questions about the manner in which classification error varies spatially within a scene. How strongly autocorrelated are errors in a classified scene? In other words, if one pixel is incorrectly classified how likely is it that surrounding pixels are also incorrectly classified? Are errors more likely for pixels in small patches than for pixels in large patches? Are errors more or less likely along the boundary between patches than in the patch interior? Does the answer to the previous question depend upon the cover types forming the boundary? For example, one might expect that a forest–water edge is more accurate than a scrub–agriculture edge. Does the shape of a patch have an effect on classification accuracy? The answers to questions like these determine how the size, shape, and spatial arrangement of patches, and hence the value of physiognomy-based landscape indexes, are affected by classification error. An error matrix, even an accurate one, does not provide information to answer these questions because it provides no information about the pattern and structure of errors. We conclude that present methods for accuracy assessment are not sufficient for assessing the variability of landscape indexes that are sensitive to landscape physiognomy.

The error matrix was designed to address issues of classification accuracy, not the precision of landscape indexes. If uncertainty in landscape ecological analyses of remotely sensed data products is to be quantified, methods used for accuracy assessment must consider the nature of those analyses and provide the information needed to quantify their uncertainty. Enhanced approaches to accuracy assessment that generate data at multiple scales, from single pixel through a range of patch sizes, would improve our ability to assess the accuracy of landscape indexes. For example, separate error matrixes might be constructed for small and large patches, or for areas at various distances from boundaries. However, given the cost of conducting even the simplest accuracy assessment, we suspect that conducting far more extensive accuracy assessments may be infeasible.

It may be possible to use other data generated during the classification process to address uncertainty in physiognomy-based landscape indexes. For example, during a maximum likelihood classification the distance to the mean vectors of all possible classification categories is calculated for each pixel. The classified map is produced by assigning each pixel to the closest mean vector. In some data analysis packages the information about the distances to the mean vectors is available in a probability file. Corves and Place (1994) generated classification reliability maps using probability files by noting that the classification of pixels distant from the mean vector of their class is less reliable than the classification of pixels close to the mean vector. A reliability map uses this information to show where errors are more or less likely to occur. Foody (1990) and Corves and Place (1994) have suggested concentrating accuracy assessment efforts in areas of low reliability and using the information to improve the classification process. Separate error matrixes for areas of different reliability would improve our ability to estimate uncertainty in composition-based landscape indexes. In conjunction with a probability map, these error matrixes would also help identify where in a scene errors are likely to occur, bringing us closer to understanding the effect of error on landscape physiognomy and physiognomy-based landscape indexes. Fuzzy logic, by quantifying the ambiguity in the classification process, may also provide more information about the spatial structure of error in remotely sensed products (Gopal and Woodcock 1994).

We believe that landscape ecologists should determine the data needed to estimate uncertainty in landscape indexes and clearly communicate those needs to data producers. Data producers should, to the extent possible, consider data and accuracy assessment as a unified product designed to meet the end user's analysis needs. By working together in this way, remote sensing experts and landscape ecologists will move closer to understanding and demonstrating relationships between landscape pattern and ecological process.

Acknowledgements

Thanks to Roger Berger, Ray Czaplewski, Steve Ellner, and Len Stefanski for valuable insights. Earlier versions of the manuscript benefited greatly from the comments of Lee Campbell, Glenn Catts, Russ Congalton, Casson Stallings, Jay Ver Hoef, Jim Wickham, and two anonymous reviewers.

This research was supported in part through a Specific Cooperative Agreement between the United States Department of Agriculture (USDA), Agricultural Research Service (ARS) and the North Carolina Agricultural Research Service. The United States Environmental Protection Agency, through its Office of Research and Development's Environmental Monitoring and Assessment Program, partially funded the extramural research described here through an Interagency Agreement with the USDA, ARS. It has not been subjected to the Agency's review and, therefore, does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

References

Aronoff, S. 1982. Classification accuracy: A user approach. Photogrammetric Engineering and Remote Sensing 48: 1299-1307.

Barrett, G. W., Rodenhouse, N., and Bohlen, P. J. 1990. Role of sustainable agriculture in rural landscapes. In Sustainable Agricultural Systems. pp. 624-636. Edited by C.A. Edwards, R. Lal, P. Madden. R.H. Miller, and G. House. Soil and Water Conservation Society, Ankeny, Iowa.

Bowen, G. W. and Burgess, R. L. 1981. Quantitative Analysis of Forest Island Pattern in Selected Ohio Landscapes. Publication ORNL-TM-7795, Oak Ridge National Laboratory, Environmental Sciences Division, Oak Ridge, TN.

Buckland, S. T. and Elston, D. A. 1994. Use of ground-truth data to correct land cover estimates from remotely sensed data. International Journal of Remote Sensing 15: 1273-1282.

Burrough, P. A. 1981. Fractal dimensions of landscapes and other environmental data. Nature 294: 240-242.

Card, D. H. 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy. Photogrammetric Engineering and Remote Sensing 48: 431-439.

Congalton, R. G. 1988a. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54: 593-600.

Congalton, R. G. 1988b. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing 54: 587-592.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of the Environment 37: 35-46.

Congalton, R. G. and Green, K. 1993. A practical look at the sources of confusion in error matrix generation. Photogrammetric Engineering and Remote Sensing 59: 641-644.

Corves, C. and Place, C. J. 1994. Mapping the reliability of satellite-derived landcover maps an example from the Central Brazilian Amazon Basin. International Journal of Remote Sensing 15: 1283-1294.

Efron, B. and Tibshirani, R. J. 1993. An Introduction to the Bootstrap. Chapman and Hall, New York.

Fisher, P. F. 1991. Modelling soil map-unit inclusions by Monte Carlo simulation. International Journal of Geographical Information Systems 5: 193-208.

Flather, C. H., Brady, S. J., and Inkley, D. B. 1992. Regional habitat appraisals of wildlife communities: a landscape-level evaluation of a resource planning model using avian distribution data. Landscape Ecology 7: 137-147.

Foody, G. M. 1990. Directed ground survey for improved maximum likelihood classification of remotely sensed data. International Journal of Remote Sensing 11: 1935-1940.

Gopal, S. and Woodcock, C. 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets. Photogrammetric Engineering and Remote Sensing 60: 181-188.

Green, E. J., Strawderman, W. E. and Airola, T. M. 1993. Assessing classification probabilities for thematic maps. Photogrammetric Engineering and Remote Sensing 59: 635-639.

Hay, A. M. 1988. The derivation of global estimates from a confusion matrix. International Journal of Remote Sensing 9: 1395-1398.

Hess, G. 1994. Pattern and error in landscape ecology: a commentary. Landscape Ecology 9: 3-5.

Hurlbert, S. H. 1971. The nonconcept of species diversity: A critique and alternative parameters. Ecology 52: 577-586.

Janssen, L. L. F. and van der Wel, F. J. M. 1994. Accuracy assessment of satellite derived land- cover data: a review. Photogrammetric Engineering and Remote Sensing 60: 419-426.

Johnson, A. R., Wiens, J. A., Milne, B. T., and Crist, T. O. 1992. Animal movements and population dynamics in heterogeneous landscapes. Landscape Ecology 7: 63-75.

Krummel, J. R., Gardner, R. H., Sugihara, G., O'Neill, R. V., and Coleman, P. R. 1987. Landscape patterns in a disturbed environment. Oikos 48: 321-324.

LaGro, J. 1991. Assessing patch shape in landscape mosaics. Photogrammetric Engineering and Remote Sensing 57: 285-293.

Li, H. and Reynolds, J. F. 1993. A new contagion index to quantify spatial patterns of landscapes. Landscape Ecology 8: 155-162.

Lunetta, R. S., Congalton, R. G., Fenstermaker, L. K., Jensen, J. R., McGwire, K. C., and Tinney, L. R. 1991. Remote sensing and geographic information system data integration: error sources and research ideas. Photogrammetric Engineering and Remote Sensing 57: 677-687.
Magurran, A. 1988. Ecological Diversity and its Measurement. Princeton University Press, Princeton, N.J.
Mandelbrot, B. B. 1977. Fractals: Form, Chance and Dimension. Freeman, San Francisco.
Milne, B. T. 1988. Measuring the fractal geometry of landscapes. Applied Mathematics and Computation 27: 67-69.
Milne, B. T. 1991. Lessons from applying fractal models to landscape patterns. In Quantitative Methods in Landscape Ecology. pp. 199-235. Edited by M. G. Turner, and R. H. Gardner. Springer-Verlag, New York.
O'Neill, R. V., Krummel, J. R., Gardner, R. H., Sugihara, G., Jackson, B., DeAngelis, D. L., Milne, B. T., Turner, M. G., Zygmunt, B., Christensen, S., Dale, V. H., and Graham, R. L. 1988a. Indices of landscape pattern. Landscape Ecology 1: 153-162.
O'Neill, R. V., Milne, B. T., Turner, M. G., and Gardner, R. H. 1988b. Resource utilization scales and landscape pattern. Landscape Ecology 2: 63-69.
Patton, D. R. 1975. A diversity index of quantifying habitat "edge". Wildlife Society Bulletin 3: 171-173.
Pielou, E. C. 1975. Ecological Diversity. John Wiley and Sons, New York.
Pielou, E. C. 1977. Mathematical Ecology. John Wiley and Sons, New York.
Prisley, S. P. and Smith, J. L. 1987. Using classification error matrices to improve the accuracy of weighted land-cover models. Photogrammetric Engineering and Remote Sensing 53: 1259- 1263.
Rex, K. D. and Malanson, G. P. 1990. The fractal shape of riparian forest patches. Landscape Ecology 4: 249-258.
Riitters, K. H., O'Neill, R. V., Hunsaker, C. T., Wickham, J. D., Yankee, D. H., Timmins, S. P., Jones, K. B. and Jackson, B. L. 1995. A factor analysis of landscape pattern and structure metrics. Landscape Ecology 10: 23-39.
Romme, W. H. and Knight, D. H. 1982. Landscape diversity: The concept applied to Yellowstone Park. BioScience 32: 664-670.
StatSci. 1993. S-Plus Guide to Statistical and Mathematical Anlaysis, Version 3.2. MathSoft, Inc., Seattle.
Story, M. and Congalton, R. G. 1986. Accuracy assessment: a user's perspective. Photogrammetric Engineering and Remote Sensing 52: 397-399.
Tenenbein, A.. 1972. A double sampling scheme for estimating from misclassified multinomial data with applications to sampling inspection. Technometrics 14: 187-202.
Turner, M. G. 1989. Landscape ecology: the effect of pattern on process. Annual Reviews of Ecology and Systematics 20: 171-197.
Turner, M. G. and Gardner, R. H. (editors). 1991. Quantitative Methods in Landscape Ecology. Springer-Verlag, New York.
Turner, M. G. and Ruscher, C. L. 1987. Changes in the spatial patterns of land use in Georgia. Landscape Ecology 1: 241-251.
Turner, M. G., Gardner, R. H., and O'Neill, R. V. 1995. Ecological dynamics at broad scales: ecosystems and landscapes. BioScience Supplement: S29-S35.
Turner, M. G., Gardner, R. H., Dale, V. H., and O'Neill, R. V. 1989. Predicting the spread of disturbance across heterogeneous landscapes. Oikos 55: 121-129.
Walsh, S. J., Lightfoot, D. R., and Butler, D. R. 1987. Recognition and assessment of error in geographic information systems. Photogrammetric Engineering and Remote Sensing 53: 1423- 1430.
Wickham, J. D. and Norton, D. J. 1994. Mapping and analyzing landscape patterns. Landscape Ecology 9: 7-23.
Wiens, J. A. What is landscape ecology, really? Landscape Ecology 7: 149-150.
With, K. A. 1994. Using fractal analysis to assess how species perceive landscape structure. Landscape Ecology 9: 25-36.
Zipperer, W. C. 1993. Deforestation patterns and their effects on forest patches. Landscape Ecology 8: 177-184.

Appendix A. Diversity indexes tested

We tested our method using the Shannon diversity and Simpson dominance indexes, each of which provides similar information about the diversity of objects (land cover classes in our discussion). The Shannon evenness and Simpson diversity indexes are simple transformations of the Shannon diversity and Simpson dominance indexes, respectively. For further information about the differences among diversity indexes, their use, and their interpretation, see Pielou (1977) and Magurran (1988).

Index	**Formula**	**Range (Interpretation)**
Shannon Diversity	$\sum_{i=1}^{s} -p_i \bullet \log p_i$	0 (dominated by one cover) ↓ log s (all covers in equal extent)

Shannon Evenness	$\left[\sum_{i=1}^{s} -p_i \bullet \log p_i\right] \Big/ \log s$	0 (dominated by one cover) ↓ 1 (all covers in equal extent)
Simpson Dominance	$\sum_{i=1}^{s} p_i^2$	$1/s$ (all covers in equal extent) ↓ 1 (dominated by one cover)
Simpson Diversity	$1 \Big/ \sum_{i=1}^{s} p_i^2$	1 (dominated by one cover) ↓ $1/s$ (all covers in equal extent)

i land cover class
p_i proportion of scene in land cover class i
s number of different land cover classes
log the natural logarithm

Cautionary note: If the classification is extremely inaccurate, it is possible for the product of a user probability matrix bootstrap sample and the scene to result in an adjusted scene that contains a zero for one of the classes. This changes the value of s for that sample of the adjusted index. If this occurs, the confidence interval for the Shannon evenness index may not be obtained by dividing the Shannon diversity confidence interval by log s. This should not be a problem in practice because users would likely discard extremely inaccurate data.

Appendix B. S-plus code for generating confidence intervals

Efron and Tibshirani (1993) provide S-Plus (StatSci 1993) code for their bias-corrected and accelerated (BCa) method of calculating confidence intervals. We downloaded and installed their code as described in Efron and Tibshirani (1993). We modified the bcanon function to calculate a landscape index. In order to use our code, one must first download Efron and Tibshirani's code.

The S-Plus code we developed to calculate confidence intervals is available by anonymous FTP at ftp.ncsu.edu in directory /pub/ncsu/forest/grhess/conf, or you may contact George Hess at grhess@ ncsu.edu for a copy. *The code is provided as a courtesy. We would like to be informed of any errors in the code, but we do not have the resources to respond to extensive queries or software support requests.*