Range of Categorical Associations for Comparison of Maps with Mixed Pixels

Robert Gilmore Pontius, Jr. and John Connors

Abstract

This paper presents a method to compare maps that contain pixels that have partial membership to multiple categories, i.e., mixed or soft classified pixels. The method quantifies ranges for associations among categories based upon possible variations in sub-pixel spatial allocation. The paper derives the mathematical equations for constructing the range of associations based on three types of cross-tabulation matrices, the greatest matrix, the random matrix, and the least matrix. We demonstrate how the analysis can be combined with multiple resolution map comparison to specify the resolution at which clusters exist on a single map or between two maps. The method produces a range that reflects the amount of uncertainty in the categorical associations. We illustrate the procedure with both a simple example and data from the Plum Island Ecosystems study site in Massachusetts, USA.

Introduction

This paper presents a method to compare raster maps in which the pixels can have partial membership to multiple categories. The method has been designed for situations where the categorical membership represents the proportion of the pixel that contains the category, in other words, for cases where the quantity of each category within the pixel is known, but the spatial allocation of the categories within the pixel is not known. The procedure is useful for maps that have mixed pixels at the native resolution of the data. Furthermore, this approach can be used in conjunction with multiple resolution map comparison to analyze how scale influences the patterns of the spatial allocation of categories within a single map, or to assess patterns of categorical associations between two maps. Therefore, the procedure is able to characterize land change, even for cases where the categories in the map of the initial time are different than the categories for the map of the subsequent time.

Historically, it has been common for maps of land-cover to show exactly one pure category for each map unit, whether the units are vector polygons or raster pixels. However, regardless of the unit, each unit of a landscape can contain multiple categories, based on the interpretation of the map maker. One benefit of assigning a single pure category to each unit is that the analysis and comparison of the resulting map(s) is much easier when the units are assigned exactly one pure category. Methods such as those described by Congalton and Green (1999) are straight forward for pure categories because the computation of the cross-tabulation matrix is clear. The cost of assigning a single pure category to each unit is that the resulting maps do not necessarily reflect the uncertainty of the classification or the nature of the landscape. Woodcock and Gopal (2000) have established methods of map comparison based on fuzzy set theory for cases where the category assignments are ambiguous. Ambiguity can derive from the fact that: (a) the units have mixed membership to more than one category, or (b) interpreters can have difficulty assigning a category to units that do not fit clearly into a single category. The method of this paper addresses the first of these two cases.

In few situations, the mixed pixel problem can be addressed by attaining information that has finer resolution; however, finer resolution information is not necessarily a cure for the mixed pixel problem, because smaller pixels can also be mixed, and in many cases, smaller pixels can introduce new problems in category definition. For example, forest and built may be reasonable categories for 1 km resolution maps, but such pixels can be partially forest and partially built. Whereas 1 m resolution maps of the same study area can reveal pixels that are parts of trees or parts of buildings. Consequently, many investigators have suggested that hard classification is inappropriate for representing remotely sensed scenes (Fisher, 1997; Foody, 2002).

There have been several efforts to address the mixed pixel problem. Pontius (2002) proposed a minimum rule to measure the agreement between pixels that contain parts of identical categories, but that paper failed to address measurement of associations between different categories. Pontius and Cheuk (2006) proposed a method to construct a matrix to compare mixed pixels by first assigning the greatest possible association to agreement on the diagonal of the cross-tabulation matrix, and then leaving the disagreement to be distributed randomly among the off diagonal entries of the matrix. Additional approaches have been offered by Fisher *et al.* (2006) and Silván-Cárdenas and Wang (2008).

Even if the pixels are fairly pure at a particular resolution, investigators are frequently interested in patterns of

Robert Gilmore Pontius Jr. is with the School of Geography, Clark University, 950 Main Street, Worcester, MA 01610 (rpontius@clarku.edu).

John Connors is with the Department of Environmental Science, Policy & Management, University of California-Berkeley, 137 Mulford Hall, Rm. 3114, Berkeley, CA 94720, and also with the School of Geography, Clark University, Worcester, MA 01610 (jconnors@nature.berkeley.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 75, No. 8, August 2009, pp. 963–969.

0099-1112/09/7508–0963/$3.00/0

pixels at multiple scales and in how selection of scale influences the measurement of maps. This paper addresses scale effects by using multiple resolution map comparison, which is a technique that addresses the famous unsolved modifiable areal unit problem (MAUP). We examine how the statistical association between maps varies as a function of modifying the unit of analysis, i.e., the pixel. This is especially relevant for remotely sensed data, since Marceau and Hay (1999) proposed that remotely sensed data represent a unique case of the MAUP, in which the grain and extent of the data define areal units for analysis. Furthermore, this paper addresses the call by Fotheringham (1989) to address the MAUP since we develop new methods, conduct sensitivity analysis, and search for relationships between fluctuations in variables with respect to scale. We extend the work of Kuzera and Pontius (2008) who examined how three different types of cross-tabulation matrices measure the agreement between maps at multiple scales.

In order to explore the relationships between resolution, pattern, and uncertainty in map comparison, this paper constructs a range of categorical associations at multiple resolutions by systematically coarsening the maps. Woodcock and Strahler (1987) used a similar approach to characterize the relationship between spatial pattern and local variance. Similar methods have been employed to explore the influence of grain size on structural parameters (Benson and Mackenzie, 1995) and landscape metrics (Wu, 2004). It is increasingly important to understand the effects of these types of scale modification, since data of different resolutions are increasingly used together for change analysis, especially where historic data tend to be at coarser scales than contemporary data.

Methods

Data

Example to Show Multiple Resolution Map Comparison

Figure 1 gives a simplified example to show how we convert from fine resolution data where each pixel contains exactly one category to coarser resolution data where each pixel can contain partial memberships to more than one category. We consider the process of pixel aggregation for two example scenes shown in the two rows of maps, at three resolutions shown in the three columns. The solid gray lines bound each pixel, whereas the dashed lines show the boundaries that dissolve as we move to coarser resolutions. The fine resolution maps contain 16 pixels with a 1 m resolution, where each pixel is classified as entirely black or entirely white. At the middle resolution, the maps have been aggregated to four coarser pixels, each with a resolution of 2 m. The values of the resulting pixels reflect the proportional contributions of each of the four contributing pixels. Thus, each pixel in Scene 1 at a resolution of 2 m still has membership to only one class because all four of the fine resolution pixels that are aggregated to the middle resolution belong to the same class. But, the pixels in Scene 2 at a resolution of 2 m are half black and half white because two out of the four aggregated pixels were black and two were white. At the coarse resolution, there is one 4 m resolution pixel for each scene, and both are identical since they have membership of one-half to both white and black. Our method gives equations to construct three different two-by-two cross-tabulation matrices to compare these two scenes at each resolution.

Plum Island Ecosystems to Show Application

The land-cover maps of the Plum Island Ecosystems study area in Massachusetts, USA, were created from categorical maps of vector polygons for 1971 and 1999, available online

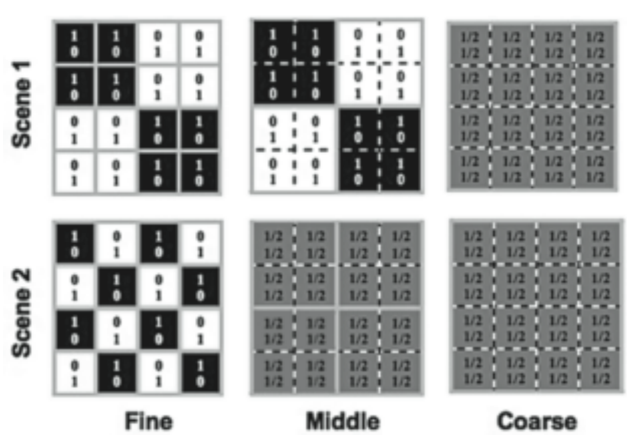


Figure 1. The scenes show two classes: black and white. The top number in each pixel indicates membership to the black class; the lower number indicates membership to the white class. Light grey lines designate the boundaries of pixels and dashed lines are the boundaries that dissolve to form the coarser resolutions.

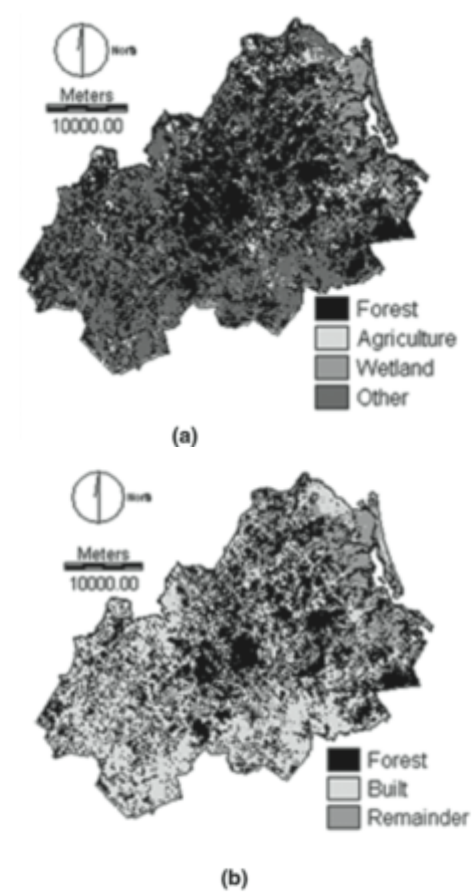


Figure 2. Land cover maps from (a) 1971, and (b) 1999 for the 26 towns of the Plum Island Ecosystems study area in northeastern Massachusetts.

at the State of Massachusetts' GIS database (MassGIS, 2005). These maps were derived from aerial photography and orthoimagery. There are neither published minimum mapping units nor published levels of precision for the vector maps. We converted the original vector maps to raster images with a 30-meter resolution. The selection of a 30-meter pixel size was strategic but was not related to the precision of the data, since the precision of the data is unknown. The 30 m pixel resolution was selected to be consistent with other types of digital information and to produce file sizes that were as detailed as possible while being still manageable in size. This is one reason why it is important to use multiple resolution map comparison methods, since the initial resolution is frequently somewhat arbitrary. Furthermore, the conversion to raster format produces pixels that are classified as exactly one category, which is a common artifact of vector to raster conversion. Figure 2 shows the resulting maps that have the categories of forest, agriculture, wetland, and other in 1971, and forest, built, and remainder in 1999. We selected aggregation to these categories for two reasons. First, our broader research project is most interested in transitions from those categories of 1971 to those categories in 1999. Second, this selection of categories shows how the methods of this paper can be used for cases where the categories of the initial time are different than the categories of the subsequent time.

Notation

This paper uses the following mathematical notation:

- $g \equiv$ grain size of pixels expressed as a multiple of initial resolution where $g = 1, \ldots, G$;
- $G \equiv$ grain size for the resolution at which the entire study area is in a single pixel;
- $n_g \equiv$ index of a specific pixel at resolution g where $n_g = 1, \ldots, N_g$;
- $N_g \equiv$ number of pixels in study area at resolution g;
- $i \equiv$ index for a particular class in first map where $i = 1, \ldots, I$;
- $I \equiv$ number of classes in first map;
- $j \equiv$ index for a particular class in second map where $j = 1, \ldots, J$;
- $J \equiv$ number of classes in second map;
- $Wn_g \equiv$ weight of pixel n_g;
- $Xn_{gi} \equiv$ membership of pixel n_g in the first map to class i where $0 \leq Xn_{gi} \leq 1$;
- $Yn_{gj} \equiv$ membership of pixel n_g in the second map to class j where $0 \leq Yn_{gj} \leq 1$;
- $Mn_{gij} \equiv$ greatest association between class i and class j for pixel n_g;
- $Rn_{gij} \equiv$ random association between class i and class j for pixel n_g;
- $Ln_{gij} \equiv$ least association between class i and class j for pixel n_g;
- $M\bullet_{gij} \equiv$ greatest association between class i and class j for study area at resolution g;
- $R\bullet_{gij} \equiv$ random association between class i and class j for study area at resolution g;
- $L\bullet_{gij} \equiv$ least association between class i and class j for study area at resolution g.

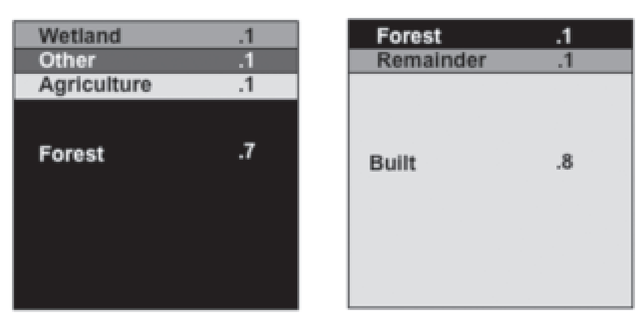
Three Different Matrices

The rows of each matrix show the categories of the first map and the columns show the categories of the second map. If each pixel belongs entirely to exactly one category, then there is no controversy concerning how to compute the entries in the cross-tabulation matrix. Each pixel position is tallied in the matrix according to the pixel's membership in each of the maps. It is common to divide all the tallies by the total number of pixels to attain a proportion of the study area for each entry in the matrix. However, if the pixels have mixed membership to more than one category, it is not obvious how to construct the cross-tabulation matrix. There are many factors to consider when deciding how to construct the matrix.

Foremost, one must be clear concerning the interpretation of the numbers that indicate the partial memberships. The methods of this paper apply to cases where the partial membership to each category represents the proportion of the pixel that is covered by the said category. Therefore, each categorical membership is bounded between 0 and 1, and the sum of all memberships in each pixel is 1. In this respect, the quantity of each category in the pixel is known, as indicated by the membership value, but the spatial allocation of the categories within the pixel is not known. Our method considers three spatial allocations among the possibly infinite number of allocations for the categories within the pixels. The technique then computes the mathematically possible range for each entry in the cross-tabulation matrix, where each entry is computed based upon a possibly different spatial allocation of the categories within the pixels.

Greatest Matrix

The first matrix that we produce is called the greatest matrix, since it calculates the greatest possible association between the categories, i.e., it leads to the largest mathematically possible entries in the matrix that are constrained by the pixel memberships. For example, consider if we need to compare a pixel of 1971 that has membership of 0.7 to the forest category and 0.1 to the agriculture, wetland, and other categories, vis-à-vis, a pixel from 1999 that has membership of 0.8 to the built category and 0.1 to the forest and remainder categories. Figure 3 shows a sub-pixel configuration that yields the greatest possible forest to built association. If the 1971 pixel were placed directly over the 1999 pixel, then there would be the maximum possible overlap between the forest of 1971 with the built of 1999. The amount of overlap between any two classes can attain but never exceed the smaller proportion of the two classes. Equation 1 calculates this relationship for a single pixel, for every position in the matrix by selecting the smaller of the two values being compared, Xn_{gi} from the first map and Yn_{gj} from the second map. Notice that we consider a new spatial reallocation to compute each entry in the matrix, therefore the sum of the entries in the matrix can exceed 1. The earliest citation we have of this matrix is Binaghi *et al.* (1999). Equation 2 calculates the greatest matrix when comparing entire maps. This equation calculates a standardized average using weights for each pixel. This allows us to compute the entries for non-square study areas when performing multiple resolution



(a) (b)

Figure 3. Sub-pixel spatial allocation that indicates the greatest possible categorical association between (a) forest in 1971, and (b) built in 1999.

map comparison. Non-square study areas lead to coarse pixels that are partially in and partially out of the study area, so those coarse pixels are weighted by their proportions that are in the study area.

$$Mn_{gij} = \mathrm{MIN}(Xn_{gi}, Yn_{gj}) \tag{1}$$

$$M \bullet_{gij} = \frac{\sum_{n_g=1}^{N_g} Wn_g \times \mathrm{MIN}(Xn_{gi}, Yn_{gj})}{\sum_{n_g=1}^{N_g} Wn_g} \tag{2}$$

Random Matrix

The next matrix that we produce is called the random matrix, since it calculates the statistically expected values for the entries assuming that the spatial allocation of the categories within the pixels is random, given the pixel memberships. Figure 4 illustrates a random sub pixel allocation for the same pixel memberships as Figure 3. If the 1971 pixel were placed directly over the 1999 pixel, then there would be a statistical expectation of the amount of overlap of each category in 1971 on each category in 1999. The random matrix gives that expected amount of overlap. Equation 3 calculates this relationship for a single pixel for every position in the matrix by multiplying the two values being compared: Xn_{gi} from the first map and Yn_{gj} from the second map. Notice that the sum of the entries in the Random matrix is always 1. Lewis and Brown (2001) is the earliest citation we have of this matrix. Equation 4 calculates the random matrix when comparing entire maps, in a manner similar to Equation 2.

$$Rn_{gij} = Xn_{gni} \times Yn_{gj} \tag{3}$$

$$R \bullet_{gij} = \frac{\sum_{n_g=1}^{N_g} Wn_g \times (Xn_{gi} \times Yn_{gj})}{\sum_{n_g=1}^{N_g} Wn_g} \tag{4}$$

Least Matrix

The third matrix that we produce is called the least matrix, since it calculates the smallest mathematically possible values for the entries, given the pixel memberships. Figure 5 illustrates a sub-pixel allocation that minimizes the forest of 1971 overlap on built of 1999, for the same pixel memberships as Figures 3 and 4. When the 1971 pixel is placed directly over the 1999 pixel, we must have some overlap of forest on built. The Least matrix gives the minimum amount of overlap. Equation 5 calculates this relationship for a single pixel, for every position in the matrix. Notice that the sum of the entries in the least matrix can be less than 1. Equation 4 calculates the least matrix when comparing entire maps, in a manner similar to Equation 2.

$$Ln_{gij} = \mathrm{MAX}(0, Xn_{gi} + Yn_{gj} - 1) \tag{5}$$

$$L \bullet_{gij} = \frac{\sum_{n_g=1}^{N_g} Wn_g \times \mathrm{MAX}(0, Xn_{gi} + Yn_{gj} - 1)}{\sum_{n_g=1}^{N_g} Wn_g} \tag{6}$$

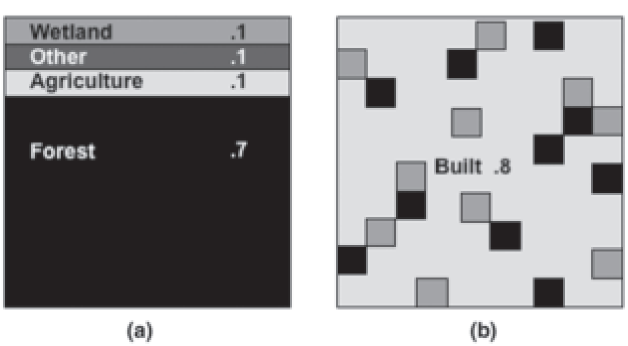


Figure 4. Sub-pixel spatial allocation that indicates random categorical association: (a) 1971 pixel, and (b) 1999 pixel.

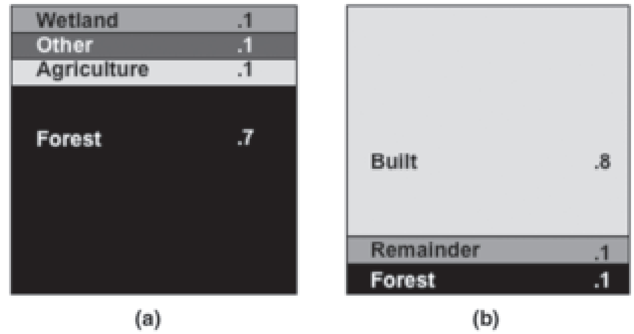


Figure 5. Sub-pixel spatial allocation that represents the least possible categorical association between (a) forest in 1971, and (b) built in 1999.

From Table to Range to Map

Table 1 gives all three matrices in one cross tabulation for the pixel used in Figures 3, 4, and 5. The upper number in each position of the matrix gives the entry for the greatest matrix, the middle number gives the entry for the random matrix, and the lower number gives the entry for the least matrix. The range for each position in the matrix is computed as the greatest entry minus the least entry.

It can be helpful to summarize information about numerous pixels in the format of a table; however, the table fails to reveal how the information is distributed across the

Table 1. Transition Matrix where Each Cell Contains the Greatest Entry on the Top, the Random Entry in the Middle, and the Least Entry at the Bottom for the Pair of Pixels in Figures 3, 4 and 5

		1999			
		Forest	Built	Remainder	1971 Membership
1971	Forest	0.10 0.07 0.00	0.70 0.56 0.50	0.10 0.07 0.00	0.7
	Agriculture	0.10 0.01 0.00	0.10 0.08 0.00	0.10 0.01 0.00	0.1
	Wetland	0.10 0.01 0.00	0.10 0.08 0.00	0.10 0.01 0.00	0.1
	Other	0.10 0.01 0.00	0.10 0.08 0.00	0.10 0.01 0.00	0.1
	1999 Membership	0.1	0.8	0.1	1.0

landscape. Therefore, we have found it useful to map the values for the three matrices and the range for each particular transition of interest. This is possible because every pixel generates all three matrices and a range matrix. The results section shows the maps for the transition from forest in 1971 to built in 1999.

Results

Example to Show Multiple Resolutions

Figure 6 shows the association between the black category in scene 1 and the white category in scene 2 for the example data at multiple resolutions. All three matrices produce the same association of one half at resolutions 1 and 2, because all the pixels are pure in scene 1 at those two resolutions. This illustrates a general principal that all three matrices produce identical results for a particular pixel location when either of the pixels belongs completely to one category.
At resolution 4, the range for the association between black in scene 1 and white in scene 2 spans from zero to one-half, since both coarse pixels have a membership of one half to both categories at resolution 4.

Figure 7 shows the association between the black category of scene 2 and the white category of scene 2. This illustrates how the method can be used to compare a map to itself. For such applications, the method reveals how intermingled the categories are within a single map. When combined with multiple resolution analysis, the results reveal the distances over which this intermingling or clustering occurs. Figure 7 shows that the results jump from zero at resolution 1 to a range that spans from zero to one-half at resolution 2. This illustrates how the multiple resolution analysis reveals the distances at which the compared categories are intermingled since the range increases at that resolution, thus indicating the relationship between scale and pattern along with their combined impact on uncertainty. At resolution 2, the black and white categories are spread uniformly across scene 2, so Figure 7 shows that the association between the categories is constant from resolution 2 to 4.

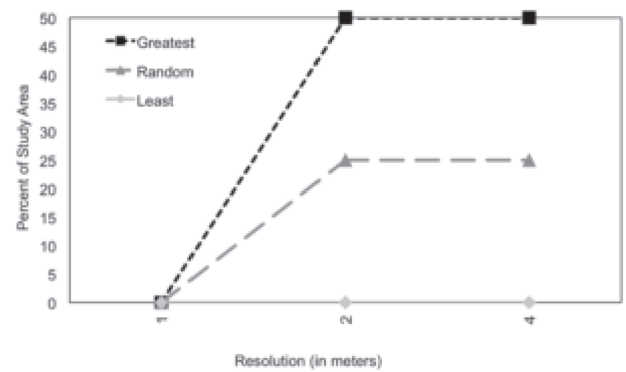


Figure 7. Categorical association between the black class in scene 2 and the white class in scene 2.

Plum Island Ecosystems

Figures 8 and 9 show results for the Plum Island Ecosystems data that are analogous to Figures 6 and 7 for the example data. Figure 8 gives the association between the forest of 1971 and the built of 1999. All three matrices indicate that eight percent of the study area transitioned from forest to built at the 30-meter resolution, since the pixels are pure at this resolution as an artifact of the data formatting process. The range defined by the greatest and least matrices increases as resolution becomes coarser in a manner that is related to the spatial pattern of the two relevant categories. The range grows more quickly for the distances at which the forest of 1971 is more intermingled with the built of 1999. The entire study area resides in a single coarse pixel at a resolution of 61,440 m, for which the range extends from zero to nearly half of the study area. Figure 9 gives the association between the forest of 1999 and the built of 1999. The range in Figure 9 grows more quickly for the distances at which the forest of 1999 is more intermingled with the built of 1999.

Figures 10, 11, 12, and 13 show the mapped results from the three matrices for the transition from forest in 1971 to built in 1999 at a resolution of 360 m, where the darker shades indicate higher proportions of the pixel that contains the transition. Figure 10 derives from the greatest matrix, Figure 11 from the random matrix, and Figure 12 from the least matrix.

Figure 13 gives the range in the transition from forest to built as defined by the greatest values in Figure 10 minus the least values in Figure 12. The circles on the map indicate areas of especially high or low uncertainty, where the corresponding boxes explain the reason for the range. The range is largest in pixels that contain half forest in 1971 and half built in 1999, in which case there is maximum uncertainty concerning the association between forest and

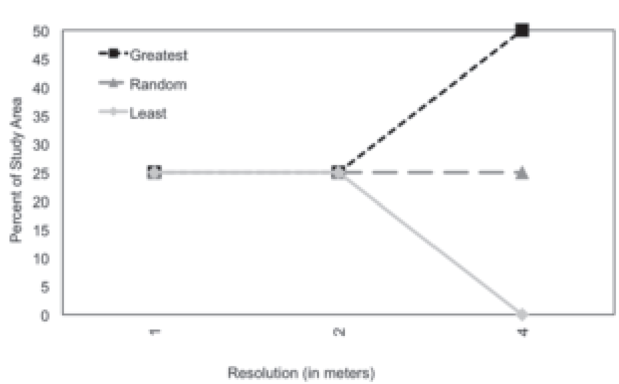


Figure 6. Categorical association between the black class in scene 1 and the white class in scene 2.

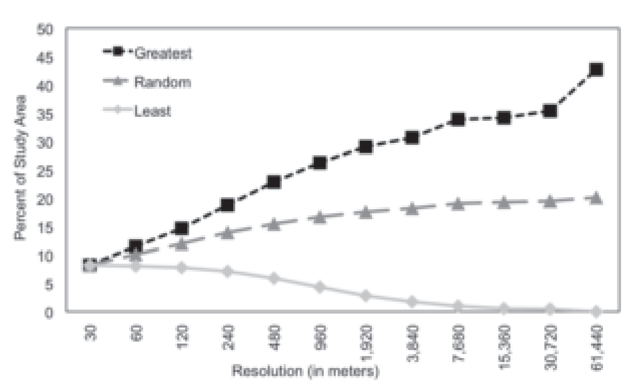


Figure 8. Range of categorical association between forest in 1971 and built in 1999 at multiple resolutions for the Plum Island Ecosystems study area.

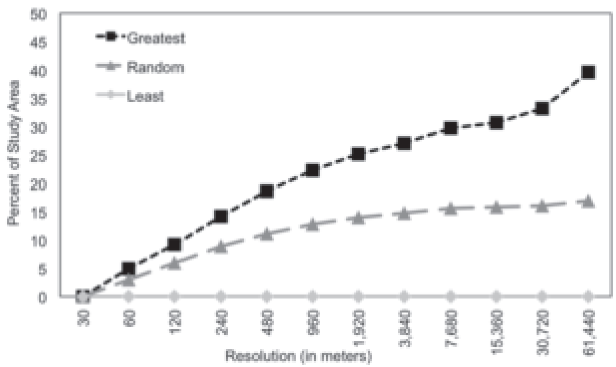


Figure 9. Range of categorical association between forest in 1999 and built in 1999 at multiple resolutions for the Plum Island Ecosystems study area.

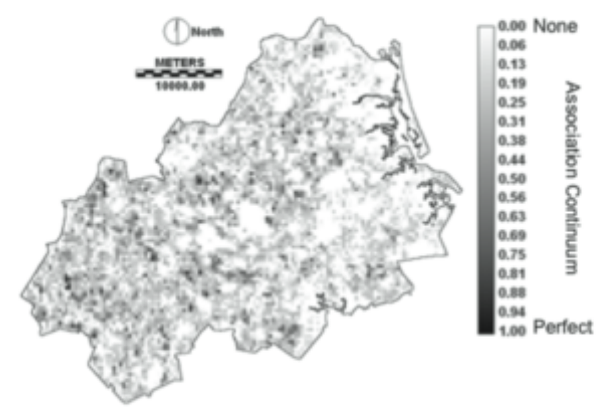


Figure 11. Random association between forest in 1971 and built in 1999 at a resolution of 360 meters for the Plum Island Ecosystems study area.

built. If a pixel lacks forest in 1971 or lacks built in 1999, then the transition is impossible, so uncertainty of the association between forest and built is zero. If a pixel contains a high proportion of forest in 1971 and a high proportion of built in 1999, then a transition from forest to built is assured, so the uncertainty is low.

Discussion

There are a variety of issues to consider when one interprets the output that derives from this paper's methods. This discussion section anticipates some of the questions that others will likely encounter when applying the methods of this paper.

The methods of this paper give the mathematically possible range for the association between the category in the first map and the category in the second map, but all values in this range are not necessarily equally likely. Consider the case where the categories of the initial and subsequent times are identical, so we can construct a square matrix where the diagonal entries record land persistence and the off-diagonal entries record land change. Many landscapes are dominated by persistence, where change is a relatively rare event. In this case, the true value of persistence would likely be between the random estimate and the greatest estimate for the diagonal entries, and the true value of change would be between the random estimate and the least estimate for the off-diagonal entries. An additional validation exercise with finer resolution data might establish more precisely where the most appropriate estimate lies within the range. However, an additional validation exercise would require additional higher quality data, which may not be available, especially for the past.

The method can be used for other applications not illustrated in this paper. For example, it is possible to use the method to characterize the patchiness of a single category by comparing a map to itself, and then examining an entry that is on the diagonal of the resulting square matrix. This approach is particularly effective when combined with multiple resolution map comparison, so one can see the resolutions at which pixels of specific categories are clustered.

This paper uses a multi-resolution analysis to assess the relationship between scale and pattern; however, the methods of this paper can be useful even if one is not interested in scale effects or in patterns across the landscape. The equations derived to generate the range of associations are still useful in the simpler case of comparing

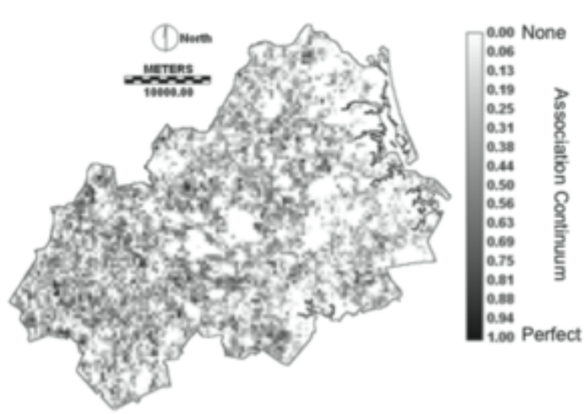


Figure 10. Greatest association between forest in 1971 and built in 1999 at a resolution of 360 meters for the Plum Island Ecosystems study area.

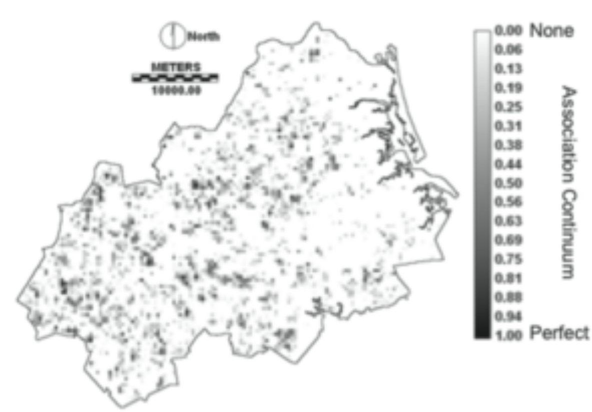


Figure 12. Least association between forest in 1971 and built in 1999 at a resolution of 360 meters for the Plum Island Ecosystems study area.

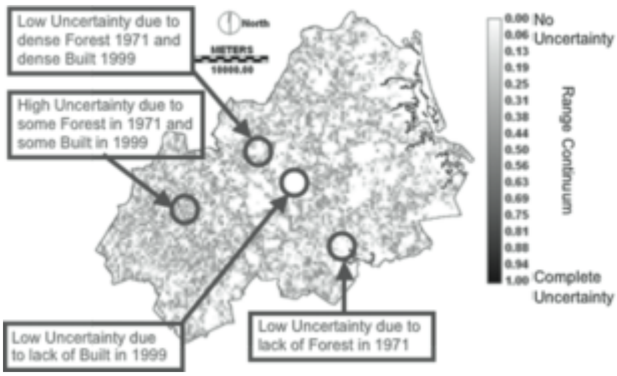


Figure 13. Range of association between forest in 1971 and built in 1999 at a resolution of 360 meters for the Plum Island Ecosystems study area.

two maps at a single resolution, where the pixels have mixed membership to more than one category.

Conclusions

This paper presents a method to address the mixed pixel problem, for which the proportion of each category in a pixel is known, but the spatial allocation of the categories within the pixel is not known. The method defines a range of possibilities for the cross-tabulation matrix to quantify the association between any two categories. The technique is appropriate to compare two different maps or to assess the patterns within a single map. When combined with multiple resolution map comparison, the method reveals the distances over which any two categories are intermingled. The technique marks another step in quantifying categorical associations at multiple scales.

Acknowledgments

The United States' National Science Foundation (NSF) supported this work through three of its programs: Human-Environment Regional Observatory program using Grant 9978052, Long Term Ecological Research using Grant OCE-0423565, and Coupled Natural Human Systems using Grant BCS-0709685. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect those of the NSF. Clark Labs facilitated this work by creating the GIS software Idrisi®. We also thank the anonymous reviewers.

References

Benson, B.J., and M.D. Mackenzie, 1995. Effects of sensor spatial resolution on landscape structure parameters, *Landscape Ecology*, 10(2):113–120.

Binaghi, E., P. Brivio, P. Ghezzi, and A. Rampini, 1999. A fuzzy set-based accuracy assessment for soft-classification, *Pattern Recognition Letters*, 20:935–948.

Congalton, R.G., and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, Boca Raton, Florida.

Fisher, P., 1997. The pixel: A snare and a delusion, *International Journal of Remote Sensing*, 18(3):679–685.

Fisher, P., C. Arnot, R. Wadsworth, and J. Wellens. 2006. Detecting change in vague interpretations of landscapes, *Ecological Informatics*, 1(2):167–178.

Fotheringham, S. 1989. Scale-independent spatial analysis, *The Accuracy of Spatial Databases* (M. Goodchild and S. Gopal, editors), Taylor & Francis, pp. 221–228.

Foody, G.M., 2002. The role of soft classification techniques in the refinement of estimates of ground control point location, *Photogrammetric Engineering & Remote Sensing*, 68(9):897–903.

Kuzera, K., and R.G. Pontius, Jr. 2008. Importance of matrix construction for multiple-resolution categorical map comparison, *GIS and Remote Sensing*, 45(3):249–274.

Lewis, H.G., and M. Brown, 2001. A generalized confusion matrix for assessing area estimates from remotely sensed data, *International Journal of Remote Sensing*, 22:3223–3235.

MassGIS, 2005. Land Use Map, URL: *http://www.mass.gov/mgis/lus.htm*, Executive Office of Environmental Affairs, Boston, Massachusetts (last date accessed: 14 May 2009).

Marceau, D.J., and G.J. Hay, 1999. Remote sensing contributions to the scale issue, *Canadian Journal of Remote Sensing*, 28 (4):357–366.

Pontius, R.G., Jr., 2002. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions, *Photogrammetric Engineering & Remote Sensing*, 68(10):1041–1049.

Pontius, R.G., Jr. and M.L. Cheuk, 2006. A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions, *International Journal of Geographical Information Science*, 20(1):1–30.

Silván-Cárdenas, J.L., and L. Wang, 2008. Sub-pixel confusion-uncertainty matrix for assessing soft classifications, *Remote Sensing of Environment*, 112:1081–1095.

Woodcock, C.E., and S. Gopal 2000. Fuzzy set theory and thematic maps: accuracy assessment and area estimation, *International Journal of Geographic Information Science*, 14(2):153–172.

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

Wu, J. 2004. Effects of changing scale on landscape pattern analysis: Scaling relations, *Landscape Ecology*, 19:125–138.