Accommodating Data Reliability in Determining Map Classes: A Heuristic Approach

Min Sun, David W. Wong and Barry Kronenfeld

ABSTRACT: Geographical data used in spatial analysis and mapping are often regarded as accurate with little error or with error levels that are of no significant concern. Typically, map readers assume that observations assigned to two classes have attribute values that are different, but in fact statistically they may not be because attribute data gathered from surveys can be severely affected by sampling error. Subsequently, spatial patterns inferred from differences in values displayed on the map may not be true. To assist cartographers in making maps that incorporate data quality information, this paper presents a heuristic approach to determine map classes with awareness of class separability, the levels of certainty that values in different classes are statistically different from each other. Unfortunately, when more classes are used, values between classes are less separable. In our heuristic approach, i) cartographers can first determine the number of classes, and our algorithm will seek break values that will give the highest levels of separability; or ii) cartographers can determine the lowest acceptable level of separability between classes, and our algorithm will identify the number of classes and the associated break values. To communicate the levels of separability between classes, we develop a new legend design, indicating the probability that values in a particular pair of class are statistical different from each other.

KEYWORDS: confidence level, class separability, class breaks, heuristic, legend design

Introduction

Both researchers and laymen often regard mapping as simply the display of spatial data. However, the process of creating a map involves analysis, requiring a thorough understanding of the data characteristics and determination of the message to present in a map. Reading a map is also an analysis process, as spatial information is extracted by comparing values shown by the maps in order to formulate spatial patterns. The processes of compiling and reading a map are in no way simple even when there is 100% confidence in the underlying data, but the presence of errors in spatial data complicates these processes even further. The complexity introduced by errors in map compilation and map reading processes are often not realized because errors in spatial data are assumed away. Most of the time, it is assumed that spatial data used and presented in maps are accurate, or at least that errors in the data are not substantial to the extent that they should be considered. These premises have been generally acceptable in many choropleth mapping exercises, using reasonably accurate data, such as the decennial census data, but nonetheless these data have errors. However, as spatial data gathered through surveys are subject to increasing data quality issues, the magnitudes of error in these data can no longer be ignored.

Unfortunately, even when data users are well aware of the presence of errors in data, they have the general tendency to ignore them. Despite the relatively large scale of the American Community Survey (ACS), users of the ACS data should not disregard the data quality information associated with the estimates (Citro and Kalton, 2007). However, many studies mapping ACS data did not take into account the quality of estimates, many of which have significant margins of error (e.g., MacDonald and Peters, 2011). Simply mapping these data and failing to consider data error has significant implications. MacEachren et al. (1998) pointed out that the “power of human vision to synthesize information and recognize pattern … can mislead investigators … if data reliability is not addressed directly…” (p. 1547). Map readers often look for the presence of spatial patterns. Spatial patterns emerge when differences between values exhibit some systematic or organized tendencies. However, if substantial errors are inherent in the data, differences between estimates may not be significant and thus patterns exhibited by the data may not be true. In other words, mapping spatial data without including error information may mislead readers to believe that “something is there”, when in fact no significant pattern exists.

While quality information of spatial data may be available (such as the margin of error associated with each estimate in the ACS data), and cartographers may know the importance of data quality information, how data quality information should be included in maps has not been clear. No standard approaches or practices have been adopted to include data quality information in a map, and no standard tools in GIS are specifically designed to handle data quality information. Dealing with data quality information in mapping generally is messy and haphazard. Therefore, few mapping projects have included data quality information (e.g., Pickle et al., 1996).

While many aspects of a map may incorporate data quality information to some degree, this article describes our effort to expand the current map classification approach by taking into consideration the quality of spatial data in determining class break values. Determination of class breaks is an important issue in choropleth mapping that has received much attention, but few studies have considered the implications of data uncertainty in the classification process. Several tasks are associated with our effort, but the primary one is to develop a map classification method that can determine classes with the highest levels of class separability given other conditioning factors. Associated with the results of the proposed classification method is to introduce a new legend design that explicitly shows the separability between classes. The overall goal is to show the map readers how certain that values in different classes are really different so that they may interpret the spatial patterns exhibited by the data more accurately with a known degree of confidence.

Attribute Error in Mapping

Dealing with error in spatial data has a long history in GIS and cartography (Beard and Buttenfield, 1991). A simplistic approach, but nonetheless better than not providing data quality information at all, is to include some statistical graphical display to show the reliability of the data. This approach was adopted in the *Atlas of United States Mortality* by including a box-plot for each map, depicting the reliability of the health statistics shown on the corresponding maps (Pickle et al., 1996). While the data quality information is provided, matching the data reliability information with the corresponding statistics between two display channels (a map and a graph) reduces the effectiveness and efficiency in assimilating the information. It has been shown that coincident display of data and reliability information is more effective (MacEachren et al., 1997; MacEachren et al., 2005).

As coincident displays are preferable, many studies have investigated various aspects of cartographic symbology in depicting data reliability information effectively (e.g., MacEachren, 1992). Visualization of attribute uncertainty can be achieved by adjusting color saturation (Burt et al., 2011) or lightness (Hengl et al., 2004), or by overlaying semitransparent cues such as hatch symbols (Xiao et al., 2007) onto the enumeration units with the greatest levels of uncertainty. While many studies have focused on the display of attribute accuracy (e.g., Leitner and Buttenfield, 2000), accuracy of spatial data may cover all dimensions of spatial data quality, such as completeness and positional accuracy. These types of metadata for spatial data can also be presented cartographically using various symbols for the map elements (Thomas et al., 2005).

Another approach to handle the attribute accuracy is to focus on map classification. With the objective to determine which map classification method was the most appropriate for epidemiological data, Brewer and Pickle (2002) tested seven map classification methods. Individual maps and a series of maps for comparison were created using these classifications to evaluate how they may affect the accuracy in interpreting spatial data. They concluded that the quantiles and minimum boundary error classification methods (Cromley, 1996) performed the best, followed by natural breaks and a modified version of the equal-interval method. Brewer and Pickle (2002) also offers a relatively concise but thorough review of studies on map classification. Most of the reviewed studies compared the performance of different classification methods (e.g., Chang, 1978). However, almost all reviewed studies treated spatial attribute data as accurate, ignoring the intrinsic errors found in these data. The study by Stegna and Csillag (1987) went as far as determining class intervals by statistically testing the differences between class values, but fell short of explicitly acknowledging that attribute data may have errors, and differences among values may not be statistical significant.

A study by Xiao et al. (2007) was the first to explicitly evaluate the effects of data uncertainty on map classification. They defined a tolerance threshold for the probability that the actual value of a given enumeration unit falls within the range proscribed by the class to which it is assigned. This tolerance can be interpreted as the minimum classification certainty level acceptable to the cartographer. Overall classification robustness was then defined as the percentage of enumeration units meeting this tolerance threshold. Using this robustness measure to evaluate equal interval, quantile and natural breaks classifications of census data, they concluded that robustness is a function of both data uncertainty and number of classes, with a smaller number of classes leading to more robust classification.

Along the approach suggested by Stegna and Csillag (1987) in creating classes with values that are significantly different, and adopting a concept similar to that of Xiao et al. (2007) in taking into consideration errors in estimates, Sun and Wong (2010) suggested a modified natural breaks algorithm in determining classes and break values. Observed values are sorted from the smallest to largest. Assuming standard error of each observed value is given, then a test of significant different can be performed for each sequential pair of observed values. If the values of a sequential pair are significantly different, then a class break can be inserted between the pair of values. The goal is to form classes such that the two values closest to the class breaks are statistically different, and therefore, observations are likely assigned to classes that are really different in values.

This approach lets the data speak for themselves, and the simple algorithm forms classes with observed values that are likely different from each other. However, this method could suffer from at least two drawbacks. The algorithm compares only sequential pairs. Even if values of the pair are significantly different, such result does not guarantee that all other values between the two classes are significantly different. Also, if the two values closest to the class break have relatively small standard errors, these two values can be significantly different, prompting the creation of a class break between them; yet values farther away from the class break could have larger standard errors and not be significantly different from each other. In other words, the modified natural breaks algorithm does not provide a fool-proof method to determine statistically separable classes.

As the determination of class breaks are entirely data driven in this method, the classification results are also completely determined by the data characteristics, and the results could be undesirable in many ways, such as too few or many classes, and too few or many observations within a class. In the New Jersey county data example provided by Sun and Wong (2010, pp. 294-295), four classes were formed from the twenty-one counties. While such number of classes is reasonable, one class has only one observation, one class has two observations, and another class has three observations. The unbalanced distribution of counties across the four classes is obvious, and may not be desirable for many applications.

Class Separability Metric

As our main objective is to determine classes such that values in different classes are statistically different to the highest significance level possible, we first need to determine how differences between values in classes are compared. In Xiao et al. (2007), the robustness measure q_α is defined as the percentage of enumeration units whose maximum classification likelihood (*p*) is greater than α. This robustness measure requires prior specification of the maximum likelihood threshold α. For example, if $\alpha=0.8$ then this indicates that the user will be satisfied with any enumeration unit that can be said with 80% confidence to fall within a particular class. Then, if 3/4 of all enumeration units meet or exceed this threshold, then the robustness level $q_\alpha = 0.75$.

The robustness measure proposed by Xiao et al. (2007) operates under the premise that class break values are provided, and thus their measure is useful to evaluate the overall performance of the classification method in determining the class breaks. However, our objective is to determine specific class break values that can produce classes among which values are as highly and statistically different as possible. Therefore, we need a measure to evaluate the statistical differentiability of values on either side of a class break. We label such a measure a *class separability* measure. Our measure differs from the metrics defined by Xiao et al. (2007) in that it provides a measure of the robustness of a class break, rather than of a single enumeration unit or an entire classification.

A separability measure should indicate to what extent values in two classes are statistically different from each other. While there are many possible ways to formulate such a measure, our formulation starts from the basic notion of determining if the difference between any two enumeration units is statistically significant. We assume that for each value in a survey context, its standard error or margin of error is also provided and the error conforms to a normal distribution. Thus, we consider the value of each enumeration unit to be an estimate based on a sample. Then, for each pair of enumeration units, we ask the question, *could the respective samples have come from the same population*? If so, then the two enumeration units may be considered to belong to a single population and should not be assigned to different classes.

Formally, we define the confidence level $CL_{i,j}$ associated with the assignment of enumeration units *i* and *j* to separate classes as:

$$CL_{i,j} = 2 \times \Phi\left(\frac{\left|\bar{x}_i - \bar{x}_j\right|}{\sqrt{SE_i^2 + SE_j^2}}\right) \qquad \text{(eq. 1)}$$

where Φ is the cumulative normal distribution function, $\left|\bar{x}_i - \bar{x}_j\right|$ is the absolute difference in estimates or (mean) values of the two enumeration units, and SE_i and SE_j are the standard errors of the estimates. Eq. 1 is derived from the standard *z*-test for comparing two means. Specifically, it is equal to one minus the two-tailed probability that the two samples could have been obtained from the same population. We use the *z*-test instead of the standard *t*-test because sample sizes are often not reported with geographic data (U.S. Census Bureau, 2008).

The confidence level will depend on the probability distribution function (PDF) of each value and their overlap (Figure 1). Theoretically, these two curves overlap, regardless of how different the two corresponding values are, and so there is a non-zero probability that the two samples actually come from a single population. Confidence in separability will increase when there is less overlap, which occurs when the standard error of each PDF is smaller and when the centers of the two PDFs are further apart.

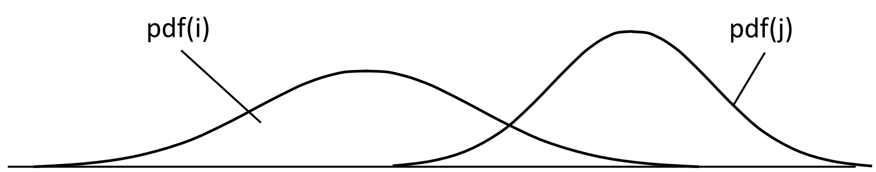

Figure 1: Example of the probability distribution functions of the values associated with two map enumeration units.

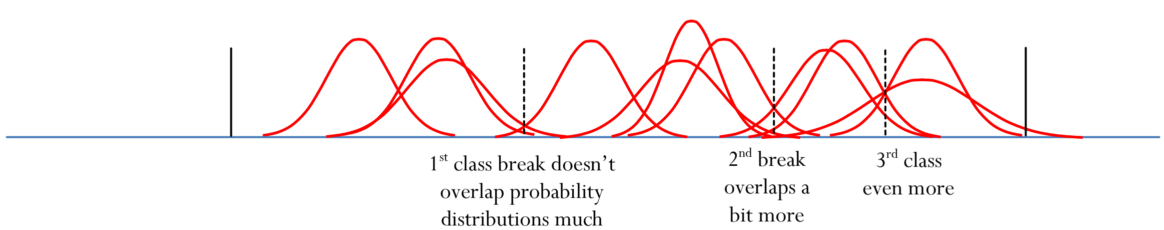

Figure 2: Sequential selection of class breaks in decreasing order of separability. The natural selection process graphically explains the trade-off between number of classes and average separability or overall robustness, since each successive class break does a slightly poorer job at separating enumeration units. The process is simplified by assuming all enumeration units to have identical error distributions (see text for details).

The above illustration considers only two values, but in practice multiple values are involved. Figure 2 shows a possible situation with multiple values with quite different error distributions, assigned to two classes. Ideally, we should be confident that each enumeration unit in class A is statistically distinguishable from each enumeration unit in class B. To determine this, we have to compare all possible pairs of values. For each possible pair, a CL of statistical difference can be derived to indicate the separability between all value pairs. Thus, we define the separability $S_{A,B}$ between two classes A and B as the minimum of the confidence levels of all pairwise combinations of individuals from each class:

$$S_{A,B} = \min_{i \in A, j \in B}(CL_{i,j}), \qquad i \neq j, \qquad \text{(eq. 2)}$$

It is also useful to have a single measure of robustness that is not dependent on an arbitrary threshold. We define robustness as the minimum separability of all pairs of classes in a classification scheme.

Determining Class Break Values and a New Legend Design

Given the above definition of class separability, an interesting problem is how to determine class breaks that maximize separability. When the numbers of classes and enumeration units are both small, all possible classifications can be evaluated and the one that maximizes overall robustness can be chosen. However, this may not be feasible for larger datasets.

We extend the method proposed by Sun and Wong (2010) to determine class breaks that are likely to maximize class separability for any number of classes. Our method takes advantage of the fact that, in most cases, separability confidence levels will be lowest for adjacent enumeration units when values in units are ordered. This suggests that a good candidate for a robustness-maximizing classification with *k* classes can be obtained by sequentially ranking enumeration units from low to high, determining the CL for each pair of adjacent units, and then placing class breaks between the *k-1* pairs of adjacent units with the highest CLs (Figure 2). The above process will often but not always maximize robustness. In particular, when standard errors of enumeration units vary substantially, the lowest confidence levels may not be between pairs of adjacent units if other units with similar values have larger standard errors. The problem of identifying and handling such cases without determining the CL for all possible pairs is left to future research, but in this regard it may be noted that the number of comparison may be reduced substantially by not comparing pairs of estimates with relatively small standard errors but large differences.

The proposed process of selecting class breaks that maximize separability and robustness begins with a sequential process: the first class break is chosen to produce the highest confidence level, the second break produces the second highest confidence level, etc. This process results in a conundrum: as we increase the number of classes, we are forced to insert class breaks between enumeration units that are less and less separable. As Xiao et al. (2007) pointed out, fewer classes provide more robust classification while more classes produces less robust classification. The reason for this is made clear when one realizes that overall robustness is a function of the average separability between all pairs of classes. Although our measure of robustness differs from that proposed by Xiao et al. (2007), the mechanism is the same. Since the class break which results in the greatest separability is likely to be selected first, there is an inherent trade-off relationship between the number of classes and the separability levels between classes.

In other words, the number of classes has to be associated with the lowest CL that one may accept. Cartographers may choose to have fewer but highly separable classes, or more but less separable classes. To facilitate such a decision process, we develop an interactive tool to implement the separability concept (the calculation of CLs) and the determination of class breaks. Figure 3 shows the interface and the graphical display of the tool.

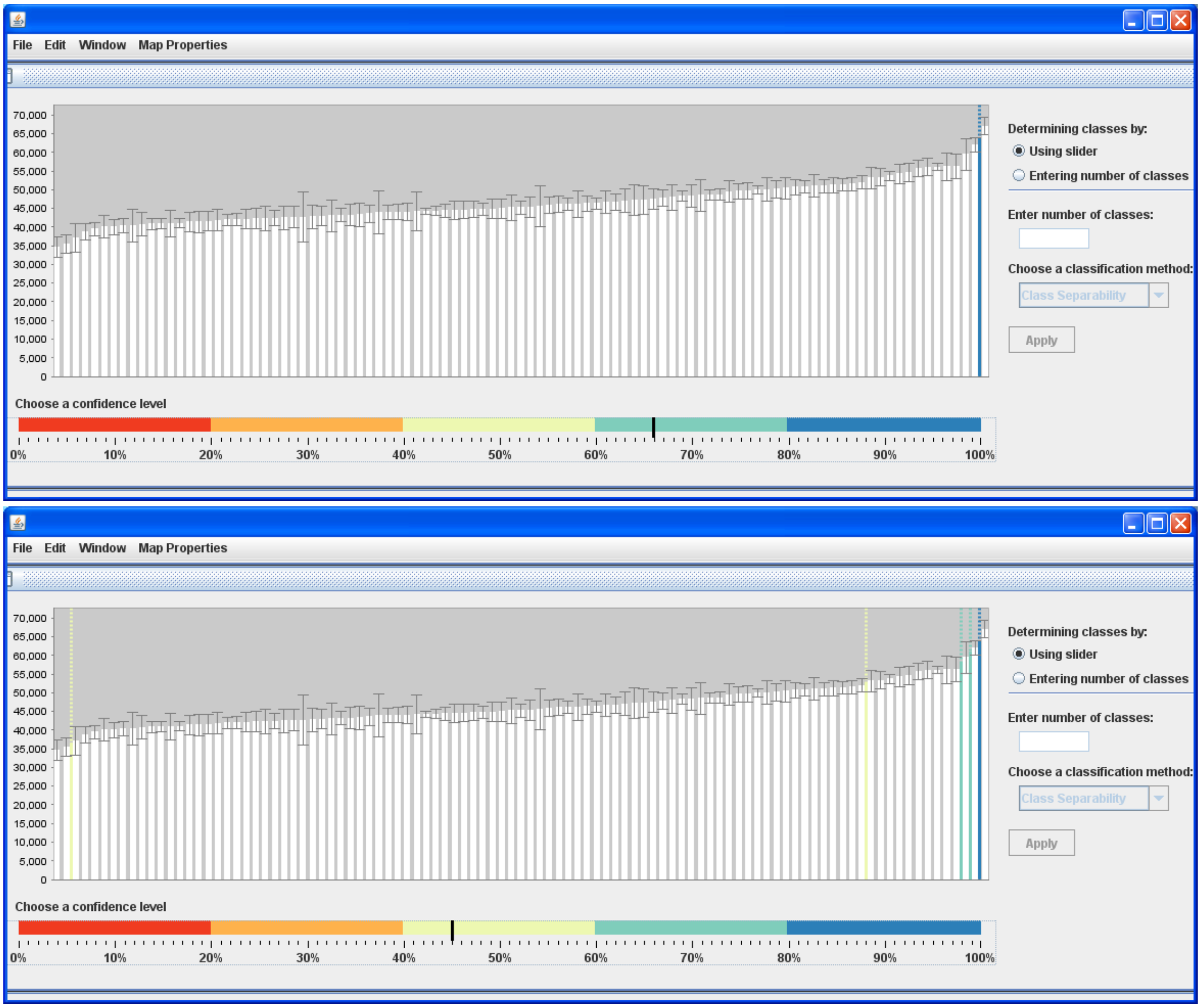

Figure 3: The graphical displays of the classification tool with values in ascending order and error bars representing the margins of error. The color-coded slider bar on the bottom allows users to choose the lowest acceptable level of CL in determining class breaks. Figure 3a (upper): only two classes can be determined with 66% of CL; Figure 3b (lower): six classes can be determined with approximately 45% CL.

On the graphical display, values for all observations are shown in ascending order in a vertical bar plot with an error bar for each value to represent its margin of error. Below the bar plot is a slider bar where users can drag the cursor to a particular CL. This CL is the minimum CL across all class break values that the user is willing to accept in determining classes. If higher CL is chosen, fewer classes can be determined. If lower CL is accepted, then more classes can be used. In this specific example in Figure 3 using the 2006 to 2010 5-year estimates of median family income from the American Community Survey (ACS) for the ninety nine counties in Iowa, even with a CL of 66%, only two classes can be determined (Figure 3a). To come up with six classes, the acceptable CL has to be lowered to 45% (Figure 3b). Note that dotted lines indicating the class break values are color coded, corresponding to the different CLs on the slider bar.

Instead of using the slider bar, users can also enter the desirable number of classes to determine the class breaks with the highest possible CL. This interactive tool also allows users to use several other common classification methods. Another feature of this tool is to take the class break values determined through the heuristic process and to create a choropleth map. Figure 4 shows the map using the Iowa counties data from ACS, corresponding to the six classes determined above. Besides showing the class interval for each class on the right as in the legend of a typical choropleth map, this legend also includes the CLs for all class break values, indicating the confidence levels that values between two consecutive classes are statistically different. These confidence levels are the minima among all pairs of comparisons and therefore, provide conservative results.

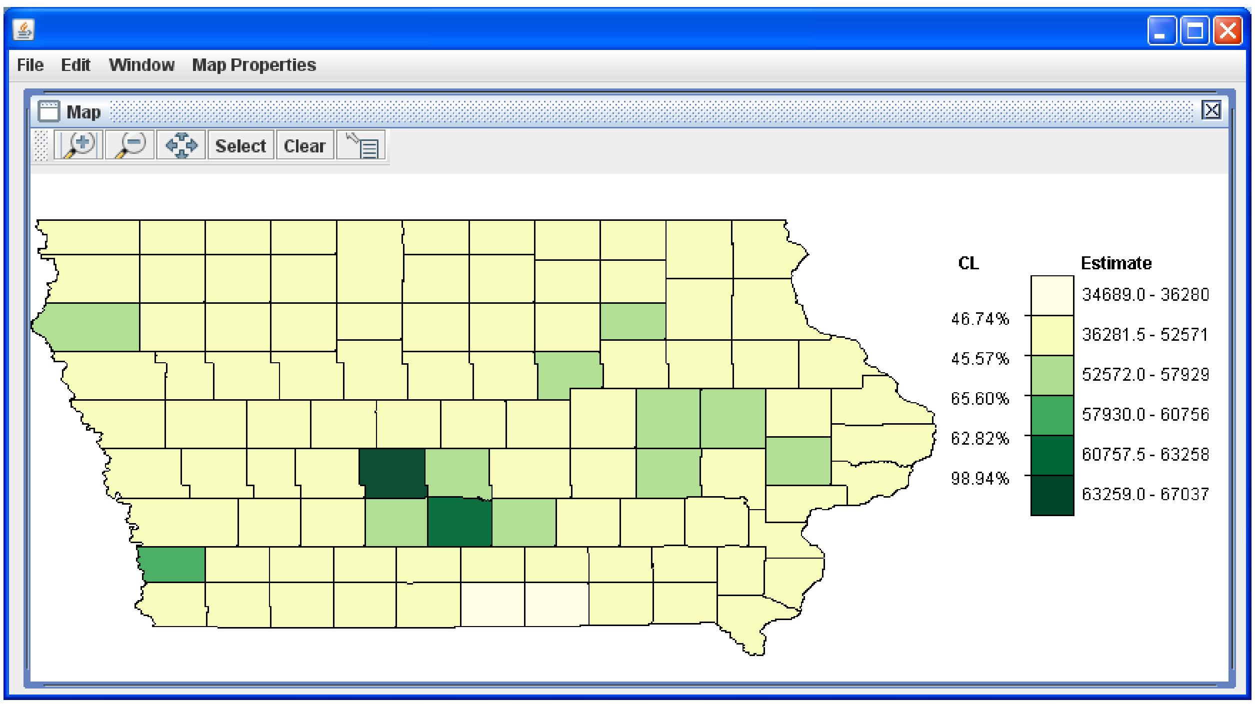

Figure 4: The 5-year estimates of median family income from 2006-2010 ACS for Iowa counties using the class separability classification method with the new legend design. For each class break, the legend includes the confidence level (CL) that values in the two classes are statistically different.

In this specific example, six classes were determined in the process above, assuming that the user was willing to accept approximately 45% CL as the minimum. By interpreting the legend, one may say that values between the two highest classes are 98% statistically different, while values in the two lowest classes are only 46% statistically different. In other words, in this new legend design, we can attach a probability or confidence level to a class break value, reflecting how values between the two classes are separable by taking into account the sampling errors in the values. These CLs are crucial in determining spatial patterns on a map. Because of the relatively high CLs between the two highest classes, we have more confidence to say that the county with the highest value in the south-central part of the state is likely different from its surrounding counties and the rest of the state. For those counties with medium values scattered across the state, their chances that they were different from the surrounding counties with were not too high at all (around 50%).

Evaluating the Class Separability Method

The above discussion proposed a classification method using class separability as the criterion to determine class breaks. Given the number of classes, the objective is to determine break values that can give the highest levels of confidence that values in different classes are statistically different even if those values are estimates with sampling errors. When more classes are needed, some classes are less separable, and vice versa. Ideally, we would like all classes to be highly separable such that spatial patterns revealed by such maps should be relatively reliable, as values in different classes are likely different statistically.

As the concept of class separability is used to show how well values in different classes are statistically different, we may use this concept to evaluate the separability of class break values determined by other map classification methods. The interactive classification tool we developed also supports other popular classification methods. Using the tool, we determined the class breaks for six classes using the natural breaks, equal interval and quantile methods. Separability levels between classes are reflected by the confidence levels that values in two successive classes are different. Results from these three classification methods together with those from the class separability method are summarized in the Table 1 below.

Table 1: Class break values and associated confidence levels for the four classification methods, using the median family income of ninety nine counties in Iowa from the 2006-2010 American Community Survey (ACS) data.

	Approximated Break Values (in US $) by Classification Methods				*Approximated Confidence Levels (in %) by Classification Methods*			
Classes	*Class Separability*	*Natural Breaks*	*Equal Interval*	*Quantile*	*Class Separability*	*Natural Breaks*	*Equal Interval*	*Quantile*
1-2	36,280	38,624	40,080	41,869	47	37	1	9
2-3	52,571	43,286	45,471	43,887	46	9	7	7
3-4	57,929	47,507	50,863	45,594	66	8	11	7
4-5	60,756	51,961	56,254	48,275	63	46	5	1
5-6	63,258	56,379	61,646	51,378	99	66	63	4
				Averaged %	64.2	33.2	17.4	5.6

In Table 1, the class breaks using the four classification methods are summarized on the left. On the right are the confidence levels for all class break values determined by the four classification methods. The averaged confidence level for each method was also reported. As can be seen, the class separability method results in the highest averaged confidence levels that values in different classes are statistically different. According to the separability criterion and using the median family income of Iowa counties from ACS data, the quantile method performs the worst. This result should not be surprising as popular classification methods do not take into consideration the errors in estimates, and therefore statistically differences between values are not of concern. To some degree, these results are consistent with the evaluation performed by Xiao et al. (2007) in which quantile method has the worst performance in their state-level data. While the equal interval method was slightly better than the natural breaks method in their evaluation, we found the opposite, based upon this particular Iowa county dataset.

Using the Iowa county dataset, we can evaluate the performances of different classification methods by measuring their robustness. We built on the robustness concept suggested by Xiao et al. (2007), but removed the requirement to specify a threshold probability. Instead, we define robustness of each areal unit as the probability that it is classified correctly given the error distribution for its estimate. Graphically, it is the area under the error curve bounded by the class interval for the correct class. This probability can be determined for each areal unit. To evaluate the classification performance, we take the average robustness for all areal units. The robustness levels of four classification methods using the Iowa county data are reported in Table 2. For comparison, the average CL levels of all classification methods are also included. According to robustness level, the proposed class separability method performs the best, and quantile the worst. The natural breaks and equal interval methods are quite similar in their performance. Although these results are not completely consistent with the evaluation using CL levels, the class separability method fares the best according to both evaluation criteria.

Table 2: Comparing the performance of the four classification methods using the Iowa county data based upon robustness measure in Xiao et al. (2007) and the proposed confidence level.

Performance Measures	*Class Separability*	*Natural Breaks*	*Equal Interval*	*Quantile*
Average robustness (Xiao et al.)	0.89	0.60	0.64	0.51
Average Confidence Level in %	64.2	33.2	17.4	5.6

Summary and Discussion

In this article, we proposed a class separability metric and a related sequential, heuristic method to determine map classes that seeks to maximize confidence that values of observations in adjacent classes are statistically different. Class break values are chosen between values with highest levels of confidence that they are statistically different, indicating their separability levels. When more classes are needed, class breaks are chosen between values that are less different with lower levels of separability. Therefore, cartographers or users have to determine the trade-off between separability levels and number of classes. We have also developed associated tools to facilitate the determinations of class breaks using the proposed class separability method. The tools use the results from the proposed classification method to create a map with a new legend design in which each class break value is associated with a confidence level, indicating to what extent the two adjacent classes are statistically separable.

To a large extent, we have adopted an optimization approach, but users can control how to optimize the map design. This approach is very much aligned with what Brewer and Pickle (2002) stated: “We expect that the role of classification beyond statistical optimization and classification options in dynamic and interactive mapping of multiple variables will continue to be a focus of choropleth map research (p. 667). Cromley (1996) suggested using multiple criteria in determining classes and Armstrong et al. (2003) provides an application of this. Here, we consider only two criteria: separability and the number of classes. Apparently, additional criteria warrant consideration. For instance, Smith (1986) suggested that homogeneity within classes should be used as a criterion in determining classes. In suggesting compactness, Cromley (1996) has demonstrated the utilities of accounting for the underlying spatial structure in the data, particularly the patterns of spatial autocorrelation, and such approach can be used for spatial data mining (Murray and Shyy, 2000).

We believe that besides separability level and number of classes, other criteria will be needed, but the roles of different criteria may be case-dependent. It is obvious that the example of Iowa counties used in our demonstration needs another criterion: balancing the numbers of observations among classes. Among the six classes in the Iowa example, most classes have only one or a few observations, and the most observations fell into the second lowest class. To achieve a perfect even distribution of observations across all classes, we can adopt the quantile classification method. But, according to results reported in Tables 1 and 2, the quantile method performs especially poorly in respect to separability between classes. Therefore, we need a rather flexible framework, including effective graphics, to depict all possible options available to the map makers when multiple criteria are involved. This is a clear direction of future research.

References

Armstrong, M. P., Xiao, N. and Bennett, D. A. (2003) Using Genetic Algorithms to Create Multicriteria Class Intervals for Choropleth Maps. *Annals of the Association of American Geographers*, 93, pp. 595–623.

Beard, M. K. and Buttenfield, B. P. (1991) *NCGIA Research Initiative 7: Visualization of Spatial Data Quality*. NCGIA, Technical Paper. pp. 91-26.

Brewer, C. A., and Pickle, L. (2002) Evaluation of Methods for Classifying Epidemiological Data on Choropleth Maps in Series. *Annals of the Association of American Geographers*, 92, pp. 662-81.

Burt, J. E., Zhu, A-X and Harrower, M. (2011) Depicting Classification Uncertainty Using Perception-Based Color Models. *Annals of GIS*, 17, 3, pp. 147-153.

Chang, K-T. (1978) Visual Aspects of Class Intervals in Choropleth Mapping. *The Cartographic Journal*, 15, 1, pp. 42-48.

Citro, C.F., and Kalton, G. (eds) (2007) *Using the American Community Survey: Benefits and*

Challenges. Washington, DC: The National Academies Press.

Cromley, R. G. (1996) A Comparison of Optimal Classification Strategies for Choroplethic Displays of Spatially Aggregated Data. *International Journal of Geographic Information Science*, 10, 4, pp. 405 - 24.

Hengl, T., Walvoort, D. J. J., Brown, A. and Rossiter, D. G. (2004) A Double Continuous Approach to Visualization and Analysis of Categorical Maps. *International Journal of Geographical Information Science*, 18, 2, pp. 183-202.

MacDonald, H. and Peters, A. (2011) *Urban Policy and the Census*. ESRI Press.

MacEachren, A.M. (1992) Visualizing Uncertain Information. *Cartographic Perspective*, 13, pp. 10-9.

MacEachren, A.M., Brewer, C. A. and Pickle, L. W. (1998) Visualizing Georeferenced Data: Representing Reliability of Health Statistics, *Environment & Planning A*, 30, pp. 1547-61.

MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., and Hetzler, E. (2005) Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*, 32, 3, pp. 139-60.

Pickle, L.W., Mungiole, M., Jones, G. K. and White, A. A. (1996) *Atlas of United States Mortality*. Hyattsville, MD, USA: National Center for Health Statistics

Smith, R. M. (1986) Comparing Traditional Methods for Selecting Class Intervals on Choropleth Maps. *The Professional Georgrapher*, 38, 1, pp. 62-67.

Stegna L., and Csillag, F. (1987) Statistical Determination of Class Intervals of Maps. *The Cartographic Journal*, 24, 2, pp. 142-46.

Sun, M., and Wong, D. W. S. (2010) Incorporating Data Quality Information in Mapping the American Community Survey Data. *Cartography and Geographic Information Science*, 37, 4, pp. 285-300.

Thomson, J., Hetzler, B., MacEachren, A., Gahegan, M., and Pavel, M. (2005) Typology for Visualizing Uncertainty. In: *Proceedings of the IS&T/SPIE Symposium on Electronic Imaging*, Conference on Visualization and Data Analysis, San Jose.

U.S. Census Bureau. (2008) *A Compass for Understanding and Using American Community Survey Data: When General Data Users Need to Know*. U.S. Government Printing Office, Washington, DC.

Xiao, N., Calder, C. A. and Armstrong, M. P. (2007) Assessing the Effect of Attribute

Uncertainty on the Robustness of Choropleth Map Classification. *International Journal of Geographical Information Science*, 21, 2, pp. 121-44.

Min Sun, Gradaute Student, Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA 22030. Email <msun@gmu.edu>

David W. Wong, Professor, Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA 22030. Email <dwong2@gmu.edu>

Barry Kronenfeld, Assistant Professor, Department of Geology and Geography, Eastern Illinois University, Charleston, IL 61920. Email < bjkronenfeld@eiu.edu>