Annals of GIS

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tagi20

Thematic uncertainty visualization usability - comparison of basic methods
Petr Kubíček & Čeněk Šašinka
Department of Geography, Faculty of Science, Masaryk University, Kotlářská, Brno, Czech Republic
Available online: 23 Nov 2011

To cite this article: Petr Kubíček & Čeněk Šašinka (2011): Thematic uncertainty visualization usability - comparison of basic methods, Annals of GIS, 17:4, 253-263
To link to this article: http://dx.doi.org/10.1080/19475683.2011.625978

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Thematic uncertainty visualization usability – comparison of basic methods

Petr Kubíček* and Čeněk Šašinka

Department of Geography, Faculty of Science, Masaryk University, Kotlářská, Brno, Czech Republic

(Received 7 September 2011; final version received 18 September 2011)

The uncertainty of geographic data and visualization can be viewed from different angles. While the conceptual and theoretical base is quite wide, visualization tools and pilot projects are rare, and there exist only a few studies dealing with the user’s ability to cope with uncertainty visualization efficiency. This article introduces two widely used uncertainty visualization methods – maps compared and maps combined – and focuses on the usability issues of these methods. A combination of the regular kriging interpolation method and interpolation uncertainty values was used for maps compared, while the whitening approach was applied for maps combined. During this study, we electronically interviewed and tested more than 100 participants with backgrounds in geography or computer science. Both the above-mentioned methods were tested on three different levels. On the first level, the intuitiveness of the whitening method was questioned. On the second level, both methods were compared taking into account the simple decoding of predicted values (interpolation results), uncertainty decoding and the decoding of both variables (uncertainty and interpolation results) at the same time. Finally, the ability to compare two or more values was tested for the same combinations. Both correctness and processing time were recorded in order to enable further statistical processing. The overall testing was performed within the web-interactive environment developed specifically for interdisciplinary (cartographic–psychology) purposes.

Keywords: uncertainty visualization; kriging; maps combined and maps compared; evaluation; cognitive processes; parallel and serial processing

Uncertainty visualization – background and objectives

Uncertainty is one of the potentially critical factors of geospatial data visualization. Users have a tendency to accept computer-created maps as more reliable and usually do not take into account the original data quality, generalization level, modelling and other issues of the whole cartographic production cycle. Using geographic information systems for geospatial data collection, analysis and visualization even raises the possibility of data uncertainty or poor quality, when more data sources are combined with insufficient or no data quality metadata.

Efforts to develop visualization methods and tools that can help users to understand and cope with information uncertainty have been under way for almost 20 years but so far without a comprehensive understanding of the parameters that influence successful uncertainty visualization.

MacEachren (1992) presented three possibilities of uncertainty representation (visualization) to be used separately or in combination:

- Map pairs in which a data map is depicted side-by-side with a map of uncertainty about the data – we call this representation maps compared in accordance with Slocum et al. (2005).
- Bivariate maps in which both the data of interest and the uncertainty estimate are incorporated in the same representation – we call these maps combined.
- Sequential representation in which the user might be informed about uncertainty with an initial map, which is followed by a map of data – this type of representation is sometimes called ‘interactive’ (Pang et al. 1997, Slocum et al. 2005) and is out of the framework of this article.

The first and second types of representation are often called ‘static’, thus applicable in digital and analogue (printed) form; the last representation is ‘dynamic’ and can be effectively used only within a digital environment.

Besides the representation issues, a number of methods for visualizing thematic and positional uncertainty have also been proposed. MacEachren (1992) has suggested the use of Bertin’s graphic variables to depict uncertainty and added even specialized variables for depicting uncertainty, including crispness, resolution and transparency. Gershon (1998) grouped these into intrinsic and extrinsic visual

*Corresponding author. Email: kubicek@geogr.muni.cz

ISSN 1947-5683 print/ISSN 1947-5691 online
© 2011 Taylor & Francis
http://dx.doi.org/10.1080/19475683.2011.625978
http://www.tandfonline.com
variables, depending on whether the variable is visually separable from the variable depicting the actual attribute. While extrinsic variables are separable, intrinsic variables are not. Another logical step is to describe how these variables including possible additions or modification might be logically matched with different components of data uncertainty (Buttenfield 2000, MacEachren 1992, Leitner and Buttenfield 2000). MacEachren (1992), for instance, stated that the graphical variables’ size and colour values are most appropriate for depicting uncertainty in numerical information, while colour hue, shape and perhaps orientation can be used for uncertainty in nominal information. He also emphasized colour saturation as important for uncertainty visualization. Saturation can vary from pure hues for very certain information to unsaturated (light grey) hues for uncertain information. The role of graphical variables in mapping ecological uncertainty was further developed and documented by Buttenfield (2000). MacEachren et al. (2005) specified that most research directed towards uncertainty visualization had focused on developing representation methods (Pang et al. 1997, Hengl et al. 2004), software applications for the display of uncertainty (Heuvelink et al. 2006) or uncertainty visualization theory (Pang 2001, Thomson et al. 2005). Much less has been done on the empirical evaluation and testing of use and usability. On the basis of their review and synthesis of literature on uncertainty visualization, MacEachren et al. (2005) identified seven core goals requiring interdisciplinary efforts to be accomplished. One of these goals deals directly with assessing the usability and utility of uncertainty representation and interaction methods and tools. Despite a reasonable amount of work done in the field of uncertainty visualization testing (Evans 1997, Leitner and Buttenfield 2000), there is still a wide gap between uncertainty visualization theory and the widely accepted use of uncertainty representation with known effects on users. Hope and Hunter (2007a, 2007b) tested the effects of positional and attribute uncertainty on spatial decision-making in both static and dynamic ways, concluding that the method used for depicting uncertainty in spatial information can have extremely significant effects on decision-making and that there exist subjective preferences for certain visual representations. They also strongly supported the need for further testing of visualization methods.

The influence of acute stress on cognitive processes and map reading was studied by Stachoň et al. (2010), who concentrated on understanding the stress mechanism and its influence on particular cognitive functions, such as short-term and long-term memory, attention, perception and decision-making. They further explored the possibilities and methods for testing the usability of maps and validated their results by means of statistical analysis tools using a combination of quantitative and qualitative methods.

Test approach
Following our previous work in which we dealt with the perception of different cartographic representations
serial models assume that within the short-term memory search task, each item in memory is compared with the target in sequence, so that a comparison on an item must be completed before the next comparison can begin. Parallel models assume, in contrast, that all items are compared with the target simultaneously, although comparisons on different items can be finished at different times.

Treisman and Gelade (1980) created ‘Feature Integration Theory’ (FIT), which states that people shift attention in visual search serially from one object to the next, looking successively for the target. In experimental situations, participants should search for some target object with specific features (e.g. with a defined colour or size) which is located among other distracters (e.g. a red circle among white, red, blue circles, triangles and squares).

Evidence from experiments shows that search time increases if the conjunction of two features occurs. Wolfe et al. (1989) argue that searches for triple conjunctions (e.g. colour × size × form) are easier than searches for standard conjunctions and assume that ‘parallel processes use information about simple features to guide attention in the search for conjunctions . . . and triple conjunctions are found more efficiently than standard conjunctions because three parallel processes can guide attention more effectively than two’. Herd and O’Reilly (2004), following the ideas of Wolf’s group and their guide search model, present their own computational model of visual search which combines serial and parallel processes. They explain that ‘visual search is parallel and in some circumstances often partly serial too. Parallel process may guide attention fixation, so that easy “pop-out” searches require only one fixation, while very difficult searches may require individual inspection of each item’. Thus, they emphasize the quality of the explored material.

It might be expected that tasks on the maps performed in our study would lead to differences in the way information was processed. Serial processing of information is expected for maps compared, where the map user is forced to decode the predicted value, hold this information within memory and subsequently identify the uncertainty level at the corresponding spatial location on the second map. Both comparison and synthesis processes are being performed. However, the second method (maps combined, whitening) offers both variables in the same place and the user has all the information available at the same moment. This configuration makes possible the parallel processing of information. The question is whether this map – the stimulus material – is easy to explore and may guide ‘pop-up’ attentional fixation or whether the participant has to inspect each item (place on the map) serially. Also, the individual cognitive strategy of the participant can play its own role in the processes of map reading.

Data and visualization methods

We used soil sampling data originally collected and interpreted by Lukas et al. (2009) for visualization purposes. Soil depth samples were collected in irregular grids, in which the points had been selected subjectively.

A combination of the regular kriging interpolation method and interpolation uncertainty values was used for maps compared. Both variables (soil depth and uncertainty) were visualized using the hue saturation graphic variable (see Figure 2) and lighter values were used for higher uncertainty. Kriging is widely recommended for uncertainty visualization because this method not only generates estimated field values (soil depth) at unsampled locations, but also provides information about their error variances (Zhang and Goodchild 2002). These properties were further explored as a source of uncertainty and used for alternative visualization source data.

An alternative approach was applied for maps combined. Whitening is a visualization method based on the hue–saturation–intensity (HSI) colour model and is supposed to be psychologically appealing (Jiang 1996) when hue is used to visualize values of thematic space and whiteness (paleness) is used to visualize uncertainty. A 2D legend was designed to accompany the visualizations. Unlike standard legends for continuous variables, this legend had two axes: (1) the vertical axis (hues) was used to visualize the predicted values and (2) the horizontal axis (whiteness) was used to visualize the prediction error (Hengl 2007).

Cognitive and usability testing environment

The formalization of cartographic knowledge usually involves some kind of empirical research such as interviews or text analysis, which is, in most cases, very tedious and time consuming. Kumar (1999) distinguished between
three main methods of primary usability data collection: observation, interviews and questionnaires. Elzaker (2004) further developed the ‘think aloud method’ and tested this approach for exploratory cartography methods.

Contemporary research on use and user issues with spatial technology employs methods that require direct interaction with one or more users. However, no multidisciplinary environment enabling both cartographic inputs and psychological measurement and testing in a digital environment is currently available. Therefore, an interactive web-based testing tool known as MuTeP (Multivariate Testing Programme) was designed and an early prototype developed. The tool was devised in order to test a wide variety of inputs from isolated cartographic symbols or symbol sets to complex map compositions, both static and dynamic.

The application’s core was developed along the lines of Google Web Toolkit and the cartographic part relies on Open Layers libraries. Its overall architecture consists of three basic modules – client, server and database. The client module communicates with the server site and its GUI works within standard web browsers (Google Chrome, IE, Firefox). The server module processes the client requirements and returns demanded data or information.

Web testing environments bring new possibilities for conducting user research at a distance. Distributed web-based testing is a topic that has received substantial attention in the human–computer interaction community. Multiple studies have sought to evaluate the differences between in-person and distributed asynchronous approaches. Robinson (2011) describes several case studies in which distributed means for eliciting user feedback and evaluating cartographic outputs were employed.

Basic test functionality includes test person identification and the pre-test calibration of individual computer and cartographic abilities. Within the test environment, there are three basic types of tasks – forms with pull-down menus with predefined answers; visual choice scenes, in which the test person is forced to choose one or more possibilities of visual variables; and localization tasks, in which the test person must place the symbol on the right position or draw the line or polygon of a certain task. After choosing an appropriate test, the overall content is downloaded from the server to the client site and initialized. The test person attempts the tasks and each interaction (mouse click) is recorded on the event log. Both reaction times and positional accuracy (mouse position can be further compared with the expected answer) are stored and further processed. The test results are stored within the database and exported via the exporting module as *.csv files for further statistical processing. The technical background of the system is further described in Kubíček and Kozel (2010).
Construction of the battery of tests and the experimental environment

The cartographic test which was focused on the exploration of uncertainty visualizations was incorporated into the more complex battery of tasks. The whole experimental battery consisted of the introduction slide, a personal data questionnaire, training tasks, general instructions, a psychological test and two cartographic tests. A standardized experimental environment was prepared for all participants. The laboratory experiment was conducted in an air-conditioned classroom. A uniform type of PC mouse was used and the resolution of LCD displays was set at 1280 × 1024 pixels. The size of the presented pictures was 900 × 675 pixels.

Testing methods

The aforementioned uncertainty visualization methods were tested on three different levels. On the first level, the intuitiveness of the whitening method was questioned, visualized as map combined (Figure 2). Users were asked to mark the area with the highest uncertainty level. Our intention was to test whether the association between uncertainty and paleness (see Jiang 1996, Hengl et al. 2004 for comments) is really intuitive or whether it is based on prior knowledge of uncertainty visualization methods. The map legend was simplified in order to inform about visualized variables (soil depth, uncertainty) but not about the scale and order.

In the Section “Test participants”, users were informed about the basic principles of uncertainty visualization and both methods (maps combined and maps compared) were presented together with examples and legends.

On the second level, both visualization methods were compared starting with maps combined followed by maps combined. Testers were asked to mark the polygon number exceeding a certain level of soil depth (e.g. 50 cm) and the uncertainty level (e.g. 70%) and find a combination of certain soil depth and uncertainty at the same time (e.g. more than 70 cm and 40%). These cases were tested in six successive tasks – three for each type of visualization.

Thus the test was focused on the following user abilities:

- Uncertainty level decoding (soil depth uncertainty)
- Simple decoding of predicted value (soil depth interpolation results)
- Comparison of combined values (both soil depth and uncertainty, Figure 3)

While the tested level of soil depth and uncertainty was identical for both visualization types, the placement of testing polygons was different for both tasks and visualization methods in order to avoid the carry-over effect (see Hope and Hunter 2007 for comments). Correctness and processing time were recorded in order to enable further statistical processing.

Test participants

There were three different groups of test participants. The first group comprised 15 students of the University of Defence (aged 19–23) with just a basic background in the field of spatial information. The second group was created by 39 geography and geoinformatics students (aged 19–23) with intermediate skills in the field of spatial information. Both student groups were relatively homogeneous as for age and skills. The third group of 50 participants was tested during the ‘open door day’ and was heterogeneous with respect to both age (15–61) and an educational (spatial skills) point of view. The gender of participants was also recorded; the group was relatively well balanced in this respect – 63 males and 41 females in total.

Results

The results gathered during the tests were processed and statistically tested. The t-test for independent samples was used for the first level, while the paired Student’s t-test or Wilcoxon signed-rank test was used for the second level of samples. The Wilcoxon signed-rank test was used as an alternative to the paired Student’s t-test when the samples were not normally distributed.

The results from the first level of testing (intuitiveness of uncertainty visualization with the HSI method) confirmed that more participants (63%) acknowledged the lighter value to be more uncertain. Those with this preference were also quicker and were able to decide within a shorter time interval (Figure 4). However, the results were not confirmed as statistically significant, either for the homogeneous or for the heterogeneous groups.

The results from the second level of testing can be divided according to the tested variables. For the purposes of simplification, only correct answers were taken into account; thus, a pair comparison of time needed for correct decision-making for both methods (maps combined and maps compared) was always available. As for uncertainty level decoding, there were significantly better results for maps combined (whitening) than for maps compared. This result was valid both for homogeneous groups (groups 1 and 2) and for heterogeneous group (1–3 together, Figure 5).

The decoding of predicted values testing (soil depth) brought slightly better results for maps combined (whitening) but without statistical significance. Again, this was valid for all three tested groups when considering only correct answers (Figure 6). It is arguable whether the slightly
better results for the whitening method are conditioned by the fact that test participants had faced the whitening method in the first level of testing (intuitiveness).

The last part of the testing dealt with the comparison of combined values (both soil depth and uncertainty at the same time) and was the most controversial as far as correct
answers are concerned. Only 43% of answers were correct for both methods at the same time; 64% of correct answers were valid for maps compared and 56% for maps combined (see Section “Discussion” for further details). Significantly better results were achieved by whitening methods: participants were not only quicker, but also more confident about the result – the standard deviation was only 0.5 in comparison with maps compared (Figure 7). The significant results were valid even if we took into account only one correct answer for each pair.

Gender differences in map reading and map use are often quoted; therefore, we compared both the time to resolve the task and the correctness of answers for group 2 (geography students). This was the only group with a sufficient gender balance since both of the other groups were either too homogeneous (group 1 – only males) or too heterogeneous (group 3 – mixed age groups). Only group 2 was thus appropriate for gender comparison, and differences in task accomplishment were primarily caused by different cognitive abilities between the genders.
Nevertheless, there were no significant differences in both measured values according to gender. Males were in general more correct (they achieved a higher percentage of correct answers) but needed more time to decide than females. This was true for both visualization methods.

We did not try to compare results from specialists (groups 1 and 2) with those from non-specialists (group 3) because of the heterogeneity of group 3 and different age categories. Both aspects could significantly influence the results. Nevertheless, we accomplished the statistical comparison of groups 1 and 2. Both groups were comparable with respect to the age of participants, their education and also their computer skills. However, we expected that group 2 would be more experienced with geoinformation and cartography. We selected only male participants from group 2 because no women were tested in group 1.

Statistically significant differences existed between both groups regarding uncertainty intuitiveness, group 2 (cartography and geoinformatics students) needing more time to make decisions. However, group 2 achieved better results for all tasks on the second level (the decoding of
values). Particular differences arose in the last task (decoding value and uncertainty at the same time), where only 27% of group 1 members were able to decode values correctly for maps combined visualization. All other results were comparable for both groups.

Discussion

Leitner and Buttenfield (2000) recommended symbolization schemes for correct decision-making and proposed graphic variable value as the most appropriate; they also supported the idea that more certain information should be visualized by lighter values. They were aware of the fact that this result appears to be counterintuitive, since darker values have been repeatedly suggested for the depiction of more certain information because they are perceived by the map reader as being more prominent. Lighter values, by contrast, are perceived as being less prominent (MacEachren 1992). However, they saw the main reason for this conflict in the use of different testing environments and stated that MacEachren’s original assumption is correct if certainty information is depicted on printed paper and colours are perceived by reflected light. On a computer screen, however, where colours are perceived with emitted light, the results might be reversed. The results achieved in our test (performed purely in a digital environment), however, have not verified their conclusions. Because of their lack of statistical significance we can only state that the use of light values for less certain information as proposed by MacEachren is more in line with human intuition according to our results. Comparison of both methods – maps combined and maps compared – did not reveal significant differences with respect to the simple decoding of predicted values (soil depth). We can thus expect that there is no difference whether we use lighter or darker values for prediction (as in the case of maps compared) or different hues across the spectrum (as in the case of whitening). This assumption has not yet been properly verified.

Another result to be discussed deals with the response times of tested subjects relative to the variable complexity of the information being processed. Leitner and Buttenfield (2000) claim that adding attribute information of any kind should slow down subject response times. However, adding attribute certainty did not increase response times in their original research. No significant differences in response times were found when comparing one class maps with attribute certainty maps. This finding implies that map readers do not assimilate attribute certainty in the same way as they assimilate added map detail. The inclusion of certainty information appeared to clarify map patterns without requiring additional time to reach a decision. The fastest response times were achieved for certainty maps showing saturation; thus, if a fast decision is the highest priority, attribute certainty should be symbolized by more pastel colours.

As for serial and parallel processing results, serial is more time consuming but yielded better quality results (67 of 104 correct answers, 64%, Figure 8a). Parallel thinking for the whitening method yielded significantly better results as for speed, but was less accurate with respect to answers (only 59 of 104 correct answers, 56%, Figure 8b). In both cases there were five possible answers in the test with only one correct – no.3 for maps compared and no.5 for maps combined. However, there were always secondary ‘champions’ close-to-correct-answers results – no.4 for maps compared (29% of answers) and no.4 for maps combined (26% of answers).

Conclusions and future plans

The test results confirmed the importance and relevancy of empirical testing and yielded preliminary comparison results concerning the two basic uncertainty visualization techniques – maps compared and maps combined. According to our testing there exist significant differences between the methods when decoding uncertainty, value and both of them at the same time. The main findings can be summarized as follows:

- Intuitiveness testing confirmed that less certain information should be visualized by light values.
- Test participants were able to make quicker decision for maps combined method when decoding simple values of uncertainty or predicted soil depth.
- Correctness of answers is lower for more complex task – decoding of both values at the same time. Maps compared achieved slightly better results than maps combined.

The results presented in this article are still of a preliminary nature and serve as a basis for further investigation, with respect to both statistical processing and more extensive psychological (cognitive) testing of uncertainty visualization. It can be concluded that the whitening technique is not appropriate for direct decoding (reading) of interpolated values because of the use of a complex legend in which people with even minor colour sense malfunction are not able to recognize correct values of uncertainty.

A usability and cognitive testing web-based environment (GP test) has proven its suitability and will serve as a basic tool for ongoing research. Further development of distributed methods for user evaluation as well as the software platform itself is fully compliant with research challenges outlined by Robinson (2011). The uncertainty visualization tests presented here are still available for more extensive user group testing, repetition of tests and the possible incorporation of other cartographic issues within the test environment.
Although the test group was large enough and homogeneous, the achieved results have certain limitations. Our tasks, which were focused on the users’ work with uncertainty, were included in more complex tests with other cartographic and psychological tasks. The number of single items in the task was limited for this reason. The inclusion of more items in the tasks can reduce the possible risk that unexpected intervening variables (e.g. impact of learning, distribution of correct answers in the map field) will affect the results. It is necessary to increase the number of items in future research.

Acknowledgement
This article was elaborated as part of Research Plan MSM0021622418 of the Ministry of Education, Youth, and Sports of the Czech Republic called ‘Dynamic Geovisualization in Crisis Management’ and Czech Science Foundation project GA205/09/1437 – Agrisensor.

References


