Dear Author,

Please, note that changes made to the HTML content will be added to the article before publication, but are not reflected in this PDF.

Note also that this file should not be used for submitting corrections.
Evaluating the use of uncertainty visualization for exploratory analysis of land cover change: A qualitative expert user study

Christoph Kinkeldey, Jochen Schiewe, Henning Gerstmann, Christian Götze, Oleksandr Kit, Matthias Lüdeke, Hannes Taubenböcke, Michael Wurm

HafenCity University Hamburg, Lab for Geoinformatics and Geovisualization (g4lab), Hamburg, Germany
Martin Luther University of Halle-Wittenberg, Department of Remote Sensing and Cartography, Halle, Germany
GeoData Institute, University of Southampton, Southampton, United Kingdom
Potsdam Institute for Climate Impact Research, Potsdam, Germany
German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, Germany

Abstract

Extensive research on geodata uncertainty has been conducted in the past decades, mostly related to modeling, quantifying, and communicating uncertainty. But findings on if and how users can incorporate this information into spatial analyses are still rare. In this paper we address these questions with a focus on land cover change analysis. We conducted semi-structured interviews with three expert groups dealing with change analysis in the field of climate research, urban development, and vegetation monitoring. During the interviews we used a software prototype to show change scenarios that the experts had analyzed before, extended by visual depiction of uncertainty related to land cover change. This paper describes the study, summarizes results, and discusses findings as well as the study method. Participants came up with several ideas for applications that could be supported by uncertainty, for example, identification of erroneous change, description of change detection algorithm characteristics, or optimization of change detection parameters. Regarding the aspect of reasoning with uncertainty in land cover change data the interviewees saw potential in better-informed hypotheses and insights about change. Communication of uncertainty information to users was seen as critical, depending on the users’ role and expertise. We judge semi-structured interviews to be suitable for the purpose of this study and emphasize the potential of qualitative methods (workshops, focus groups etc.) for future uncertainty visualization studies.

1. Introduction and background

Uncertainty is an inherent characteristic of geodata and can play an important role during their analysis (Zhang and Goodchild, 2002). Thus, a research effort in GIScience is to develop methods to incorporate uncertainty into geodata analysis. In the last decades, a wide number of user studies have been conducted to assess potential benefits of uncertainty visualization for this purpose (MacEachren et al., 2005). While the vast majority of studies focus on the impact of uncertainty visualization on decision making (Griethe and Schumann, 2005) only very few deal with potential effects on reasoning with geodata.

This research contributes to filling this gap with a user study about if and how geodata uncertainty can be utilized in land cover change analysis. The study is based on a concept to utilize uncertainty in change analyses that includes a measure for uncertainty in change (Kinkeldey, 2014b), a technique to visualize uncertainty (noise annotation lines, Kinkeldey et al., 2014), and a software prototype for change analysis (ICchange, Kinkeldey, 2014b). We report upon interviews with three expert user groups utilizing the software prototype to analyze land cover change data and discuss the concept. Topics include the use of uncertainty in change analysis, as well as potential and benefits of the software prototype and the uncertainty visualization technique.

This article is based on a workshop paper that summarized preliminary results of this study (Kinkeldey and Schiewe, 2014). It extends the paper by detailed descriptions of the study method and the change scenarios used in the interviews, and by presenting an in-depth discussion of the method and findings, as well as

Please cite this article as: Kinkeldey, C., et al., Evaluating the use of uncertainty visualization for exploratory analysis of land cover change: A qualitative expert user study. Computers & Geosciences (2015), http://dx.doi.org/10.1016/j.cageo.2015.08.007
recommendations for future work.

2. Method

The goal of this research was to assess our concept for uncertainty-aware land cover change analysis (Kinkeldey, 2014a). The main questions being if expert users would find it useful for their work and where they see benefits and limitations. In past uncertainty evaluation research, the majority of user studies applied quantitative methods, i.e., mainly experiments in laboratory settings or over the Internet (Kinkeldey et al., 2014). Exceptions include a number of qualitative studies, for instance, a focus group study by Roth (2009a) to investigate the impacts of uncertainty visualization on decision making in the context of floodplain mapping. Other authors conducted interviews to evaluate the usability of a tool utilizing uncertainty visualization (Slocum et al., 2003) and the usefulness of different visualization techniques to describe uncertainty (Gerharz and Pebesma, 2009). Apart from this, mixed methods (combining quantitative and qualitative methods) have been applied, but remain very rare (e.g., Štěrba et al., 2014). For our study we needed to make sure that several topics were covered. At the same time, we wanted to leave room for a discussion of new aspects and ideas. We identified the method of semi-structured interviews as suitable for our purposes because it connects these requirements.

2.1. Interviews

To evaluate the usability of the concept we conducted three semi-structured interviews with expert groups that are concerned with land cover change analysis. The core idea was to utilize the software prototype for the interviews to present change scenarios the interviewees had already worked on. We found three groups of two to four experts dealing with change analysis who were interested in taking part in the interview. The groups covered the areas of climate research, urban remote sensing, and vegetation monitoring. The interviews had four parts:

1. Introduction: In the first part we explained the concept and the software prototype showing an exemplary change dataset, not yet the data for the discussion, to keep the focus on the prototype and the visualization technique. The participants were free to ask questions.

2. Uncertainty: The main part of the interview started with questions about the role of uncertainty in the specific dataset. First, we showed the experts their change scenario without uncertainty and asked them about insights they had gained from it so far. We then added the uncertainty display to let them explore uncertainty related to the changes. Instead of operating the software prototype themselves, participants were asked to give instructions to us. This idea is adapted from pair analytics that involves a visual analytics expert operating the tool and a subject matter expert posing the questions (Arias-Hernandez et al., 2011). This was done to ensure that the discussion stays focused on the data and to avoid discussions about the usability of the prototype, an aspect that had already been assessed during its development (Kinkeldey, 2014b). The questions were about, if and how the uncertainty display helps to confirm, reject, or modify the insights they had reported on before the uncertainty display was added. In addition, we were interested in their opinion about the significance of uncertainty in change analysis from a general view, i.e., not related to the presented dataset.

3. Tool and visualization: Subsequently, we asked the participants about their opinion on the ICC change software prototype and on noise annotation lines, the technique we used to display uncertainty in the map (Kinkeldey et al., 2014). We talked about the potential of the prototype to support them in their work compared to the software they currently use. Regarding noise annotation lines, we asked them whether they find this technique usable for their tasks.

4. Open questions: In the last part the interviewees had the opportunity to make comments about the topics covered in the interview, and to express ideas and criticism.

The introduction took 10–20 min depending on the number of questions from each group. With all three groups the interviews took about one hour (excluding the introduction). The division of the discussion into the four parts was not strict but served as a rough guideline. We recorded the discussion with two separate voice recorders (notebook and smartphone). After transcribing the recordings in writing we categorized the findings related to ‘change detection and analysis’, ‘reasoning with uncertainty’, ‘communication of uncertainty’, and ‘tool and visualization’.

![Fig. 1. Change uncertainty measure derived from class membership values \( \mu \) [reprinted from Kinkeldey (2014a)].](image-url)
2.2. Change uncertainty measure

We defined a measure for change uncertainty based on the work of Fisher et al. (2006). The computation is as follows (Fig. 1): for each pixel, the minimum of the two class membership values $\mu_0$ and $\mu_1$ are subtracted from 1.0, yielding an uncertainty value between 0.0 and 1.0. For most applications, it is recommended to classify the values to decrease complexity, for example, using two classes of low and high uncertainty (Fig. 1, ’u classed’). Please refer to Kinkeldey (2014a) for a detailed description and discussion of the measure.

2.3. Software prototype ICchange

As part of our concept for uncertainty-aware change analysis we developed a software prototype for change analysis (ICchange) as proof of concept and as vehicle for discussion. It provides two linked views: the map view showing changes and related uncertainty and the info view, an abstract overview on occurring changes (Fig. 2). Both views include a visual depiction of change uncertainty: noise annotation lines in the map view and an uncertainty glyph in the info view. For a detailed description and discussion of the prototype and its development, refer to Kinkeldey (2014b).

3. Change scenarios

In this section we describe the change scenarios provided by each expert group. They cover the areas of climate research, urban remote sensing, and vegetation monitoring.

3.1. Group 1 (climate research, Potsdam Institute for climate impact research, Potsdam, Germany)

The first scenario was concerned with the change of informal settlements in Hyderabad, India between 2003 and 2010. The change data was derived from high-resolution optical satellite imagery (QuickBird, WorldView 2). In general, the task was to project impacts caused by future anthropogenic climate change. These depend on both, the projected change in climate and the future sensitivity of the system under investigation. For the fast growing urban agglomerations of the Global South the future extent of informal slum areas will be decisive for their sensitivity towards climate change. The present spatio-temporal dynamics of slum areas is the basis for these assessments but official data is strongly biased by political interest. Therefore, a remote sensing based approach was taken – exemplarily for the city of Hyderabad/India (about 8 million inhabitants in 2010) – to identify the extent of and location of slum areas for 2003 and 2010 using satellite data, in this case 11 bit cloudless mosaics from QuickBird full swath (acquired on 27 May and 11 June 2003) and WorldView 2 full swath (acquired on 3 and 14 February 2010). The results showed two major zones of slum growth in the north and the south of the city (Fig. 3). For further details see Kit and Lüdeke (2013).

For this scenario, the membership values needed for the computation of the uncertainty measure were derived from a lacunarity interval that defines the land cover class ‘informal settlements’ (Kit and Lüdeke, 2013). Instead of crisp interval boundaries from the initial analysis we defined a fuzzy interval to represent the uncertainty in the class definition (Fig. 4). This way, we derived fuzzy class membership values for each pixel, for both the ‘informal settlement’ and the ‘no informal settlement’ classes, between 0.0 and 1.0.
change uncertainty measure.

3.2. Group 2 (urban earth observation, German Aerospace Center (DLR), Oberpfaffenhofen, Germany)

The second change scenario concerns the rapid growth of the metropolitan region of Shanghai, China between 1987 and 2004. The metropolitan area of Shanghai is characterized by extreme urbanization dynamics. Between 1987 and 2004, the population has almost doubled from about 7 million to 14 million. Changes from non-urban to urban areas have been detected from three medium resolution optical satellite images (Landsat TM) acquired in 1987, 1995, and 2004. For land cover classification, an object-based approach was applied using the commercial software package eCognition by Trimble Geospatial. A maximum likelihood classification, based on stable spectral features such as the normalized difference vegetation index (NDVI), provided urban footprints of the investigated area. An urban footprint is here defined as the physical representation of the urban area, which can be discriminated by satellite imagery (Taubenböck et al., 2012). Three change maps are shown in Fig. 5, depicting urban growth between 1987 and 1995, 1995 and 2004, and over the whole time period (1987–2004).

For this change scenario, the class membership values for the uncertainty measure were delivered by the classification algorithm, expressing classification confidence between 0.0 and 1.0. Fig. 6 shows the map view including the change map and related uncertainty represented by noise annotation lines.

3.3. Group 3 (vegetation monitoring, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany)

In the third scenario, vegetation change in a post-mining area in central Germany was analyzed between 2000, 2003, and 2009, based on high-resolution hyperspectral data (HyMap). The former mining landscape Goitzsche is situated in the eastern part of the German federal state of Saxony-Anhalt near the town Bitterfeld-Wolfen. After approximately 100 years of mining activities, the area is being restored after these activities ended in the early 1990s. During the mining activities, the surface layers were completely degraded and the ground water table was decreased artificially, and since the mining has stopped, the ground water table has risen again. Thus, the area is characterized by very dynamic vegetation development.

We monitored the changing composition of the vegetation by applying image change detection methods on three hyperspectral HyMap datasets acquired in 2000, 2003, and 2009. Seven land cover classes (deciduous and conifer trees, species-rich and species-poor xerothermic grasslands, vegetation-free areas, water bodies, vegetation affected by water logging) were distinguished using a Support-Vector-Machines (SVM) classification algorithm (Gerstmann, 2013, Gerstmann et al., 2014). The separation of the different grassland types turned out to be the most uncertain class separation. Refer to Fig. 7 for the land cover datasets for 2000, 2003, and 2009.

For the post-mining area scenario, analogous to the one from group 2, the class membership values for the uncertainty measure were delivered by the classifier, in this case, as rule images from the SVM algorithm.

4. Findings

In this section we summarize findings from the three expert group interviews, divided into subsections with respect to change detection and analysis, reasoning with uncertain change, communication of change uncertainty, as well as the software prototypes and uncertainty visualization technique.

4.1. Change detection and analysis

In the interviews most comments were related to the use of uncertainty during change detection and analysis. Generally, all groups were highly interested in the visual depiction of geographically varying change uncertainty. For example, an expert from group 2 remarked that the step from class-specific uncertainty measures to geographically varying uncertainty information was beneficial for several applications he had in mind. For instance, to explore the characteristics of their change detection algorithms. A member of group 3 realized that the magnitude of uncertainty displayed in the map expressed the implicit knowledge they had about their data. For example, based on their experience they rated a specific change (from species-poor to species-rich xerothermic grasslands and back) as unrealistic, The prototype revealed that high uncertainty was associated with this change type, which supported their hypothesis that this is likely to
be an error. All groups were convinced that the information about areas with a high degree of uncertainty could help them identify changes that may need to be examined further.

A member of group 2 came up with the hypothesis that in their urban change map of Shanghai, uncertainty would increase in low-density rural areas because the detection of urbanized areas is more difficult there. Visual comparison of the respective areas in the prototype, however, could not verify this assumption. Nonetheless, this remains an interesting question for them that could potentially be supported by uncertainty information. They were also interested in deconstructed areas (i.e., change from built up to non built up areas) that were contained in the Shanghai dataset. The reason for their increased interest in these areas was that they usually disregard deconstruction areas in their analyses. Showing these changes in the prototype revealed that uncertainty in these areas was uniformly high and its visualization in the map made the experts believe that the changes in the dataset describing deconstruction were errors. Thus, in this case information about uncertainty helped confirm their hypothesis.

An application that all groups were interested in was to utilize uncertainty for the optimization of change detection parameters. For instance, group 1 used two parameters to detect informal settlements, one for the line detection step and one for the lacunarity interval defining the ‘informal settlement’ class (see Section 3). The members of this group suggested that for the mere visual
detection of informal settlements in the high-resolution imagery they would not need uncertainty information. However, they believed that uncertainty could help find suitable parameters for the automation, especially when a tool like ICchange would be used that immediately displays the results when parameters are modified. Yet, they pointed out that when parameters are modified locally (to see more details on the map) it could be challenging to keep the overview of the results in other areas. This led to the idea that uncertainty could be displayed in an overview first to identify areas where the current parameter set leads to high uncertainty. After that, parameters could be modified for these areas to optimize the results. A similar idea was brought up by group 2 who suggested optimizing parameters for change detection by examining areas with low and high uncertainty to find out where the parameter set leads to good results and where it does not. One expert from group 3 remarked that he could imagine using uncertainty to enhance land cover class definitions and to support the decision as to whether or not further classes are needed to increase the quality of the change data.

4.2. Reasoning with uncertainty

The question as to whether the process of generating insights from change maps can be supported by uncertainty visualization was another topic brought up in the interviews. Group 1 stated that they wanted to utilize information about change uncertainty and that this could be interesting information for their work. For instance, the information that the disappearance of informal settlements within the city of Hyderabad was more uncertain than the growth of settlements in the surroundings would add value to the data. They stated that this could have helped them with the interpretation and claimed that they are used to interpreting ‘soft’ data. Another aspect they mentioned was that when they estimate the population of informal settlements, they use predicted population densities that vary a lot between different studies. In their opinion, it could potentially help to know about the uncertainty of detected informal settlements.

As a member of group 2 suggested, he could imagine that when deriving insights from geodata they may express higher confidence with changes that are more certain and doubts with interpretations of more uncertain changes. However, he found it challenging to involve multiple types of uncertainty, e.g., uncertainty related to population density and the location of a settlement’s boundary. In this case, it could be a complex task to consider both types simultaneously during reasoning.

4.3. Communication of uncertainty

A number of aspects regarding the communication of uncertainty have been addressed during the interviews. The first one was the question as to if uncertainty should be communicated to users, e.g., decision makers. In the three interviews this aspect was seen in different ways. For instance, members of group 2 and 3 were skeptical about communicating uncertainty to non-experts in remote sensing. A participant from group 2 claimed they were willing to use ‘soft’ data for the analysis but as soon as insights are communicated they need to be deterministic, for instance, ‘city A grows faster than city B’. On the one hand, if uncertainty is communicated along with the insights it could raise doubts about the general quality of the data. On the other hand, they were not convinced that decision makers would be able to effectively use this information and that ‘it will be hard in the beginning to create acceptance for this’. They also saw the problem that discussions about the data might focus too much on the related uncertainty instead of the content of the data itself. However, all groups agreed that the question of whether to communicate uncertainty information to users depends on the users’ role and expertise. Group 3 remarked that they would only use this information internally and communicate it to colleagues from research but not to external users. Similar to group 2, they saw potential pressure in justifying uncertainty in the data to users. In contrast, the experts from group 1 were convinced that people who make decisions based on their data should be provided with uncertainty information (‘they need information about what is certain and what is uncertain’). They also pointed out that their typical users (decision makers dealing with climate scenarios) are already used to dealing with ‘soft’ information.

Apart from communicating uncertainty related to change data, group 2 suggested to utilize uncertainty to illustrate characteristics of change detection algorithms, e.g., in publications or presentations. In their opinion, showing spatially varying uncertainty in a map could be a means to report on the characteristics in a more attractive and graspable way than showing statistical charts to explain the properties of a specific algorithm.

4.4. Tool and visualization

In the third part of the interviews, we asked the participants if they could imagine using a tool like ICchange and what they would need to use it in their practical work. Generally, all subjects found that the biggest barrier that prevents them from using uncertainty information is the lack of tool support. The majority suggested integrating tools like ICchange as plugins into standard GIS (ArcGIS, QGIS) to establish its seamless integration into existing workflows. However, one member of group 1 would prefer a standalone solution independent from standard GIS software.

The spatially varying depiction of uncertainty in the map view was seen as a clear advantage over commonly used class-specific quality measures that only provide one value per change type and do not show its spatial distribution. One expert suggested using the table from the info view, which shows all changes as a kind of legend for change maps, instead of the commonly used change matrix.

All groups expressed positive opinions about the noise annotation lines technique and found it intuitive and useful. Most experts quickly identified the advantages of the technique, e.g., that the maps in the background do not have to be altered and that they could be shown using the original color schemes. Another potential advantage mentioned by group 2 was that it prevents possible ‘salt-and-pepper’-effects when uncertainty varies a lot because of its smoothing characteristic (Kinkeldey et al., 2014). Some participants recognized limitations, e.g., that readability may be low with a bright or visually complex background.

5. Discussion

In this section we discuss selected findings from the summary in Section 4. In addition we highlight the benefits and limitations of the method used for the study.

5.1. Findings

After summarizing the findings form the interviews we noticed that it was easier for most participants to imagine using uncertainty for analytical purposes, e.g., for calibration of change detection algorithms or for the quality assessment of the data. Making suggestions as to how to use uncertainty for reasoning with the data seemed more difficult for them because the groups do not primarily deal with interpretation of change data. Still a number of ideas about supporting the generation of hypotheses and insights were suggested.
Regarding reasoning with uncertain change data, the experts could imagine using uncertainty as additional information to better judge hypotheses about change. They were convinced that insights could be better-informed and more differentiated upon incorporating uncertainty information. However, in the end, the majority wanted to generate ‘hard’ insights from the uncertain data because they do not want to confront their users, e.g., decision makers, with uncertain information. This is related to the finding that the success of incorporating uncertainty depends on the users’ role and expertise, confirming results from past uncertainty visualization studies (Cliburn et al., 2002, Hope and Hunter, 2007, Roth, 2009b). All this stresses the importance of carefully defining types and levels of expertise required for the use of uncertainty measures and visualizations.

During the discussions about ICchange, the software prototype, the reactions towards both the prototype and the uncertainty visualization technique were positive. This goes in line with our findings from the usability study (Kinkeldey, 2014b), yet we had expected more critical feedback because the prototype was intentionally kept simple, and it was likely that experts would have expected more functionality from the prototype.

All in all, the three interviews mainly yielded similar conclusions towards the use of uncertainty in change analysis, which strengthens the validity of the findings. One of the aspects often neglected in uncertainty visualization research, acceptance of uncertainty by users, will be further discussed in the conclusion.

5.2. Method

For the purpose of this study, semi-structured interviews offered a number of advantages. Since uncertainty is a fuzzy concept and includes various aspects (modeling, quantification, communication, etc.) discussions can digress very easily. The method helped keep the focus on the utility of uncertainty in change analysis without too much digression towards other topics such as issues with the software tool. At the same time the loose structure did not prevent people from discussing freely and expressing their opinions and ideas.

For the study we had to select the change scenarios from those the groups had already analyzed before. This can be seen as a limitation towards external validity because two of the scenarios utilized in the study cover urban applications and one covers vegetation monitoring. Yet, since the results we selected here are not specific to the application field, we do not expect that a wider coverage of fields would have been advantageous.

In addition, using existing change scenarios provided by the expert groups required additional effort prior to the interviews. Suitable scenarios had to be found and legal issues needed to be clarified, but the fact that the interviewees had already worked with the data helped the discussion from becoming too theoretical and it helped them imagine using uncertainty information for their work. However, it might have been reasonable to conduct a training session before the interview to brush up their knowledge about the dataset.

6. Conclusion

In this article, we reported on three expert studies assessing the role of uncertainty in exploratory land cover change analysis. The semi-structured interviews had four parts: an introduction, a second part about the role of uncertainty information in change analysis, a third part about the software prototype (ICchange) and the uncertainty visualization technique (noise annotation lines), as well as a concluding part for comments. We interviewed three groups of two to four experts using the prototype to present change scenarios provided by each group, complemented by visually depicted uncertainty. Each session took about one hour (excluding the introduction).

The experts were interested in geographically varying change uncertainty, information they usually do not have, and were curious about seeing uncertainty displayed for their data in order to assess the quality of detected change in different areas. Potential applications of uncertainty were suggested during the interviews, such as optimizing change detection parameters, assessing the characteristics of different detection algorithms, or identifying erroneous change. Regarding reasoning with land cover change data, better-informed hypotheses and insights were seen as possible when information about uncertainty is available. Most participants agreed that an important requirement for incorporating uncertainty into change analysis is the support in standard software tools. Thus, for the future it would be meaningful to integrate ICchange into standard GIS software packages such as ArcGIS2 or QGIS to facilitate integration into existing workflows.

One of the main findings from the interviews was that the first step towards the use of uncertainty in practical work is to establish acceptance that having uncertainty depicted can be beneficial. As long as data is seen as inferior when uncertainty is communicated and users are not willing to invest additional time and effort the widespread use of uncertainty in geodata analysis will remain theory. But this is not the only challenge to be countered; as Cliburn et al. (2002) suggested, non-scientific users do not only need the information about uncertainty but ‘to maximize the effectiveness of visualization, uncertainty must be represented and ways to deal with it must be provided’ (p. 948). This is an issue we see supported by our results. For future research, we recommend paying more attention on how to assist users in utilizing uncertainty. Furthermore, it would be reasonable to conduct further studies such as workshops or focus groups in which participants conduct actual analyses to gain knowledge about how uncertainty can be used to support reasoning with uncertain geodata (rather than decision making). Generally, we see great potential in qualitative methods (in addition to quantitative methods) to help understand what is needed to come closer to the goal of successfully using uncertainty visualization to support geodata analysis.

References


