Published in final edited form as:

Cartogr Geogr Inf Sci. 2017 ; 44(3): 246–258. doi:10.1080/15230406.2016.1145072.

A heuristic multi-criteria classification approach incorporating data quality information for choropleth mapping

Min Sun^a, David Wong^a, and Barry Kronenfeld^b

^aDepartment of Geography and Geoinformation Science, George Mason University, Fairfax, VA, USA

^bDepartment of Geology/Geography, Eastern Illinois University, Charleston, IL, USA

Abstract

Despite conceptual and technology advancements in cartography over the decades, choropleth map design and classification fail to address a fundamental issue: estimates that are statistically indifferent may be assigned to different classes on maps or vice versa. Recently, the class separability concept was introduced as a map classification criterion to evaluate the likelihood that estimates in two classes are statistically different. Unfortunately, choropleth maps created according to the separability criterion usually have highly unbalanced classes. To produce reasonably separable but more balanced classes, we propose a heuristic classification approach to consider not just the class separability criterion but also other classification criteria such as evenness and intra-class variability. A geovisual-analytic package was developed to support the heuristic mapping process to evaluate the trade-off between relevant criteria and to select the most preferable classification. Class break values can be adjusted to improve the performance of a classification.

Keywords

Class separability; multi-criteria classification; data reliability; choropleth maps

1. Introduction

Maps are often designed to reveal spatial patterns. Using classed maps can improve the accuracy of information recovery and reduce the time to process information (Gilmartin and Shelton 1989; Mersey 1990; Armstrong, Xiao, and Bennett 2003). Thus, choropleth maps with classes (versus unclassed maps suggested by Tobler 1973) are useful in exploratory spatial data analysis (e.g., Anselin 1999), and determining classes is critical in revealing the information captured in the data. Many methods for determining classes have been proposed and they serve different objectives. However, a highly important factor in choosing a classification method, regardless of the objective of a map, is that the map should be as truthful as possible, even though a map can never be perfectly accurate (Monmonier 1991).

CONTACT Min Sun msun@gmu.edu Department of Geography and Geoinformation Science, 4400 University Drive, MS 6C3, Fairfax, VA, 22030-4444, USA.

Disclosure statement

No potential conflict of interest was reported by the authors.

Less serious than lying intentionally, ignoring data error nevertheless often misleads map readers into believing something is true, when in fact it is not. Therefore, considering error in data is essential to produce more truthful maps.

Despite advances in cartographic concepts and map-making technologies, how to make more “truthful” maps has not received a great deal of attention. Recently, Sun, Wong, and Kronenfeld (2014) pointed out that most choropleth maps tend to mislead readers because values being mapped are often statistical estimates, and therefore have error (or uncertainty), which is often reflected by the margin of error (MOE) or standard error (SE). When these estimates are assigned to different classes, they may be statistically indifferent from each other (for instance, the confidence level that they are different may be less than 50%). Map readers may find spatial patterns, which are the results of differences between classes appearing in a systematic manner over space, but in fact the patterns may not exist if values in different classes are not really different (MacEachren, Brewer, and Pickle 1998).

Sun, Wong, and Kronenfeld (2014) also proposed the “class separability” concept, which can be used to determine class breaks to maximize the likelihood that estimates in different classes are statistically different. Resultant map patterns are dependent upon the differences in estimates with certain levels of confidence. Unfortunately, this method often produces highly uneven classes and the resultant map may not be very useful. While assigning statistically different values to different classes is important, determining more evenly distributed classes is also desirable (e.g. Brewer 2006; Brewer and Pickle 2002). Therefore, in addition to class separability, other criteria should be considered to create more balanced classes. In other words, multiple criteria, including the class separability criterion, should be used in determining classes. The resultant classes may not have the maximum separability levels, but the compromise classification should produce more balanced maps.

This article describes our effort to develop an approach to determine class breaks by considering multiple criteria, including the class separability concept. The approach evaluates different options and the associated trade-offs in order to determine highly separable but more balanced classes. Our approach assumes that the map data, which can be any variable at an interval or ratio scale, include statistics indicating the reliability of the data. To facilitate the heuristic nature of the proposed approach, a geovisual-analytic environment was developed that enables users to explore and experiment with different classification schemes and evaluate corresponding map realizations with respect to multiple criteria in an interactive manner (e.g., Andrienko et al. 2007). This approach supports evaluation and experimentation of different classification schemes involving multiple criteria, and stands in contrast to the single-criterion approach in which the map classification is determined according to one measure or statistic.

2. How to map data with error?

Errors in spatial data are often ignored in compiling choropleth maps, as data used in maps are assumed to be accurate, or the errors are assumed to be insignificant. Unfortunately, most spatial data are estimates, which are derived from samples and are associated with certain error levels. Magnitudes of error have to be derived from the original sample data.

Therefore, data providers should have access to the data quality information, but such information sometimes is not provided to data users.

Even when information about spatial data quality (such as the MOE in the American Community Survey data) is available, how data quality information should be incorporated into maps has not been clear. Most studies have overlaid certain cartographic symbols, such as graduated symbols or symbols of different shapes that represent the reliability information onto mapped values (e.g., MacEachren 1994; Leitner and Buttenfield 2000; Sun, Kronenfeld, and Wong 2013). Aligning statistical plots with maps by geographical features (the linked-micromaps design) has also been adopted (Carr and Pickle 2010; Pickle et al. 1996). Showing the reliability information together with estimates in a bivariate legend was proven to be effective (MacEachren, Brewer, and Pickle 1998, 2005), and the value-by-alpha map is a specific implementation of this type of cartographic design (Roth, Woodruff, and Johnson 2010). However, these methods just display error information, warning readers about the reliability of the maps, but they cannot offer more reliable maps. In other words, the current practice of compiling choropleth maps ignores the reliability of estimates and thus statistically indifferent estimates may be put into different classes. Therefore, a logical approach to incorporating data reliability information into maps is to report the extent to which estimates between classes are statistically different and determine map classes such that estimates between classes are different to the largest possible extent. This type of information can assist map readers in properly determining the reliability of the spatial patterns, if any, reflected by the data.

Stegen and Csillag (1987) determined breaks by statistically testing the differences between class values, but without considering errors in the values. Xiao, Calder, and Armstrong (2007) evaluated the probability that the estimate of an area is assigned to the correct class given the distribution of the errors of estimates. Their method only evaluates the accuracy of an existing classification, but not how class breaks should be determined. Following the general idea of the natural breaks classification method, Sun and Wong (2010) suggested inserting class breaks between two sequential estimates (in ascending or descending order) that are significantly different at a given confidence level (such as $\alpha = 0.05$). However, a weakness of their algorithm is that it does not compare non-sequential pairs in neighboring classes. In addition, their method may not be able to create any class break if estimates have large errors (unless the α level is raised).

Recently, Sun, Wong, and Kronenfeld (2014) proposed a classification method that inserts class breaks between estimates that are different to the highest statistical levels. For each potential break value, the associated confidence level (CL) reflects the likelihood that the two estimates of areal units i and j ($\bar{x}_i$ and $\bar{x}_j$) on two sides of a potential break are statistically different. Formally,

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

where SE_i is the SE of estimate i and the function Φ returns the probability of a z-score of the estimate difference. The minimum CL of all possible pairs of estimates between two classes A and B is defined as the separability measure associated with the respective class break. The separability between the two classes is

$$S_{A,B} = \min_{i \in A, j \in B} (CL_{i,j}), i \neq j, \quad (2)$$

The separability measure is a rather conservative measure as it uses the CL of the least separable pair of estimates to reflect how well a given class break value can separate estimates above and below the break. Estimates are different only to a certain level (as indicated by CL).

The class separability concept and measure can be used to insert class break values along a set of estimates arranged in either ascending or descending order. Estimates for all possible pairs are compared and their CL s indicating the probabilities that any two estimates are different can be determined (Equation 1). These CL s can be sorted in descending order such that the pair with the largest CL has the most different estimates and therefore a break point should be inserted between these estimates. The pair of estimates corresponding to the second largest CL will be the next pair to have a class break inserted in between. The process continues until the desirable number of classes is created. Thus, successive class breaks inserted during this process have lower CL s and the corresponding pairs of estimates are less differentiable. As more class breaks are inserted, the new classes are less different statistically. Therefore, the trade-off is between having higher separability levels but fewer classes, or lower separability levels but more classes. Class breaks with higher separability levels tend to be at the extremes of the distribution. Thus, resultant classifications are likely unbalanced with one or two classes in the middle that include most estimates, but classes at the extremes include very few estimates. This unbalanced classification result is applicable to both normal and skewed distributions.¹

Since using class separability as the only criterion may produce relatively separable but highly unbalanced classes, additional criteria should be considered to create more balanced classes. Many cartographers support using multiple criteria in determining map classes (e.g., Brewer 2001; Cromley 1996; Declercq 1995; Slocum et al. 2009). However, no consensus has been reached about which and how classification criteria should be considered. Popular map classification criteria have been discussed thoroughly (e.g., Slocum et al. 2009). Additional classification methods have also been suggested (e.g., Scriptor 1970; Smith 1986; Cromley 1996; Murray and Shyy 2000). Unfortunately, some of these criteria conflict with each other, and therefore when using them together, we have to evaluate the trade-offs (Armstrong, Xiao, and Bennett 2003).

In general, combinatorial methods are needed to enumerate all possible classification options when using multiple criteria (Armstrong, Xiao, and Bennett 2003). These criteria may be

¹We generated distributions of different skewness levels to reach such a conclusion. Space does not allow us to report the details of the experiment and results.

combined and formulated as one objective function using a set of additive weights (Cohon 1978). Then the objective function will be optimized to identify the “best” classification. This approach has been implemented in geographic information systems (GIS) to solve multi-criteria problems (e.g., Carver 1991; Jankowski 1995; Malczewski 1999). A drawback of this method is that the set of weights, which is often specified arbitrarily, cannot be easily determined in an objective manner. With the set of weights, searching for the optimal solutions using, for instance, linear programming methods is often the strategy (Cromley 1995, 1996; Cromley and Mrozinski 1999). In contrast, approximation methods can identify a set of feasible solutions and exploratory graphical tools can be used to examine the trade-offs among alternative classification criteria (see, e.g., Armstrong, Xiao and Bennett 2003).

None of the aforementioned studies using combinatorial methods considered error in data. They all assume that mapped values have no error, and therefore if values are different numerically, then they are different. Our proposed approach treats data quality as a required criterion in determining map classes while other criteria are also considered. The goal is to determine classes that are more balanced than those provided by using only the separability criterion, but produce reasonably separable estimates between classes.

3. Determining highly separable classes using multiple criteria

In addition to class separability, criteria such as intra-class variability and evenness of estimates across classes may be considered to create more balanced classes. But creating highly separable classes is still warranted as such maps likely reveal reasonably reliable spatial patterns. Equation 2 ($S_{A,B}$) evaluates the separability between only two classes of a classification. In order to evaluate the separability of the entire classification, we will take the *average separability*, that is

$$S = \frac{\sum S_{i,i+1}}{k} \quad (3)$$

where k is the number of class breaks, and i and $i + 1$ are the two sequential classes. Besides S , other relevant classification criteria have to be determined in order to evaluate all possible combinations of classification options.

3.1 Additional mapping criteria

Besides the minimum class separability, the approach can conceptually accommodate any criterion to evaluate a classification. For demonstration purposes, only several popular criteria are summarized below, but other criteria can be used or added to the procedure.

- (1) *Number of classes*: in general, humans may not be able to differentiate more than nine classes effectively (e.g., Brewer and Suchan 2001; Slocum et al. 2009). On the other hand, too few classes will fail to show the spatial variation of a phenomenon. But when more classes are used with the class separability method, some classes will have low separability levels. This is a major trade-off that the heuristic approach needs to consider.

- (2) *Unevenness of observations across classes*: in general, a classification is undesirable if observations distribute highly unevenly across classes (Slocum et al. 2009).² To evaluate the degree of unevenness across classes, a possible unevenness measure (*UE*) is:

$$\text{Unevenness } (UE) = \sqrt{\frac{\sum_{i=1}^k (n_i - \bar{n})^2}{k}} \quad (4)$$

where n_i is the number of observations in class i , $\bar{n}$ is the average number of observations across all classes, and k is the number of classes. This unevenness measure is actually the standard deviation of the numbers of the observations assigned to different classes. The quantile method essentially minimizes the unevenness of distribution across classes.

- (3) *Within-class variability*: within-class variability is one of the criteria encompassed in the Jenks's natural breaks (JNB) method to group observations with similar attribute values to the same classes (Slocum et al. 2009). Overall within-class variability for all classes may be summarized by:

$$\text{Variability } (V) = \frac{1}{k} \sum_{j=1}^k \sqrt{\frac{\sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{n_j}} \quad (5)$$

where x_{ij} is observation i in class j , $\bar{x}_j$ is the mean of observations in class j , n_j is the number of observations in class j , and k is the number of classes.

3.2 Multi-criteria heuristic classification procedure

Given an initial classification method, such as JNB or quantile, a map can be created with a number of classes (e.g., 4–9 classes). A classification scheme is defined by a class number and associated performance levels of all evaluation criteria. For instance, using the JNB method and selecting 5 classes, this classification scheme is associated with certain levels of separability, intra-class variability, and unevenness. The trade-offs between separability and other criteria are the major issues to consider in order to determine the most desirable classification. Our proposed heuristic procedure will involve human intelligence, facilitated by various graphical aides. Evaluating the trade-offs involves repeating several steps: choosing an initial classification method, examining the performance of the selected classification scheme through statistical graphics, testing the results by creating a map based upon the selected classification scheme, and repeating these steps by choosing another classification method. After experimenting with different classification schemes, the

²On p. 62 in Slocum et al. (2009), the authors pointed out that classes with no observations are wasted. Thus, by extension, creating classes with very few observations is not effective in using of these classes.

scheme/map best suited to the mapping purpose will be chosen (Monmonier 1972; Schultz 1961). Using this general concept, we designed a heuristic mapping procedure (Figure 1).

An initial classification method has to be chosen to start the procedure. Previous experiments showed that the quantile method produced highly inseparable classes, while equal interval and JNB methods performed better in terms of separability (Sun, Wong, and Kronenfeld 2014). Therefore, either the JNB or equal interval classification method may be chosen as the initial classification method to start the procedure (#1 in Figure 1). Using the selected initial method (say JNB), the average separability levels (S) and measures of other selected criteria are computed from two to nine classes, covering the reasonable numbers of classes of a classification (#2 in Figure 1). A classification scheme is defined by a combination of class number and corresponding values for the chosen criteria.

To facilitate the heuristic approach, a geovisual-analytic environment (a standalone application developed using Java and open-source libraries such as GeoTools³) was designed to display measures of selected mapping criteria and to enable interactive selection of elements on the graphical display (#3 in Figure 1). Such a setup allows users to experiment with different classification schemes and helps expose the trade-off relationships among multiple criteria in different classification schemes (e.g., Andrienko et al. 2007; Tukey 1977) (#4 in Figure 1). The visualization tools are also supported by real-time computations responding to the user's interactive operations (Kohonen 2001). Figure 2 shows the three major components of the geovisual-analytic environment: a star plot, a bar plot, and a map display.

A star plot, which is also known as a radar and spider graph, is effective at describing multiple variables (e.g., Chambers, Cleveland, and Tukey 1983). In this research, each classification criterion is treated as a variable in the star plot. Each axis of a star plot represents one selected criterion. Figure 3 shows only two star plot examples with three and four classification criteria, but more criteria can be added. Based on specific applications and needs, the user determines the relevant of criteria to be included in the classification process. One of the axes (Class Number) refers to the number of classes. All reasonable numbers of classes are shown along the axis. Each number of classes is associated with a level of separability represented by the second axis, a level of intra-class variability indicated by the third axis (Figure 3, left plot for the three-criterion situation) and a degree of unevenness shown by the fourth axis (Figure 3, right plot for the four-criterion situation). Given a selection of class number, lines are used to connect the associated levels in respective axes of other criteria, indicating their associations and the performances according to these criteria. Thus, a classification scheme, which is defined by a class number and associated performances according to selected criteria, is represented by a polygon (Figure 3). In general, the more desirable values are at the periphery of the plot. As criteria have trade-off relationships, not all potential options of each criterion are feasible when multiple criteria are considered simultaneously. For instance, it is often impossible to construct a map that has both low unevenness (i.e., relatively even distribution) across classes and relatively high separability level.

³<http://www.geotools.org>.

To experiment with different classification schemes (polygons in the star plot) and the trade-off relationships, lines forming polygons are linked to the map display such that clicking part of a polygon on the star plot will trigger a (re)rendering of the map using the parameters of the chosen scheme (#5b in Figure 1). Note that the choropleth map legend is modified to indicate the separability level of each class break (Equation 2), a feature that is not found in previous map designs and classification methods. The bar plot arranges the estimates in ascending order, with heights representing the estimate values and error bars indicating roughly the statistical differences between estimates. Vertical lines on the bar plot represent the class breaks inserted between estimates. Clicking a line of a polygon on the star plot will also trigger the bar plot to change, matching the classification selected on the star plot to show the trade-off between class numbers and separability levels, as additional classes are inserted among estimates that are less differentiable (with successively lower *CLs* or separability levels). As the map display renders a new map after a new scheme is explored (selected through the star plot), if the new map is better than the previous one, as determined by the user, it can be kept (#6 in Figure 1); otherwise, the previous map can be retained and other schemes can be explored subsequently (#7a in Figure 1). Note that when a classification scheme is chosen, a set of weights has been implicitly applied to different criteria by the user. Thus, the difficulty in determining weights for multi-criteria decision, mentioned in Section 2, is handled by the user's decision after evaluating the trade-offs among multiple criteria.

Sometimes, a classification scheme cannot satisfy the user because just one of the considered criteria performs poorly. For instance, the “example” polygon in the star plot in Figure 4 (upper left) has six classes and acceptable within-in class variability, but the separability level is the lowest among other schemes. The reason is that the break value between the second and third classes (indicated by the second vertical line from the left on the upper bar plot) has a very low class separability level (0.08 or 8%). Instead of exploring another classification scheme, the user may choose to adjust that single break value to improve the separability between classes. The bar plot in the geovisual-analytic environment offers the interactive function that allows the user to select an existing break value and move it to a new location or to insert a new class break in a chosen location. As shown in Figure 4 (upper bar plot), the break value between the second and the third classes was moved to the right (lower bar plot in Figure 4). The vertical lines are color-coded, corresponding to the levels of separability of break values shown underneath the horizontal axis. As shown in the two bar plots in Figure 4, the color for the class break was red before the move (i.e., having a separability level between 0% and 20%, while the actual level was 8%); the break line became yellow after the move (i.e., a separability level between 40% and 60%, as the actual level was 47%). Thus the level of separability of this break point was higher than the previous break point. The star plot (upper right in Figure 4) was also updated synchronously showing the new trade-off relationships and a more desirable classification scheme. Note that the map was also updated simultaneously as soon as the break value was moved.

The two star plots show all classification schemes based on an initial classification method. In this case, the “example” polygon was the chosen scheme. The two bar plots show the distributions of the estimates with break values inserted as vertical lines. These vertical lines

are color-coded, corresponding to the horizontal color bars of confidence levels underneath the y-axes.

When all schemes of a classification (say JNB) are exhausted, another classification method (say equal interval) can be explored (back to #1 from #8 in Figure 1). Thus, this geovisual-analytic interactive environment facilitates the exploration of different mapping schemes and classification methods, and eventually, the most preferable classification method and scheme can be determined to generate the desirable map.

4. Demonstrating the multi-criteria heuristic approach

To demonstrate the proposed approach, we used the mortality rates for Whites for all causes of death for 768 health service areas (HSAs) in the continental U.S. from 1969 to 2011. The data were collected and maintained by the National Center for Health Statistics⁴ and are available for downloading through the Surveillance, Epidemiology, and End Results database.⁵ Each estimate (mortality rate of an HSA) has an associated SE. Because using only the separability criterion will likely create unbalanced classes, the purpose of this demonstration is to create highly separable classes while considering other criteria, including the number of classes, within-class variability, and unevenness, that will improve the balance in the distribution of estimates across classes. While we use mortality rate with SE as the data quality indicator in this example, the method is applicable to many other variables, including public health statistics and large-scale survey data accompanied by data quality information such as the MOE in the case of the American Community Survey.

As the JNB method seems to be generally effective in revealing spatial patterns, it was used here to show how the class separability concept can assist in determining class number. Figure 5 shows the trade-offs between classification schemes with 2–9 classes. While most schemes have moderate to undesirable levels of average separability (S), the 6-class scheme produces the highest average separability of 0.33. Although the 7-, 8-, and 9-class schemes have lower variability and unevenness levels, they have relatively low separability levels, likely creating more unreliable maps. The 6-class scheme is thus likely the most desirable classification.

Although the 6-class scheme provides the best balance between competing evaluation criteria among the classifications produced using popular classification methods, for demonstration purposes, the popular 5-class scheme was selected for further exploration. If the map creator has a preference for a 5-class scheme, but the separability or evenness level of the classification schemes produced by popular single-criterion methods is too low, the two options are to change the number of classes (from 5 to 6 or 7 classes) and re-run the classification process to optimize a single criterion, or to manually adjust one or more class break values to balance the multiple competing criteria. The latter (multi-criteria) option was used to identify an alternative 5-class scheme that would achieve reasonably high levels of both separability and evenness.

⁴<http://www.cdc.gov/nchs/>.

⁵<http://seer.cancer.gov/seerstat/>.

Maps created using the JNB, class separability (CS), and multi-criteria (MC) methods with 5 classes are shown in Figure 6. Results were evaluated according to the separability level (S) for each class break, within-class variability (V) for each class, averages of separability and variability, and overall unevenness (UE). These evaluation statistics are reported in the first three columns of Table 1. Under the “Classes” column, rows with single digits (1, 2, 3, ...8) report the variability levels (V) of the corresponding classes while the rows with the two-digit pairs (e.g., 1–2, 2–3, 3–4, etc.) report the separability levels (S) of the class breaks corresponding to the class pairs. The JNB method produces classes that are acceptable in terms of unevenness and variability, but the separability levels (S) of certain classes are very low (8%). As expected, the separability method (CS) produces classes with the highest separability levels (from 93% to 100% – the 100% is due to rounding). However, as most units were assigned to the highest classes, the classification is highly unbalanced (high unevenness) and has high variations within classes. Such maps are not effective in revealing the spatial distribution of the phenomenon (Figure 6, middle).

In mapping mortality rates, the last class with the highest mortality rates points to the population having the highest risk. When the JNB method with the 5-class scheme was used, the last break value has very low separability level (8%), indicating that those observations (areal units) in the last class, the group with the highest mortality levels, were not very different from the rest. In other words, we cannot be sure that the map using the JNB method identifies those areas with the most vulnerable population groups. Using the multi-criteria approach, the separability level of the last break can be improved by slightly adjusting its value (location). Figure 6 (bottom) shows the 5-class map generated by the JNB initiated multi-criteria approach with the last break point being moved to the right (from 1293 to 1324). The new separability level between the two highest classes is improved from 8% to 64% while the performances on other criteria are not negatively affected.

Comparing the separability levels of the top and bottom maps in Figure 6, the multi-criteria approach (bottom) gives us much higher levels of confidence about the presence of the stretch of high mortality region from Iowa to Louisiana than the one using the JNB (top) method (refer to the separability levels on the map legends). Further, by incorporating unevenness as a criterion, the multi-criteria approach also preserves more information about the spatial pattern than the class separability method (middle). Thus, using the multi-criteria mapping procedure can determine map classes revealing more reliable and informative spatial patterns than those using the class separability and JNB methods alone.

We also compared the classification results from the multi-criteria approach using different initial classification methods. Figure 7 shows the classification schemes (polygons in the start plots) using JNB, equal interval (EI), and quantile (Q) as the initial methods. All these initial classification methods performed poorly with the 5-class scheme. The 8-class scheme initiated by the equal interval method performs better than others on separability and variability, but not on unevenness (Table 1 and Figure 8, top). Schemes initiated by the quantile method score well on unevenness and variability in general, but the 6-class scheme is undesirable for separability (Figure 8 middle and also in Table 1). Note that the map using the quantile method as the initial method (Figure 8 middle) highlights the extensive, and almost spatially continuous high mortality region in the middle of the country, but the

separability level of the last class break is very low (~8%). On the other hand, the multi-criteria approach initiated by the equal interval and JNB methods produce maps with relatively high separability levels for the last class breaks, indicating the high mortality areas are still extensive, but not as spatially continuous as the one using quantile method as the initial method. These classification schemes corresponding to different initial classification methods were selected and discussed here because they performed better than other possible schemes based upon our subjective judgment after evaluating the star plots in Figure 7.

5. Conclusion

The class separability method has been proven to be effective in creating class breaks to maximize the statistical differences between estimates of observations in different classes. However, estimates distributed across classes are highly unbalanced on such maps (Sun, Wong, and Kronenfeld 2014). The proposed multi-criteria approach considers the trade-offs between separability levels and other relevant criteria. The geovisual-analytic environment with dynamically linked plots and maps provides a wealth of information to users and reveals the trade-offs between classification criteria. After a classification method is chosen, the environment shows how different class numbers are associated with the performances in selected classification criteria. The environment also allows users to adjust class break values to improve classification performances such that the resultant maps will have reasonably high separability levels between classes, but also score satisfactorily in other relevant classification criteria such as within-class variation and unevenness across classes. We demonstrated, using the mortality rates of Whites at the HSA level in the US, that this approach can produce classifications with reasonably separable classes, but also perform quite well on other classification criteria. The package encompassing the proposed classification-mapping tools can be downloaded from <http://geospatial.gmu.edu>.

Maps produced using the proposed approach not only reveal spatial patterns more effectively than those using the separability method, but also provide certainty information about the spatial patterns. Such information is crucial in making decisions, particularly in identifying spatial clusters, such as hot-spots of specific health outcomes, crime incidences, or areas that deserve special attention. Maps have been offering decision support in this capacity in the past, but without considering error in data. Thus, the certainty of the identified patterns or clusters was not known. We have been operating under the assumption that what we see on choropleth maps is true! The separability concept tackles the data reliability issue, but usually cannot produce maps showing spatial patterns effectively. The proposed heuristic approach improves upon the separability classification method by compromising on the separability levels but producing maps that may be more informative about the spatial distribution of the phenomenon.

While the conceptual foundation of the multi-criteria approach to improve separability levels is sound, the current implementation through the geovisual-analytic environment still has room for improvement. The star plot design enables the comparison of different classification schemes after an initial classification method is selected and allows users to adjust break values flexibly to improve the classification performance. However, the environment will be more desirable if it can compare different classification schemes,

including all different initial classification methods without selecting them one at a time at the beginning of step #1 in Figure 1. To create such capability, all initial classification methods have to be executed and associated classification schemes have to be enumerated at the beginning. Then highly effective summary statistics and graphics are needed to assist the user to evaluate all these feasible classification schemes.

Acknowledgments

We are very grateful to the detailed and meticulous edits provided by one of the reviewers. The comments provided by all reviewers and the assistance of the editor are gratefully acknowledged.

Funding

This work was partly supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Grant R01HD076020. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Andrienko G, Andrienko N, Jankowski P, Keim D, Kraak M-J, MacEachren A, Wrobel S. Geovisual Analytics for Spatial Decision Support: Setting the Research Agenda. *International Journal of Geographical Information Science*. 2007; 21(8):839–857. doi:10.1080/13658810701349011.
- Anselin, L. Interactive Techniques and Exploratory Spatial Data Analysis. In: Longley, PA, Goodchild, MF, Maquire, DJ., Rhind, DW., editors. *Geographical Information Systems*, vol. 1, Principles and Technical Issues. Wiley; New York: 1999. p. 253–266.
- Armstrong MP, Xiao N, Bennett DA. Using Genetic Algorithms to Create Multicriteria Class Intervals for Choropleth Maps. *Annals of the Association of American Geographers*. 2003; 93:595–623. doi:10.1111/1467-8306.9303005.
- Brewer CA. Reflections on Mapping Census 2000. *Cartography and Geographic Information Science*. 2001; 28(4):213–235. doi:10.1559/152304001782152982.
- Brewer CA. Basic Mapping Principles for Visualizing Cancer Data Using Geographic Information Systems (GIS). *American Journal of Preventive Medicine*. 2006; 30(2):S25–S36. doi:10.1016/j.amepre.2005.09.007. [PubMed: 16458787]
- Brewer CA, Pickle L. Evaluation of Methods for Classifying Epidemiological Data on Choropleth Maps in Series. *Annals of the Association of American Geographers*. 2002; 92:662–681.
- Brewer, CA., Suchan, TA. Mapping Census 2000: The Geography of U.S. Diversity. US Government Printing Office; Washington: 2001. Census Special Report, Series CENSR/01-1
- Carr, DB., Pickle, LW. Visualizing Data Patterns with Micromaps. Chapman and Hall/CRC; Boca Raton, FL: 2010.
- Carver SJ. Integrating Multicriteria Evaluation with Geographical Information Systems. *International Journal of Geographical Information Systems*. 1991; 5(3):321–339. doi:10.1080/02693799108927858.
- Chambers, J., Cleveland, W., Tukey, P. Graphical Methods for Data Analysis. Wadsworth International Group; Belmont, CA: 1983.
- Cohon, JL. Multiobjective Programming and Planning. Academic Press; New York: 1978.
- Cromley RG. Classed versus Unclassed Choropleth Maps: A Question of How Many Classes. *Cartographica*. 1995; 32(4):15–27. doi:10.3138/J610-13NU-5537-0483.
- Cromley RG. A Comparison of Optimal Classification Strategies for Choropleth Displays of Spatially Aggregated Data. *International Journal of Geographic Information Science*. 1996; 10(4):405–424.
- Cromley RG, Mrozinski RD. The Classification of Ordinal Data for Choropleth Mapping. *The Cartographic Journal*. 1999; 36(2):101–109. doi:10.1179/caj.1999.36.2.101.

- Declerq, FAN. Choropleth Map Accuracy and the Number of Class Intervals; Proceedings of the 17th Conference and the 10 General Assembly of the International Cartographic Association; Barcelona: Institut Cartogràfic de Catalunya. 1995; p. 918-922.
- Gilmartin P, Shelton E. Choropleth Maps on High Resolution CRTs: The Effects of Number of Classes and Hue on Communication. *Cartographica*. 1989; 26(2):40–52. doi:10.3138/W836-5K13-1432-4480.
- Jankowski P. Integrating Geographical Information Systems and Multiple Criteria Decision-Making Methods. *International Journal of Geographical Information Systems*. 1995; 9(3):251–273. doi:10.1080/02693799508902036.
- Kohonen, T. Self-Organizing Maps. 3rd ed.. Springer; Berlin: 2001.
- Leitner M, Battenfield BP. Guidelines for the Display of Attribute Certainty. *Cartography and Geographic Information Science*. 2000; 27(1):3–14. doi:10.1559/152304000783548037.
- MacEachren, AM. Some Truth with Maps: A Primer on Symbolization and Design. Association of American Geographers; Washington, DC: 1994.
- MacEachren AM, Brewer CA, Pickle LW. Visualizing Georeferenced Data: Representing Reliability of Health Statistics. *Environment & Planning A*. 1998; 30:1547–1561. doi:10.1068/a301547.
- MacEachren AM, Robinson A, Hopper S, Gardner S, Murray R, Gahegan M, Hetzler E. Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*. 2005; 32(3):139–160. doi:10.1559/1523040054738936.
- Malczewski, J. GIS and Multicriteria Decision Analysis. Wiley; New York: 1999.
- Mersey JE. Colour and Thematic Map Design: The Role of Colour Scheme and Map Complexity in Choropleth Map Communication. *Cartographica*. 1990; 27(3):1–157.
- Monmonier, M. How to Lie with Maps?. University of Chicago Press; Chicago: 1991.
- Monmonier MS. Contiguity-Biased Class-Interval Selection: A Method for Simplifying Patterns on Statistical Maps. *The Geographical Review*. 1972; 62(2):203–228. doi:10.2307/213213.
- Murray AT, Shyy T-K. Integrating Attribute and Space Characteristics in Choropleth Display and Spatial Data Mining. *International Journal of Geographical Information Science*. 2000; 14(7):649–667. doi:10.1080/136588100424954.
- Pickle, LW., Mungiole, M., Jones, GK., White, AA. Atlas of United States Mortality. National Center for Health Statistics; Hyattsville, MD: 1996.
- Roth RE, Woodruff AW, Johnson ZF. Value-By-Alpha Maps: An Alternative Technique to the Cartogram. *The Cartographic Journal*. 2010; 47(2):130–140. doi:10.1179/000870409X12488753453372. [PubMed: 21927062]
- Schultz GM. An Experiment in Selecting Value Scales for Statistical Distribution Maps. *Surveying and Mapping*. 1961; 21:224–230.
- Scripter MW. Nested-Means Map Classes for Statistical Maps. *Annals of the Association of American Geographers*. 1970; 60:385–392. doi:10.1111/j.1467-8306.1970.tb00727.x.
- Slocum, TA., McMaster, RB., Kessler, FC., Howard, HH. Thematic Cartography and Visualization. 3rd ed.. Prentice Hall; Upper Saddle River, NJ: 2009.
- Smith RM. Comparing Traditional Methods for Selecting Class Intervals on Choropleth Maps. *The Professional Geographer*. 1986; 38(1):62–67. doi:10.1111/j.0033-0124.1986.00062.x.
- Stegen L, Csillag F. Statistical Determination of Class Intervals for Maps. *The Cartographic Journal*. 1987; 24(2):142–146. doi:10.1179/caj.1987.24.2.142.
- Sun M, Kronenfeld BJ, Wong DW. Cartographic Techniques for Communicating Class Separability: Alternative Choropleth Maps of Median Household Income, Iowa, USA, 2006-2010. *Journal of Maps*. 2013 doi:10.1080/17445647.2013.768183.
- Sun M, Wong DW, Kronenfeld BJ. A Classification Method for Choropleth Maps Incorporating Data Reliability Information. *Professional Geographer*. 2014 doi:10.1080/00330124.2014.888627.
- Sun M, Wong DWS. Incorporating Data Quality Information in Mapping American Community Survey Data. *Cartography and Geographic Information Science*. 2010; 37(4):285–299. doi:10.1559/152304010793454363.
- Tobler W. Choropleth Maps without Class Intervals? *Geographical Analysis*. 1973; 3:262–265.
- Tukey, J. Exploratory Data Analysis. 1st ed.. Pearson; London: 1977.

Xiao N, Calder CA, Armstrong MP. Assessing the Effect of Attribute Uncertainty on the Robustness of Choropleth Map Classification. *International Journal of Geographical Information Science*. 2007; 21(2):121–144.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Workflow of the heuristic multi-criteria mapping approach. Numbers associated with different tasks represent their general sequence, but some tasks (e.g., 4–7) can be repeated multiple times in order to explore more desirable classifications.

Figure 2.

Interface components of the geovisual-analytic environment. The environment includes three main components: an interactive star plot window showing different numbers of classes (2–9 classes) and corresponding values in other classification criteria (upper left); a bar plot interface which shows the estimates in ascending order and allows users to choose various classification parameters such as the confidence level and number of classes (bottom); and a map display with legend showing the map after a specific classification scheme is selected either using the star-plot or the bar plot (upper right).

Figure 3.

Two star plot examples designed to represent the trade-off relationships between criteria and different feasible classification schemes. The left plot includes three classification criteria (class number, separability, and variability). The right plot includes four (the three used in the left plot plus unevenness).

Figure 4.

Adjusting break values to improve classification performance. Upper left star plot and the upper bar plot show the original situation with one class break value having very low separability level (below 20%, and the vertical line of the class break in red). If the break value was moved to the right slightly (the lower bar plot), the separability level increased to over 20% (indicated by a change in the color of the vertical line to orange). The star plot (upper right) also reflected the improvement.

Figure 5.

The star plot for selected classification criteria using the mortality rate of white for all causes of death. The Jenk's natural breaks classification method was used. The star plot shows the trade-offs in performance according to the four criteria (including the number of classes).

Figure 6.

Maps of mortality rates of white for all causes of death by Health Service Areas, using the Jenks's natural breaks (JNB, top), class separability (middle), and JNB initiated multi-criteria approach (bottom). The separability of the last class break (high mortality rates) was very low using the JNB method (top). After adjusting the last break value, separability improved to over 64% (bottom).

Figure 7.

Start plots for classification schemes generated by the multi-criteria approach with different initial methods (left: Jenks's natural breaks, same as Figure 4; middle: equal interval; right: quantile). The "selected schemes" corresponding to the three initial classification methods were chosen by the authors to be the more desirable ones after evaluating the trade-offs among classification criteria for different schemes.

Figure 8.

Maps of mortality rate for whites using the multi-criteria approach but with different initial classification methods (top: equal interval method with 8-class scheme; middle: quantile method with 6-class scheme; bottom: Jenks's natural breaks method with 6-class scheme).

Table 1

Separability ($S_{i,i+1}$) and within-class variability (V) for each class or class break and unevenness (UE) by classification methods.

Classes	JNB (Figure 6 top)		CS (Figure 6 middle)		MC (starting with JNB, Figure 6 bottom)		MC (starting with JNB, Figure 8 bottom)		MC (starting with EI, Figure 8 top)		MC (starting with Q, Figure 8 middle)	
	$S_{i,i+1}$	V	$S_{i,i+1}$	V	$S_{i,i+1}$	V	$S_{i,i+1}$	V	$S_{i,i+1}$	V	$S_{i,i+1}$	V
1	75	0	0	75	75	69	40	83				
(1-2)	56	100	56	56	44	80	25					
2	50	0	0	50	41	43	27					
(2-3)	8	99	8	18	47	28						
3	46	8	8	46	34	40	21					
(3-4)	18	95	18	17	17.4	5						
4	50	0	0	56	39	41	24					
(4-5)	8	93	64	23	9.5	3						
5	80	194	74	43	40	34						
(5-6)				64	10.7	8						
6				74	64	104						
7					60	35						
8					44							
Mean	22	60	97	101	37	60	33	50	41	40	14	49
UE	77	315	80	58	80	2						

Notes: CS: class separability; JNB: Jenks's natural breaks; MC: multi-criteria; EI: equal interval; Q: quantile.

Under the "Classes" column, the single digits (1, 2, 3, ...8) refer to the classes, and the numbers inside the parentheses (e.g., (1-2), (2-3), etc., refer to the break values between the corresponding classes. Averaged values of all $S_{i,i+1}$ and V values for each column were reported as "mean". The 100% value is due to rounding, and the 0 value in V was because there was only one estimate in that class.