Published in final edited form as:

J Vis Lang Comput. 2018 February ; 44: 89–96. doi:10.1016/j.jvlc.2017.11.007.

Geovisualizing attribute uncertainty of interval and ratio variables: a framework and an implementation for vector data

Hyeongmo Koo,

School of Economic, Political and Policy Sciences, The University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas 75080-3021, USA

Yongwan Chun, and

Associate Professor, School of Economic, Political and Policy Sciences, The University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas 75080-3021, USA

Daniel A. Griffith

Ashbel Smith Professor, School of Economic, Political and Policy Sciences, The University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas 75080-3021, USA

Abstract

Geovisualization of attribute uncertainty helps users to recognize underlying processes of spatial data. However, it still lacks an availability of uncertainty visualization tools in a standard GIS environment. This paper proposes a framework for attribute uncertainty visualization by extending bivariate mapping techniques. Specifically, this framework utilizes two cartographic techniques, choropleth mapping and proportional symbol mapping based on the types of attributes. This framework is implemented as an extension of ArcGIS in which three types of visualization tools are available: overlaid symbols on a choropleth map, coloring properties to a proportional symbol map, and composite symbols.

Keywords

Uncertainty; geovisualization; bivariate mapping; GIS

1. Introduction

Spatial data inevitably contain uncertainty in representations of their locations and attributes. The representation and measuring of the world by a data model should have a discrepancy with the real world because the processes of abstracting and generalizing are inevitable attempts to express the real world in a discrete and finite data model [1]. The U.S. Federal Geographic Data Committee (FGDC) lists the following five fundamental components of data quality by adhering to recommendations of the Spatial Data Transfer Standard (SDTS):

Correspondence to: Yongwan Chun.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

lineage, positional accuracy, attribute accuracy, logical consistency, and completeness [2]. Among these components, location uncertainty, which refers to inaccurate locations of geographic features, often results from global positioning system (GPS) and geocoding errors. Because attributes describing non-locational characteristics of geographic features usually are estimated with samples, attribute uncertainty is caused by sampling errors as well as measurement errors. Furthermore, uncertainty in spatial data can propagate through an aggregation of spatial units, which frequently is necessary due to confidentiality issues, data management issues, computational issues, and representation concerns. Especially, when location and attribute uncertainties simultaneously occur for individual observations, this spatial aggregation process can further increase the level of uncertainty in spatial data [3]. For example, when a sample point is assigned to an incorrect areal unit due to its locational error in an aggregation process [4], the location errors further contribute to the attribute uncertainty for an aggregated spatial unit.

Understanding data uncertainty is important and often necessary in various research and practical activities [5]. Researchers consider uncertainty that effects data or a modeling process, such as differences between measured or predicted values and actual or true values. Decision makers usually contend with uncertainty based on how current conditions or policies affect the future results in their decision processes [6]. Visualization can provide leverage for a data source to meet social and scientific needs [7]. But, a comprehensive visualization of uncertainty is difficult, because it often requires multiple representations of data from different perspectives. Especially, uncertainty visualization commonly requires a simultaneous representation of data in terms of their corresponding attributes on a map. Furthermore, uncertainty visualization can become more complex and more difficult to read and interpret for map readers [8].

Previous research has investigated visualization methods to effectively and efficiently represent spatial data uncertainty (e.g., [9–11]). However, uncertainty visualization is still not popular in a geographic information system (GIS) environment. One reason is the unavailability of uncertainty information for attributes (e.g., decennial census), and, hence, the demand for uncertainty visualization in common GIS software is not high, with several exceptions, such as [12,13]. The purpose of this paper is to propose a framework for uncertainty visualization methods for vector data, and to prototype implementation of these methods in a standard GIS environment. It focuses on attribute uncertainty visualization of interval and ratio data. Specifically, the coloring properties for proportional symbols (CPPS) and composite symbol (CS) strategies can represent polygon, point, and line features, while the overlaid symbols on a choropleth map (OSCM) strategy is restricted to only polygon features. In this study, the proposed methods are illustrated with census block group level variables for the city of Plano, Texas using only polygon features.

2. Related work

Uncertainty visualization can be considered the same as the display of any other information in conventional cartography [14]. However, it often requires further attention because it needs to be used for a supplementary purpose to better understand its corresponding attribute [3]. For example, uncertainty information commonly is required to be presented with

attribute information in uncertainty visualization. Technically, attribute uncertainty is depicted with visual variables in cartographic representations. Hence investigations in the literature address how Bertin's [7] graphic variables can be logically matched to, and practically utilized for, different uncertainty types. For example, empirical studies (e.g., [9,15]) investigate the effectiveness of specific visual variables, such as color value, saturation, and opacity. MacEachren [3] discusses how size, color value, and grain (texture) are useful for depicting uncertainty in numerical information, and how color hue, shape, and orientation can be used for depicting uncertainty in nominal information within the context of Bertin's original graphic variables. Table 1 summarizes previous studies with visual variables investigated for attribute uncertainty visualization.

Color components, including hue, color value (or lightness), and saturation, have been widely used to represent attribute uncertainties among these visual variables. Goodchild et al. [16] and Davis and Keller [18] suggest that hue, color value, and/or texture can be used for depicting uncertainty. Also, MacEachren [3] argues that low uncertainty should be represented by pure hues, while high attribute uncertainty should use a less saturated color (i.e., graying out uncertainty features). Jiang et al. [17] suggest visualization techniques for fuzzy spatial analysis, where lightness and saturation are used to represent uncertainties. Hengl [19] uses a similar method based on the hue, saturation, and intensity (HSV) color model to depict uncertainty. These two methods rely on both saturation and value to increase whiteness based on the degree of uncertainty. Kubí ek and Šašinka [22] conducted an intuitiveness testing for Hengl's method, and their empirical test confirmed high uncertainty should be visualized with higher whiteness. More recently, Deitrick and Wentz [5] merged a visual variable color value with saturation to depict attribute and uncertainty using a bivariate mapping technique.

Transparency and fog also are viable candidates for representing uncertainty information. The metaphor of fog has been proposed by Beard et al. [25], and an initial application of uncertainty visualization using transparency by MacEachren [3] was developed based on the metaphor of fog. In this perspective, foggy elements, which make identifying the level of attribute difficult, indicate high uncertainty [10]. Transparency, which also is called opacity, was proposed for integral symbolization. Drecki [15] asserts that elements with higher transparency (i.e., more opaque objects) are logically considered as more certain things. MacEachren et al. [26] confirm through empirical tests that transparency is one of the feasible visual variables for uncertainty representation.

Texture has been popularly utilized when attribute uncertainties are represented with classifications. Texture is useful for a binary classification, but also can be used for ordered or numerical data with spacing between hatching in texture symbols [3]. Goodchild et al. [16] also discuss that texture is a proper visual variable for uncertainty visualization, and can introduce visual noise as a metaphor for noisy data. MacEachren et al. [9] found that texture overlay symbols effectively represent the reliability of death rates, and Xiao et al. [20] illustrate the effect of attribute uncertainty on map classification using texture. Recently, Sun and Wong [11] and Wong and Sun [21] also utilized texture overlay symbols to visualize the quality of American Community Survey (ACS) data.

Statistical maps for uncertainty visualization are constructed in a form of either adjacent display or bivariate mapping [3]. In adjacent display, which also is called a map pair strategy, attribute uncertainty is represented with an additional map accompanying the corresponding attribute map (e.g., [8,27,28]). The bivariate approach also has been popularly used for uncertainty visualization, because an attribute and its associated uncertainty information may be considered to be a pair of variables. Previous studies conclude that the bivariate approach is more effective than adjacent display in terms of cluster detection [9,22,29]. A major advantage of this approach is that readers do not need to switch their focus back and forth between two maps to build connections between the estimates and their uncertainties [11].

3. A framework for uncertainty visualization based on bivariate mapping

An attribute and its associated uncertainty generally are visualized using either an adjacent display or a bivariate mapping method [3]. Adjacent display refers to uncertainty being visualized separately from its corresponding attribute, whereas bivariate mapping represents both data and uncertainty in a single map. McGranaghan [14] argues that the complexity of symbols in a bivariate mapping is beyond a threshold indicting the amount of information decoded at one time. However, uncertainty visualization without interfering with the decoding ability of users has been investigated (e.g., [32,33]). Furthermore, studies [9,33] show with empirical assessments that bivariate mapping also is more effective than adjacent display. With such support, bivariate mapping for uncertainty visualization primarily has been discussed and implemented as a visual tool in GIS. Implementation of bivariate mapping for uncertainty visualization can be supported by a framework for attribute uncertainty visualization. While previous research, including [25,30,31], attempted to provide a comprehensive framework for uncertainty visualization with considerations of location and attribute uncertainty, research summarized in this paper focuses solely on attribute uncertainty visualization, and suggests a framework for geovisualization of attribute uncertainty based on the types of attribute and visual variables.

When uncertainty visualization based on bivariate mapping is implemented, two issues need to be cautiously decided. First, an appropriate thematic mapping technique needs to be selected to visualize a corresponding attribute considering both the nature of the underlying phenomenon and the purpose of maps [34]. Choropleth and proportional symbol maps commonly are used to display data collected for areal units, such as census tracts and counties. A choropleth map is appropriate to display numerical values that are summarized based on polygon boundaries [35], and to focus on typical values for each individual unit [34]. Importantly, data standardization (e.g., population density) often is necessary in choropleth mapping to adjust for varying sizes of areal units. In contrast to the choropleth map, a proportional symbol map is used to represent raw counts or totals (e.g., number of people). A proportional symbol map also can be complementary to a choropleth map. Hence, when attributes are visualized in a thematic map with uncertainty, different thematic mapping techniques should be used, depending on the types of attributes.

Second, uncertainty visualization with a bivariate mapping approach can be categorized into either a visually integral method or a visually separable method [9]. The visually integral

method modifies one visual variable to represent both attribute and uncertainty. For example, color can be modified with hue and saturation; hue usually represents an attribute, and saturation is suitable to represent uncertainty. The visually separable method adds another visual variable to represent uncertainty. For instance, usually hue and lightness of color are used to represent an attribute, and an additional visual variable, such as spacing between hatching in overlaid texture symbols, is combined to describe uncertainty. Researchers [9,36] comment that the visually separable method is better than the visually integral method in terms of extracting information. Thus, uncertainty should be represented by other visual variables or symbols when bivariate mapping is employed for uncertainty visualization.

Table 2 presents a framework for uncertainty visualization using a bivariate mapping technique that is proposed based upon the two preceding considerations. In this framework, standardized attributes are visualized with choropleth mapping, and their uncertainty is represented with overlaid symbols, such as texture and bars. Because hue and lightness typically are utilized for portraying an attribute in a choropleth map, other variables, including spacing, height, and size, can serve to effectively represent uncertainty. When a proportional symbol map is used for an unstandardized attribute, visually separable variables (e.g., color value, saturation, and size) can be used to display attribute uncertainty.

4. An implementation of uncertainty visualization

Attribute uncertainty visualization based on bivariate mapping is implemented in ArcGIS in accordance with the preceding framework. First, an OSCM method is implemented to visualize attribute uncertainty. A choropleth map is suitable to represent attributes at the ratio scale, and the colors of the symbols (usually hue and lightness) in the choropleth map are used to represent uncertainty values. Additional overlaid symbols to visualize attribute uncertainty include textures (spacing), circles (size), and bars (size). A texture symbol is one of the most widely used visual variables to represent uncertainty in previous results (e.g., [16], [20,21]). Texture symbols are the best in a binary classification [3], but also can be employed for numerical variables using spacing between hatching [11]. Symbol sizes also can be utilized for depicting uncertainty (e.g., [5]), although they can be confusing when appearing with attribute values [34]. Table 3 illustrates uncertainty mapping with these three overlaid symbol maps.

Second, CPPS is implemented to represent uncertainty in a GIS environment. A proportional symbol map is more appropriate to represent raw counts or frequencies [34]. The graphical sizes of symbols can be defined by a mathematical scaling based on the minimum and maximum values of attributes. Attribute uncertainty can be represented with color saturation and color value in the HSV color model of proportional symbols. Table 4 graphically illustrates this approach. The saturation and color value linearly decrease from those of an original color to its minimum value (for which zero or a predefined quantity can be used). Thus, in terms of saturation, an attribute value with a high level of certainty is represented with pure hues, whereas those with low levels of certainty are visualized with less saturated colors (i.e., graying out uncertainty features) [3]. With regard to color value, highly certain attribute values are represented with pure hues, and, in contrast, highly uncertain attribute values are represented with low color values (e.g., black) [17,18].

Standard errors, which measure a certain type of attribute uncertainty, enable visualizing a range of attribute values with a given degree of confidence, based upon a specific probability model. This confidence interval can be depicted with an individual symbol. Table 5 illustrates composite symbols in which both attribute values and such corresponding uncertainties are visualized. Conceptually, these composite symbols are constructed with three different proportional symbols overlaid for each individual location. Two of these three proportional symbols represent uncertainty by visualizing the upper and lower limits of attribute values at a given confidence level. Accordingly, this range, which is represented by the thickness of the circle symbols in Table 5, portrays a possible range of an attribute value. Corresponding attribute values appear in the range constructed by the upper and lower bound lines (e.g., the center line in the circle symbols in Table 5). The sizes of three symbols can be proportionally defined with corresponding values. With bar chart symbols, first and second items on the top of the symbol represent the lower and upper confidence limits of an attribute value. The common boundary of lower and upper confidence areas represents an attribute value. These two approaches using a composite symbol allow users to directly compare uncertainties with corresponding attribute values and their confidence intervals. Moreover, with the circle composite symbols, a thicker bound around the center line indicates a wider confidence bound; in other words, a thick symbol indicates that its standard error or attribute uncertainty is large. For a bar chart composite symbol, the level of uncertainty is represented by the combined heights of the first and second items.

The proposed methods have been developed as an extension in ArcGIS 10.1 using C# on the Microsoft .Net Framework 4.¹ This geovisualization extension of attribute uncertainty provides three visualization functions: OSCM, coloring properties to a proportional symbol map, and composite symbols. Figure 1 portrays the menu interface of the extension. Each function has its own graphic user interface (GUI), and the functions allow users to select fields for an attribute and its uncertainties with proper visual variables. The GUIs are presented in the Appendix.

5. An application

The application presented in this section illustrates the implemented methods in ArcGIS using the 5-year ACS data (2009–2013). For the interval/ratio measurement scale, by definition the standard error is the standard deviation based upon sampling error. This quantifies the attribute uncertainty due to sampling [11]. These data include a margin of error (MOE) that is based upon the standard error of the sampling distribution for an attribute variable, as well as survey estimates. Because MOEs represent the lower and upper bounds for a 90-percent confidence level, the standard error of an estimate can be calculated directly from its MOE: that is, $MOE = 1.645 \times \text{standard error}$ [37]. However, some studies [11,38] discuss that a visual representation of absolute error measures (e.g., the standard error of an estimate) may lead to an inappropriate conclusion and interpretation because large attribute values tend to have large standard errors. Furthermore, Francis et al. [29] discuss the appropriateness of using relative measures of errors, such as the coefficient of variation (CV), for counts (e.g., the total number of households and of occupied houses) and

¹<https://hdl.handle.net/10735.1/5609>

median or mean values (e.g., median age and mean household income). Nevertheless, an absolute measure of error is still appropriate for proportional or ratio values (e.g., percentage of Asians), which are bounded by a specific value range: i.e., between 0 and 1 for proportional values and between 0 and 100 for percentages [29].

Two variables with their corresponding MOEs were obtained from the 5-year ACS data (2009–2013) at the census block group level for the city of Plano, Texas, which is located in the northeastern part of the Dallas-Fort Worth metropolitan area (Figure 2). Median age was chosen to illustrate a choropleth map with overlaid symbols, and ‘the number of housing units’ was chosen to illustrate proportional symbols with coloring properties and composite symbols. A relative measure of errors, the CV, is used to represent attribute uncertainty for both variables.

5.1 Overlaid symbols on a choropleth map (OSCM)

Figure 3 portrays visual representations of 2013 median age with its CV at the census block group level for the city of Plano using the OSCM method. Color symbols (specifically, hue and brightness) of the choropleth map represent the estimated median ages. The estimates are classified with four categories and the natural break algorithm. Attribute uncertainties are visualized with three different types of overlaid symbols; textures, bars, and circles. Spacing between hatching in the texture overlay, heights of the bar symbols, and circle sizes furnish a way to represent the degrees of the CV of median age estimates; thus small spacing in texture, tall bars, and large circles describe a high level of uncertainty in estimated median age values. Because median age estimates and their uncertainties are integrated into one display frame, map readers need to concentrate on only one frame. Thus, users can easily build connections between the estimates and their corresponding CV levels. In the median age estimate maps, census block groups in the southern part of the city have relatively higher median age estimates. However, mapping only the estimates may not provide sufficient information with which map readers can easily recognize the credibility of the differences across census block groups. Census block groups with relatively higher CV levels also are observed in the southern part of the city, which indicates that the census block groups in this region contain a higher level of uncertainty in terms of their CV levels.

However, these maps show some potential disadvantages of this visualization method. First, determining corresponding classes from texture overlay symbols is difficult, especially when small areal units have only a small number of lines because of large line spacing separations. There is no general guide to determine the largest spacing value, but each areal unit should be displayed with at least two lines to utilize spacing for a texture symbol. Hence, the largest spacing value should be judiciously chosen according to sizes of areal units and their corresponding attribute uncertainty levels. Second, the large size of overlaid circles or the tall height of overlaid bars can partially cover color on a choropleth map, and can cause problems in interpreting the distribution of estimates. The sizes of overlaid symbols may need to be restricted so that the underlying geographic distribution of estimates can be effectively displayed. In most cases, attributes are still more important than their accompanying uncertainties.

5.2 Coloring properties on a proportional symbol map (CPPS)

Figure 4 visualizes the estimated numbers of housing units and their uncertainty at the census block group level in the city of Plano. The sizes of the circle symbols represent the numbers of housing unit estimates, and their uncertainties are portrayed with saturation and color value in terms of the HSV color model. The census block groups with low saturation and color values represent high uncertainty for the number of housing unit estimates. With a visually separable bivariate mapping technique adopted for this visualization tool, the additional visual variables (saturation and color value) are combined into proportional symbols. Hence, this map should help map readers to easily build linkages between estimates and their corresponding CVs, as MacEachren et al. [9] claim.

This map also has some disadvantages that stem from common problems of proportional mapping. When the difference between the maximum and minimum values is extremely large, the map pattern can be difficult to read due to massive overlaps of symbols [34]. Although the drawing order of the symbols is controlled with regard to the symbol size range (i.e., the largest symbol to the smallest symbol) in order to prevent small symbols from being concealed by larger symbols, the minimum symbol size needs to be selected carefully in order to achieve communicative and visual effectiveness of a map. In addition, the legend for such maps may become complicated. Such a legend has two legend groups. Each of them shows, respectively, the distribution of estimates and uncertainties (CVs) with their own minimum and maximum values. But, because the CVs show only relative uncertainty for estimates, the legend does not necessarily clarify actual differences between estimates.

5.3 Composite symbols (CS)

Figure 5 portrays composite symbol maps for the number of housing unit estimates with 99 percent confidence intervals based on a normal distribution assumption. In the top map, the center lines in the proportional symbols represent the estimates, and the bounds around the center lines portray corresponding confidence intervals. A thick bound indicates high uncertainty of an attribute value. For the bar chart symbol map, the two areas on the top of stacked bars respectively represent the upper and lower bounds of the attribute values. The common boundary between the first and second items on the top of a symbol in the symbols represents attribute values. That is, the number of housing unit estimates is represented by the combined height of the second and third items. The level of uncertainty is represented by the combined heights of the first and second items. The variable used in the composite symbols method in Figure 6 is identical to the one in the map displayed with CPPS in Figure 5. Although a relative comparison is only possible in the map using CPPS, this composite symbol map allows map readers to compare actual differences of estimates for a given confidence level, based on the sizes of the symbols.

Composite symbol maps are particularly useful for simultaneously identifying uncertainty information in specific locations and general spatial patterns of an attribute. When a spatial pattern of an attribute is emphasized, an increase in symbol sizes can enhance the spatial pattern. When a user focuses on local data exploration, symbols can be adjusted accordingly. The implemented tool furnishes an easy way to change symbols as well as exploration

functions, including zoom in/out. However, a composite symbol map has several potential issues that originate with proportional symbol mapping. These include an excessive number of overlaps and legibility of small size symbols [34]. Distinguishing the border between the estimate and confidence bounds is difficult especially when symbol size is small and/or uncertainty is relatively small for a corresponding estimate. Also the level of legibility can vary depending on a user's familiarity with symbols used.

6. Summary and conclusions

Geovisualization of spatial data can be enhanced by providing a visual representation of their attribute uncertainty, which in turn can assist users to more easily recognize an underlying pattern of data. However, it still lacks such tools in a GIS environment, and a general framework for uncertainty visualization. This paper proposes a framework for attribute uncertainty visualization based on bivariate mapping techniques, and provides such tools implemented in the ArcGIS environment. This framework principally utilizes choropleth maps and proportional symbol maps for thematic mapping of attributes based on the following types of variables: count and ratio estimates. This implementation contains three types of visualization functionalities: overlaid symbols on a choropleth map, coloring properties for a proportional symbol map, and composite symbols. Many of the visual variables that have been investigated in the literature are available to provide flexibility to enhance communication and visualization effectiveness.

In this paper, the visualization tools were examined with a single dataset comprising 84 areal units. In addition, these visualization methods may need to be fine-tuned for specific datasets. For example, when a proportional symbol map is used, a large proportion of symbols can overlap with each other, and, for another example, texture symbols with hatchings in a choropleth map may not be easily recognized for small areal units. In these circumstances, minimum and/or maximum symbol sizes need to be carefully set. Similarly, the maximum spacing between hatch lines also needs to be wisely set. Although a composite symbol map allow map readers to directly compare differences between estimates and their uncertainties, small differences still are difficult to identify. Furthermore, a focused group survey conducted as future research would be useful to evaluate the effectiveness of the implemented tools with various geographical scales and resolutions.

Acknowledgments

This research was supported by the National Institutes of Health grant 1R01HD076020-01A1; any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors, and do not necessarily reflect the views of the National Institutes of Health.

References

1. Chrisman, NR. The error component in spatial data. In: Maguire, DJ, Goodchild, MF., Rhind, DW., editors. *Geographical Information Systems: Principles and Applications*. Vol. 1. Longman Scientific and Technical; Harlow: 1991. p. 165-174.
2. A.N.S. Institute. Logical specifications, ANSI NCITS 320–1998. Washington, D.C: 1998. *Spatial Data Transfer Standard (SDTS) - Part 1*.
3. MacEachren AM. Visualizing uncertain information. *Cartographic Perspective*. 1992; 13:10–19.

4. Hay G, Kypri K, Whigham P, Langley J. Potential biases due to geocoding error in spatial analyses of official data. *Health & Place*. 2009; 15:562–7. [PubMed: 19004662]
5. Deitrick S, Wentz EA. Developing Implicit Uncertainty Visualization Methods Motivated by Theories in Decision Science. *Annals of the Association of American Geographers*. 2015:1–21.
6. Erskine RH, Green TR, Ramirez JA, MacDonald LH. Comparison of grid-based algorithms for computing upslope contributing area. *Water Resources Research*. 2006; 42:1–9.
7. Bertin, J. *Semiology of graphics: diagrams, networks, maps*. University of Wisconsin press; Madison: 1983.
8. Leitner M, Battenfield BP. Guidelines for the Display of Attribute Certainty. *Cartography and Geographic Information Science*. 2000; 27:3–14.
9. MacEachren AM, Brewer CA, Pickle LW. Visualizing georeferenced data: representing reliability of health statistics. *Environment and Planning A*. 1998; 30:1547–1561.
10. 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; 32:139–160.
11. Sun M, Wong DWS. Incorporating Data Quality Information in Mapping American Community Survey Data. *Cartography and Geographic Information Science*. 2010; 37:285–299.
12. Heuvelink GBM, Brown JD, van Loon EE. A probabilistic framework for representing and simulating uncertain environmental variables. *International Journal of Geographical Information Science*. 2007; 21:497–513.
13. Pebesma EJ, de Jong K, Briggs D. Interactive visualization of uncertain spatial and spatio-temporal data under different scenarios: an air quality example. *International Journal of Geographical Information Science*. 2007; 21:515–527.
14. McGranaghan M. A cartographic view of spatial data quality. *Cartographica*. 1993; 30:8–19.
15. Drecki, I. Visualisation of uncertainty in geographical data. In: Shi, W.Fisher, P., Goodchild, M., editors. *Spatial Data Quality*. Taylor & Francis; London: 2002. p. 140–159.
16. Goodchild, MF., Battenfield, BP., Wood, J. Introduction to visualizing data validity. In: Hearnshaw, HM., Unwin, DJ., editors. *Visualization in Geographical Information Systems*. John Wiley & Sons; New York: 1994. p. 141–149.
17. Jiang, B., Ormeling, F., Kainz, W. Auto-Carto 12. Charlotte, North Carolina: 1995. Visualization support for fuzzy spatial analysis; p. 291–300.
18. Davis TJ, Keller CP. Modelling and visualizing multiple spatial uncertainties. *Computers & Geosciences*. 1997; 23:397–408.
19. Hengl, T. Visualisation of uncertainty using the HSI colour model: computations with colours. *Proceedings of the 7th International Conference on GeoComputation*; Southampton, United Kingdom. 2003.
20. 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; 21:121–144.
21. Wong DW, Sun M. Handling Data Quality Information of Survey Data in GIS: A Case of Using the American Community Survey Data. *Spatial Demography*. 2013; 1:3–16.
22. Kubí ek P, Šašinka C. Thematic uncertainty visualization usability - Comparison of basic methods. *Annals of GIS*. 2011; 17:253–263.
23. Slingsby A, Dykes J, Wood J. Exploring geodemographics with interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*. 2011; 17:2545–2554. [PubMed: 22034376]
24. Kaye NR, Hartley A, Hemming D. Mapping the climate: Guidance on appropriate techniques to map climate variables and their uncertainty. *Geoscientific Model Development*. 2012; 5:245–256.
25. Beard MK, Leader I, Battenfield BP, Clapham SB. NCGIA Research Initiative 7 Visualization of Spatial Data Quality. Database. 1991
26. MacEachren AM, Roth RE, O'Brien J, Li B, Swingley D, Gahegan M. Visual semiotics & uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics*. 2012; 18:2496–2505. [PubMed: 26357158]

27. MacEachren AM. Accuracy of Thematic Maps/Implications of Choropleth Symbolization. *Cartographica: The International Journal for Geographic Information and Geovisualization*. 1985; 22:38–58.
28. Aerts JCJH, Clarke KC, Keuper AD. Testing Popular Visualization Techniques for Representing Model Uncertainty. *Cartography and Geographic Information Science*. 2003; 30:249–261.
29. Francis, J., Tontisirin, N., Anantsuksomsri, S., Vink, J., Zhong, V. Alternative Strategies for Mapping ACS Estimates and Error of Estimation. In: Hoque, MN., Potter, LB., editors. *Emerging Techniques in Applied Demography*. Springer; Netherlands: 2015. p. 247-273.
30. Gahegan M, Ehlers M. A framework for the modelling of uncertainty between remote sensing and geographic information systems. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2000; 55:176–188.
31. Thomson, J., Hetzler, B., Maceachren, A., Gahegan, M., Pavel, M. Typology for visualizing uncertainty. *Proceedings of Conference on Visualization and Data Analysis 2005 (part of the IS&T/SPIE Symposium on Electronic Imaging 2005)*; San Jose, California. 2005.
32. MacEachren, AM. *How maps work: representation, visualization, and design*. Guilford Press; New York: 1995.
33. Edwards L, Nelson E. Visualizing data certainty: A case study using graduated circle maps. *Cartographic Perspectives*. 2001:19–36.
34. Slocum, TA., McMaster, RB., Kessler, FC., Howard, HH. *Thematic cartography and geovisualization*. 3. Prentice Hall; Upper Saddle River, New Jersey: 2009.
35. MacEachren AM, DiBiase D. Animated maps of aggregate data: Conceptual and practical problems. *Cartography and Geographic Information Systems*. 1991; 18:221–229.
36. Slocum TA, Cliburn DC, Feddema JJ, Miller JR. Evaluating the Usability of a Tool for Visualizing the Uncertainty of the Future Global Water Balance. *Cartography and Geographic Information Science*. 2003; 30:299–317.
37. U.S. Census Bureau. *A Compass for Understanding and Using American Community Survey Data: What Researchers Need to Know*. U.S. Government Printing Office; Washington, DC: 2009.
38. Li, XR., Zhao, Z. Relative error measures for evaluation of estimation algorithms. 2005 7th International Conference on Information Fusion, FUSION; 2005. p. 211-218.

Appendix

A1. GUI for overlaid symbols on a choropleth map (OSCM)

The graphical user interface (GUI) has three major sections. The first section is to specify a target layer, an attribute field, and a corresponding uncertainty field in the target layer. This part is identical in the GUIs for CPPS and CS. The second section is to configure an output choropleth map: that is, the number of classes and symbols. Note that the color ramp is automatically created with a start and an end color. In the third section, a user can choose a type of overlaid symbol with corresponding tab pages, each of which has its own options for uncertainty symbolization.

A2. GUI for coloring properties to proportional symbols (CPPS)

The first section is the same as the one in the GUI for OSCM. The second section is to specify attributes and their uncertainty visualization. Symbol sizes are determined with mathematical scaling based on a minimum symbol size, which a user should provide. Uncertainty is symbolized with two different visual variables, color saturation and color value in the hue-saturation-value (HSV) color model, which are listed in a combo box. The saturation or color value is linearly calculated from that of a selected color to a chosen minimum level.

A3. GUI for composite symbols (CS)

The first section is the same as the one for the other GUIs. An uncertainty field needs to contain standard errors. With the second section, different composite symbols can be chosen with tab pages; i.e., a circle or a bar. A confidence level should be specified to determine a symbol's size of the confidence interval.

Figure 1.
Menu interface of the attribute uncertainty geovisualization extension

Figure 2.
Study Area

Figure 3.
OSCM for 2013 median age and its coefficient of variation by census block group in Plano, TX

Figure 4. Coloring properties on proportional symbol maps for the 2013 number of housing units and its coefficient of variation by census block group in Plano, TX

Figure 5.
Composite symbol maps for the 2013 number of housing units with 99 percent confidence intervals by census block group in Plano, TX

Table 1

Investigations of visual variables for attribute uncertainty visualization

Studies	Visual variables
MacEachren [3]	Color saturation, crispness, transparency, and resolution
Goodchild et al. [16]	Fog and Texture
Jiang et al. [17]	Lightness, and saturation
Davis and Keller [18]	Hue, color value, and texture
Dreki [15]	Opacity
Hengl [19]	Saturation and Lightness
Xiao et al. [20]	Textures
Sun and Wong [11], Wong and Sun [21]	Textures
Kubí ek and Šašinka [22]	Lightness and saturation
Slingsby et al. [23]	Hue and lightness
Kaye et al. [24]	Saturation
Deitrick and Wentz [5]	Saturation and size

Table 2

A framework for attribute uncertainty visualization based upon bivariate mapping

Type of Attribute	Thematic mapping techniques	Additional visual variable for uncertainty representation
Standardized data	Choropleth mapping	Spacing, sizes, height
Unstandardized data	Proportional symbol maps	Saturation, color value, size

Table 3

Overlaid symbols for uncertainty on a choropleth map

Attribute	Uncertainty (Low ↔ High)		
	Spacing	Size	
Choropleth map

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Coloring properties for uncertainty on a proportional symbol map

Attribute	Uncertainty (Low ↔ High)	
	Saturation	Color value
Size (proportional symbols)

Table 5

Composite symbols for uncertainty visualization

Attribute	Uncertainty (Low ↔ High)	
	Size of circles	Height of charts
Size (proportional symbol or chart)