

Expanding the conceptual, mathematical and practical methods for map comparison

Robert Gilmore Pontius Jr and John Connors

Clark University
School of Geography
950 Main Street
Tel.: + 001 508 793 7761; Fax: + 001 508 793 8881
rpontius@clarku.edu; jconnors@clarku.edu

Abstract

Conventional methods of map comparison frequently produce unhelpful results for a variety of reasons. In particular, conventional methods usually analyze pixels at a single default scale and frequently insist that each pixel belongs to exactly one category. The purpose of this paper is to offer improved methods so that scientists can obtain more helpful results by performing multiple resolution analysis on pixels that belong simultaneously to several categories. This paper examines the fundamentals of map comparison beginning from the elementary comparison between two pixels that have partial membership to multiple categories. We examine the conceptual foundation of three methods to create a crosstabulation matrix for a single pair of pixels, and then show how to extend those concepts to compare entire maps at multiple spatial resolutions. This approach is important because the crosstabulation matrix is the basis for numerous popular measurements of spatial accuracy. The three methods show the range of possibilities for constructing a crosstabulation matrix based on possible variations in the spatial arrangement of the categories within a single pixel. A smaller range in the possible spatial distribution of categories within the pixel corresponds to more certainty in the crosstabulation matrix. The quantity of each category within each pixel constrains the range for possible arrangements in subpixel mapping, since there is more certainty for pixels that are dominated by a single category. In this respect, the proposed approach is placed in the context of a philosophy of map comparison that focuses on two separable components of information in a map: 1) information concerning the proportional distribution of the quantity of categories, and 2) information concerning the spatial distribution of the location of categories. The methods apply to cases where a scientist needs to compare two maps that show categories, even when the categories in one map are different from the categories in the other map. We offer a fourth method that is designed for the common case where a scientist needs to compare two maps that show the same set of categories. Results show that the methods can produce extremely different measurements, and that it is possible to interpret the differences at multiple resolutions in a manner that reveals patterns in the maps. The method is designed to present the results graphically in order to facilitate communication. We describe the concepts using simplified examples, and then apply the methods to characterize the change in land cover between 1971 and 1999 in Massachusetts, USA.

Keywords: accuracy, fuzzy, error, matrix, uncertainty

1 Introduction

The goal of this paper is to propose useful intuitive general statistical methods to compare any two raster maps where the pixels can have partial membership to multiple categories. The

proposed methods are designed to be applicable even for cases where the categories in one map are different than the categories in the other map.

Our proposed method relies on the crosstabulation matrix, also known as the contingency table. The crosstabulation matrix is the bedrock on which a mountain of statistical analyses for categorical variables is based. Crosstabulation matrices are used regularly to measure the spatial accuracy of raster maps (Congalton and Green 1999) and more generally to quantify the association between two categorical maps for a variety of reasons (Pontius et al. 2004a, Pontius et al. 2004b, Pontius and Spencer 2005). The columns of the matrix show the categories of the variable in the reference map for a study area and the rows show the categories of the variable in the comparison map for the same study area. If the reference map and the comparison map are realizations of a single categorical variable, then the category labels for the columns are the same as the category labels for the rows, in which the diagonal entries of the matrix indicate map agreement. The complete matrix that compares the entire reference map to the entire comparison map is the sum of the tallies of all the individual pixel positions in the study area. If each pixel position in the study area is classified as exactly one category in the reference map and exactly one category in the comparison map, i.e. if the pixels are hard classified, then the pixel position is tallied in a single obvious column and row within the matrix. One reason why scientists are tempted to classify each pixel in the map as exactly one category is because the statistical analysis of such data is straight forward. However, the decision to hard classify the data introduces a fundamental problem because scientists frequently know that each pixel really contains partial membership to multiple categories. Consequently, the data processing step to hard classify the pixels corrupts the underlying information in the maps. Nevertheless, scientists regularly classify each pixel as its single dominant category, because there are not widely recognized intuitive statistical methods available to compare maps that have partial membership to multiple categories. Specifically if the pixels do not belong to exactly one category, i.e. if the pixels are soft classified, then the construction of the crosstabulation matrix is not immediately obvious (Pontius 2002, Pontius and Cheuk 2006).

Consider the following riddle that shows how tricky it can be to analyze soft classified pixels. Assume the reference pixel contains membership of 0.7 to the black category and 0.3 to the white category, while the comparison pixel contains membership of 0.9 to the black category and 0.1 to the white category. How should we measure the association between the categories of the reference pixel to the categories of the comparison pixel? (Figure 1)



Figure 1. The reference pixel and comparison pixel with partial membership to the black category as specified by the top number and to the white category as specified by the bottom number within the pixel.

If we consider the partial memberships as probabilities, then it makes sense to multiply the memberships in order to quantify the categorical associations. If we follow this logic, then the

black-to-black association is $0.7 \times 0.9 = 0.63$ and the white-to-white association is $0.3 \times 0.1 = 0.03$. So the overall agreement is $0.63 + 0.03 = 0.66$. At first this might seem reasonable, because the overall agreement is less than 1 for the two pixels that are not identical.

If we were to compare the reference pixel to itself using the same logic, then what would the resulting associations? The black-to-black association would be $0.7 \times 0.7 = 0.49$ and the white-to-white association would be $0.3 \times 0.3 = 0.09$. So the overall agreement would be $0.49 + 0.09 = 0.58$. This result is unsatisfying for at least two reasons. First, the agreement between a pixel and itself is less than 1. Second, the agreement between the reference pixel and itself is less than the agreement between the reference pixel and a different pixel.

If we think more deeply about the problem, we find that there can be a range of reasonable answers. To envision this range, it is helpful to think of the two categories as black and white paint within the pixels, where the proportion of the pixel's area covered by each color of paint is equal to the pixel's partial membership to the corresponding category. We can then consider various possibilities for the spatial arrangement of the paint within each pixel. With this analogy, we can envision possibilities for how the paint within the reference pixel could overlay on the paint within the comparison pixel, so that the amount of overlap for a particular combination of paint is the measurement of association between the corresponding combination of categories (Pontius and Suedmeyer 2004). For example, if both the reference pixel and the comparison pixel were to have its black paint concentrated in the northern part of the pixel, then the black-on-black overlap would be as large as possible. In this case, given the memberships in Figure 1, the greatest possible black-to-black association would be 0.7, which is equal to $\text{MIN}(0.7, 0.9)$. If the black paint were distributed randomly within either pixel, then the black-to-black association would be $0.7 * 0.9 = 0.63$, as initially computed in the riddle. The least possible association would occur if the black paint were in the northern part of the reference pixel and in the southern part of the comparison pixel, in which case there would still be an overlap of 0.6, which is equal to $\text{MAX}(0, 0.7+0.9-1.0)$.

Clearly, scientists need to think more deeply about how to compare a single pair of pixels that have partial membership to multiple categories, because this information is necessary to understand the level of certainty we place in maps (Foody and Atkinson 2002). First, we need to create a solid foundation to compare two pixels, then we can expand the methods to compare entire maps, because maps are collections of pixels. The purpose of this paper is to articulate general concepts and to lay a unifying mathematical foundation for categorical map comparison.

2 Methods

2.1 Four matrices for a single pixel position

This first subsection (2.1) uses the riddle of Figure 1 to motivate the mathematical principles to compare two pixels. The second subsection (2.2) gives the equations to compute the crosstabulation matrix to compare two entire maps. The third subsection (2.3) illustrates those principles with example maps. A fourth subsection (2.4) presents data for a practical application, which we analyze with the proposed methods.

This paper uses the following mathematical notation:

g – grain size of the pixels, i.e. resolution,

G – maximum resolution where the entire study area resides in one coarse pixel,

n – index for each pixel position in study area,
 N_g – number of pixels of data in study area at resolution g ,
 i – index for each category in comparison map X ,
 I – number of categories in comparison map X ,
 j – index for each category in reference map Y ,
 J – number of categories in reference map Y ,
 W_{gn} – weight of pixel n at resolution g ,
 X_{gni} – membership to category i of pixel n at resolution g in comparison map X ,
 Y_{gnj} – membership to category j of pixel n at resolution g in reference map Y ,
 L_{gnij} – entry in row i column j of the Least matrix for pixel n at resolution g ,
 R_{gnij} – entry in row i column j of the Random matrix for pixel n at resolution g ,
 M_{gnij} – entry in row i column j of the Most matrix for pixel n at resolution g and diagonal entry of the Composite matrix for pixel n at resolution g ,
 C_{gnij} – off-diagonal entry in row i column j of the Composite matrix for pixel n at resolution g ,
 L_{g+ij} – entry in row i column j of the Least matrix for study area at resolution g ,
 R_{g+ij} – entry in row i column j of the Random matrix for study area at resolution g ,
 M_{g+ij} – entry in row i column j of the Most matrix for study area at resolution g and diagonal entry of Composite matrix for study area at resolution g ,
 C_{g+ij} – off-diagonal entry in row i column j of the Composite matrix for study area at resolution g .

The memberships are constrained to conform to inequalities (1) and (2) for each pixel. The memberships must sum to 1 such that equations (3) and (4) hold for each pixel.

$$0 \leq X_{gni} \leq 1 \quad (1)$$

$$0 \leq Y_{gnj} \leq 1 \quad (2)$$

$$\sum_{i=1}^I X_{gni} = 1 \quad (3)$$

$$\sum_{j=1}^J Y_{gnj} = 1 \quad (4)$$

Comparison of the pair of pixels at a single position n in the map generates an entire crosstabulation matrix. We consider four methods to compute the entries in the crosstabulation matrix. The four matrices are called: the Least matrix, the Random matrix, the Most matrix, and the Composite matrix.

2.1.1 Least matrix

The entries in the Least matrix give the least possible association between the categories in the reference pixel and the categories in the comparison pixel. We consider possible arrangements of the categories within pixel n such that category i in the comparison pixel overlaps as little as possible with category j in the reference pixel. If the sum of the pair of memberships is less than 1, then it is possible to arrange the categories within the pixels such that there is no overlap. If the sum of the pair of memberships is greater than 1, then category i in the

comparison pixel must have some positive overlap with category j in the reference pixel. Equation 5 gives the least possible association between the comparison category i and the reference category j within pixel n . We consider a different possible rearrangement for each entry in the matrix, so it is possible that the sum of all values in the matrix is less than 1.

$$L_{gnij} = \text{MAX}(0, X_{gni} + Y_{gnj} - 1) \quad (5)$$

2.1.2 Random matrix

The entries in the Random matrix give the statistically expected association between the categories in the reference pixel and the categories in the comparison pixel, assuming those categories are distributed randomly within the pixels. The membership to each category is the proportion of that category contained in the pixel, so the amount of overlap between a pair of categories is the mathematical product of the memberships of those categories. Equation 6 gives the expected association between comparison category i and reference category j within pixel n , assuming random distribution of the categories within pixel n . Equation 6 produces the Random matrix such that the sum of all values in the matrix is equal to 1. Lewis and Brown (2001) propose an equivalent procedure.

$$R_{gnij} = X_{gni} \times Y_{gnj} \quad (6)$$

2.1.3 Most matrix

The entries in the Most matrix give the greatest possible association between the categories in the reference pixel and the categories in the comparison pixel. We consider possible arrangements of the categories within pixel n such that category i in the comparison pixel overlaps as much as possible with category j in the reference pixel. The maximum overlap is constrained by the minimum of the memberships to the two categories. Equation 7 gives the greatest possible association between comparison category i and reference category j within pixel n . We consider a different possible rearrangement for each entry in the matrix, so it is possible that the sum of all entries in the Most matrix is greater than 1. Binaghi et al. (1999) propose this same minimum rule for all entries in the matrix.

$$M_{gnij} = \text{MIN}(X_{gni}, Y_{gnj}) \quad (7)$$

2.1.4 Composite matrix

The entries of the Composite matrix are computed according to a mathematical rule that is a composite of the concepts expressed by the minimum rule of equation 7 and the multiplication rule of equation 6. The Composite matrix is designed specifically for the special case where the categories in the reference map are the same as the categories in the comparison map. For this common situation, the matrix is square, where the diagonal entries show agreement between the categories, while the off-diagonal entries show disagreement. Equation 7 gives the diagonal entries, i.e. when $i = j$. Equation 8 gives the off-diagonal entries, i.e. when $i \neq j$. Pontius and Cheuk (2006) give the full derivation of the Composite matrix.

$$C_{gnij} = [X_{gni} - \text{MIN}(X_{gni}, Y_{gnj})] \times [Y_{gnj} - \text{MIN}(X_{gni}, Y_{gnj})] / \sum_{j=1}^J [Y_{gnj} - \text{MIN}(X_{gni}, Y_{gnj})] \quad (8)$$

The Composite matrix has conceptual advantages over the other three matrices when the reference map and the comparison map both show the same categorical variable (Kuzera and Pontius 2004). Specifically, when a pixel is compared to itself, the Composite matrix gives zeroes for all off-diagonal entries, which is not necessarily the case for the Random and Most matrices. Furthermore, the entries of the Composite rule sum to 1, which is not necessarily the case for the Least and Most matrices. The Discussion section compares the relative merits of the four matrices in more detail.

2.2 Four matrices to compare two maps

The previous subsection (2.1) defines four methods to compute a crosstabulation matrix for a single pixel position, while this section describes how to compute a crosstabulation matrix for an entire study area composed of many pixels. The approach to generate the map-level matrix is to compute a weighted average of all pixel-level matrices. Each pixel has a weight (W_{gn}), which is usually the proportion of the pixel that resides in the study area. The most common case is where a pixel is either entirely within the study area, in which case its weight is 1, or entirely outside the study area, in which case its weight is 0. Some pixel weights are likely to be between 0 and 1 when the study area does not conform to a raster grid structure. Furthermore, if the projection of the map is not an equal area projection, then the weight of each pixel should be proportional to its area on the ground. For example, if each pixel is one degree latitude by one degree longitude, then pixels near the equator should have appropriately larger weights than pixels near the poles. Equations 9-12 give the map-level matrix entries that correspond to the pixel-level matrix entries expressed in equations 5-8 respectively. On the left side of equations 9-12, the + sign is the second subscript because the equations sum the pixel-level matrix entries over all pixels.

$$L_{g+ij} = \frac{\sum_{n=1}^{N_g} W_{gn} \times \{\text{MAX}(0, X_{gni} + Y_{gnj} - 1)\}}{\sum_{n=1}^{N_g} W_{gn}} \quad (9)$$

$$R_{g+ij} = \frac{\sum_{n=1}^{N_g} W_{gn} \times \{X_{gni} \times Y_{gnj}\}}{\sum_{n=1}^{N_g} W_{gn}} \quad (10)$$

$$M_{g+ij} = \frac{\sum_{n=1}^{N_g} W_{gn} \times \{\text{MIN}(X_{gni}, Y_{gnj})\}}{\sum_{n=1}^{N_g} W_{gn}} \quad (11)$$

$$C_{g+ij} = \frac{\sum_{n=1}^{N_g} W_{gn} \times \left\{ \left[X_{gni} - \text{MIN}(X_{gni}, Y_{gnj}) \right] \times \left[Y_{gnj} - \text{MIN}(X_{gni}, Y_{gnj}) \right] \right\}}{\sum_{n=1}^{N_g} W_{gn}} \quad (12)$$

We include a subscript g for the resolution, because it is usually useful to consider how the results vary as a function of scale. Our approach is to compute the matrices at the resolution of the raw data, then to aggregate the data to coarser resolutions in order to compute the results at multiple coarser resolutions. The coarsest resolution is the resolution at which the entire study area is in one coarse pixel. We average the fine pixels' weights and their category memberships in order to construct maps of coarser pixels. The categorical membership for each coarse pixel is the average of the categorical memberships of the finer resolution pixels that contribute to the coarse pixel, so the coarse pixels have partial memberships to multiple categories when the fine resolution pixels show multiple categories contained within the borders of the coarse pixel, as the next subsection's example illustrates.

2.3 Example data

Figure 2 presents nine maps that illustrate the main concepts to demonstrate multiple resolution analysis for a non-square study area. The top row of Figure 2 shows the reference maps, which have two categories called black and white. The middle row shows the comparison maps, which have three categories called A, B, and C. The bottom row shows the weight maps, which mask the lower right quadrant by assigning a weight of 0 to the four fine resolution pixels in that quadrant. The diagonal stripes in the reference and comparison maps indicate that the four pixels in that quadrant have no data, so there are 12 fine resolution pixels in the study area. The Fine Resolution column on the left shows three maps of pixels arranged in four rows and four columns. The solid horizontal and vertical lines show the borders of the pixels, where the length of the side of each pixel is 1 meter, i.e. resolution = 1. The numbers within the pixels indicate the categorical memberships. For example, the pixel in the upper left corner of the reference map at the fine resolution has a membership of 1 to the black category and 0 to the white category. The pixel in the upper left corner of the comparison map at the fine resolution has a membership of 1 to category A, 0 to category B, and 0 to category C. The Medium Resolution column in the middle shows the maps where clusters of four contiguous fine resolution pixels are aggregated as indicated by the conversion of the pixel borders from solid lines to dotted lines. The aggregation spreads the category memberships more coarsely and uniformly within the reference map, thus the resolution of the information is 2 meters. The aggregation to the medium resolution does not influence the information in the comparison and weight maps because the fine resolution information is already clustered by quadrant for those two maps. The Coarse Resolution column on the right shows the maps at the coarsest resolution where the categorical memberships are spread uniformly across the entire map, thus the resolution is 4 meters. The weights are averaged at the coarsest resolution so that the single coarse pixel has a weight of $12/16$, which illustrates how the technique addresses non-square study areas.

Figure 3 shows the association between the black category in the reference map and category A in the comparison map. At the fine resolution of 1 meter, the association is $2/12 = 17\%$, because the black category overlays category A for 2 of the 12 pixels in the study area. All the matrices produce identical results when the pixels in the reference map or comparison map are hard classified, as they are for the maps in the Fine Resolution column and for the comparison map in the Medium Resolution column. The matrices produce different results at the coarse resolution because the pixels for both the reference map and the comparison map contain partial membership to multiple categories in the Coarse Resolution column. The association ranges from 0 to $4/12 = 33\%$ at the coarse resolution of 4 meters. The Composite matrix is not

relevant to the comparison in Figure 3 because the categories in the reference map are different than the categories in the comparison map.

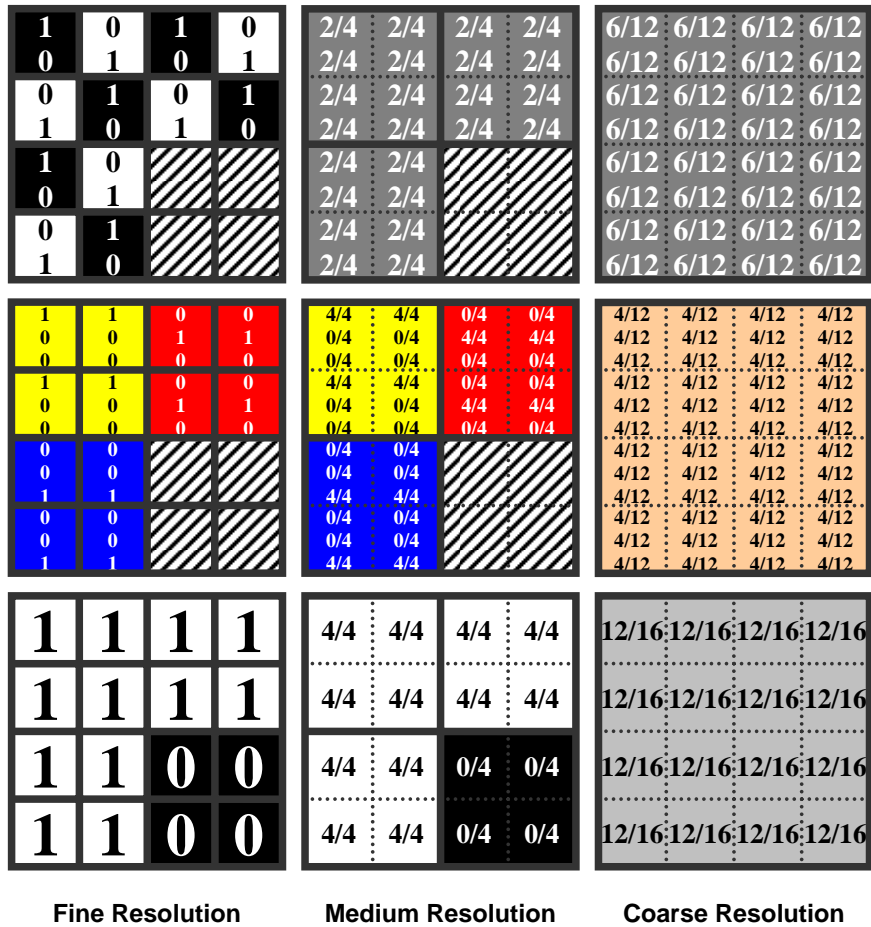


Figure 2. Example data where the fine resolution maps have four rows and four columns of pixels. The top row of maps shows the reference data which have two categories. The middle row of maps shows the comparison data which have three categories. The bottom row of maps shows the pixel weights which mask the lower right quadrant. The three columns show three grain sizes of fine, medium, and coarse.

It can also be useful to compare a map to itself in order to characterize the patterns and uncertainties within a single map. Figure 4 shows the association between the black category in the reference map and the white category in the reference map. At the fine resolution, the association is 0, because the pixels are hard classified so that there is no overlap between black and white. At the medium resolution of 2 meters, the matrices produce different results because the pixels become soft classified. The association ranges from 0 to $6/12 = 50\%$ at the medium and coarse resolutions. As resolution becomes coarser, the range in the association

increases at the resolution that corresponds to the size of the patches in the map. The Composite matrix gives the same results as the Least matrix in Figure 4.

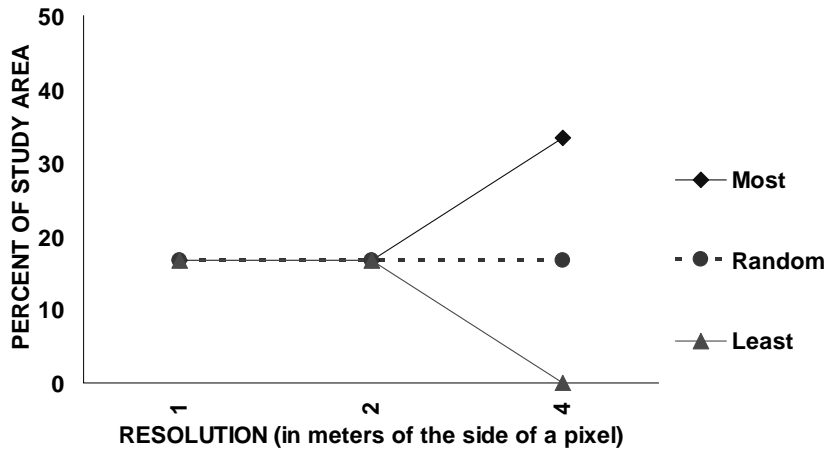


Figure 3. Association of the black category in Reference map to the category A in the comparison map at three resolutions for the example data.

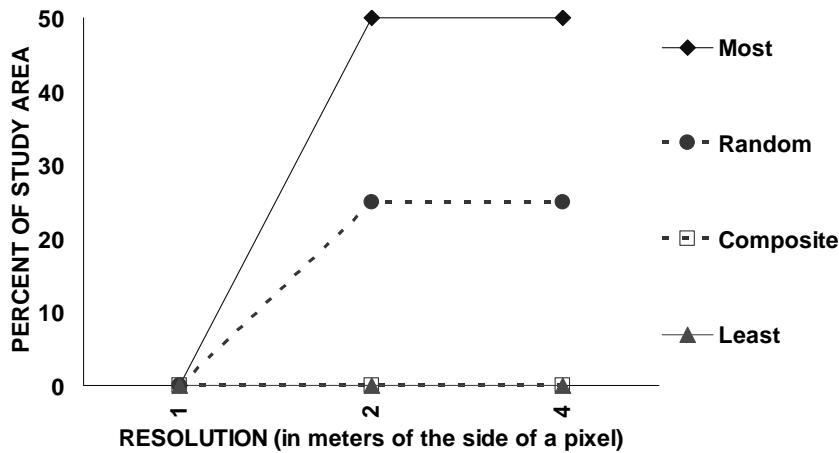


Figure 4. Association of the black category in reference map to the white category in the reference map at three resolutions for the example data.

2.4 Plum Island Ecosystems (PIE) – Long Term Ecological Research data

Figures 5 and 6 show maps of the 26-town region that comprises the Plum Island Ecosystems (PIE) - Long Term Ecological Research site of the United States' National Science Foundation. The State of Massachusetts (MassGIS 2005) supplied the original maps in vector

format, and then Clark University's Human-Environment Regional Observatory (HERO 2006) reformatted the maps to raster format, such that each 30-meter pixel is hard classified as exactly one of three categories: Forest, Built, or Other. The decision to harden the data to 30-meter pixels allows for direct comparison with other case studies where pixels are hardened to a 30-meter resolution because 30 meters is a popular resolution of satellite based data. Figure 5 shows the year 1971 and Figure 6 shows 1999. Therefore, the crosstabulation matrix of these maps indicates the land change from 1971 to 1999.

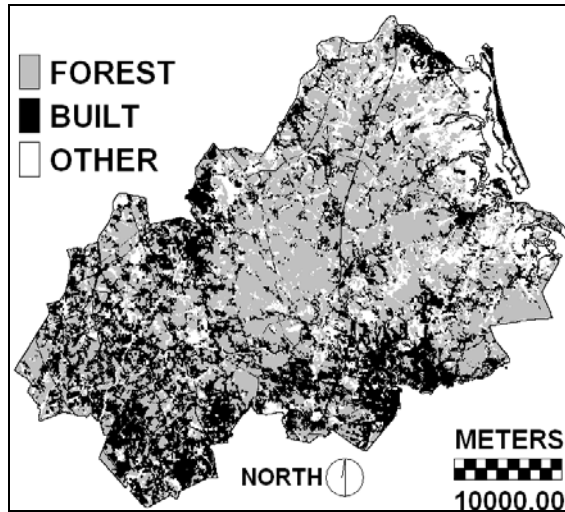


Figure 5. Land cover in 1971 for pixels hardened at a 30-meter resolution for the PIE data.

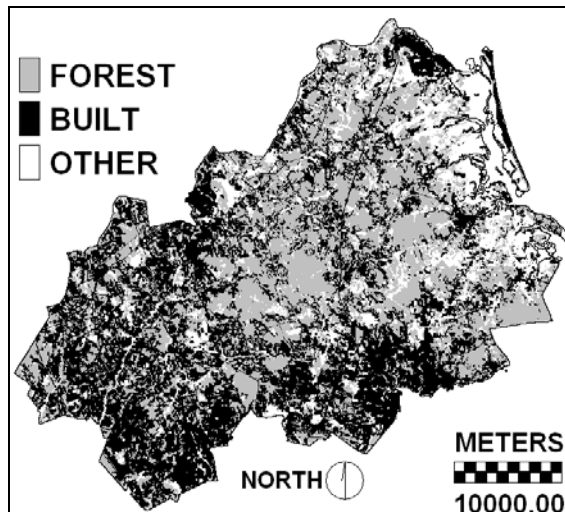


Figure 6. Land cover in 1999 for pixels hardened at a 30-meter resolution for the PIE data.

3 Results

Table 1 shows the crosstabulation matrix that compares the 1971 map to the 1999 map for the PIE study area at the 30-meter resolution. The Least, Random, Most, and Composite matrices produce the same results, because the pixels are hard classified at the 30-meter resolution. The diagonal entries show persistence on the landscape while the off-diagonal entries show differences. Forest is the largest category in 1971, covering 47% of the study area; while Built is the largest category in 1999, covering 43% of the study area. There is a transition from Forest in 1971 to Built in 1999 for 8% of the study area. Table 2 gives the results at the resolution of 960 meters for the Least, Random, Most, and Composite matrices. At each matrix entry position in Table 2, the top number in bold is the Most matrix, the second number in normal font is the Composite matrix, the third number in italics is the Random matrix, and the bottom number in bold italics is the Least matrix. Table 3 gives the range of results at the 960-meter resolution, which is the Most matrix minus the Least matrix. The association of Forest of 1971 to Forest of 1999 has the largest range, i.e. 24 percentage points, while the association from Built of 1971 to Other of 1999 has the smallest range, i.e. 8 percentage points.

Table 1. Crosstabulation matrix in percent of study area for the PIE data at 30-meter resolution.

| | | 1999 | | | Total |
|------|--------|--------|-------|-------|-------|
| | | Forest | Built | Other | |
| 1971 | Forest | 39 | 8 | 0 | 47 |
| | Built | 1 | 32 | 0 | 33 |
| | Other | 0 | 3 | 17 | 20 |
| | Total | 40 | 43 | 18 | 100 |

Table 2. Crosstabulation matrix in percent of study area for the PIE data at 960-meter resolution. The top numbers give the Most matrix, the second numbers give the Composite matrix, the third numbers give the Random matrix, and the bottom numbers give the Least matrix.

| | | 1999 | | | Total |
|------|-----------|-----------|-----------|------------|-----------|
| | | Forest | Built | Other | |
| 1971 | Forest | 39 | 26 | 12 | 77 |
| | | 39 | 8 | 0 | 47 |
| | | <i>24</i> | <i>17</i> | <i>6</i> | <i>47</i> |
| | Built | 15 | 4 | 0 | 19 |
| | | 18 | 32 | 8 | 58 |
| | | 0 | 32 | 0 | 33 |
| | Other | 9 | 20 | 3 | 33 |
| | | 0 | 12 | 0 | 12 |
| | | 13 | 12 | 17 | 42 |
| | Total | 0 | 3 | 17 | 20 |
| | | 6 | 6 | 8 | 20 |
| | | 0 | 0 | 4 | 5 |
| | 70 | 70 | 37 | 177 | |
| | 40 | 43 | 18 | 100 | |
| | <i>40</i> | <i>43</i> | <i>18</i> | <i>100</i> | |
| | 15 | 17 | 4 | 36 | |

Table 3. The Most matrix minus the Least matrix for the PIE data at 960-meter resolution.

| | | 1999 | | | |
|------|--------|--------|-------|-------|-------|
| | | Forest | Built | Other | Total |
| 1971 | Forest | 24 | 22 | 12 | 58 |
| | Built | 18 | 10 | 8 | 46 |
| | Other | 13 | 12 | 13 | 37 |
| | Total | 55 | 53 | 33 | 141 |

Figure 7 examines the association between the Forest category of 1971 and the Built category of 1999. Figure 7 confirms that the transition from Forest to Built occurs in 8% of the PIE study area according to all the matrices at the fine resolution where the data are hard classified. The range in the association grows as resolution grows coarser. The rate at which the range grows is an increasing function of the level of interspersed between the patches of Forest and Built.

Figure 8 gives the association between the Built category of 1999 and the Built category of 1999. The rate at which the range widens as resolution grows coarser is an increasing function of the level of patchiness of the Built category, relative to the size of the pixels. The range is small where pixels are either dominated by Built or dominated by non-Built. The range is large where the amounts of Built and Non-Built in the pixels are close to one half. The Composite matrix gives the same results as the Most matrix in Figure 8.

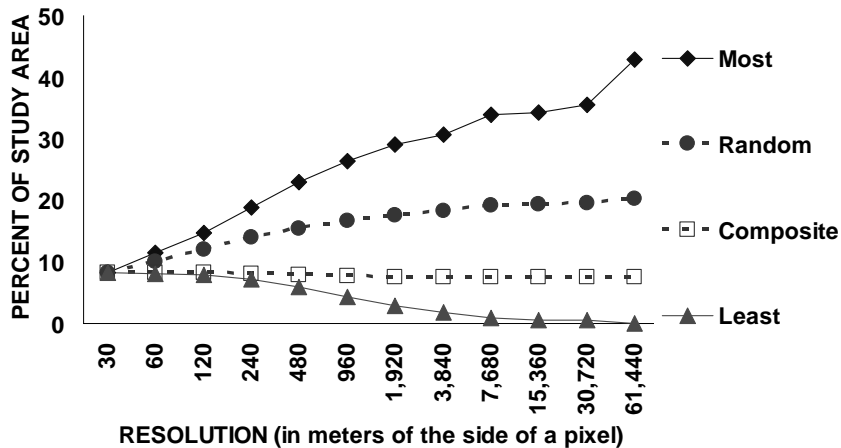


Figure 7. Association of the Forest category in 1971 map to the Built category in the 1999 map at multiple resolutions for the PIE data.

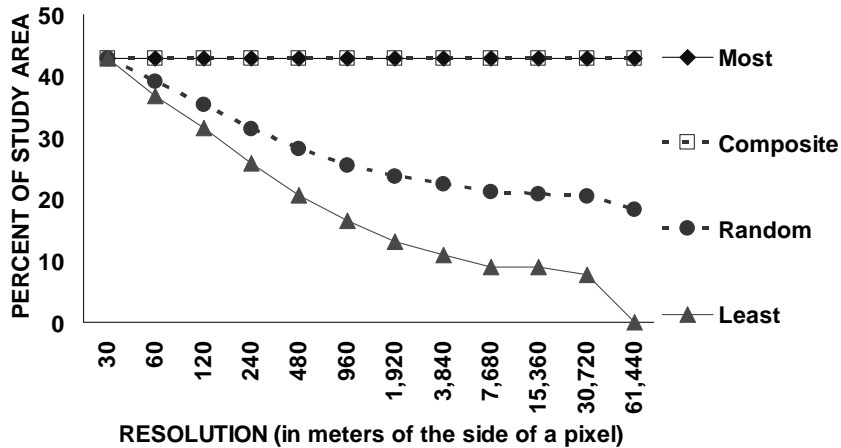


Figure 8. Association of the Built category in 1999 map to the Built category in the 1999 map at multiple resolutions for the PIE data.

4 Discussion

4.1 Interpretation of results

Table 2 shows the four matrices at a resolution of approximately 1 kilometer. The Least matrix gives 0 for five transitions: Forest to Other, Built to Forest, Built to Other, Other to Forest, and Other to Built. This means that the 1-kilometer resolution is probably not sufficiently precise to detect these transitions, because the available fine resolution data indicate that these transitions occur in patches that are substantially smaller than 1 square kilometer on the landscape.

If we were to use 1-kilometer data to examine the transition from Forest to Built, then we would be able to establish some positive estimate of the transition, however that estimate would have substantial uncertainty, since Table 2 shows that the range for that transition spans from a low of 4% in the Least matrix to a high of 26% in the Most matrix. The increase in the range in Figure 7 as resolution grows coarser reflects the interspersion of the patches of Forest in 1971 among patches of Built in 1999. The range is large where Forest of 1971 and Built of 1999 are found in the same coarse pixels.

The variation of the range in Figure 8 reflects the spatial distribution of Built within the 1999 study area. If the Built in 1999 were distributed in a chessboard pattern of 30-meter square patches, then the Least matrix would give an association of 0 at all resolutions coarser than 30 meters and the Most matrix would give a large stable association similar to Figure 8. At the other extreme, if the Built were distributed in one large patch in the northeast quadrant of the map, then all matrices would give an identical large stable association for resolutions from 30 meters to much coarser. Figure 8 shows how the Built-to-Built association is somewhere between these extremes.

4.2 Comparison of four matrices

Table 4 compares characteristics of the four proposed matrices. All four matrices give entries that are expressed as proportions of the study area. The sum of the entries is 1 for the Random and Composite matrices, therefore those two matrices show plausible internally consistent complete descriptions of the associations between the maps. The Random matrix is the only one of the four matrices that gives entries that sum to 1 when comparing maps that show different categories. The Composite matrix requires that the categories in the reference map be identical to the categories in the comparison map. The sum of the entries is not necessarily 1 for the Most and Least matrices, so the interpretation of those two matrices must be somewhat different than for the Random and Composite matrices. The entries for the Most and Least matrices should be interpreted one entry position at a time, because each entry is computed according to a potentially different arrangement of the categories within the pixels, which is why the entries do not sum to 1. It can be useful to subtract the entries of the Least matrix from the corresponding entries in the Most matrix in order to construct the range for each entry in the matrix, as in Table 3. A larger range demonstrates a larger uncertainty concerning the particular association between the two categories.

Table 4. Characteristics of four methods to compute the crosstabulation matrix.

| Characteristic | Least | Random | Most | Composite |
|---|-------|--------|------|-----------|
| Gives entries that are proportions of the study area | Yes | Yes | Yes | Yes |
| Gives entries that sum to 1 | No | Yes | No | Yes |
| Can compare two maps that show different variables | Yes | Yes | Yes | No |
| Can compare two maps that share a single variable | Yes | Yes | Yes | Yes |
| Gives zeroes in off-diagonal entries for identical pixels | No | No | No | Yes |

4.3 Comparison to other approaches

There are many other methods to compare maps that show categorical variables, because there are many ways to compare categorical variables. We have examined some of those alternative methods and have found them to be more challenging conceptually and mathematically compared to this paper's methods. The additional complexity of the alternative methods is no guarantee for improvement in practical usefulness.

For example, information theory forms the conceptual foundation of some alternative approaches. Finn (1993) and Foody (1996) use information theory in a procedure that computes the average mutual information shared between maps. Similarly, Wear and Bolstad (1998) use information theory to compare categorical maps in a hypothesis testing framework. Information theory computes statistics in units of entropy, which is a concept we find challenging to grasp intuitively when relating to a map. The mathematics necessary to compute the entropy in maps is certainly much more complex than this paper's mathematics.

Fuzzy set theory forms the conceptual foundation of another popular set of approaches. Woodcock and Gopal (2000) propose a fuzzy based method to compare maps for which a given pixel can have various intensities of agreement to multiple categories. The total agreement to all categories is not necessarily constrained mathematically, just as the entries in the Most matrix are not constrained to sum to 1 (Binaghi et al. 1999). The reason for this lack of constraint is related to the fact that the source of the uncertainty in fuzzy set theory is ambiguity in the class membership. In fuzzy set theory, the class memberships are not necessarily proportions of the classes in the pixels.

Pattern metrics form a third popular set of approaches to evaluate maps that show categorical variables. There are many pattern metrics that can be computed for raster maps where the pixels are hard classified, and many of these metrics are related to the overall patchiness of the map (Ritters et al. 1999). These metrics are not computable when the pixels are soft classified. This paper's methods produce results that reflect the patchiness in a map when the pixels are soft classified, as Figures 3 and 4 illustrate. When the sizes of the patches are smaller than the grain of the resolution of the maps, then the range between the Most and Least matrices can be large. When the sizes of the patches are larger than the grain of the resolution of the maps, then the range between the Most and Least matrices is likely to be small. Figures 7 and 8 show smooth growth in the range because the PIE data contain patches of several different sizes within the same study area. This type of multiple resolution procedure is important to consider when deciding the scale at which to perform an analysis, because it allows one to quantify the uncertainty of the mapped information as a function of patch size and resolution (Woodcock and Strahler 1987).

5 Conclusion

This paper proposes an approach to compare pixels that contain partial membership to multiple categories, i.e. pixels that are soft classified. It then extends the approach to compare maps at multiple resolutions. We can envision the memberships to various categories within the pixels as proportions of various colors of paints within the pixels. The underlying conceptual approach to compare categories in pixels is analogous to examining the range of possibilities for how much of a particular color of paint in a reference pixel can overlap with another color of paint in the corresponding comparison pixel, depending on the spatial arrangement of the paint within the pixels. Hopefully, scientists will find this approach intuitive and useful, especially because the paint analogy is applicable also to real variables (Pontius et al. in press). Scientists should find this approach mathematically accessible, because division is the most complicated mathematical operation. The methods can be used to examine uncertainty in maps as a function of the patterns in the maps, whether or not the two maps share the same categories. Uncertainty concerning a particular categorical association is larger where the Most and Least matrices show a larger range. Hopefully this technique and its paint analogy will open the doors for additional conceptual development of techniques for categorical map comparison that are intuitive, useful, and rigorous.

Acknowledgements

The United States' National Science Foundation supported this research via three of its programs: Plum Island Ecosystems - Long Term Ecological Research (PIE-LTER), Research Experience for Undergraduates (REU), and Human-Environment Regional Observatory (HERO). Grant numbers are OCE-0423565 with subaward 24123 and SES-0243772 with subaward 2911-CU-NSF-3772. Clark Labs facilitated this work by creating the GIS software Idrisi.

References

- Binaghi, E., Brivio, P., Ghezzi, P. and Rampini, A., 1999, A fuzzy set-based accuracy assessment for soft classification. *Pattern Recognition Letters*, 20, pp. 935-948.
- Congalton, R.G. and Green, K., 1999, *Assessing the Accuracy of Remotely Sensed Data: principles and practices*, New York: Lewis Publishers.

- Foody, G. M., 1996, Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data. *International Journal of Remote Sensing*, 17 (7), pp. 1317-1340.
- Foody, G. M. and Atkinson, P. M. (Eds.), 2002, *Uncertainty in Remote Sensing and GIS*, Hoboken NJ: Wiley & Sons.
- Finn, J. T., 1993, Use of average mutual information index in evaluating classification error and consistency. *International Journal of Geographical Information Science*, 7 (4), pp. 349-366.
- Human-Environment Regional Observatory (HERO), 2006, Digital Datasets. Available online at: www.clarku.edu/departments/hero (last accessed 1 May 2006).
- Kuzera, K. and Pontius Jr., R.G., 2004, Categorical coefficients for assessing soft-classified maps at multiple resolutions. In *Proceedings of the joint meeting of the 15th annual conference of The International Environmetrics Society and the 6th annual symposium on Spatial Accuracy Assessment in the Natural Resources and Environmental Sciences*, 28 June – 1 July 2004, Portland, ME.
- Lewis, H.G. and Brown, M., 2001, A generalized confusion matrix for assessing area estimates from remotely sensed data. *International Journal of Remote Sensing*, 22 (16), pp. 3223-3235.
- Massachusetts Geographic Information System (MassGIS), 2005, Download Free Data. Available online at: www.mass.gov/mgis (last accessed 1 June 2005).
- Pontius Jr., R. G., 2002, Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering & Remote Sensing*, 68 (10), pp. 1041-1049.
- Pontius Jr., R. G. and Cheuk, M. L., 2006, A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *International Journal of Geographical Information Science*, 20 (1), pp. 1-30.
- Pontius Jr., R. G., Huffaker, D. and Denman, K., 2004a, Useful techniques of validation for spatially explicit land-change models. *Ecological Modelling*, 179 (4), pp. 445-461.
- Pontius Jr., R. G., Shusas, E. and McEachern, M., 2004b, Detecting important categorical land changes while accounting for persistence. *Agriculture, Ecosystems & Environment*, 101(2-3), pp. 251-268.
- Pontius Jr., R. G. and Spencer, J., 2005, Uncertainty in extrapolations of predictive land change models. *Environment and Planning B*, 32, pp. 211-230.
- Pontius Jr., R.G. and Suedmeyer, B., 2004, Components of agreement and disagreement in categorical maps at multiple resolutions. In *Remote Sensing and GIS Accuracy Assessment*, R. S. Lunetta and J. G. Lyon (Eds.), pp. 233-251, Boca Raton FL: CRC Press.
- Pontius Jr., R. G., Thontteh, O. and Chen, H., in press, Components of information for multiple resolution comparison between maps that share a real variable. *Environmental and Ecological Statistics*.
- Ritters, 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., 1999, A factor analysis of landscape pattern and structure metrics. *Landscape Ecology*, 10 (1), pp. 23-39.
- Wear, D. and Bolstad, P., 1998, Land-use changes in southern Appalachian landscapes: spatial analysis and forecast evaluation. *Ecosystems*, 1, pp. 575-594.
- Woodcock, C.E. and Strahler, A. H., 1987, The factor of scale in remote sensing. *Remote Sensing of Environment*, 21, pp. 311-332.
- Woodcock, C.E. and Gopal, S., 2000, Fuzzy set theory and thematic maps: accuracy assessment and area estimation. *International Journal of Geographical Information Science*, 14 (2), pp. 153-172.