HCU HafenCity Universität Hamburg

Incorporating uncertainty information into exploratory land cover change analysis: a geovisual analytics approach

A Dissertation by
Christoph Kinkeldey

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doktor-Ingenieur (Dr.-Ing.)

October 2014

Reviewers:
Jochen Schiewe
Alan M. MacEachren
Alexander Klippel

“Although our intellect always longs for clarity and certainty,
our nature often finds uncertainty fascinating.”

Carl Philipp Gottlieb von Clausewitz
(1780-1831)

Abstract

During the last decades, research in the field of geographic information science (GIScience) has been conducted to describe, quantify, and utilize information about uncertainty in geodata. However, practically usable models and methods for this purpose are still rare and the supposed positive effect of communicating uncertainty during spatial analyses is still subject to research. This dissertation deals with these challenges, focusing on land cover (LC) change analysis based on remote sensing data, a field in which uncertainty plays an important role. It has been shown that when uncertainty is ignored during change analysis, misleading results can be the consequence.

The project described in this dissertation aims to establish uncertainty-aware change analysis by following a Geovisual Analytics (GVA) approach. This approach utilizes workflows that integrate automated algorithms with interactive visual interfaces to combine the strengths of computation and human reasoning. The central question of this work is if users are able to benefit from information on change uncertainty using a GVA tool: Does it help with hypotheses they make and with insights they gain during change analysis? This question is addressed in the following stages: (1) development of a concept for uncertainty-aware land cover change analysis, (2) literature review on uncertainty visualization user studies, (3) design and implementation of a GVA software prototype including a usability assessment, and (4) a user study with expert analysts who apply the prototype to analysis tasks and evaluate the concept:

(1) A concept to incorporate uncertainty into land cover change analysis is the basis of this work. Here, the most important steps include the definition of an uncertainty measure for LC change, the selection, implementation, and evaluation of a technique to portray uncertainty in LC change maps (*noise annotation lines*), as well as the creation of a task categorization for exploratory change analysis. A case study with real-world land cover change data demonstrates the applicability of the concept.

(2) A systematic review of uncertainty visualization user studies is a fundament for the selection of a suitable visualization technique. Findings from the studies are summarized and lessons learned, as well as recommendations for future studies in the field, are provided. This includes a novel categorization model (*Uncertainty Visualization Cube, UVis^3*) for more systematic selection and evaluation of uncertainty visualization techniques in the future.

(3) A software prototype for uncertainty-aware change analysis (*ICchange*) serves as a proof-of-concept and as a vehicle for an expert user study to assess the practicability of the concept developed in this doctoral research. The prototype development includes a verbal protocol analysis (VPA) study to assess its usability. The description of the task-based development

approach is intended to support developers creating GVA software tools for land cover change analysis or related fields.

(4) An expert user study assesses the practicability of the developed concept with three groups of remote sensing experts who conduct land cover change analyses in practice. In semi-structured group interviews, the software prototype is utilized to demonstrate real-world change scenarios that the experts have analyzed before, enriched by information about uncertainty in the results. The participants generally see the concept as useful for supporting change analysis and a number of potential applications are discussed. Barriers for the use in practice are addressed as well, such as user acceptance issues and the lack of support in standard GIS. The experts' opinions about the ICchange software prototype and the noise annotation lines visualization technique were predominantly positive. From the results of the study and experience from prototype development, recommendations are given for the implementation of the concept in practice and for questions to be addressed in future research.

This dissertation provides various novel findings regarding geodata uncertainty visualization and uncertainty-aware exploratory analysis. It contributes to bridging the gap between theory and practice and represents a crucial step to come closer to the goal of making uncertainty usable in real world geodata analysis.

Zusammenfassung

Im Fachgebiet der Geographischen Informationswissenschaft (GIScience) wurde in den letzten Jahrzehnten intensiv Forschung betrieben, um Unsicherheiten in Geodaten zu beschreiben, zu quantifizieren und nutzbar zu machen. Praktisch einsetzbare Modelle und Methoden für diesen Zweck sind allerdings noch selten, und der vermeintliche positive Effekt, wenn Informationen zu Unsicherheiten bei räumlicher Analysen kommuniziert werden, ist noch Teil der Forschung. Die vorliegende Arbeit stellt sich diesen Herausforderungen und konzentriert sich dabei auf Veränderungsanalysen von Landbedeckung auf Basis von Fernerkundungsdaten. In der Vergangenheit wurde gezeigt, dass bei solchen Analysen Unsicherheiten eine wichtige Rolle spielen, und dass falsche und irreführende Ergebnisse die Folge sein können, wenn diese ignoriert werden.

Das in dieser Dissertation beschriebene Projekt zielt darauf ab, ein Konzept nach dem Geovisual Analytics-(GVA)-Ansatz zu entwickeln, um Unsicherheiten bei Veränderungsanalysen zu berücksichtigen. Dieser Ansatz ermöglicht Arbeitsabläufe, die automatische Algorithmen mit interaktiven visuellen Schnittstellen vereinen, um die Stärken von maschineller Berechnung mit denen der menschlichen Interpretation zu verbinden. Die zentrale Frage dieser Arbeit ist, ob Nutzer in der Lage sind, bei Verwendung eines GVA-Tools von Unsicherheitsinformationen zu profitieren: Helfen diese beim Aufstellen von Hypothesen und beim Erlangen von Erkenntnissen während Veränderungsanalysen? Dieser Frage wird sich in folgenden Schritten gewidmet: (1) Entwicklung eines Konzepts für Veränderungsanalysen unter Einbeziehung von Unsicherheiten, (2) Erstellung einer systematischen Übersichtsarbeit zu Nutzerstudien im Bereich Unsicherheitsvisualisierung, (3) Design, Implementierung und Evaluation eines GVA-Software-Prototypen und (4) Durchführung einer Nutzerstudie mit Experten-Nutzern, welche den Prototyp einsetzen und das Konzept bewerten:

(1) Ein Konzept für die Einbeziehung von Unsicherheiten in Veränderungsanalysen der Landbedeckung wird entwickelt. Dabei sind die wichtigsten Schritte die Definition eines Unsicherheitsmaßes für Landbedeckungsveränderungen, die Auswahl, Implementierung und Evaluierung einer Technik zur Darstellung von Unsicherheiten in Veränderungskarten (*Noise Annotation Lines*), sowie die Erstellung einer Kategorisierung von Tasks in explorativen Veränderungsanalysen. Eine Fallstudie mit realen Veränderungsdaten zur Landbedeckung demonstriert die Anwendbarkeit des Konzepts.

(2) Eine systematische Übersichtsarbeit von Nutzerstudien zur Visualisierung von Unsicherheiten dient als Basis für die Auswahl einer geeigneten Visualisierungsmethode. Ergebnisse der Studien werden zusammengefasst, „Lessons Learned“ formuliert, sowie

Empfehlungen für zukünftige Studien gegeben. Ein neuartiges Modell für die Kategorisierung von Visualisierungsmethoden (*Uncertainty Visualization Cube, UVis³*) wird vorgeschlagen, als Basis für eine systematischere Auswahl und Evaluation von Methoden zur Unsicherheitsvisualisierung in der Zukunft.

(3) Als Machbarkeitsnachweis (Proof-of-Concept), sowie als Grundlage für eine Expertenstudie, welche die Praktikabilität des hier entwickelten Konzepts überprüft, wird ein Software-Prototyp für die Änderungsanalyse mit Unsicherheiten (*ICchange*) entwickelt. Im Rahmen der Entwicklung wird eine Verbal Protocol Analysis (VPA)-Studie durchgeführt, um die Gebrauchstauglichkeit des Prototypen zu überprüfen. Die Beschreibung des task-basierten Entwicklungsansatzes für den Prototpen soll als Unterstützung dienen für Entwickler, die GVA-Tools für Veränderungsanalysen oder ähnliche Anwendungen entwickeln.

(4) In einer Studie mit drei Gruppen von Experten, die Landbedeckungsanalysen in der Praxis durchführen, wird die Praktikabilität des entwickelten Konzepts überprüft. In semi-strukturierten Gruppeninterviews wird der Software-Prototyp eingesetzt, um verschiedene, den Experten bereits vertraute Änderungsszenarien zu demonstrieren, angereichert mit Informationen über Unsicherheiten in den Ergebnissen. Generell sehen die Teilnehmer das Konzept für die Unterstützung von Änderungsanalysen als nützlich an, und eine Vielzahl potenzieller Anwendungen wird diskutiert. Barrieren für den Einsatz in der Praxis werden ebenfalls angesprochen, wie beispielweise die fehlende Unterstützung durch gängige GIS-Software, sowie Vorbehalte, Unsicherheitsinformationen an die Nutzer der Daten zu weiterzugeben. Die Meinungen der Experten zum ICChange-Software-Prototyp und zur Noise Annotation Lines-Methode sind vorwiegend positiv. Aus den Ergebnissen der Studie und aus Erfahrungen während der Entwicklung des Prototyps werden Empfehlungen für die Umsetzung des Konzepts in die Praxis gegeben, sowie zukünftige Forschungsfragen formuliert.

Diese Doktorarbeit liefert neuartige Erkenntnisse im Hinblick auf Visualisierung von Unsicherheiten in Geodaten, sowie bezüglich der Unterstützung explorativer Analysen durch Unsicherheiten. Sie trägt dazu bei, die Kluft zwischen Theorie und Praxis zu schließen und stellt einen wichtigen Schritt dar in Richtung des Ziels, Unsicherheiten für Analysen von Geodaten in der Praxis nutzbar zu machen.

Acknowledgments

This PhD project was a long, exciting, sometimes exhausting journey with a lot to learn and to be curious and enthusiastic about. This great experience would not have been possible without several people whom I would like to thank.

Thank you,
Jochen Schiewe, my advisor, mentor, and valued colleague from the g²lab, HafenCity University Hamburg, for all his guidance, for countless discussions, hints and tips, and for being a patient, balancing person, getting me back on track whenever I was stuck again.

Thank you,
Alan MacEachren and Alexander Klippel, for a wonderful time at PennState University, for our inspring joint work, for being my thesis co-advisors, and – last but not least – for coming to my defense in Hamburg.

Thank you,
Beate Weninger, Sarah Goodwin, Anna-Lena Kornfeld, and Tonio Fincke, my office mates at the g²lab, for countless discussions, mutual support, enjoyable times at doctoral meetings and the "HCU visualizes!" group, coffee breaks, conference visits etc. – all that made my time at HCU unforgettable.

Thank you,
Jenny Mason, David Retchless, Alexander (Sasha) Savelyev, Paolo Raposo, and my other colleagues at PennState University, for a great time during my research stay in 2012, for inspiration from the uncertainty reading group, and for fun at conferences.

Thank you,
Till Nagel, for all the discussions, support, and geovis gossip, Michael Rigby, for helping me with the final works and for being my wingman in "Raiden", Christian Rudeloff, for our "439" sessions and chit-chat about science, life and everything, and Raphael Drechsel, for my GREAT temporary workspace in Vienna, for printing, help and support, and Schnitzel.

Thank you,
Brigitte, Christine, Andreas, my close family, for supporting me morally and for never asking "why aren't you finished yet?".

Thank you,
Juliane, for being the most important person in my life, and for your love and support over all the years.

Contents

List of figures

List of tables

1 Introduction

Geodata can be defined as "[d]igital data that represent the geographical location and characteristics of natural or man-made features, phenomena and boundaries of the Earth", including "abstractions of real-world entities, such as roads, buildings, vehicles, lakes, forests and countries"[1]. They are based on models that are abstracted, generalized, and approximated, thus geodata are always to some extent uncertain. This extent can be made transparent for the user through modeling, quantification, and communication of geodata uncertainty (Foody and Atkinson 2002, Zhang and Goodchild 2002). Related to land cover data, the assignment of geographic objects to land cover classes is ambiguous and vague and can be expressed using metadata. For instance, instead of labeling the object in Figure 1.1 as 'deciduous forest' and ignoring other probable land cover classes, we can provide more differentiated information that by 60% chance the class could be 'deciduous forest', but by 40% chance it could also be 'mixed forest'. This depicts the uncertainty in the classification of this object, caused by ambiguous class assignment and vagueness in the definition of land cover classes.

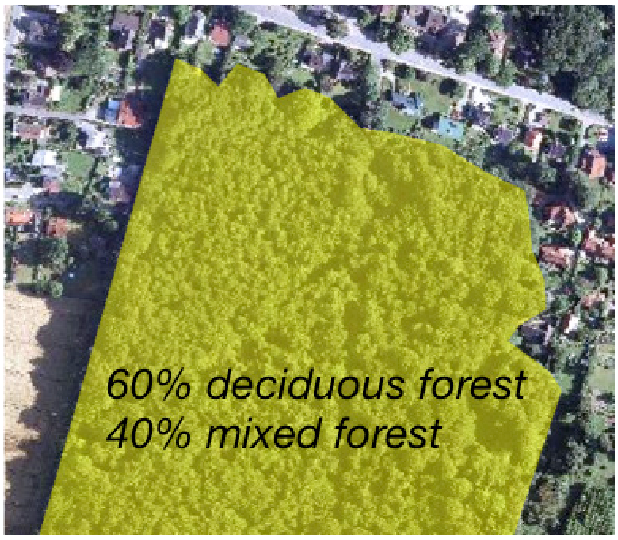

Figure 1.1. Uncertain classification of a land cover object.

Communication of uncertainty through visualization has been attracting attention in geographic information science (GIScience) research for decades. Research in this field faces the challenge of finding suitable ways to portray different types of uncertainty in a way that users can easily read,

[1] http://www.opengeospatial.org/ogc/glossary/g

understand, and interpret (MacEachren et al. 2005). One example is the so-called 'cone of uncertainty' (also called 'error cone') used by the U.S. National Hurricane Center (NHC). The cone is used in maps to indicate the possible paths of a storm, e.g., those of hurricane 'Sandy' at the East Coast of the USA in 2012 (Figure 1.2). The predicted tracks are uncertain in their geographic location. On the map, the current location of the eye of the storm is shown as a black dot with a red boundary. The solid white cone depicts the predicted storm track area in the following three days. The other cone, shown as a light white dotted pattern represents the same, but for the following 4 to 5 days. The difficulty with this depiction is that the cone shows the region with the most probable storm tracks, which is often misunderstood (Broad et al. 2007). Many map users assume that the cone depicts the size of the storm and have the impression that only the black dots represent its track – in fact, the actual track could be anywhere inside of the cone. To address this uncertainty, a textual warning is often added to this type of map, however in trying to explain this textually it reveals the weakness of the visualization (see Figure 1.2, black box in the upper part of the map). A detailed critique of the cone of uncertainty is provided in Broad et al. (2007) and studies about alternative visualization techniques can be found in Ash et al. (2013) and Cox et al. (2013).

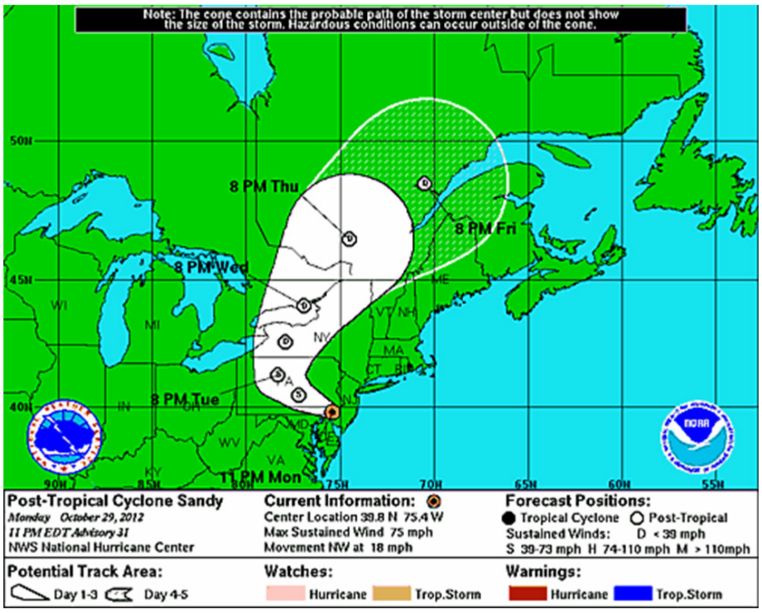

Figure 1.2. 'Cone of uncertainty' (source: http://www.nhc.noaa.gov, Oct 2012)

The example of the cone shows that visualizing uncertainty can be a challenging task, especially when probabilities come into play that are generally hard to understand by humans (Micallef 2012) and there is a risk of the user misunderstanding the information. Furthermore, it highlights the importance of user evaluation to learn about how users can be supported by uncertainty information.

As of today, numerous evaluation work has been conducted to address the challenges of uncertainty visualization. One part focus on how to communicate uncertainty visually, i.e., on answers to the question what techniques are best to convey this information graphically (Drecki 2002, Edwards and Nelson 2001, Evans 1997, Pang 2001, Sanyal et al. 2009). Another part assesses the impacts of communicating uncertainty to users, mostly the effect of uncertainty representations on decision making (Deitrick and Edsall 2006, Griethe and Schumann 2005, Leitner and Buttenfield 2000). Here questions addressed include whether decision performance (in terms of accuracy or speed) is affected, different outcomes are generated, or how confidence in a decision changes when users are provided with information about uncertainty in geodata.

However, research about how uncertainty visualisation can support exploratory analysis of (rather than decision making based on) geodata is still rare. In their article about the status of uncertainty visualization research, MacEachren et al. (2005) identify : "Most of the empirical evaluations [...] address visualization used to present information for a particular decision, such as the location of an airport. However, visualization is commonly used for exploratory analysis where there is no single question to be answered; rather the user is seeking to glean insights from the data. A core question here is how the portrayal and interaction with uncertainty can help the user better find and assess these insights" (p.156).

This dissertation addresses the above mentioned research challenge with a focus on exploratory land cover change analysis from remote sensing (RS) data where analysts generally have to cope with a high degree of uncertainty. Uncertainty from the RS data accumulates from dataset to dataset, and during the detection of changes to be analyzed, more uncertainty is introduced (Figure 1.3). However, commonly used GIS are mainly designed to handle precise data, which makes it difficult to make uncertainty transparent during analyses. Apart from this, change analysis often has an exploratory character, i.e., the exact questions are not clear before the analysis. Thus, the main challenge of this research is to develop a concept for supporting uncertainty-aware exploratory change analysis, including appropriate methods for modeling, quantifying and communicating uncertainty in land cover change maps.

This research aims to address the challenges discussed above by developing a new concept for remote sensing change analysis tools. The approach first addresses the closer integration of the change detection and analysis steps and second, the incorporation of uncertainty during analysis. For both, the use of a geovisual analytics (GVA) approach is promising because GVA integrates algorithms with interactive visual interfaces to facilitate iterative workflows that are not equally possible in standard GIS.

The implementation of the concept requires the following steps: definition of a measure to quantify uncertainty in land cover (raster) change maps, identification of an uncertainty visualization technique that meets the requirements of such maps, and categorization of tasks during exploratory change analysis under uncertainty. Based on this, a software prototype is developed that serves not only as a proof-of-concept, but also as a vehicle for a later expert user study. The study assesses if and how expert users can incorporate uncertainty information into change analyses using the prototype. Findings from the study are anticipated to help assess the potential of this concept for its use in practice and, together with experience from the development process, help derive lessons learned and recommendations for the development of uncertainty-aware tools. The following subsections provide information on the background of this research and present four research questions and the methods used to address them.

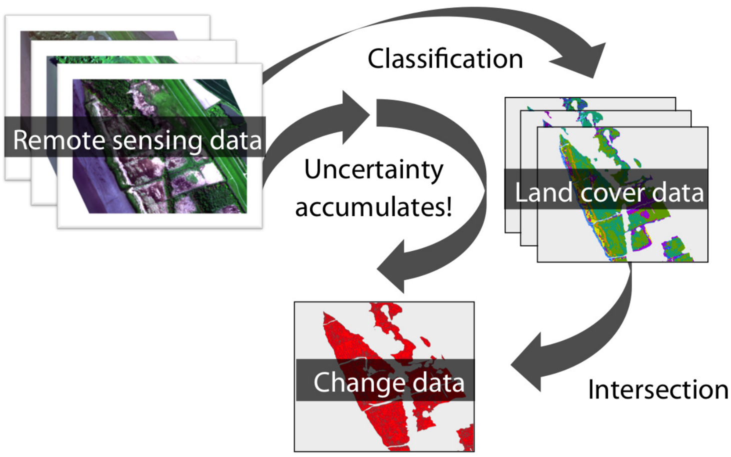

Figure 1.3. Post-classification change detection workflow. Uncertainty in the RS datasets accumulates and processing the data introduces further uncertainty.

1.1 Background

This work cuts across several fields of research, from uncertainty research and geovisual analytics, to remote sensing change analysis (Figure 1.4). The following subsections provide a brief introduction to the three fields.

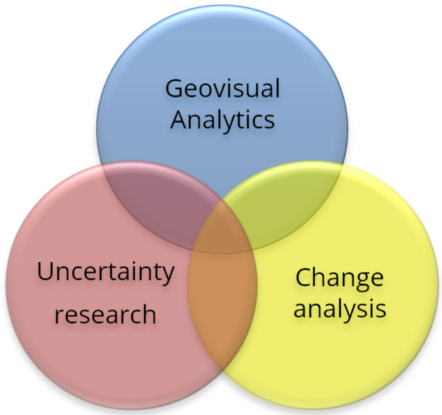

Figure 1.4. The three fields of research this dissertation relates to.

1.1.1 Uncertainty research

As discussed above, geodata refer to models of reality which inherently contain simplifications and abstractions. The reason is that the world can be considered as being infinitely complex and that mapping is a matter of interpretation. Thus, geodata may contain uncertainty of different types, stemming from various possible sources: "(a) inherent uncertainty in the real world, (b) the limitation of human knowledge in cognition of the real world, (c) the limitations of measurement technologies for obtaining spatial data, and (d) the potential of generating and propagating uncertainty in the spatial data processing and analysis, based on captured spatial data." (Shi 2010, p.8).

A number of terms are used to describe uncertainty related concepts, including 'reliability', 'data quality', 'data validity', or 'imperfection / imperfect knowledge'. Since the definition of these concepts is ambiguous, they are not further addressed here. Analogous to this, numerous definitions of the term 'uncertainty' can be found in literature. From a general point of view, uncertainty can be seen as "[t]he state of knowledge about a relationship between the world, and a statement about the world." (Goodchild and Case 2001, p.8). Here uncertainty may be defined as the 'gap' in knowledge which may exist between a statement we make about the world and how it really is. Relating specifically to geodata, van der Wel (2000) defines uncertainty as "a useful concept to express the inability to be confident of, and knowledgeable about the *truth value* of a particular data characteristic" (p. 45) and Zhang and Goodchild (2002) suggest that "uncertainty may be defined as a measure of the difference between the data and the meaning attached to the data by the current user." (p. 5). Thus, the term does not describe the difference between the data and an objective truth but between the data and their meaning as interpreted by the user. This means, and this is the basic agreement of most definitions, that uncertainty serves as an "umbrella term for errors, randomness, and vagueness" (Zhang and Goodchild 2002, p.7), and that the "[u]se of the term uncertainty seems to signal an appreciation of vagueness as well as randomness in geographical information, while error may send a misleading signal that true values are definable and retrievable" (Zhang and Goodchild 2002, p.7). This research follows the definition of uncertainty coined by Leyk et al. (2005) that "uncertainty in GIS is defined as the lack of knowledge about: (1) objects of the real world due to erroneous measurement, vague definitions and concepts or unknown and ambiguous meaning; (2) effects of transformations performed on the data; and (3) the latter's suitability for the intended application" (p. 294). Since we describe a lack of knowledge when depicting uncertainty, we never know to what extend this description may be correct; thus, there can be no 'optimal', 'correct', or 'true' uncertainty (Caers 2011).

Classified land cover data (that post-classification change detection is based on) are an example why the concept of uncertainty is needed. This type of data is related to land cover classes that are inherently vague and classifying a land cover object is usually ambiguous. Thus, there is no 'true' land cover class to compare to – the class definition is itself uncertain. Even if remotely sensed data is compared to reference data captured on-site, i.e., ground truthed, it is still a matter of interpretation as to what part of the landscape belongs to 'mixed forest', for example, and where exactly the forest's boundary is drawn (Fisher et al. 2006). Therefore, boundaries between land cover objects are inherently uncertain due to geometric inaccuracies, ambiguity in classification, and vagueness in class definitions (Figure 1.5).

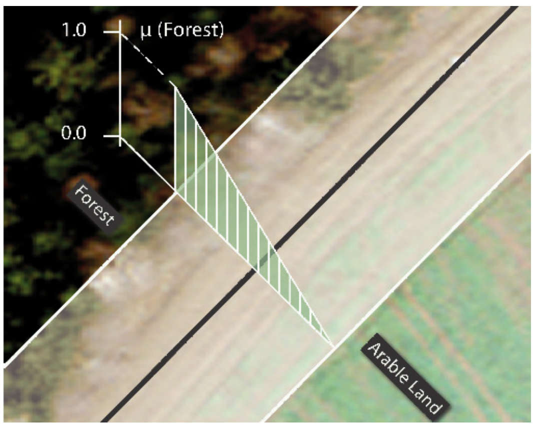

Figure 1.5. Uncertain boundary ('transition zone') between forest and arable land, described by a fuzzy membership function (Schiewe et al. 2009).

Historically, there has been a distinction between three types of uncertainty relating to the core information concepts from geodata modelling: attribute (also called thematic), positional (geometric / locational), and temporal. Combining data from a number of sources, uncertainty typologies have been created (Gahegan and Ehlers 2000, Thomson et al. 2005). These can serve as a conceptual basis for the choice or development of uncertainty models and visualization techniques for specific applications. An example is the typology shown in Table 1.1, presented by MacEachren et al. (2005), as extension of the one by Thomson et al. (2005). It is based on categories used for describing spatial data quality such as accuracy/error, precision, completeness, etc. (ISO 2006), complemented by the more abstract categories credibility, subjectivity, and interrelatedness. This can be read as uncertainty from *a lack of* accuracy, a lack of precision etc. For each category, examples for uncertainty regarding attribute, location, and time are provided, resulting in 27 types of uncertainty. This typology can serve as a framework for identifying, describing and visualizing different types of uncertainty, supporting the development of geodata analysis applications.

Table 1.1. Typology of uncertainty of geospatial information (MacEachren et al. 2005).

Category	**Attribute Examples**	**Location Examples**	**Time Examples**
Accuracy/error	counts, magnitudes	coordinates, buildings	+/- 1 day
Precision	nearest 1000	1 degree	once per day
Completeness	75% of people reporting	20% of photos flown	2004 daily/12 missing
Consistency	multiple classifiers	from / for a place	5 say Mon; 2 say Tues
Lineage	transformations	#/quality of input sources	# of steps
Currency	census data	age of maps	$C = T_{present} - T_{info}$
Credibility	U.S. analyst interpretation of financial records <…> informant report of financial transaction	direct observation of training camp <…> e-mail interception with reference to training camp	time series air photos indicating event time <…> anonymous call predicting event time
Subjectivity	fact <…> guess	local <…> outsider	expert <…> trainee
Interrelatedness	all info from same author	source proximity	time proximity

1.1.2 Geovisual analytics

As discussed in the beginning of this section, this research utilizes a geovisual analytics (GVA) approach to counter the challenges of remote sensing change analysis. GVA is an interdisciplinary field following a "new paradigm for how information technologies can be used to process complex geospatial information to facilitate decision making, problem solving, and insight into geographical situations" (De Chiara 2012, p. 23). It "integrates perspectives from Visual Analytics (grounded in Information and Scientific Visualization) and Geographic Information Science (growing particularly in work concerning geovisualization, geospatial semantics and knowledge management, geocomputation, and spatial analysis)" (Tomaszewski 2007, p. 174). The core principle is the "linkage of visual and computational methods and tools for extracting hypotheses and information from spatial data" (Schiewe 2013, p.126). In other words, GVA combines human capabilities such as vision and cognition with computation by the use of interactive visual interfaces and supports analytical reasoning for applications with which computation or visual interpretation alone cannot generate necessary insights from the data.

Visual Analytics is "the science of analytical reasoning supported by interactive visual interfaces" (Thomas and Cook 2005, p.4) which has successfully been applied in other fields such as health research (Wang et al. 2011) or financial analysis (Schreck et al. 2007), where neither automated nor manual analysis can provide the needed insight (Keim et al. 2010). Figure 1.6 depicts the relationship between cartography, geovisualization, and GVA after Schiewe (2013). In this interpretation, GVA is defined as an extension of Visual Analytics with its roots in human computer interaction (HCI), crossing the boundaries to cartography, information visualization, and scientific visualization.

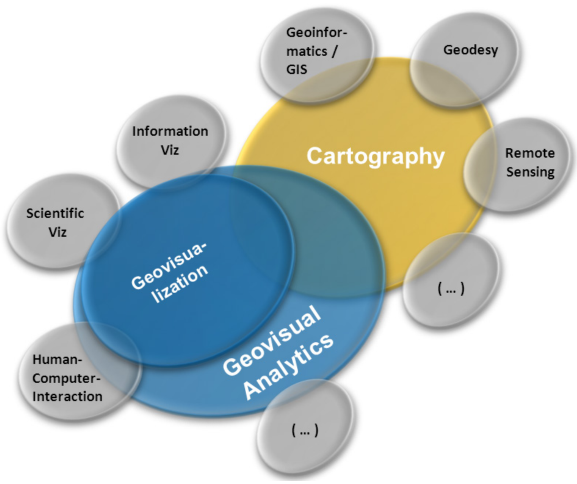

Figure 1.6. Schematic representation of the interplay of Geovisual Analytics and Cartography / Geovisualization (Schiewe 2013)

A general description of analysis workflows according to GVA is shown in Figure 1.7. They include iterative loops in which the user generates hypotheses and revises them on the basis of information communicated through visual interfaces (refer to chapter 2 for a more detailed discussion). Therefore, relating to the challenges in change analysis, GVA is a promising approach to establish a close integration of change detection and analysis as well as visually communicating uncertainty during analysis.

GVA workflows cannot easily be conducted in standard GIS that traditionally do not offer the high degree of interaction and responsiveness that is needed to establish highly interactive analysis workflows. However, there are GIS extensions utilizing multiple views and a high degree of interactivity for a special purpose, e.g., the Earth Trend Modeler[2], an IDRISI GIS toolbox for the analysis of remote sensing image time series. Extensions of this kind serve a specific application but do not give the user the possibility to freely define their own workflows. Apart from this, there are software packages and tools explicitly tailored to GVA, such as the V-Analytics[3] toolkit by Andrienko and Andrienko, also known as CommonGIS. It extends traditional GIS functionality by providing coordinated views and interactive query tools. Further examples for tools supporting (Geo-)Visual Analytics can be found in chapters 6.2 and 6.3.

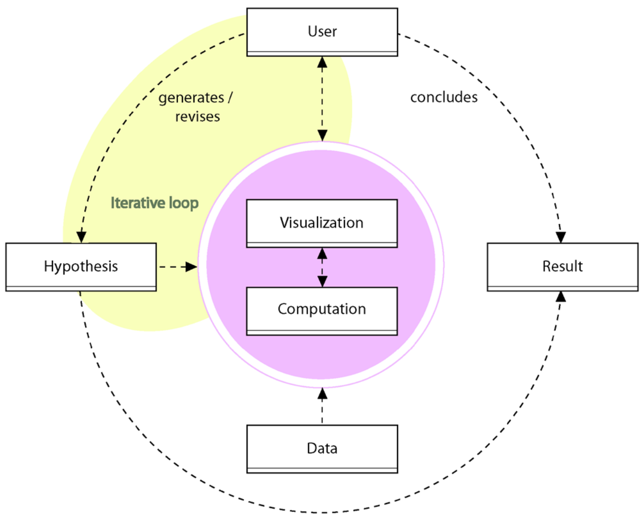

Figure 1.7. Schematic description of Geovisual Analytics workflows.

[2] http://www.clarklabs.org/products/Earth-Trends-Modeler.cfm

[3] http://geoanalytics.net/V-Analytics/

1.1.3 Land cover change analysis

This dissertation is concerned with uncertainty-aware exploratory analysis, focused on analysis of changes detected from remotely sensed imagery. According to the definition in the 'Encyclopedia of GIS' (Shekhar and Xiong 2008, p.77) change detection can be defined as "the process of identifying differences in the state of an object or phenomenon by observing it at different times" and is "usually applied to earth surface changes at two or more times". To perform change detection, a variety of methods exist from simple image differencing to post-classification and multidate methods (Coppin et al. 2004).

The term 'change analysis' depicts the analysis of change detection results. Here, the challenge is that changes detected often contain a high amount of false-positives, i.e., areas falsely detected as change. Change detection outputs are typically very sensitive towards geometric misregistration of the imagery or misclassifications. In sensitivity analyses of detected change, Pontius and Lippitt (2008) observed that about half of the observed difference could be explained by error whilst user accuracy of the single classified images was high. With multi-temporal detection (including more than two scenes) these negative effects can become even stronger because uncertainty can accumulate from scene to scene.

Traditionally in GIS, change detection and analysis are conducted as separate steps in a linear workflow (Figure 1.8). To counter the challenges described above, this research defines the goal of improving change analysis through a closer integration of the two steps. For this, a concept is developed based on GVA that makes iterative user-centered workflows possible (Figure 1.7).

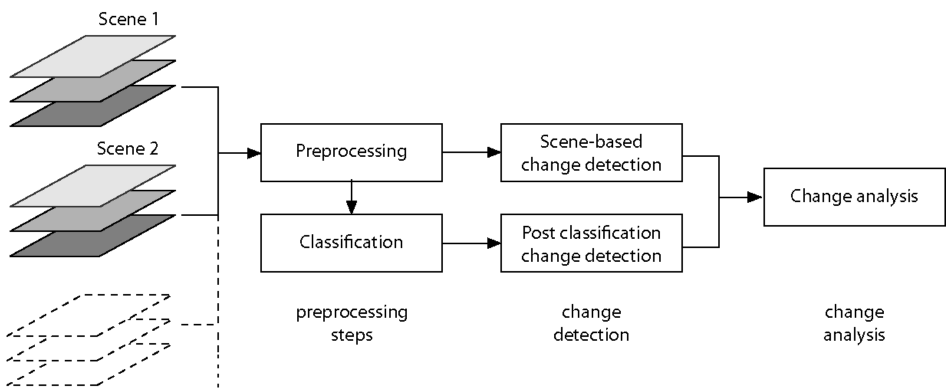

Figure 1.8. Common workflow including change detection and analysis as separate steps.

1.2 Research questions and methodology

The overall goal of this dissertation is to develop and assess a concept for uncertainty-aware exploratory change analysis with a focus on land cover change from RS data. Drawing from the research goal to develop a concept for uncertainty-aware change analysis using a GVA approach, four research questions (RQ1 to RQ4) are identified that will be addressed in six peer-reviewed journal papers. Each section is comprised of one or more articles contributing to the respective research question.

RQ1. How can information about uncertainty be incorporated into land cover change analysis?

The paper in chapter 2 addresses the first research question. The main hypothesis is that a concept using a GVA approach can help counter the challenges of exploratory change analysis identified above. Possible applications of the concept are highlighted: enabling better informed analysis, optimizing change detection parameters, and reducing false-positive change. The first step is to introduce a measure to describe uncertainty in land cover change. A simple prototype is developed that shows change detection results on a map and lets the user filter the changes by uncertainty thresholding. In a case study utilizing a change scenario in an area near Hamburg, Germany, the prototype is used to remove erroneous change using filtering by uncertainty. Potentials and limitations of the concept are discussed and recommendations for further development are given.

Chapter 2: Peer-reviewed journal paper
Kinkeldey, C. 2014. "A Concept for Uncertainty-Aware Analysis of Land Cover Change Using Geovisual Analytics." *ISPRS International Journal of Geo-Information* 3(3): 1122-1138.

RQ2. What is a suitable method for the visual representation of uncertainty in land cover change maps?

A comprehensive review of user evaluation in the field of uncertainty visualization is the first step to address this research question. The review consists of two parts, one with a focus on studies assessing communication aspects, i.e., how best to communicate uncertainty to the user (chapter 3). The second part is on studies about the effect that communicated uncertainty can have, for example, during decision making (chapter 4). Both review articles summarize lessons learned and give recommendations for future studies in the field.

Based on the findings from the review, an uncertainty visualization method called *noise annotation lines* is identified to meet the requirements of exploratory analysis in land cover change maps. However, its usability has not been assessed, thus, the paper in chapter 5 reports on two web-based user studies assessing the basic usability of the visualization technique. The studies compare different designs of the technique in terms of user accuracy through comparison between the levels of uncertainty in two areas. The number of levels is varied between four and eight to assess the limits of the technique with respect to visual discrimination.

Chapter 3: Peer-reviewed journal paper
Kinkeldey, C., A. M. MacEachren, and J. Schiewe. 2014. "How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies." *The Cartographic Journal* 51(4): 372–386.

Chapter 4: Peer-reviewed journal paper
Kinkeldey, C., A. M. MacEachren, M. Riveiro, and J. Schiewe. 2015. "Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations." *Cartography and Geographic Information Science*, published online.

Chapter 5: Peer-reviewed journal paper
Kinkeldey, C., J. Mason, A. Klippel, and J. Schiewe. 2014. "Evaluation of Noise Annotation Lines: Using Noise to Represent Thematic Uncertainty in Maps." *Cartography and Geographic Information Science* 41(5): 430–439.

RQ3. How can a software tool support uncertainty-aware land cover change analysis?

Using the concept from RQ1 and the visualization technique from RQ2, the development and implementation of a prototypical software tool (*ICchange*) are described in the paper included in chapter 6. It presents a list of elementary tasks that occur during exploratory change analysis. The prototype is designed to support these tasks with a focus on uncertainty visualization. In a user study, participants perform a number of tasks from the list to assess the usability of the prototype. We summarize the insights from the study and discuss the development of the prototype.

Chapter 6: Peer reviewed journal paper
Kinkeldey, C., 2014. "Development of a prototype for uncertainty-aware geovisual analytics of land cover change." *International Journal of Geographical Information Science* 28(10): 2076–2089.

RQ4. Does the developed concept help experts with land cover change analysis?

To answer this research question expert interviews are conducted that assess potentials and limitations of the concept. In three semi-structured interviews, experts answered questions about the following aspects: use of uncertainty for change analysis, reasoning a bout change under uncertainty, communication of uncertainty to users, as well as the *ICchange* prototype and *noise annotation lines*, the uncertainty visualization technique. The interviews were based upon change scenarios the experts were familiar with. The scenarios (+ uncertainty) were demonstrated using the software prototype introduced in chapter 6. Findings from the interviews are provided, along with a discussion of the methodology, and recommendations for the use of uncertainty visualization in the field of change analysis.

Chapter 7: Peer reviewed journal paper
Kinkeldey, C., J. Schiewe, H. Gerstmann, C. Götze, O. Kit, M. Lüdeke, H. Taubenböck, and M. Wurm. 2015. "Evaluating the use of uncertainty visualization for exploratory analysis of land cover change: A qualitative expert user study." *Computers & Geosciences* 84: 46–53.

1.3 Thesis Structure

The thesis is structured in eight chapters. This chapter provided an introduction to the research described in this thesis, chapters 2 to 7 consist of peer-reviewed journal articles addressing the four research questions. All six papers in chapters 2 to 6 have been accepted and published online and thus fulfill the formal requirements established by HafenCity University Hamburg leading to a cumulative doctoral dissertation. The last section summarizes the main findings, derives implications of this research for the fields of uncertainty visualization and exploratory change analysis, and provides recommendations for future work. Due to the nature of a cumulative dissertation, which incorporates unchanged original research papers, a certain amount of recurrence in this thesis is unavoidable. In some chapters, heading numbering and citation styles were changed to establish consistency throughout the document.

2 A Concept for Uncertainty-Aware Analysis of Land Cover Change Using Geovisual Analytics

This chapter was previously published as peer-reviewed journal paper:

Kinkeldey, C., 2014. "A Concept for Uncertainty-Aware Analysis of Land Cover Change Using Geovisual Analytics." *ISPRS International Journal of Geo-Information*. 2014, 3(3), p. 1122-1138.

The candidate is the sole author of this work.

Abstract

Analysis of land cover change is one of the major challenges in the remote sensing and GIS domain, especially when multi-temporal or multi-sensor analyses are conducted. One of the reasons is that errors and inaccuracies from multiple datasets (for instance caused by sensor bias or spatial misregistration) accumulate and can lead to a high amount of erroneous change. A promising approach to counter this challenge is to quantify and visualize uncertainty, i.e., to deal with imperfection instead of ignoring it. Currently, in GIS the incorporation of uncertainty into change analysis is not easily possible. We present a concept for uncertainty-aware change analysis using a geovisual analytics (GVA) approach. It is based on two main elements: first, closer integration of change detection and analysis steps; and second, visual communication of uncertainty during analysis. Potential benefits include better-informed change analysis, support for choosing change detection parameters and reduction of erroneous change by filtering. In a case study with a change scenario in an area near Hamburg, Germany, we demonstrate how erroneous change can be filtered out using uncertainty. For this, we implemented a software prototype according to the concept presented. We discuss the potential and limitations of the concept and provide recommendations for future work.

Keywords

remote sensing; change analysis; uncertainty; geovisualization; geovisual analytics

2.1 Introduction

Uncertainty is inherent in geospatial data and can have severe impacts on spatiotemporal analysis (Zhang and Goodchild 2002). However, it is still common to assume that data is error free although it has been shown that "[e]rror-laden data, used without consideration of their intrinsic uncertainty, are highly likely to lead to information of dubious value" (p.3). This is especially true for change detection and analysis since multiple remote sensing (RS) scenes are involved. Uncertainty accumulates over the scenes and further uncertainty is introduced, e.g., during the classification step in post-classification change detection. Thus, analysis of detected change has to cope with a high degree of uncertainty. Sensitivity analyses, performed by Pontius and Lippitt, demonstrated that half of all detected "changes" were caused by errors, although the overall accuracy for each of the input data sets was determined to be 91% Pontius and Lippitt 2008. This is one of the reasons why we see the need for a concept to incorporate uncertainty into change analysis instead of just ignoring it.

Traditionally, GIS are created to utilize precise geodata, which makes it difficult to use them for uncertainty-aware analysis. One reason for this is that uncertainty is stored as a separate data layer and not as an integrated part of a dataset. Another reason is that change detection and the analysis of changes are conducted as separate steps (Figure 2.1). In the past, much effort was devoted to enhancing the detection step, often with the goal of full automation (Coppin et al. 2004). We follow a different approach and suggest a closer integration of the change detection and analysis steps and the incorporation of uncertainty through visualization.

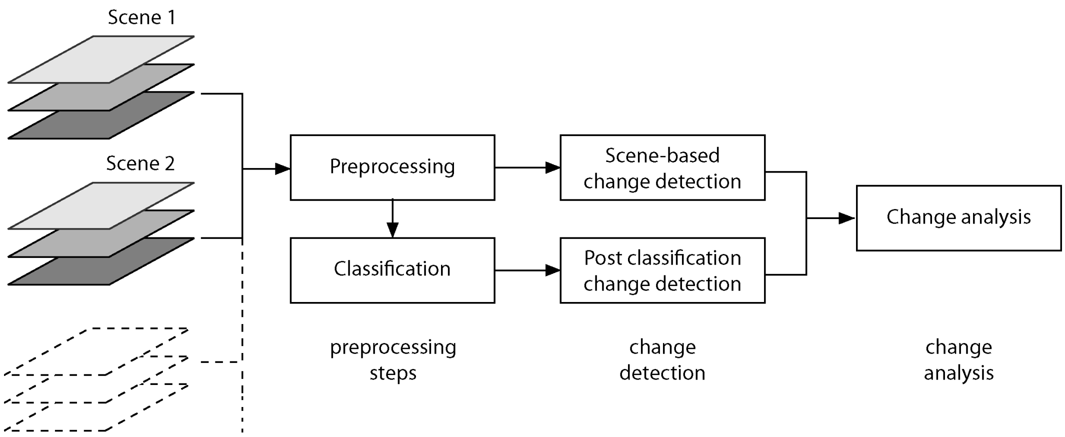

Figure 2.1. Common workflow including change detection and analysis as separate steps.

This article is structured as follows: we introduce the concept in Section 2.2, focusing on geovisual analytics and on a measure for change uncertainty. Possible applications are presented in Section 2.3. In Section 2.4, a case study shows how the concept can be implemented and discusses the potential benefits and drawbacks. From this, we derive a conclusion and provide an outlook on future work in Section 2.5.

2.2 Concept

In this section we present a concept to counter the challenges in change analysis described in Section 2.1. The basic idea is to facilitate uncertainty-aware change analysis by using iterative workflows with a high degree of user interaction. In the following subsections we introduce the concept following a geovisual analytics (GVA) approach and define an uncertainty measure for post-classification change.

2.2.1 Geovisual Analytics

We propose a geovisual analytics (GVA) approach to deal with the challenges discussed above. GVA is an interdisciplinary field following a “new paradigm for how information technologies can be used to process complex geospatial information to facilitate decision making, problem solving, and insight into geographical situations” (De Chiara 2012, p. 23). GVA “integrates perspectives from Visual Analytics (grounded in Information and Scientific Visualization) and Geographic Information Science (growing particularly in work concerning geovisualization, geospatial semantics and knowledge management, geocomputation, and spatial analysis)” (Tomaszewski et al. 2007, p. 174). By integrating algorithms and user input with the help of visual interfaces, GVA tools establish the “linkage of visual and computational methods and tools for extracting hypotheses and information from spatial data” (Schiewe 2013, p.126). GVA is based on Visual Analytics that has successfully been used in other fields such as health research (Wang et al. 2011) or financial analysis (Schreck et al. 2007), where neither automated nor manual analysis alone can provide the needed insight.

This research builds upon related work from two different categories: first, visual tools which make uncertainty usable in the analysis of RS data; and second, change analysis through visual analysis/analytics. Most work from the first category focuses on the classification of RS data, for example, a visual tool by Arko Lucieer named *Parbat* supports the optimization of segmentation parameters by visualizing uncertainty during classification (Lucieer 2004). Other work deals with the enhancement of class definitions, e.g., Ahlqvist used visualization of semantic similarity and overlap between class definitions to address incompatibilities of class definitions for land cover and land use (Ahlqvist 2008). Based on the work from the *FLIERS* project (Fuzzy Land Information from Environmental Remote Sensing), Bastin et al. developed a toolkit named

VTBeans to enhance the classification of RS data with the help of uncertainty (Bastin et al. 2002). Multiple linked views can be combined to visualize uncertainty from different sources and, for instance, explore fuzzy spectral signatures to enhance class definitions.

In the second category regarding visual change analysis from RS data, there have also been a number of promising approaches that serve as a basis for our work. Zurita-Milla et al. extended the GIS software package *ILWIS* (Integrated Land and Water Information System) by adding a toolbox called SITS (Satellite Image Time Series, Zurita-Milla et al. 2012). It facilitates the use of animation and interaction to analyze changes based on the imagery and provides filtering and aggregation functionality to conduct analyses at different levels of granularity. Another example is *Change Matters*, a web-based application for interactive change analysis of Landsat satellite imagery (Green 2011). A change overlay is displayed on top of the imagery and the user can modify thresholds, e.g., for vegetation gain and loss. Based on this, change maps can be created interactively and distributed over the Internet. An interesting approach for multi-temporal change analysis was presented by Hoeber et al. They use spatiotemporal difference graphs to display change over multiple time points in *GTDiff*, a system that facilitates interactive exploration of multi-temporal changes (Hoeber et al. 2010).

All in all, there has been substantial work regarding the use of visual analysis and GVA to either explore uncertainty in geospatial data or to analyze change visually without incorporating uncertainty. However, we see the need for a concept that integrates the two approaches, merging the strengths of GVA and uncertainty visualization. Thus, the first goal of this concept is to use GVA to integrate detection and analysis of change and the second goal is to use potential benefits of uncertainty by visually communicating this information during analysis.

For a systematic design and documentation of workflows we created a workflow description concept in the form of a graph. It contains the following categories (Figure 2.2):

- Data: The data used during analysis (e.g., RS imagery, GIS layers, etc.)
- User: User interaction (e.g., choosing a threshold)
- Hypothesis: A hypothesis about change (e.g., "this area was falsely detected as change")
- Computation: Computational steps in the workflow (e.g., classification of RS data)
- Visualization: Visual communication to the user (e.g., display of change uncertainty in a map)
- Result: The resulting change set

As already mentioned, the core of this concept is the combination of automated algorithms ("Computation") and visual interfaces ("Visualization"). In close interplay with these two components the user generates hypotheses, evaluates their plausibility with the help of visualization tools and revises them iteratively. When the user confirms a hypothesis the end results can be exported. Example workflow descriptions are provided in Section 2.3.

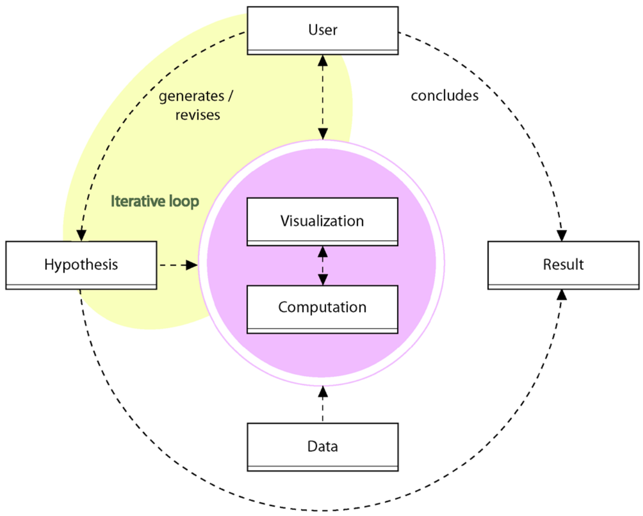

Figure 2.2. Basic concept of iterative analysis facilitating geovisual analytics.

2.2.2 Change Uncertainty

Change detection based on remotely sensed data typically involves a high degree of uncertainty. Ignoring this fact can make further analysis of changes questionable (Pontius and Lippitt 2008). During change detection, uncertainty from multiple RS scenes (errors and inaccuracies from sensor bias, misregistration, etc.) accumulate during the process, especially when multi-temporal (including more than two scenes) or multi-sensoral (using data from different sensors) detection is applied.

In the field of RS and GIS, research on modeling uncertainty in geospatial data has been conducted for decades. Different types of uncertainty (e.g., attribute, positional, and temporal) and their formal descriptions have been discussed in literature (Zhang and Goodchild 2002, Foody and Atkinson 2002, Shi et al. 2002). Substantial work regarding the description of uncertainty in land cover change was provided by Fisher and colleagues (2006). They extended the widely used change matrix to include fuzzy change values and provided a model for describing change uncertainty. Using a case study, they highlighted how fuzzy change data can reveal subtle changes that would not have been detectable with boolean change detection. At the same time they pointed out that "[w]hether the fuzzy mappings better reflect the landscape character than the Boolean remains an open question, and as is the reality of the differences noted here between the Boolean and fuzzy matrices of real change." (Fisher et al. 2006, p.176). Generally, there are at least three ways to quantify uncertainty in land cover change:

- from accuracy assessment (class-specific, Zhang and Goodchild 2002),
- from classification confidence, e.g., class membership probabilities (pixel-/object-specific, Brown et al. 2006), and,
- from expert knowledge, e.g., estimation of class similarities (class-specific, Lowry et al. 2008).

Since our goal is to depict geographically varying uncertainty, class-specific measures are not taken into account since they only provide one uniform uncertainty value for all instances (pixels or objects) of a change type. Thus, we focused on uncertainty measures that are quantified by classification confidence. Based on the work by Fisher and his colleagues, we defined a straightforward uncertainty measure for land cover change on the basis of fuzzy membership values. The intersection of membership values μ_i from each scene is conducted by applying the minimum operator (Fisher et al. 2006). The complement of the minimum membership value yields a value for change uncertainty:

$$u = 1.0 - \min(\mu_i) \text{ with } \mu_i \in [0.0, 1.0] \qquad (1)$$

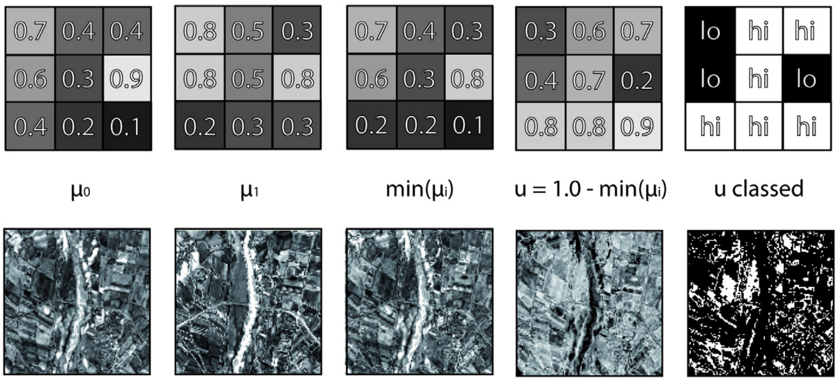

Figure 2.3. Example for change uncertainty measure (here: per-pixel and bi-temporal). Schematic (top row) and real data example (bottom row). Uncertainty is represented by a grayscale from black (0.0) to white (1.0).

This uncertainty measure ranges from 0.0 (no uncertainty) to 1.0 (maximum uncertainty). It is straightforward to compute and can be applied to different spatial units (pixels, objects) and to any number of scenes (two or more). It is a compound measure reflecting uncertainty from all scenes resulting from errors and inaccuracies in the imagery, vagueness in land cover class definitions, ambiguity in class determination during classification, etc. Figure 2.3 illustrates the computation of the measure for two RS datasets, 0 and 1, in a schematic way (top row) and with real data (bottom row). The first two columns show the membership values of each dataset followed by the minimum membership values from both datasets. Subtraction from 1.0 yields the magnitudes of uncertainty. An exemplary division into two classes, "high" and "low", is provided in the last column. This is an optional step, but in many cases unclassified uncertainty data may be too complex to interpret. A suitable number of uncertainty classes may depend on the task, yet for many applications two classes, i.e., "low uncertainty" and "high uncertainty", will suffice. An alternative is percentage steps, for instance, the use of the three classes 0%, 50%, and 100%.

2.3 Applications

As already mentioned, we see several potential benefits of incorporating uncertainty in change analysis. In the following we discuss three specific applications to illustrate this: enabling a better informed analysis of detected change, optimizing the parameters used for change detection, and filtering by uncertainty to reduce false-positive change.

2.3.1 Enable Better Informed Analysis

When a human analyst explores change detection results, he or she creates hypotheses about land cover change. For instance, when informal settlements have been detected from satellite imagery the analyst may come up with the hypothesis that a specific settlement has grown over time. However, informal settlements are usually hard to delineate and have fuzzy boundaries, therefore the exact size of a detected settlement cannot be determined. In such cases, change uncertainty can provide further information to help interpret the detected change. For instance, a detected informal settlement may show lower uncertainty in its interior (making this part more reliable) and higher uncertainty at its boundaries, where a transition zone towards formal urban settlements may exist. When the analyst wants to determine the area of a settlement (and the number of people living there) uncertainty can provide information about the range of possible locations of the area's boundary and thus can help derive a better estimate of the actual area. This way, information about uncertainty can make a hypothesis better defined and more plausible.

2.3.2 Optimize Change Detection Parameters

Another application of uncertainty that can potentially help with analyzing change is the optimization of the parameters used for its detection. With post-classification change detection, which we focus on here, the reliability of detected changes is highly dependent on the quality of the individual classifications on which the change detection step is based on. It was shown that misclassification can greatly impact change error (Burnicki 2011). Thus, often the only way to minimize change error is to increase the quality of the classified datasets. However, the optimization of classification parameters is not trivial and iterative modifications of the parameters are necessary to find a reasonable parameter set. To support this process, information on change uncertainty can help. The workflow depicted in Figure 2.4 shows that the user can visually explore a result after the initial classification of the single scenes and the change detection step. While iteratively modifying the parameters, the analyst gets immediate visual feedback on the resulting classified datasets and, thus, the detected changes. When the outcome is satisfying the current changes can be exported as final results. During the process, it is also possible to show the results for different parameter sets simultaneously so that the analyst is able to compare them and pick the best set.

This workflow is also applicable to other change detection approaches, e.g., when image differencing or ratioing are conducted, a threshold between change and no-change must be defined. In an analogous way, as described above, the choice of the change threshold can be optimized iteratively.

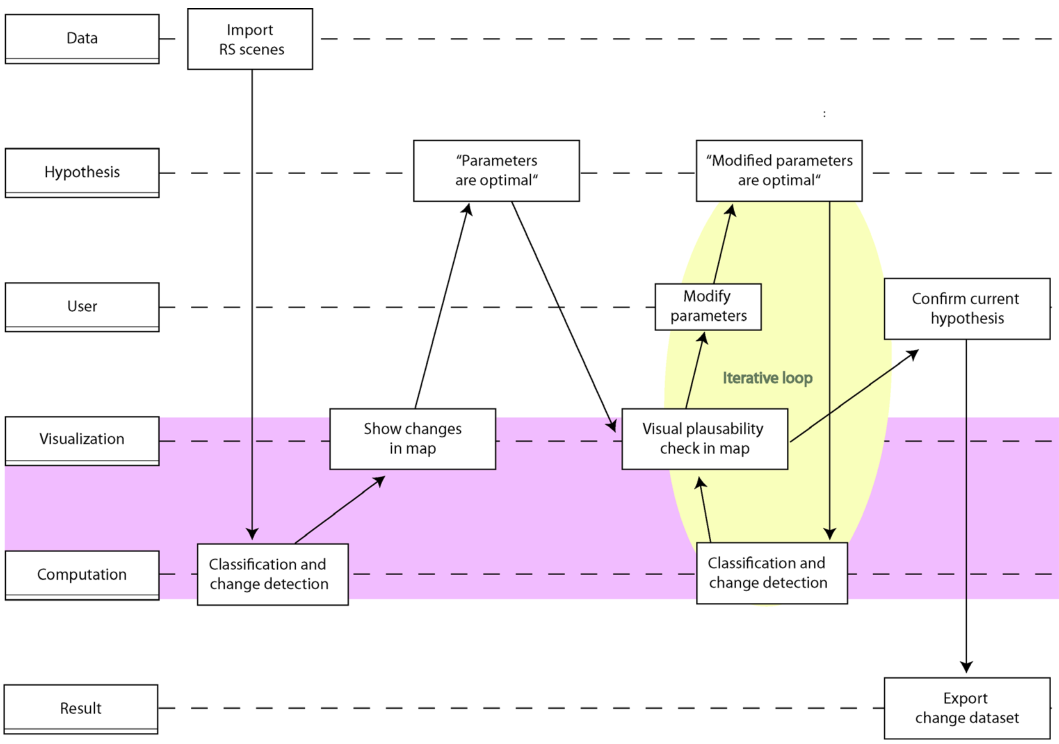

Figure 2.4. Workflow: Optimizing change parameters.

2.3.3 Reduce False-Positive Change

The third potential benefit of using change uncertainty lies in the reduction of falsely detected change. The basic idea is to use high uncertainty as an indicator for erroneous change so that filtering by uncertainty can improve the results. In the workflow presented in Figure 2.5, the user imports the result from post-classification change detection and the related uncertainty. The next step is to apply filtering so that only changes of a certain type remain. The reason is that it is more likely to find a suitable threshold for a single change type than for changes pertaining to all types. This subset of changes is now visualized along with the connected uncertainty. The initial hypothesis is that this is the optimal change set (in terms of correctness). Now, the user can apply a filter so that only changes with uncertainty less than a certain threshold (e.g., <80%) remain (which can be seen as a revised hypothesis). While changing the filter threshold the user gets immediate visual feedback to help him or her judge if the result has improved or not. When the user is satisfied with the result it can be exported as the final result.

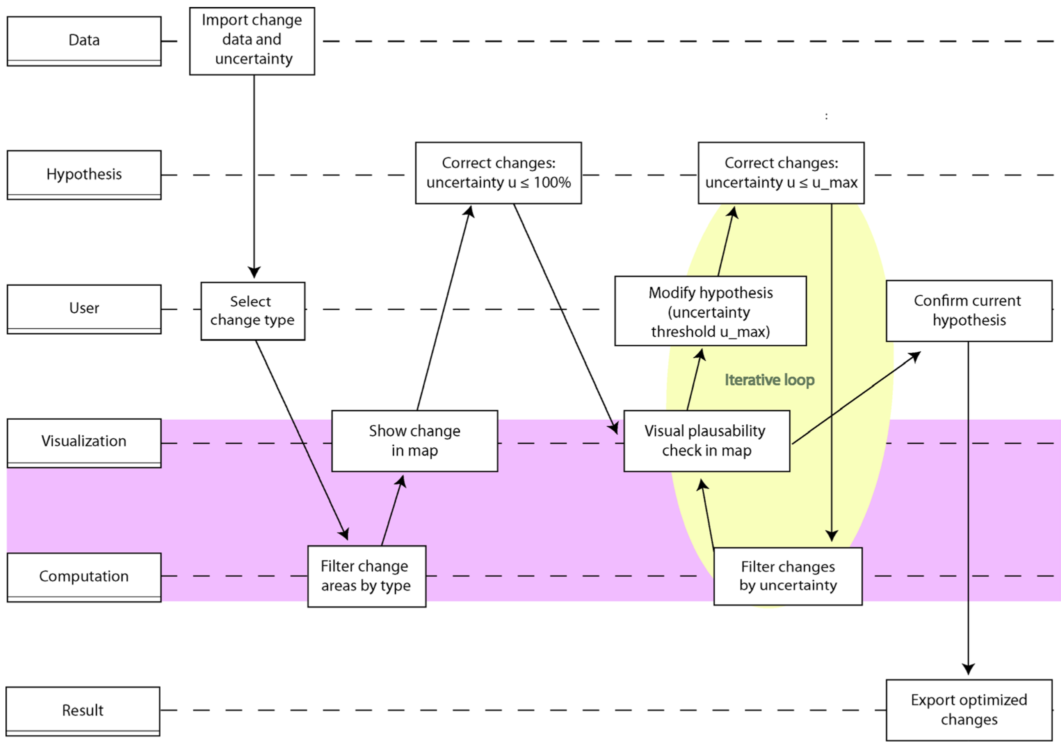

Figure 2.5. Workflow: Filtering change by uncertainty to reduce false-positive change.

2.4 Case Study

The following case study serves as proof-of-concept and highlights potential benefits and limitations of the approach for uncertainty-aware change analysis we present here. We created a change scenario based on two satellite scenes from an area in the vicinity of the Elbe River, northwest of Hamburg, Germany. The imagery was taken in 2010, on 5 June and on 16 July, by RapidEye, a satellite system that provides imagery with a geometric resolution of 5 m (pansharpened) and a temporal frequency of up to one image every 24 hours under ideal circumstances (http://www.blackbridge.com/rapideye/). The sensor covers five spectral channels, including the so-called "Red Edge" channel between red and near-infrared, giving the sensor potential advantages for vegetation mapping compared to other optical sensors (Schuster et al. 2012). The imagery we used here was radiometrically corrected and orthorectified (RapidEye Ortho-Level 3A). It was acquired via RapidEye Science Archive (RESA), which provides imagery free of cost for scientific use (http://resa.blackbridge.com/).

2.4.1 Change Detection

The area covered by the satellite imagery is a 25 km × 25 km rural agricultural area between Stade and Pinneberg in the vicinity of the Elbe River. Since the images were taken in summer and there were just six weeks between their acquisition we expected that most changes would be related to vegetation and water bodies (due to vegetation growth/harvesting and tidal changes). Therefore, we classified the two scenes separately using ISODATA unsupervised classification with eight classes, and aggregated them in both datasets yielding three classes: "water", "vegetated area" (i.e., vegetated arable land, meadows, forest, etc.), and "non-vegetated area" (settlements, roads, non-vegetated arable land, etc.). For both scenes most of the land cover is vegetated area, mainly consisting of agricultural crops, meadows, and forests (Figure 2.6). Non-vegetated area is comprised of bare soil or settlements and roads, but also riverbanks. Most of the area classified as water belongs to the river—a minor part consists of a few small water bodies.

From the aggregated dataset with the three classes described above, we created spectral signatures and rerun the classification using the Maximum Likelihood algorithm to get additional information about the membership values of each pixel. By intersecting the two classified datasets we created a change dataset containing changed areas and their type (Figure 2.7). On the basis of the class membership values we computed the change uncertainty measure (refer to Section 2.2). We removed a no-data area in the southeastern corner of the scene and filtered out all changes that occurred in areas smaller than the minimum mapping unit of 1000 m^2.

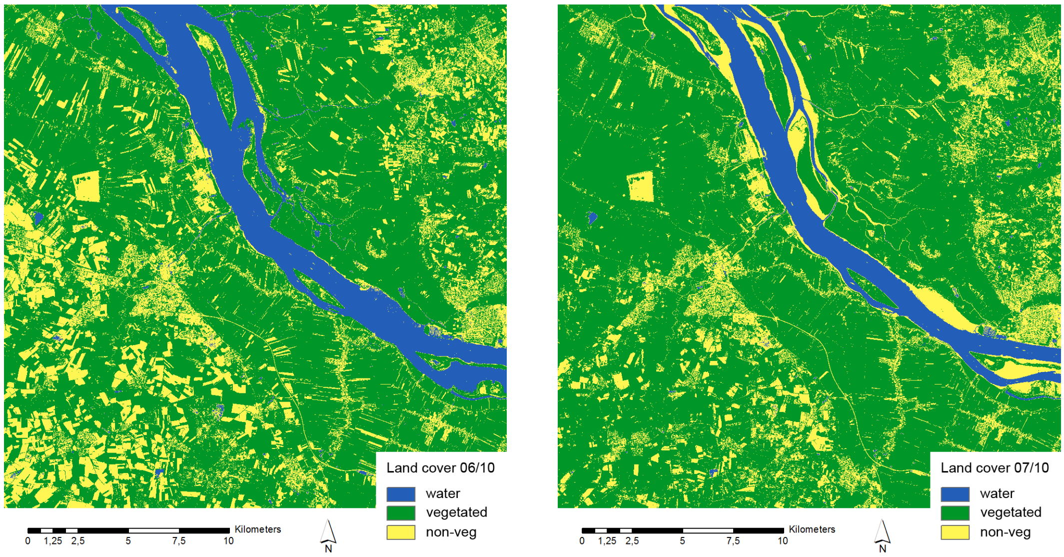

Figure 2.6. The two classified datasets from RapidEye imagery we used in this case study.

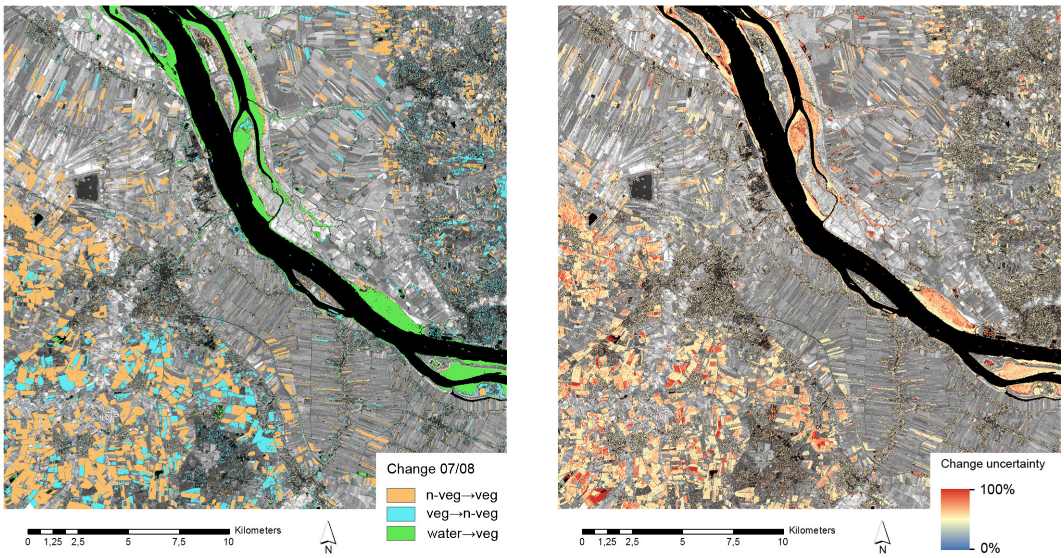

Figure 2.7. Change dataset (left) and related uncertainty (right).

2.4.2 Change Analysis

When analyzing the result, the first observation is that 22% of the study area was detected as having changed. Most changes belong to the type "non-vegetated to vegetated" (58% of the changed area), followed by "vegetated to non-vegetated" (28.5%), and "water to non-vegetated" (13%). There was also a very small amount of change from water to vegetated area (about 0.5%)

that we did not consider in this analysis. Generally, the detected change types and their resulting area proportions seem plausible considering the two scenes are from June and July, when vegetation growth and harvesting of some crop types is common and therefore may explain most of the detected change.

In order to assess the accuracy of the detected change we conducted a visual assessment with stratified random point sampling over the whole area. We determined the number of points after Tortura's method of approximation (Khorram 1999). With an estimated average accuracy of 70% and an error tolerance of 5%, we obtained a required sample size of 323 points. We focused on false positive changes, i.e., changes that have been falsely detected, since they are a common challenge in change analysis (see Section 2.1). Thus, we only assessed the areas detected as changes and not the no-change areas (false-negatives). Reference data from a different source was not available for the exact two dates so we visually examined each of the points based on the available imagery and estimated the correctness of the change in each position. Consequently, since the assessment conducted here does not rely on independent data, it does not fulfill the requirements of a statistically sound accuracy assessment. However, for the purpose here we see it as sufficient in order to get an impression about the quality of the change set. Table 2.1 shows the change types, their proportions of the overall area, and their corresponding number of sample points. The last two columns contain the result from the visual assessment.

Table 2.1. Stratified point sampling and results of the visual assessment of change correctness.

Change type	**Proportion of overall area**	**Number of sample points**	**Correct change**	**Erroneous change**
All	100%	323	-	-
No change	77.7%	251	-	-
Water to non-vegetated	2.9%	9	66.7%	33.3%
Vegetated to non-vegetated	6.3%	21	80.9%	19.1%
Non-vegetated to vegetated	12.9%	42	85.7%	14.3%

In the following we focused on the change type "water to non-vegetated area" because it showed the highest amount of erroneous change (33.3%). The high error rate is explainable by the uncertainty in separating water bodies from bare areas. The changes from water to non-vegetated areas are due to the different water levels of the river at the two dates. Regarding change uncertainty, the first observation was that all changes of this type that did not occur directly along the bank of the river were highly uncertain. Some small areas of this nature were detected. This corresponds to the assumption that changes from water to non-vegetated areas would only occur along the riverbanks. Small water bodies could potentially dry out in summer, but since this is unlikely in such a short period of time (at least for the area under research) we hypothesize that these small area changes were caused by misclassification. A second observation is that most areas situated directly at the edge of the water seem to be homogeneously certain and become more heterogeneous the further away from the water's edge they are. Figure 2.8 shows an exemplary change area at the eastern part of the river. The distribution of uncertainty seems plausible, because directly at the river bed the change is more likely to have happened than when we move further away from the water's edge. It is obvious that the boundary between the riverbed (belonging to the class "water") and mud flats on the banks (belonging to the class "non-vegetated areas") plays an important role for this change type and is highly uncertain at the same time.

For a closer look at the changes from water to vegetated areas we generated 41 additional random sample points over the area covered by this change type to attain 50 random points (Figure 2.9), which follows the rule of thumb suggested by Congalton and Green (2009). Again, we conducted a visual assessment without independent reference data. The result was that the detected changes at 72% of the points were correct, while 28% were incorrect due to misclassification. For most applications this level of accuracy does not seem acceptable and the change dataset would have to be improved. In the following, we would like to demonstrate how information on change uncertainty can help increase the accuracy of the change dataset.

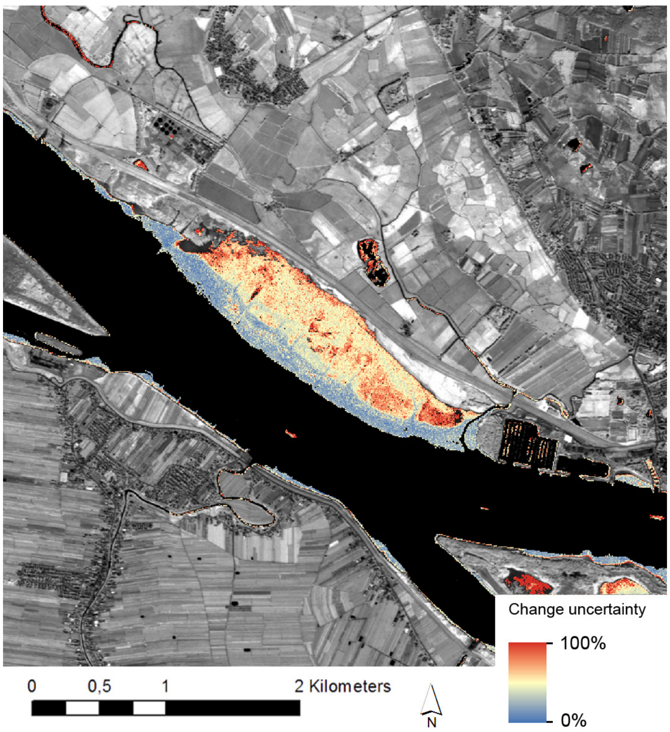

Figure 2.8. Change uncertainty for "water to non-vegetated area".

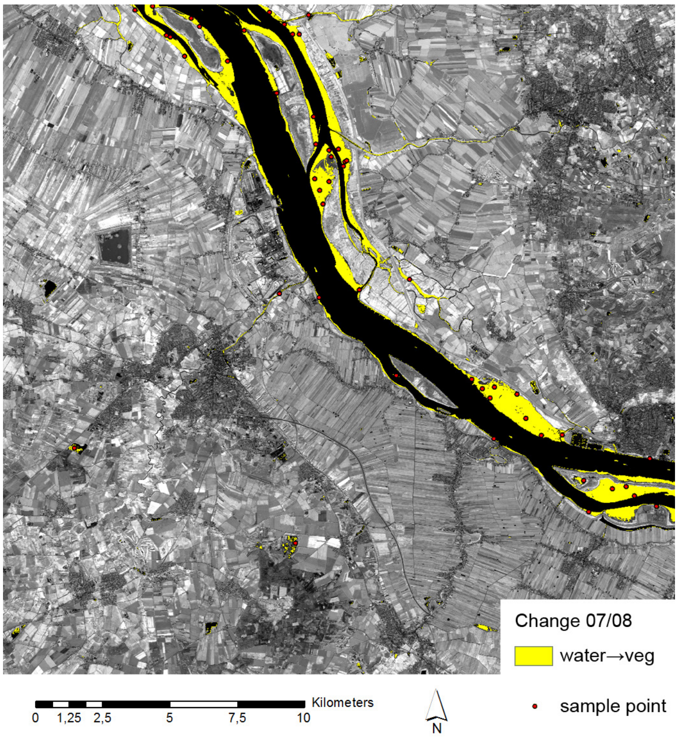

Figure 2.9. Sample points (red) in the area of change from water to non-vegetated area (yellow).

2.4.3 Reduce False-Positive Change

In Section 2.3 we hypothesized that the accuracy of detected change could be increased if we filter out uncertain change by applying a suitable threshold. We tested this hypothesis using our change scenario. For this, we utilized a simple prototype including a map client and a slider to filter change by uncertainty (Figure 2.10). The prototype is written in Java and is based on the geotools map client (http://www.geotools.org/) that provides standard map functionality (pan, zoom, etc.) under a free and open source license. The RS imagery is shown in the background and yellow pixels on the map represent changes. Change uncertainty is visualized using a color scheme from blue (0% uncertainty) to yellow (50%) to red (100%). The slider at the bottom of the window can be used to interactively filter the changes by uncertainty, starting with an initial threshold of 100% (meaning that no changes are filtered out).

Figure 2.10. Software prototype for iterative filtering by uncertainty.

In order to show how erroneous change can be filtered out by thresholding we chose an area of 5 km × 5 km with a large zone of change from water to non-vegetated area (Figure 2.11). To get started we lowered the threshold so that changed pixels with an uncertainty of more than 50% were filtered out. A visual check showed that this removed a number of questionable areas of change, but not all of them. After modifying the threshold several times and visually checking the outcome, we found that setting the uncertainty threshold to 40% filtered out most of the outliers we had identified on the map, clearly improving the result compared to the initial state. While the threshold was only determined for the small area presented here, we were interested to see if it would work for the whole scene. After taking a closer look at our sample points, it was revealed that all 17 misclassified points were filtered out, i.e., all erroneous change could be eliminated through simple thresholding. Furthermore, only one point with a correct change was (falsely) removed. This shows that a threshold chosen for a local area can work for a larger area such as a whole satellite scene. But it is likely that this does not hold true with scenes that show greater spectral variation as those we used here.

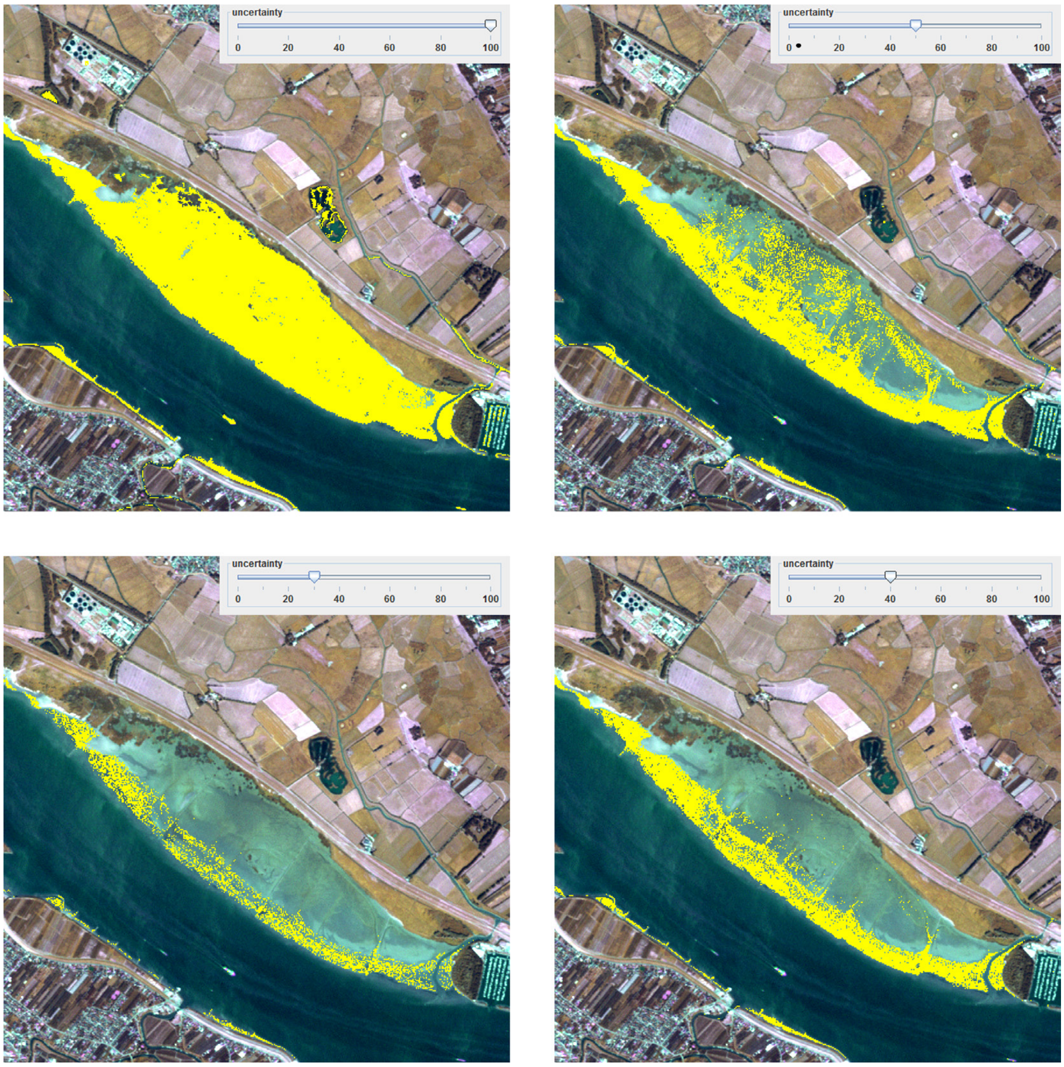

Figure 2.11. Iterative filtering of change by 100% (upper left), 50% (upper right), 30% (lower left), and 40% (lower right) uncertainty.

2.4.4 Discussion

The case study presented here highlighted the potential of uncertainty-aware change analysis using a GVA approach. First, it could be shown that better-informed change analysis is possible and that uncertainty can provide valuable information during the interpretation of detected change. Apart from this, as shown in the case study, uncertainty can serve as an indicator for erroneous change and can be used for filtering. It became clear that it is not a fully reliable indicator, but that it can serve as a criterion when deciding whether a change is correctly detected or not. Further criteria should be taken into account to validate detected change, e.g., the size or spectral signature of a changed area. It became clear that a suitable threshold can be determined locally, e.g., for a single area of change, but that one overall threshold for a whole scene (although it is only for one change type) often does not seem realistic, depending on the change type and the variability in the imagery. One downside of this concept that we mentioned is that by including uncertainty into the analysis and thus adding another data dimension, change analysis becomes more complex. This stresses the importance of minimizing the complexity, for instance by keeping the visual load for the user low.

In this case study we performed an analysis with the help of a simple prototype. In comparison to standard GIS, it facilitates the straightforward modification of the filter threshold and provides immediate feedback on the map. The simple tool we used here makes iterative analysis more fluent and more intuitive—however, this assumption has not yet been assessed in user studies. We are convinced that with more sophisticated workflows, e.g., when more information about a change is visualized (for example, area, spectral signature, and uncertainty), the advantages compared to common GIS analysis will become more apparent.

2.5 Conclusions

In this article we presented a concept to enhance the analysis of change derived from remote sensing (RS) data. The idea was to incorporate information about uncertainty concerning detected changes into the analysis. We suggested a geovisual analytics (GVA) approach that combines manual and automated analysis with the help of visual interfaces. This contributes to a better integration of change detection and analysis as well as to an enhanced visual communication of uncertainty during the analysis. We defined a measure for change uncertainty and identified potential applications of the concept. In a case study we used a simple software prototype that showed changes on a map and provided filtering by uncertainty. The study included a bi-temporal change scenario with RS data in the vicinity of the Elbe River near Hamburg, Germany, and showed how false positive changes could be successfully filtered out by the use of uncertainty thresholding. We pointed out that generally, uncertainty can be an indicator for erroneous change, however, it has limitations since the critical level of uncertainty (for the distinction between correct and erroneous change) may vary within a dataset. Thus, we recommended the use of other additional indicators to help decide whether changes are correct, e.g., the area of a change or its spectral signature.

All in all, it was shown that tools implementing the concept have the potential to counter challenges in change analysis such as the high amount of false positive changes. At the same time, analysis naturally becomes more complex since uncertainty adds another dimension of data that has to be taken into account. This fact stresses the importance of well-crafted visual interfaces and interaction functionality to minimize user burden. User studies will be necessary to evaluate if analysts can use uncertainty information when it is visually depicted and how they cope with the complexity and visual load that is added when incorporating uncertainty.

The case study we presented here involved a bi-temporal analysis of change. But the concept allows analysis of more than two RS datasets at a time and we hypothesize that this is one of the strengths of the approach. This will have to be tested in future studies involving more than two RS scenes.

The software prototype we used in the case study shall serve as a starting point for more complex change analysis tools of this kind. Thus, an important part of future work will be the derivation of guidelines and recommendations to support the development of change analysis tools based on this concept. These should include recommended techniques for uncertainty visualization and user interaction. All in all, we see the support for GVA tool development as a crucial step to get closer to the goal of establishing tools for uncertainty-aware analysis of change that can be used in practice.

Acknowledgments

This work was partly funded by the research program "KLIWAS—Impacts of climate change on waterways and navigation" financed by the German Federal Ministry of Transport, Building and Urban Development. Helpful comments on the article provided by Jochen Schiewe (HafenCity University) were greatly appreciated. Maya Donelson (HafenCity University) has been of great help proof-reading the article.

Conflicts of Interest

The authors declare no conflict of interest.

3 How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies

This chapter was previously published as peer-reviewed journal paper:

Kinkeldey, C., MacEachren, A.M., and Schiewe, J., 2014. How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies. *The Cartographic Journal* 51(4), 372–386. doi:10.1179/1743277414Y.0000000099

The candidate was the primary author and responsible for the systematic review of the papers, the summary of the findings, and contributed to lessons learned and recommendations (70% of the overall work).

Abstract

For decades, uncertainty visualisation has attracted attention in disciplines such as cartography and geographic visualisation, scientific visualisation and information visualisation. Most of this research deals with the development of new approaches to depict uncertainty visually; only a small part is concerned with empirical evaluation of such techniques. This systematic review aims to summarize past user studies and describe their characteristics and findings, focusing on the field of geographic visualisation and cartography and thus on displays containing geospatial uncertainty. From a discussion of the main findings, we derive lessons learned and recommendations for future evaluation in the field of uncertainty visualisation. We highlight the importance of user tasks for successful solutions and recommend moving towards task-centered typologies to support systematic evaluation in the field of uncertainty visualisation.

Keywords

uncertainty, geovisualisation, information visualization, scientific visualization, evaluation, user studies

3.1 Introduction

All geospatial data contain uncertainty and ignoring this fact can have severe consequences for spatial analysis and decision making (Zhang and Goodchild, 2002). Past research has suggested that communicating information about data uncertainty has the potential to increase trust in the results when analyses are conducted (Fisher et al., 2012) and to support decision making that uses the data (Aerts et al., 2003; Deitrick and Edsall, 2006; Leitner and Buttenfield, 2000). Visualisation of uncertainty has attracted substantial attention over more than two decades. Much of the work has focused on developing typologies of uncertainty that represent various aspects of data and how it might be signified (Buttenfield and Weibel, 1988; Pang et al., 1997; Sanyal et al., 2009; Thomson et al., 2005) and on developing methods to depict uncertainty visually (e.g. Cedilnik and Rheingans, 2000; Ehlschlaeger et al., 1997; Sanyal et al., 2010; Wittenbrink et al., 1996). A comprehensive review of uncertainty typologies is provided by MacEachren et al. (2005) and a review of uncertainty visualisation across science by Brodlie et al. (2012). From the broad literature, five common dichotomous categories for uncertainty visualisation can be identified:

- explicit/implicit

  This category distinguishes between directly expressing uncertainty, e.g. using glyphs that signify levels of uncertainty (explicit) or signifying it indirectly, e.g. through multiple visualisations showing different possible outcomes (implicit) (Deitrick, 2012). Explicit depiction of uncertainty graphically is most common. Implicit uncertainty depiction is given less attention in the uncertainty visualisation literature.

- intrinsic/extrinsic

  This commonly used distinction was introduced by both Howard and MacEachren (1996) and Gershon (1998) in cartographic and information visualisation contexts, respectively. Intrinsic techniques alter the existing symbology to represent uncertainty, basically through manipulation of visual variables, e.g. colour value. In contrast to this, extrinsic approaches add new objects to the display to depict uncertainty, e.g. glyphs or grids.

- visually integral/separable

  The third dichotomy focuses on the visual cognitive response of the viewer: A visually integral signification of uncertainty cannot be perceptually separated from the data signification while a visually separable signification can be read independently (MacEachren et al., 1998). These categories show some overlap with intrinsic/extrinsic because intrinsic methods tend to be visually integral and extrinsic ones visually separable, but there are exceptions.

- coincident/adjacent

This categorisation refers to view organisation, i.e. if data and uncertainty are represented in an integrated view (*coincident*) or in separate views (*adjacent*) (MacEachren, 1992).

- static/dynamic

The distinction here is between a classical *static* map versus a *dynamic* map using animation and/or interactive controls. One example of the latter is dynamic alternation in which a map depicting the data is alternated with one depicting data uncertainty, often with user control ('toggling').

Based on these categories, we propose a structure to describe uncertainty visualisation approaches in a systematic way using the following three main dichotomies:

- coincident/adjacent;
- intrinsic/extrinsic; and
- static/dynamic.

We represent these three dichotomies as axes of an *Uncertainty Visualisation cube* (UVis³, Figure 3.1). We left out 'explicit/implicit' and 'integral/separable' from the list, since most approaches are explicit. And, as already discussed, 'visually integral/separable' on the one hand corresponds to 'intrinsic/extrinsic' in most cases and on the other hand is a distinction focused on human visual processing rather than signification. Thus, the cube distinguishes eight main combinations that help us to discuss uncertainty visualisation approaches in a systematic way in the remainder of the paper.

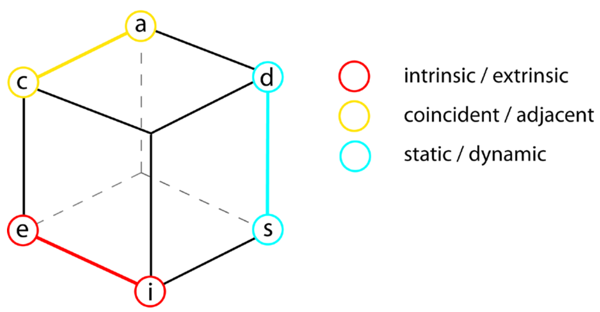

Figure 3.1. UVis³ ('Uncertainty Visualisation cube') for categorisation of uncertainty signification in visualisations

As the range of methods for signifying uncertainty has grown, so has the need to assess their usability and applicability for various use contexts. This need has stimulated a range of studies focused on evaluating uncertainty visualisation methods, but the studies have been idiosyncratic, thus difficult to compare and develop generalisations from. To address this gap, we present a review of user studies directed to evaluation of uncertainty visualisation methods and tools. The first step is to give an overview of past research and summarize the state of the art in uncertainty visualisation assessment. In the second step, after a critical review of existing studies and their findings, we derive lessons learned and recommendations for future work. Our focus lies on user studies from the early 1990s until the present that involve uncertainty visualisation in the domains of cartography and geovisualisation, scientific visualisation and information visualisation. We concentrate on visualisation of geospatial uncertainty, thus, this review is more comprehensive for studies that include some geographic component than those that are primarily aspatial. Apart from that, we focus on empirical studies involving users, typically using metrics such as map reading accuracy and speed or user confidence. Studies that do not appear in our list include, e.g. conceptual evaluations that use heuristics derived from general guidelines and rules for data visualisation coined by Bertin, Tufte, Ware and Chambers (Riveiro, 2007a; Riveiro, 2007b; Wittenbrink et al., 1996; Zuk and Carpendale, 2006). Such studies, while not our focus here, do provide basic statements on the usability of different methods and therefore can help to choose suitable visualisation techniques. Another type of study we did not consider are case studies (without users involved) that demonstrate the basic usability and utility of a method (e.g. Allendes Osorio and Brodlie, 2008; Dooley and Lavin, 2007), sometimes in connection with assessing display performance (Rhodes et al., 2003).

Owing to the volume of literature, differences in goals and diversity of methods, this paper reviews and analyses studies focused on communicating uncertainty (together or separately from communicating data) and not on those that investigate uncertainty visualisation impacts on reasoning and decision making (which we will address in a follow up paper). Some studies contribute to both aspects and will thus appear in both papers. This paper is organized as follows: In the section on 'Analysis of the literature', we describe the methodology of the review and outline the main characteristics of the studies. This includes the distribution of studies over the years, the types of uncertainty and the visualisation techniques that have been assessed, the kind of application domains the studies deal with, the groups of participants that were involved and the tasks that were conducted. In the section on 'Discussion of findings', we summarize and discuss the main findings of the studies, organized using the UVis³ we introduced above. From the discussion, we derive lessons learned and identify open questions. In the conclusion, we summarize the main findings, discuss the limitations of the review and suggest future directions for evaluation of uncertainty visualisation.

3.2 Analysis of the literature

In this section, we analyse the main characteristics of the user studies we included in the review; we have identified 44 studies described in 34 publications. Every sub-study that involved a different group of subjects was treated as a separate study, e.g. the Boukhelifa et al.'s (2012) paper contained six studies. We analysed each study by summarizing its main characteristics such as the study methodology, the visualisation techniques used, the data and scenario that were involved and the reported findings. Based on this description, we were able to make comparisons and thus identified commonalities and differences.

We have included study reports published between 1992 and 2014. Over the years, there has been an increase in the number of studies that evaluate (rather than just categorize or suggest methods for) uncertainty visualisation, but the numbers are generally small with considerable fluctuation (Figure 3.2). As shown in Figure 3.3, the majority of studies apply quantitative methods (38 out of 44), i.e. controlled lab experiments (22) and web-based experiments (16). Web-based experiments have emerged since 2002 and their number has become comparable to traditional laboratory experiments (16 web-based versus 15 lab studies from 2002 to 2014). Just a small fraction of studies (6 of 44) are based solely on qualitative methods, i.e. interviews (4) and focus groups (2).

In the remainder of this section, we analyse five specific aspects of the studies: uncertainty categories, visualisation techniques, application domains, participants and tasks, organized in subsections.

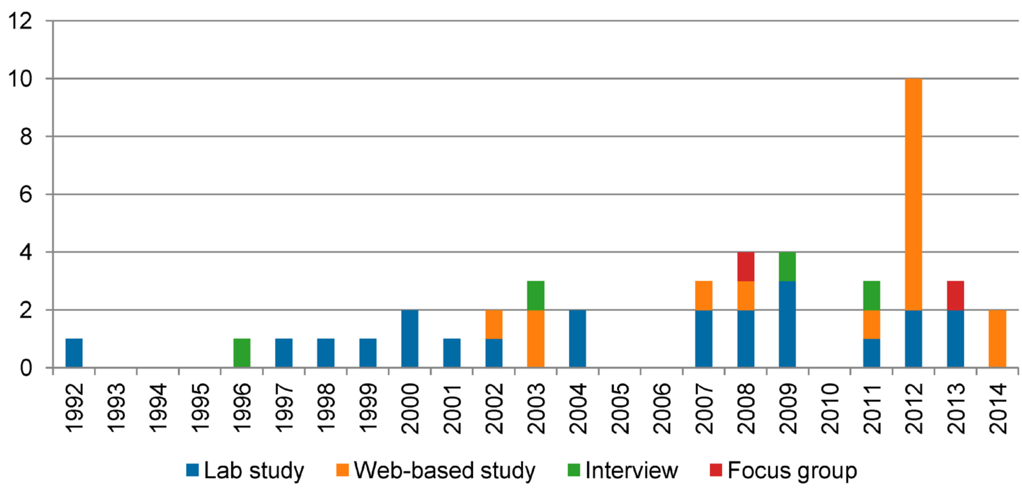

Figure 3.2. Number of studies over time separated by their type. The peak in 2012 is due to the six studies from the Boukhelifa et al.'s (2012) paper.

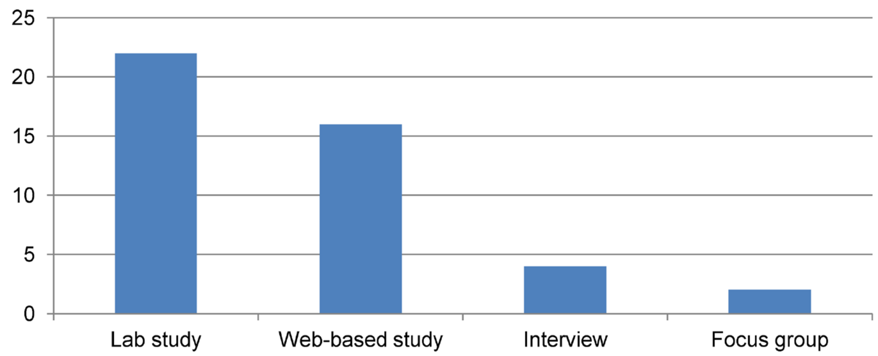

Figure 3.3. Number of studies per type.

3.2.1 Uncertainty categories

When uncertainty of geospatial data is depicted, it can be quantified and represented for each of the three core information components: attribute (what), positional (where) and temporal (when) uncertainty (MacEachren et al., 2005). Studies dealing with uncertainty information mainly cover the first component, attribute (also called thematic) uncertainty. Typically, they relate to uncertain model outputs (Aerts et al., 2003; Alberti, 2013), classification uncertainty (Blenkinsop et al., 2000; Drecki, 2002; Kinkeldey et al., 2014b) or reliability (e.g. variance) of statistical data (MacEachren et al., 1998). Just one study (at least from those focusing on communication of uncertainty) deals with positional uncertainty visualisation exclusively (Grigoryan and Rheingans, 2004). There are no studies in the set assembled here that solely deal with temporal aspects of uncertainty. Less than a fourth of the studies (10/44) involve multiple types of uncertainty. These studies fall into two categories: either different types of uncertainty are evaluated separately, such as the symbol sets for attribute, positional, and temporal and uncertainty in MacEachren et al. (2012). Or a combination of types is involved in the evaluation, as in Kardos et al. (2003; 2007; 2008) or Zhang et al. (2008) with a focus on attribute and positional aspects of uncertainty at the same time.

3.2.2 Visualisation techniques

One of the basic criteria for the selection of studies for this review was that the studies reported upon involve a visual representation of uncertainty. To systematize the approaches used in the studies, we use the three dichotomies from the UVis[3] (refer to the introduction).

Intrinsic/extrinsic

Most of the studies involve intrinsic approaches, i.e. when visual variables of existing map content are manipulated to represent uncertainty. Examples include colour hue and colour value (a term under which we subsume value, lightness and brightness) (e.g. Aerts et al., 2003; Edwards and Nelson, 2001; Leitner and Buttenfield, 2000; MacEachren et al., 1998; Nadav-Greenberg et al., 2008; Retchless, 2012; Schweizer and Goodchild, 1992; Slocum et al., 2003), transparency (e.g. Drecki, 2002; Newman and Lee, 2004; Slocum et al., 2003; Viard et al., 2011), or colour saturation that is used in a number of studies (e.g. Drecki, 2002; Kubíček and Šašinka, 2011; Kunz et al., 2011; Leitner and Buttenfield, 2000; Retchless, 2012; Sanyal et al., 2009).

While in the minority, we also identified multiple studies that focus on extrinsic techniques, i.e. when additional graphical objects are used to represent uncertainty, typically approaches using glyphs or error bars (Alberti, 2013; Drecki 2002; Sanyal et al., 2009; Slocum et al., 2003), grid-based techniques (Kardos et al., 2007; 2008; Kinkeldey et al., 2014b) or contouring (Senaratne et al., 2012).

Coincident/adjacent

Starting with the earliest research on uncertainty visualisation, coincident approaches (with data and uncertainty integrated in the existing display) have been contrasted with adjacent approaches with data and uncertainty in separate views (MacEachren, 1992). While most studies assess coincident approaches, there are a number of studies that involve a direct comparison between adjacent and coincident views (Aerts et al., 2003; Edwards and Nelson, 2001; Evans, 1997; Gerharz and Pebesma, 2009; Kardos, 2003; Kardos, 2007; Kubíček and Šašinka, 2011; Kunz et al., 2011; MacEachren et al., 1998; Retchless, 2012; Senaratne et al., 2012; Viard et al., 2011).

Static/dynamic

The majority of studies deal with traditional static visualisation. As the display typically can become complex when uncertainty is added to data depictions, there have also been numerous attempts to utilize dynamic views. Some of these use non-interactive animation (Aerts et al., 2003; Blenkinsop et al., 2000; Evans, 1997; Kardos et al., 2003; Kardos et al., 2007; Zhang et al., 2008) and some incorporate interactive interfaces (Alberti, 2013; Blenkinsop et al., 2000; Evans, 1997; Gerharz and Pebesma, 2009; Slocum et al., 2003; Senaratne et al., 2012).

3.2.3 Application domains

The majority of studies (29 out of 44 studies) use applications from a defined domain, e.g. from environmental science or health research (Table 3.1). An advantage of this strategy is that it increases logical validity for the focus domain. However, by not systematically attempting to pick tasks that are general and representative of applications that cross domains, it is difficult to know the extent to which generalisation from the results is valid and it is difficult to relate results among studies. For example, when a study assesses the intuitiveness of an approach the results are often domain-specific: A symbol set used in aviation may be highly intuitive for a pilot, but this result cannot be easily transferred to other domains – for an expert with a different background, the symbol set may not be intuitive at all.

Table 3.1. Domains used in the reviewed studies.

Domain	**Studies**
Aviation	Kolbeinsson (2013)
Land-use planning, spatial planning, urban planning	Aerts et al. (2003), Leitner and Buttenfield (2000), Senaratne et al. (2012), Vullings et al. (2013), Zhang et al. (2008)
Remote sensing, land cover/land use	Blenkinsop et al. (2000), Drecki (2002), Evans (1997), Kinkeldey et al. (2014b), Senaratne et al. (2012), Wray (2007)
Health research, medical imaging, disease reporting	Edwards and Nelson (2001), Grigoryan and Rheingans (2004), MacEachren et al. (1998), Wray (2007)
Environmental modeling, water management, climate change, soil mapping, geology	Alberti (2013), Gerharz and Pebesma (2009), Kubíček and Šašinka (2011), Retchless (2012), Senaratne et al. (2012), Slocum et al. (2003), Viard et al. (2011)
Natural hazard management	Kunz et al. (2011)
Meteorology, weather forecast	Nadav-Greenberg et al. (2008), Wittenbrink et al. (1996)
Demography, census	Kardos et al. (2007), Kardos et al. (2008), Schweizer and Goodchild (1992)
No specific domain	Bisantz et al. (1999), Bisantz et al. (2009), Boukhelifa et al. (2012), Finger and Bisantz (2002), Kardos et al. (2003), MacEachren et al. (2012), Newman and Lee (2004), Sanyal et al. (2009)

3.2.4 Participants

The reported number of participants per study differs substantially (Figure 3.4); one important factor is whether evidence is quantitative or qualitative. Owing to requirements of statistical analysis, the number is much higher in studies using quantitative approaches. Lab studies, the most common evaluation type, range from 9 (Zhang et al., 2008) to 123 (Viard et al., 2011) participants – the median is 31. Compared to studies carried out in a laboratory, web-based studies have the advantage that high numbers of subjects can easily be recruited; the participant numbers in webbased studies reviewed reflect this as they range from 32 (Kinkeldey et al., 2014b) to 274 (Retchless, 2012) with a median of 82. As noted above, studies involving qualitative methods (i.e. focus groups, and interviews) use fewer participants – all have less than 15 subjects.

Another crucial aspect regarding participants is their expertise. Expertise is described in many ways that are usually not directly comparable across studies, e.g. experience in using geographic information (Gerharz and Pebesma, 2009; Kardos et al., 2008), experience with the concept of uncertainty and its visualisation (Kardos et al., 2008; Kinkeldey et al., 2014b), experience in maps and mapping (Evans, 1997, MacEachren et al., 2012), training or knowledge in the application domain (Aerts et al., 2003; Kolbeinsson, 2013; Kunz et al., 2011; Senaratne et al., 2012) or computer literacy more generally (Newman and Lee, 2004). Self-assessment was often used to determine the subjects' expertise, especially when participants were recruited via the web (Aerts et al., 2003; Kinkeldey et al., 2014b; Senaratne et al., 2012). A number of studies define groups with different levels of expertise (e.g. novices versus experts) to assess its impact on the results (Evans, 1997; Kubíček and Šašinka, 2011; Nadav-Greenberg et al., 2008; Schweizer and Goodchild, 1992; Slocum et al., 2003). All in all, the type and level of expertise of participants often remain unclear and groups are usually heterogeneous in expertise.

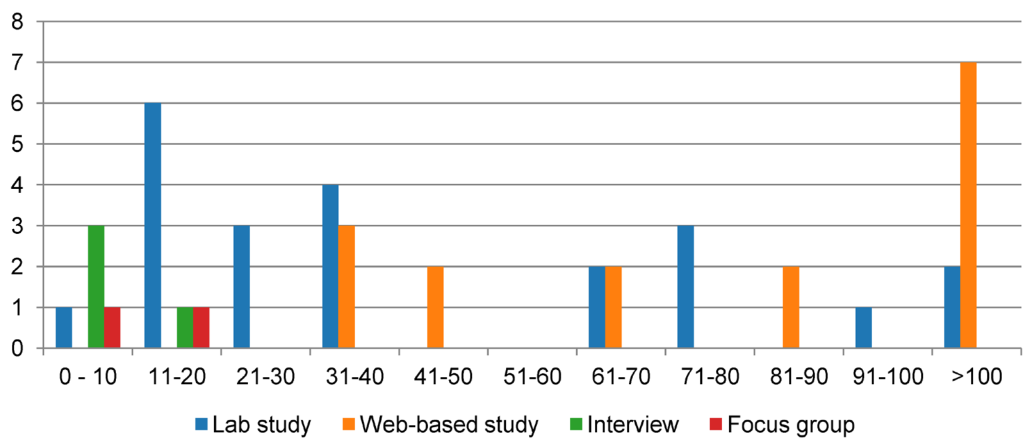

Figure 3.4. Number of participants.

3.2.5 Tasks

Assessing effectiveness of map communication, whether of the data in the map or uncertainty of those data, typically involves having participants complete some task and measuring accuracy and speed with which they do so. From the papers reviewed, we identified two major categories we discuss below: objective assessment that includes tasks with measurable correctness of results such as value retrieval, ratings, comparisons or rankings, and subjective assessment, a category of tasks for evaluating the intuitiveness of an approach, the preference compared to other options or the subjects' confidence in their responses when using it. As noted above, this review focuses on studies directed to the communication of uncertainty in visualisations and a complementary paper about impacts of uncertainty visualisation is being prepared. Thus, we do not report on decision tasks based on uncertain data here, although some papers included in this review involve such tasks in addition to communication tasks.

Few authors explicitly justify the choice of tasks, e.g. Sanyal et al. (2009) asked domain experts which tasks they would find important for their application. For this reason, they chose the search of hotspots in uncertainty ('[the expert said he] would be interested in looking at regions of extreme (high or low) uncertainty', p. 1213) and the count of features of a specific combination of data and uncertainty ('He also wanted to be able to discern features in the data, in the presence of uncertainty', p. 1213).

Objective assessment

Not surprisingly, *value retrieval* was found to be a common task in the uncertainty visualisation studies we reviewed. It is a task with a long history in cartographic communication research, tracing back to at least Flannery's 1956 dissertation on graduated symbol map interpretation (results of which were reassessed more than a decade later and published in Flannery, 1971). We identified two kinds of value retrieval tasks that are used. In the first, data and/or uncertainty values have to be retrieved separately (Aerts et al., 2003; Alberti, 2013; Kubíček and Šašinka, 2011; MacEachren et al., 1998; Nadav-Greenberg et al., 2008; Wittenbrink et al., 1996). In the second, retrieval of both data values and uncertainty happens simultaneously (Blenkinsop et al., 2000; Drecki, 2002; Kolbeinsson, 2013; Kubíček and Šašinka, 2011). While generally, value retrieval tasks help assess basic map reading, the first kind corresponds to univariate and the second one to bivariate map reading.

A rarely used extension of value retrieval tasks is *aggregation* of uncertainty over an area (Drecki, 2002; Evans, 1997; Kinkeldey et al., 2014b). This task type assesses the users' ability to retrieve an overall estimation from a spatial distribution of uncertainty.

Some studies use *rating* tasks in which levels of uncertainty have to be estimated, typically on a continuous scale (e.g. from 0 to 100) (Bisantz et al., 2009; Boukhelifa et al., 2012; Finger and Bisantz, 2002) or on a Likert scale (Drecki, 2002). The difference from value retrieval is that no legend is provided or needed. These tasks emphasize accuracy of relative judgments rather than precise value estimation or legend matching.

Another type of tasks includes comparisons of the uncertainty of different entities (e.g. 'which entity is more uncertain?') (Alberti, 2013; Blenkinsop et al., 2000; Evans, 1997; Kinkeldey et al., 2014b; MacEachren et al., 1998; MacEachren et al., 2012; Schweizer and Goodchild, 1992; Viard et al., 2011). Some studies involve comparisons for which subjects have to aggregate uncertainty over an area or several entities first. An example for this is the second experiment in

MacEachren et al. (2012) where the overall degree of uncertainty of two sets with nine icons each have to be compared.

Additionally, there are ranking tasks that let subjects assign an order to a number of entities by their data value (Bisantz et al., 1999; Finger and Bisantz, 2002) or their uncertainty (Bisantz et al., 1999; Bisantz et al., 2009; Blenkinsop et al., 2000; Boukhelifa et al., 2012). Ranking tasks by combined data value and uncertainty require interpretation and are thus not included since we focus on the communication aspects here. In contrast to the tasks mentioned so far that were dealing with specified map objects, there are other tasks including the search for entities that fulfil a certain characteristic, e.g. extremely high or low values. Several studies incorporate such tasks, for example, the search for the highest data value (Viard et al., 2011), the lowest and/or the highest uncertainty (Sanyal et al., 2009; Wray, 2007), or the identification of patterns such as clusters in the data (Edwards and Nelson, 2001; MacEachren et al., 1998) or in uncertainty only (Drecki, 2002; Edwards and Nelson, 2001; Sanyal et al., 2009). Sanyal et al. (2009) extend this task type further to multiple entities using a counting task for data values and uncertainty meaning that the number of clusters in data and uncertainty has to be determined.

Subjective assessment

Beside tasks used to measure accuracy and speed of users, there is another category of tasks used to let subjects directly assess different aspects of usability, typically by choosing from a list of options (on a Likert scale or similar) or by giving a rating (e.g. on a scale from 0 to 100). For instance, this is used to assess the confidence with a response (Alberti, 2013; Blenkinsop et al., 2000; Edwards and Nelson, 2001; Evans, 1997; Grigoryan and Rheingans, 2004; Kolbeinsson, 2013; Kubíček and Šašinka, 2011; Leitner and Buttenfield, 2000). Other studies determine the users' preference for a certain technique (Boukhelifa et al., 2012; Gerharz and Pebesma, 2009; Kardos et al., 2003; Retchless, 2012; Senaratne et al., 2012), assessment about ease-of-use of a map or visualisation (Grigoryan and Rheingans, 2004; MacEachren et al., 1998) or judgment of map attractiveness (MacEachren et al., 1998). Subjective assessment is also used to determine or the intuitiveness of different symbols (MacEachren et al., 2012), the difficulty to identify data and uncertainty (Newman and Lee, 2004), or the degree of visual overload (Newman and Lee, 2004) perceived by participants. Alternatively, open questions are used to collect different interpretations without the influence of predefined answers. For instance, Boukhelifa et al. (2012) posed open questions to find out how users interpret the meaning of the sketchy line technique.

3.3 Discussion of findings

In this section, we summarize findings from the collection of studies reviewed here. The following subsections contain findings related to five visual representation method success metrics. The first, and the main part, focuses on objective measures of user performance (accuracy, speed) that are subdivided into the visualisation categories as defined in the UVis[3] (see the section on 'Introduction'). This is followed by study results on the general acceptance of uncertainty visualisation and those from user confidence as a subjective measure of task performance. The last two subsections deal with users' judgments about representation forms, i.e. preference for and intuitiveness of different representations.

3.3.1 User performance

In this subsection, we discuss findings related to user performance that is typically measured as accuracy and response time. This discussion is divided into subsections represented by the three main dichotomies: intrinsic/ extrinsic, coincident/adjacent and static/dynamic.

Intrinsic/extrinsic

In the studies we reviewed, most uncertainty visualisation techniques under assessment were intrinsic ranging from manipulation of colour hue, value or saturation to other visual variables, such as transparency, blur or resolution.

A straightforward approach to depict uncertainty is to use *colour hue* and/or *value*. As an example, Leitner and Buttenfield (2000) compared the representation of attribute uncertainty in base maps using colour value, saturation and texture. They found that darker colour value for high uncertainty yielded the highest accuracy, followed by coarser texture and lower saturation and recommended colour value as the first choice in terms of response accuracy. Boukhelifa et al. (2012) contributed a number of studies focused on uncertainty with line features, comparing known signification methods (greyscale, blur, dashing) to a novel representation called sketchiness (an imitation of hand-drawn lines). They found that in terms of response accuracy, a greyscale representation performed better than blur, dashing and sketchiness. While up to four levels of uncertainty could be distinguished for greyscale and blur, only three were discriminable for dashing and three to four for sketchiness (depending on the task).

Since colour hue and value are often already used for representing the data itself, and 'purity' of colour has been hypothesized to be intuitive as a method to signify uncertainty, the manipulation of *colour saturation* to represent uncertainty has been subject to a number of studies. Sanyal et al. (2009) compared different uncertainty representations for line charts (1D) and surfaces (2D data in a 3D display) using artificial data. Colour-mapping from saturated blue for low and unsaturated blue for high uncertainty was compared to coloured glyphs, glyphs of different size and error bars to represent uncertainty. Response accuracy for colour saturation was not consistently higher with all tasks, but all in all the authors encouraged its use. However, in other studies, colour saturation was found to be less effective than other approaches (e.g. Drecki, 2002; Kunz et al., 2011). In direct comparison to colour value, Leitner and Buttenfield (2000) recommended saturation only if colour value or texture are not available – with less saturation for higher uncertainty (unless short response times are more important for which they recommended more saturated colours for higher uncertainty). In an extensive evaluation of intrinsic representations of uncertainty in symbols, MacEachren et al. (2012) suggested that saturation was amongst the techniques with lower accuracy, together with colour hue, orientation and shape. Thus, from current knowledge, colour saturation cannot be recommended to represent uncertainty. Instead, colour hue and value as well as transparency are better alternatives (Figure 3.5).

As an alternative to colour value and saturation, *whitening* can be used, the representation of uncertainty by whiteness in the HSI colour model (Hengl, 2003). Kubíček and Šašinka (2011) found that the combination of hue and whiteness was not suitable for continuous uncertainty with an unclassed bivariate display, because the legend was too complex to read. Supporting this, Gerharz and Pebesma (2009) measured low performance for whiteness during retrieval of uncertainty values from a coincident uncertainty display, compared to colour-coded adjacent maps. But all in all, evidence is still rare to make well-founded assumptions about the effectiveness of whitening.

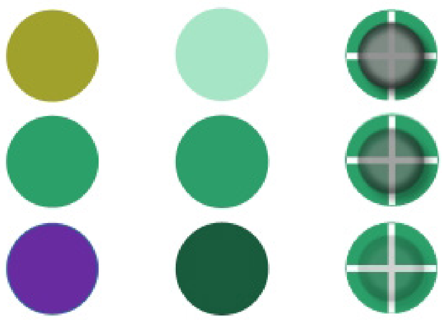

Figure 3.5. Three of the recommended intrinsic techniques w.r.t. user performance: colour hue, color value and transparency (from certain=bottom to uncertain=top).

The visual variable *transparency* (or opacity) is a popular alternative for intrinsic uncertainty representation. In a study involving land cover maps, Drecki (2002) measured higher effectiveness for transparency signifying classification uncertainty than for colour saturation.

Newman and Lee (2004) support this observation; they compared transparency to colour mapping and to a number of extrinsic techniques in static 3D scenes. Regarding ease of identification, subjects ranked transparency among the best techniques for both data and uncertainty. All in all, there is evidence that transparency generally has higher potential for uncertainty depiction than colour saturation.

In order to depict data plus uncertainty for area data, an alternative to manipulating colour attributes is to integrate *texture* and colour. While texture is often added as an overlay on top of a data depiction (thus could be considered extrinsic), the visual result is that colour and texture are integrated into the areas, thus becoming intrinsic. There is some evidence that texture on colour fill leads to good results (Kunz et al., 2011; Leitner and Buttenfield 2000; MacEachren et al. 1998; Retchless, 2012).

Another intrinsic approach to represent uncertainty is to vary *resolution*. It was used for uncertain symbols in several studies by Bisantz and colleagues who called the technique 'icon degradation'. In an early study (Bisantz et al., 1999), they compared five sets of symbols (from abstract to iconic) representing uncertain hostile and friendly identities. Using resolution to represent uncertainty (with coarser resolution depicting higher uncertainty), they found that subjects could appropriately sort both the abstract and iconic versions of the symbols representing identity ('friendly' or 'hostile') combined with six levels of uncertainty, resulting in 13 different symbols. This result was basically supported by the first experiment in Finger and Bisantz (2002) with a similar setup, but they observed lower user performance in the 'hostile' than in the 'friendly' condition. However, this observation was not made in other studies by the group. In a study by Kolbeinsson (2013) referring to the work by Bisantz and colleagues the icon degradation technique led to decreased user performance compared to symbol shape when users needed to read a data value and uncertainty in combination. These findings suggest that resolution can be a viable alternative to the manipulation of colour attributes and transparency.

The study by Bisantz et al. (2009), unlike the studies before, did not use resolution to represent uncertainty, but colour saturation, value and transparency. The authors compared two different backgrounds the symbols were placed on: a uniform grey area and a map. Surprisingly, they did not observe a significant effect of the background on user performance. More insight regarding the use of symbols on maps is provided by the Edwards and Nelson's (2001) study that assessed bivariate circle symbols with size representing data and colour value representing uncertainty. They compared bivariate symbols to an approach using univariate symbols and an additional depiction of uncertainty as a reliability diagram in the legend. Alternatively, uncertainty depiction through verbal statements in the legend was utilized. The authors found that generally, circle symbols on a map were much more effective than a verbal description of spatially-varied uncertainty and also more effective than the uncertainty diagram in the legend. In particular, 'focus-size' (colour value of the boundary of unfilled circles depicts uncertainty) surprisingly resulted in higher accuracy and confidence rates than 'value-size' (the same approach with filled circles), but the difference was not statistically significant.

Extrinsic methods for representing uncertainty have a long tradition in scientific visualisation, e.g. Pang and colleagues designed *glyphs* to represent data and uncertainty in combination. In an early study by Wittenbrink et al. (1996), the authors compared arrow glyphs including uncertainty information ('verity visualisation') to common arrow glyphs without uncertainty and found that they encoded bearing, magnitude, and uncertainty with almost the same effectiveness as with the simple arrow glyphs. In the study by Newman and Lee (2004) already mentioned above, three extrinsic techniques received the highest ratings regarding effectiveness: multi-point, ball and arrow glyphs. In the static 3D displays they used, these techniques outperformed intrinsic techniques such as colour mapping, transparency and aliasing (a combination of transparency and blur). In related research, Grigoryan and Rheingans (2004) found that with a spatial task in a 3D display (tell if a marker is inside of the error margin of a surface), subjects were significantly more accurate and faster when the error margin was represented by points than with pseudo-colouring of the surface. In the above mentioned study by Sanyal et al. (2009) subjects performed well with two kinds of spherical glyphs (varying size and colour value) in a 2D and a 3D display. Regarding the use of extrinsic methods in maps, Drecki (2002) showed that what he labeled as the 'squares' technique (square glyphs coloured by land cover type that are varied in size with smaller squares representing higher uncertainty) performed as well in terms of effectiveness as transparency and better than a display where uncertainty was represented by heights of a 3D surface.

Another set of extrinsic approaches that have rarely been assessed are *grid-based techniques*. In two recent studies, Kinkeldey et al. (2014b) assessed 'noise annotation lines', a technique that signifies classification uncertainty by a noise grid. The web-based studies incorporated qualitative comparisons of attribute uncertainty between areas in land cover maps with four to eight uncertainty classes, using different grid designs. The authors recommend the technique for up to six uncertainty classes when the most salient grid design is used. They point out that noise annotation lines can be a viable alternative to intrinsic methods, especially with complex map content. Another grid-based approach is the 'trustree' technique developed by Kardos et al. (2007) that varies the level of detail of a tree-structured grid to represent uncertainty. Since the evaluation did not include measurement of user performance, findings are limited to subjective measurement (see subsection on user preference).

All in all, results on extrinsic displays discussed here highlight the potential of glyph- and grid-based techniques for uncertainty representation in maps as alternatives to intrinsic techniques. On the question of whether to choose intrinsic or extrinsic techniques, the type of uncertainty to be displayed is deemed to play a role: Kunz et al. (2011) suggested that intrinsic approaches may be more suitable for communicating quantitative and extrinsic approaches for qualitative information. This was supported by Alberti (2013) and Kinkeldey et al. (2014b) who conclude that the extrinsic displays they used were especially successful for the communication of qualitative uncertainty. But there are also other observations suggesting that types of tasks decide if intrinsic or extrinsic approaches are more suitable: Slocum et al. (2003) evaluated the effectiveness of intrinsic vs. extrinsic visualisation techniques as part of a usability engineering approach. They compared intrinsic RGB colour coding, transparency, as well as extrinsic line glyphs and 'gcm glyphs' (vertical bars and pyramids). Based on the interviews the study relied upon, they found that subjects with a scientific background preferred glyphs and the less experienced preferred colour coding and transparency. Their explanation for this is that intrinsic techniques they used gave a better overview of uncertainty, but in-depth analysis was easier with extrinsic techniques. This suggests that there may be tasks for which intrinsic techniques are more appropriate and others for which extrinsic approaches work better.

Coincident/adjacent

This subsection deals with findings about comparison of adjacent to coincident (integrated) maps. The obvious difference between the two approaches is that adjacent maps require more eye movements (saccades) to retrieve information than coincident maps. But the latter tend to become complex and cluttering is a bigger problem when using a single view for data and uncertainty, compared to adjacent views.

Amongst the studies involving a direct comparison, some do not suggest general differences between adjacent and coincident views (Kunz et al., 2011; Retchless, 2012). Most studies report on non-significant differences between the two approaches, e.g. Kubíček and Šašinka, (2011) who reported that when retrieving data value and uncertainty at the same time, users were slightly more successful using adjacent views than a coincident display. In a study already discussed above, MacEachren et al. (1998) compared adjacent maps to coincident displays that are visually integral (colour and hue shift) and visually-separable (colour and texture). Adjacent maps were judged to be 'more pleasant and easier to use'. But the coincident, visually-separable texture overlay yielded user performance comparable to the adjacent views. In the study mentioned above, Viard et al. (2011) compared adjacent maps to coincident maps using a texture with varying transparency to represent the degree of uncertainty. A simple comparison task yielded similar results for both approaches, but with a more complex ranking task, adjacent views led to less accurate answers than the coincident view. Coincident versus adjacent results for identification of spatial patterns were reported by Edwards and Nelson (2001). The use of a small uncertainty display in the legend of a map was less successful than the coincident alternative using circle symbols. Gerharz and Pebesma (2009) compared colour-coded adjacent maps and a bivariate coincident map using whitening and reported that uncertainty retrieval was more successful with adjacent maps but no difference existed for data retrieval. In addition, all ten subjects found the tasks easy to accomplish with adjacent maps and only five subjects had this impression with the coincident map.

There are very few results regarding response time. The study by Kubíček and Šašinka (2011) is an exception, reporting that in a map reading task of either data or uncertainty values, coincident maps led to quicker responses than adjacent maps. This may again be explained by saccades that are necessary when retrieving values from two adjacent maps.

All in all, past research suggests that both coincident and adjacent approaches have their applications. There is evidence that adjacent views may be usable for retrieval of single values, but less usable when tasks become more complex and more saccades are needed. Generally, coincident maps can be seen as preferable because the integration of uncertainty into the display makes it easier to retrieve data and uncertainty simultaneously. However, they naturally become more complex than adjacent views and the map content is more likely to be obstructed by the additional uncertainty display. For instance, the use of bivariate colour schemes for data and uncertainty can be challenging. So, as shown in several studies, adjacent maps can be a viable alternative to avoid clutter.

Static/dynamic

A number of studies directed to representation of data uncertainty deal with the use of dynamic approaches such as animated displays or user interaction. The range of possible approaches is wide because elements from animation and interaction can be combined in numerous ways. This makes it even more difficult to come to consistent conclusions about the effectiveness of dynamic approaches. In an early study by Evans (1997) an animated (noninteractive) 'flicker' map and an interactive version ('toggling') were compared to static maps. The results showed that static and (non-interactive) dynamic approaches did not differ significantly in terms of user accuracy or speed. This result was not supported by Aerts et al. (2003). since they observed significantly higher accuracy for uncertainty estimation with a static adjacent view than with toggling. Referring to these results, Blenkinsop et al. (2000) reported that a static grey scale display showed better user performance than serial animation (series of animated maps) or random animation (animated display of possible outcomes). In another study comparing static and dynamic, Drecki (2002) found that blinking (variations in display time of map entities according to their uncertainty; Fisher, 1993) was less effective than static representation through glyphs or opacity, but more effective than a 3D surface or the manipulation of colour saturation. Senaratne et al. (2012) also observed higher user performance with static than with dynamic approaches. All in all, there is evidence that animated views have a potential to successfully represent uncertainty when static solutions are not feasible but little evidence that they perform better (or even as well) as more traditional static depictions. However, as Blenkinsop et al. (2000) suggest, animated approaches may be suitable for specific tasks – for instance, for exploration of uncertainty of a dataset in an early stage of analysis because they provide 'a very effective first impression of uncertainty' (p. 11).

3.3.2 Acceptance

An important overarching evaluation question is how users generally react when uncertainty is depicted visually. A number of studies report on this aspect and the findings are not consistent: some authors report that adding uncertainty information to a map had negative effects on map readability (Schweizer and Goodchild, 1992; Slocum et al., 2003) and that subjects wanted the display to remain unobstructed (Kardos et al., 2003). One obvious reason for this is that displays become more complex when uncertainty is added. Besides, users tend to be overwhelmed by the additional information when they make analyses or decisions – an aspect we discuss in the second paper on effects of uncertainty visualisation. But there are also findings suggesting that visualisations including uncertainty were not judged as too complex or even that addition of uncertainty clarifies the view instead of cluttering it (Aerts et al., 2003; Alberti, 2013; Edwards and Nelson, 2001; Kunz et al., 2011; Leitner and Buttenfield, 2000; MacEachren et al., 1998; Viard et al., 2011). This suggests that, when appropriate solutions are found, users do not necessarily see depicted uncertainty as a burden.

3.3.3 User confidence

When assessing the usability of uncertainty visualisations, the level of confidence that subjects have with their answers can be an important aspect. However, in many studies, confidence was not measured at all. From those that did, most studies reported that user performance and confidence were in agreement, e.g. in the study by Blenkinsop et al. (2000) cited above they observed higher confidence as well as better user performance with a greyscale display compared to an animation approach. Edwards and Nelson (2001) found that bivariate symbols depicting data and uncertainty (size combined with either focus or colour value) yielded higher confidence (along with more accurate results) than verbal and graphical depiction of uncertainty in the legend. Grigoryan and Rheingans (2004) suggested significantly higher confidence (as well as higher accuracy and shorter response times) when a point-based representation of positional uncertainty was used, compared to colour coding. Kolbeinsson (2013) reported higher confidence for shape changes than for icon degradation that also corresponded to higher response accuracy. All in all, the majority of studies measuring confidence provide evidence for the assumption that the successful use of a technique (in terms of accurate and/or fast answers) at the same time leads to high user confidence.

3.3.4 User preference

A number of studies measured user preference for the visual techniques under evaluation. Generally, it can be stated that in contrast to confidence, user preference did not always correspond to user performance, i.e. subjects often preferred techniques that did not necessarily work best for them. For instance, although users were not successful with colour saturation they had a preference for using it (Drecki, 2002). This effect was also measured in the above mentioned study by Boukhelifa et al. (2012) with respect to uncertain lines: Dashing was preferred over blur, greyscale and sketchiness, but this did not correspond with user performance in which dashing yielded only three discriminable levels of uncertainty (fewer than with blur or greyscale).

However, there are also results suggesting a match between performance and preference. Gerharz and Pebesma (2009) reported that most participants preferred adjacent maps over a coincident view with whiteness representing uncertainty and they could also retrieve uncertainty values most accurately using the adjacent display. In a study focused on similar datasets but in a web-based environment, Senaratne et al. (2012) also found a strong correspondence between preference and user performance but only for the static techniques they assessed (contouring, symbols, adjacent maps), not for the dynamic approaches.

All in all, these findings show that measuring preference does not suffice to determine the effectiveness of uncertainty visualisation techniques, but that it can give hints about what approaches are popular with different user groups. This information can be taken into account for choosing useful methods to depict uncertainty when user acceptance is essential.

3.3.5 Intuitiveness

The intuitiveness of different techniques was assessed by a few studies only. The first part of the study reported by MacEachren et al. (2012) included three common metaphors suggested as appropriate to uncertainty signification including: colour purity (manipulating colour saturation), fog (transparency) and blur (fuzziness). Fuzziness ranked as the most intuitive of all signification methods, while transparency was above average (but with wide variation in reactions) and colour saturation was not judged as intuitive. Boukhelifa et al. (2012) tested the intuitiveness of sketchiness as a representation of uncertainty. Overall, it was perceived as less intuitive than blur, greyscale and dashing; since blur corresponds to fuzziness in the MacEachren et al. (2012) study and greyscale to colour value, these results support each other. This provides further evidence for the assumption made above (from objective measurement of user performance) that manipulating colour saturation is not a recommendable way to depict uncertainty, compared to other techniques such as transparency and blur.

Besides the intuitiveness of a method itself, a number of studies address the question of 'which end is up': should high uncertainty be matched with low or high colour value and low or high colour saturation? Lighter values signifying uncertainty result in a 'fading out' effect, whereas darker colours make regions with high uncertainty more prominent. Bisantz et al. (2009) measured a tendency to assign lower uncertainty to darker colours when using colour value, and to more saturated colours when using saturation to represent uncertainty. This was supported by Kubíček and Šašinka (2011) as the majority of participants picked the lighter values as the best choice for higher uncertainty (thus darker value corresponds to lower uncertainty). In addition, subjects were faster when using this alternative. In the most recent 'which end is up' research, MacEachren et al. (2012) reported that for colour value, light for uncertain and dark for certain was much more intuitive than the reverse. Unsaturated colours were also more intuitively associated with uncertainty than saturated colours, but only slightly. Colour saturation in either order scored near the mean of all visual variables tested, while colour value (with light depicting uncertain) scored near the top (just below fuzziness and location depicted as a point in a coordinate space – see Figure 3.6).

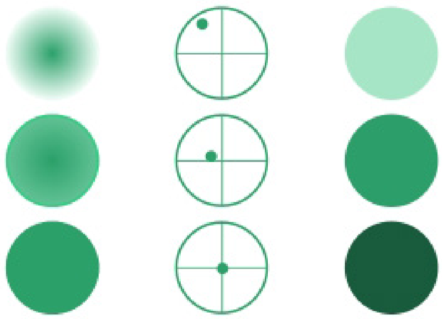

Figure 3.6. Three best options w.r.t. intuitiveness from MacEachren et al. (2012): fuzziness, position and colour value (the top depiction in all cases was interpreted to be most uncertain).

3.4 Lessons learned

Several lessons learned and recommendations can be derived from the discussion, the first ones referring to study design in general. While empirical studies of map reading and use have been carried out for 60 years or more and uncertainty visualisation has been addressed explicitly for at least half of that time, the methods used in uncertainty visualisation evaluation remain ad hoc. Studies are often approached more from a usability engineering perspective of assessing and improving a specific product than from a cognitive psychology or science perspective of developing general understanding of how and why representations work or do not work. However, even if considered from a usability engineering perspective, studies often do not follow any methodology commonly agreed upon. This lack of formalisation and rigour in empirical methods is an issue that is much broader than the study of uncertainty visualisation focused on here; it is an issue that cuts across research in cartography, information visualisation and related domains. Empirical research focused on geographically varied uncertainty (geo-uncertainty) visualisation is probably no less formalized or rigorous than that in other aspects of geovisualisation, but it is clearly no better. Thus, all comparisons and generalisations that we offer here, based on our analysis of the 44 studies reported on in 34 papers, must be considered as starting points toward developing a deep understanding of geo-uncertainty visualisation, not a definitive summary.´

3.4.1 Evaluation goals

From the publications included here, two major goals for existing research can be identified. The first is the assessment and improvement of visual displays representing uncertainty. Here, the question is often 'does method A work better than method B' (with 'better' defined in ways we outlined above). These studies (while not always saying so explicitly) are essentially following a usability engineering approach where the goal is to improve a particular product using a summative assessment rather than to create general principles. One problem in many of these studies is that the authors tend to attempt generalisation beyond the specific constraints of the test, even though the conceptual framing of the test was not designed to do this.

The second goal is to advance understanding of the cognitive processes involved in using visual displays (both static and dynamic) for interpreting information that contains uncertainty. Here, the question is often: how/why does method A work better than method B or how does any method of interest change the cognitive process? Such studies usually are grounded in perceptual and cognitive theory and as a result have the potential to provide a framework for relating the generally ad hoc results from studies that adopt a usability engineering approach. Studies from the second category are in the minority of those reviewed.

3.4.2 Uncertainty visualisation techniques

When evaluating uncertainty visualisation the appropriate choice of techniques to be assessed and compared can be a challenge. As discussed in the introduction, typologies can help to choose from the universe of possible techniques and heuristics derived from general guidelines and rules for data visualisation can support the choice. But when different techniques are to be compared, it is important that the scenarios and datasets are informationally equivalent, i.e. according to Larkin and Simon (1987, p. 67) '[t]wo representations are informationally equivalent if all of the information in the one is also inferable from the other, and vice versa'. A goal in testing, then, is often to determine whether they are also computationally equivalent, or whether one depiction has an advantage over another, i.e. '[t]wo representations are computationally equivalent if they are informationally equivalent and, in addition, any inference that can be drawn easily and quickly from the information given explicitly in the one can also be drawn easily and quickly from the information given explicitly in the other, and vice versa' (Larkin and Simon 1987, p. 67).

Another crucial aspect is the role of visual metaphors that have been used to depict uncertainty since the beginning of this field of research, e.g. fog or blur (MacEachren, 1992). The contention is that fog and blur are metaphors for lack of clarity or focus (as in a camera) and thus directly signify uncertainty. These metaphors have been suggested to have the potential to enable a better understanding of uncertainty (Gershon, 1998) and we make the assumption that the use of metaphors can lead to more intuitive approaches (Kinkeldey et al., 2014b). But the usefulness of metaphors generally has rarely been investigated. It would be worthwhile to evaluate whether metaphors increase the intuitiveness of uncertainty visualisation, the understanding of geo-uncertainty and the success of reasoning and decision-making under uncertain conditions. One start toward addressing this goal is offered in MacEachren et al. (2012), in which the authors assess the intuitiveness of several strategies for signifying uncertainty (see the section on 'Discussion of findings'). Follow-up work is needed to assess the sensitivity of these findings to the specifics of the visual signification and the experimental design and to then address the more challenging questions of the relationship between metaphor-grounded intuitiveness of uncertainty signification and subsequent information interpretation, reasoning with that information and decision-making.

3.4.3 Is uncertainty 'just another variable'?

A question that has not been extensively discussed in the literature is: Do we treat uncertainty as 'just another variable' to be visually represented or does it need to be treated differently? For example, when we picture a map showing air pressure distribution in combination with temperature, these two variables are certainly dependent on each other (in physical terms). The same is true with air pressure and its uncertainty, but we see a stronger dependency: Uncertainty can be seen as metadata of airpressure which we argue makes a difference. Thus, we support Edwards and Nelson (2001, p. 35) who stated that '[p]erhaps data certainty information is unique and will require a new type of framework for designing symbolization'. Most studies that assess the usability of uncertainty visualisations do not contribute to this aspect since they test the retrieval of data and uncertainty separately, but from the perspective that there is nothing special about uncertainty. Traditionally, studies in this field focus on the ability to read both the map content and its uncertainty at the same time. This may be the mandatory criterion for a successful use of uncertainty, but the question that remains is whether this is sufficient to ensure that a user does not only have two separate values in mind but an integrated uncertain data value.

3.4.4 Classed/unclassed representations

Another aspect that is often neglected in past studies is whether to use classed or unclassed schemes for uncertainty categories. When classed uncertainty is used, the choice of the number of classes is rarely explained. In the studies reviewed here, classification schemes range from binary classed, i.e. certain/uncertain (or reliable/unreliable), over six classes (e.g. Bisantz et al., 1999) to 15 classes (Schweizer and Goodchild, 1992). All in all, the colour scheme used in the latter study used 15 × 15 classes; thus, subjects had to distinguish 225 classes. It is not surprising that the authors reported low user accuracy of judgments. They justify the high number of classes 'to give the appearance of continuous shading' (Schweizer and Goodchild, 1992, p. 689). Their results raise the question of whether a high number of classes are more complex than a continuous uncertainty distribution or not; to our knowledge, this question has not been addressed. In contrast to studies that use a constant number of uncertainty classes, in the study by Kinkeldey et al. (2014b) about noise annotation lines the number of classes was defined as an experimental factor. In this way, they could measure the impact of the number of uncertainty classes on user performance. More work is needed on the question of necessary level of detail in uncertainty representation to support different tasks using uncertainty visualisation (Smith et al., 2013).

3.4.5 Types of uncertainty

Traditionally, a distinction is made between attribute (or thematic), positional (or geometric) and temporal uncertainty (see the section on 'Analysis of the literature'). Although these categories seem logical and match with conceptualisations of data used in geographic database research and development (Peuquet, 1994), their use may be limited for the majority of applications of uncertainty visualisation. In practice, it is hard to clearly distinguish between these categories: '[t]he categories of uncertainty are often interdependent, and the category boundaries are often hard to delineate' (MacEachren et al., 2005, p. 156). For instance, when a land cover map is created from a remotely sensed image, the boundaries shown in the map are uncertain. This can be a result of various sources, e.g. the vagueness and ambiguity in the definition of land cover classes (attribute uncertainty), measurement errors (positional uncertainty) and the use of images from different capture dates (temporal uncertainty). If all three types could be estimated, it might help an expert analyst, but a domain specialist who is not a remote sensing expert might be confused. This raises the issue of how the complexity of uncertainty relates to the categories of user and task – if someone is trying to create a better satellite system, they might need the full range of uncertainty information; if they are trying to decide what crop to plant, they might need a simple composite depiction of uncertainty.

3.4.6 User issues

A number of lessons learned are related to participants and recruiting for empirical studies. A general objective should be to recruit participants who are representative for the target user group with respect to age, background, skills, experience, etc. However, the majority of studies we review here recruited students since they are easily available in the university context. Thus, most descriptions of expertise have to be judged critically. Students can be suitable participants for studies focusing on perceptual issues where only expertise in terms of visual literacy and experience with visualisation or maps is important. But even here, students may not reflect the general population well, if that is the target audience. Students have different levels of theoretical expertise typically combined with very limited or no work experience. Despite this fact, they are often described as domain or map 'experts' without a more detailed specification of their experience. Further, if domain expertise plays a role, e.g. when it is needed to understand the symbology of a map (e.g. with a geological map), students (even those studying within the domain of interest) are not yet experts. Thus, in recruiting and selecting participants for studies, it is important to differentiate between types of expertise, including at least:

- expertise in use of maps and related visual displays;
- expertise in design of maps and related visual displays;
- expertise in statistics – thus in understanding probabilities and related uncertainty metrics;
- expertise in the application domain used for test scenarios (if any);
- expertise in using uncertainty estimates in the application domain; and
- expertise in using any technology that might be relevant (e.g. if the focus is on interactive interfaces, expertise in that).

Beyond recruitment of appropriate participant groups, it has also been shown that *training* is important for effective empirical analysis of complex information display interpretation, but it is usually not carried out to a sufficient extent. An extensive training phase is often necessary, not only to clarify the scenarios, data and tasks, but also visualisation techniques used (especially when they are still unknown to the users) and measures of uncertainty that are needed. This is an aspect that has rarely been considered by the studies from our review.

3.4.7 Task dependency

Some studies provide evidence that the usability of uncertainty representations can be highly user and task dependent. For instance, in the study conducted by Sanyal et al. (2009), search tasks for the lowest and highest uncertainty values resulted in different user performance (although using the same data). From their study about weather forecasting uncertainty, Nadav-Greenberg et al. (2008) suggest that 'it is extremely important that designers of such displays consider both user and task demands because the usability and usefulness of uncertainty information depends on these factors' (Nadav-Greenberg et al., 2008, p. 44). In their study, a box plot was successful for precise information whereas the colour-coded maps worked better for relative comparisons. Similar to this, in a study discussed above, Slocum et al. (2003) reported that participants who wanted the 'big picture' preferred intrinsic techniques (RGB-encoding or transparency) whereas others who were aiming for detailed information tended to prefer extrinsic methods (i.e. glyphs). These findings show how tasks can differ between groups with a different level of expertise in the problem domain. Further evidence for task dependency of visual depiction of uncertainty is provided by Blenkinsop et al. (2000) who support the hypothesis from MacEachren et al. (1998) that visually separable representation of uncertainty is preferable for exploratory use, i.e. when it is not clear what questions will be asked exactly during analysis. In a focus group initiated by Zhang et al. (2008), domain experts were unable to articulate what their preferred uncertainty visualisation methods were because they were convinced that this strongly depends on the task.

All this supports the hypothesis that the nature of tasks plays an important role for the usability of uncertainty visualisation techniques. This may explain many of the inconsistent outcomes from the studies under review; two studies assessing a specific technique are likely to yield different results when tasks and user groups are not comparable. We further address this aspect in the next section.

3.5 Conclusion

In this review, we systematically analysed 44 user studies from 34 publications dealing with uncertainty visualisation. More precisely, we focused on uncertainty of geospatial data and geographic displays. The first step was a description of characteristics of the study under review (types of studies, visualisation techniques under assessment, number and type of participants, etc.), as well as a summary of the main findings regarding user performance, acceptance of uncertainty visualisation, user confidence, preference and intuitiveness of techniques. From this, we derived lessons learned and identified gaps in past evaluation research. Furthermore, a number of recommendations and open research questions for future studies were discussed.

This article focused on evaluating how uncertainty can be communicated. It did not include issues with reasoning and decision-making based on uncertainty visualisations – we will address these aspects in a follow-up publication. Since this work focused on visualisation of uncertainty in geospatial data we did not try to be exhaustive for fields such as information visualisation or scientific visualisation (although we did include publications from those domains when representation of geographic information uncertainty was a component). The fact that we dealt with visualisation of uncertainty means that we did not review literature on decision-making under uncertainty. We only included literature on non-visual communication of uncertainty in cases where that topic is included in a paper having a visual focus (e.g. studies that compare uncertainty visualisation to a control of verbal description or numerical specification of uncertainty).

Generally, the most important outcome is that we need to systematize future empirical studies on uncertainty visualisation to better enable comparison and generalisation of the findings. As mentioned above, one way to advance this goal is the use of uncertainty visualisation typologies. However, as also discussed above, existing typologies are focused on data types, uncertainty categories and representation types, i.e. they map a description of the data being displayed to a recommendation of techniques. Based on the discussion about task-dependency of uncertainty visualisation usability (see the section on 'Lessons learned'), we suggest that future typologies should additionally take different categories of tasks into account. We propose that at least the following three high-level task categories deserve attention:

- communication tasks

  This category comprises map reading tasks involving data and uncertainty value retrieval (which location is most uncertain?). For tasks from this category, visualisation techniques can be chosen following the traditional rules from cartography.

- analytical tasks

  Tasks that occur during analysis fall into this category, meaning that defined analytical questions have to be answered (what area is most suitable to build a power plant?). Approaches for uncertainty visualisation should be tailored to these tasks (e.g. by choosing uncertainty representations with the number of classes and level of detail needed to conduct the task).

- exploratory tasks

  Tasks from this category occur during exploration of the data, i.e. tasks for which the strategies for and outcomes from use of visual displays can hardly be foreseen. Because of this, visualisation methods need to be versatile. This can be accomplished through adaptable and adaptive approaches. Dynamic approaches, especially those involving interactivity, play an important role here.

But this is only one piece of the complex characterisation of visual depiction of uncertainty. Complementary typologies are needed to characterize static and dynamic methods for signifying uncertainty visually, for user tasks related to uncertainty signification, and for the ways in which interactivity can apply to enable user access to data, its uncertainty and their combination. The key goal in developing and applying such typologies is to support repeatability of, make comparisons among, and make generalisations from empirical studies. Thus, we propose the following main topics for future research in the field:

- the role of intuitiveness and metaphors;
- special requirements of visual depiction of uncertainty compared to other data;
- systematic description of expertise and investigation of the role of training; and
- development of task-centred typologies and guidelines.

Dealing with these topics will help advance the goal of systematic evaluation of uncertainty visualisation, and, as a result, will facilitate the development of practical guidelines in this field. Such guidelines are needed to approach the goal of a wide application of uncertainty visualisation in geospatial analyses, reasoning and decision making.

4 Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations

This chapter was previously published as peer-reviewed journal paper:

Kinkeldey, C., A. M. MacEachren, M. Riveiro, and J. Schiewe. 2015. "Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations." *Cartography and Geographic Information Science*, published online.

The candidate was the primary author and responsible for the systematic review of the papers, the summary of the findings, and contributed to lessons learned and recommendations (70% of the overall work).

Abstract

For many years, uncertainty visualization has been a topic of research in several disparate fields, particularly in geographical visualization (geovisualization), information visualization, and scientific visualization. Multiple techniques have been proposed and implemented to visually depict uncertainty, but their evaluation has received less attention by the research community. In order to understand how uncertainty visualization influences reasoning and decision-making using spatial information in visual displays, this paper presents a comprehensive review of uncertainty visualization assessments from geovisualization and related fields. We systematically analyze characteristics of the studies under review, i.e., number of participants, tasks, evaluation metrics, etc. An extensive summary of findings with respect to the effects measured or the impact of different visualization techniques helps to identify commonalities and differences in the outcome. Based on this summary, we derive "lessons learned" and provide recommendations for carrying out evaluation of uncertainty visualizations. As a basis for systematic evaluation, we present a categorization of research foci related to evaluating the effects of uncertainty visualization on decision-making. By assigning the studies to categories, we identify gaps in the literature and suggest key research questions for the future. This paper is the second of two reviews on uncertainty visualization. It follows the first that covers the communication of uncertainty, to investigate the effects of uncertainty visualization on reasoning and decision-making.

Keywords

Uncertainty visualisation, literature review, evaluation, user studies, decision-making

4.1 Introduction

Geodata uncertainty, also called geospatial or geographic data uncertainty, has been a Geographical Information Science (GIScience) research topic for several decades with authors arguing that geodata analysis needs to consider uncertainty (Caers 2011; Shi 2010; Zhang and Goodchild 2002). Uncertainty can be defined as an umbrella term for concepts like inaccuracy, imprecision, ambiguity, vagueness, subjectivity, or error (unknown or not quantified error). In spite of the difficulties associated to the visualization of such a complex concept, a wide range of uncertainty visualization solutions have been presented in the past decades across different visualization domains, see examples included in the taxonomy by Potter, Rosen, and Johnson (2012) or in the review by Brodlie, Osorio, and Lopes (2012). However, many research challenges remain concerning uncertainty visualization (Griethe and Schumann 2006; Johnson and Sanderson 2003; MacEachren et al. 2005). One of these challenges is to better understand the effects of visual depiction of uncertainty on reasoning and decision-making. In relation to geodata uncertainty visualization, a number of studies have specifically focused on measuring the effect of uncertainty (its presence and its depiction) on the outcome of data usage, but generalized findings are still rare. In order to contribute to addressing these challenges, we present a comprehensive review of these studies, outline gaps in this field, and formulate questions for future research. Furthermore, we suggest a categorization of the main components of uncertainty visualization "effect" evaluation with respect to decision-making. By mapping the studies from this review to the categories, we identify gaps and recommend research topics for future work.

This work is the second of two complementary papers reporting on a systematic review of uncertainty visualization studies in the fields of geographical visualization, information visualization, and scientific visualization (complemented by literature from related fields such as cognitive science). While the other part of the review (Kinkeldey, MacEachren, and Schiewe 2014) focuses on findings about how to visually communicate uncertainty in an optimal way, this part concentrates on assessment of the effects that uncertainty depiction may have on decision-making or risk assessment.

This paper is structured as follows: In the "Method" section we describe how we selected and categorized the studies for this review and the visualization techniques they use. The subsequent section ("Analysis of the literature") contains an overview of the studies with respect to their number over time, evaluation methodology, types of uncertainty, and visualization techniques, as well as application domains, number and type of participants, and tasks. A summary of findings is presented in the fourth section ("Findings"), related to acceptance of uncertainty depictions, effects of uncertainty visualization, the impact of uncertainty visualization techniques on decision-making and risk assessment, and the role of expertise. In the subsequent section, we derive "Lessons learned" regarding study focus and design, evaluation methodology, effects of uncertainty visualization, tasks, expertise, and decision-making research. Furthermore, we present the above-mentioned categorization of the main components of uncertainty visualization "effect" evaluation. In the "Conclusion", we summarize the main findings of this review and suggest future work.

4.2 Method

We collected the studies for this review from journal articles, conference papers, and other review papers in the fields of geovisualization, information visualization, and scientific visualization (with a number of studies from related fields such as cognitive science). First, we searched for reports on user studies with a focus on uncertainty. Since the terminology in the field tends to be ambiguous, we additionally included search terms such as certainty, reliability, data quality, or imperfection that are used as synonyms (or antonyms) for uncertainty or for sub-concepts. We limited the collection to user studies involving visual signification of uncertainty, meaning that studies without users or those assessing non-visual (e.g., sonification) approaches were disregarded. Furthermore, we focused on studies including geographical data, for instance, displayed in a map or a spatial three-dimensional display. We made an exception with studies that use a map but communicate uncertainty indirectly through non-graphical methods such as textual information (Kobus, Proctor, and Holste 2001; Shattuck, Miller, and Kemmerer 2009). Studies implementing aspatial displays (such as charts) were added to the list when they contributed findings potentially interesting for geovisualization, such as studies by Dong and Hayes (2012), Ferreira, Fisher, and König (2014), Fisher et al. (2012), or Riveiro, Helldin, and Falkman (2014a).

4.2.1 Study categorization

As mentioned above, in the two complementary reviews we make a basic distinction between two categories of studies: on "communication" and "effect" of uncertainty. A representative question in studies on communication is "How well can you determine the degree of uncertainty in a certain area?" As reported in our initial paper (Kinkeldey, MacEachren, and Schiewe 2014), these studies mostly use quantitative metrics such as response accuracy or speed, as well as self-reported user confidence or preference. Complementary research, addressed in this paper, considers the effect of visualized uncertainty information on map use, thus on reasoning or decision-making (e.g., for decisions related to questions such as "Taking uncertainty into account, where would you site an airport?"). A typical research question about the effect of uncertainty visualization is whether choices are different with uncertainty depicted versus without. "Effect" studies are sometimes carried out in combination with studies assessing aspects of communication, such as in Finger and Bisantz (2002). To clarify the distinction between the two categories, we use the code list for the evaluation of visualization proposed by Isenberg et al. (2013): our first category ("communication") corresponds to three of their codes: "User Performance" (UP), "User Experience" (UE), and "Evaluating Communication Through Visualization" (CTV) including "evaluations that assess the communicative value of a visualization or visual representation in regards to goals such as teaching/learning, idea presentation, or casual use." (2820). Our second category ("effect") corresponds to the code "Visual Data Analysis and Reasoning" (VDAR) that represents "evaluations that assess how a visualization tool supports analysis and reasoning about data and helps to derive relevant knowledge in a given domain."

4.2.2 Categorization of visualization techniques

A variety of visualization techniques were used in the studies under review. To systematically summarize the techniques, we follow the three dichotomies that we introduced in the first review paper (Kinkeldey, MacEachren, and Schiewe 2014) as the "uncertainty visualization cube" (UVis³) categorization (Figure 4.1):

1. intrinsic/extrinsic (uncertainty depiction through manipulation of the existing display or through addition of objects distinct from the data depiction),
2. coincident/adjacent (data and uncertainty in one integrated display or in separate ones), and
3. static/dynamic (uncertainty depiction without or with the use of animation and/or user interaction).

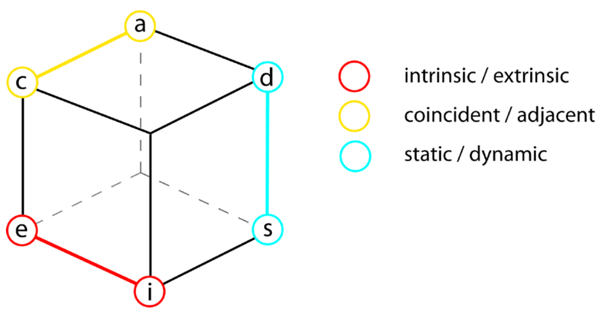

Figure 4.1. Uncertainty visualization cube (UVis³) classification [reprinted with permission from Kinkeldey, MacEachren, and Schiewe (2014)].

4.3 Analysis of the literature

Overall, we identified 87 studies (described in 68 publications) having an evaluation component. Roughly half of them (43 studies from 34 publications) contribute findings on the effect of uncertainty visualization. Most of these studies deal with effects on decision-making, e.g., on accuracy and speed of decisions. Exceptions include recent studies on the effect of displayed uncertainty on risk assessment or perceived ambiguity or fear (Ash, Schumann, and Bowser 2014; Retchless 2012; Roth 2009b; Severtson and Myers 2013). Figure 4.2 shows how over the years from 1994 to 2014 the number of studies has increased from 3 (1994–1999), through 7 (2000–2004), 14 (2005–2009), to 19 studies between 2010 and 2014. This may indicate a trend towards growing interest in evaluation of uncertainty visualization, but the numbers are low and vary much from year to year. Regarding evaluation methodology, the overwhelming majority of studies apply quantitative methods (40 out of 43), i.e., controlled laboratory experiments (33) and web-based studies (7). Thus, studies conducted via the web still remain rare compared to traditional laboratory experiments. Just a small fraction of studies (3 out of 43) are based solely or partly on qualitative methods, i.e., interviews and verbal protocol analysis ("think aloud" protocol).

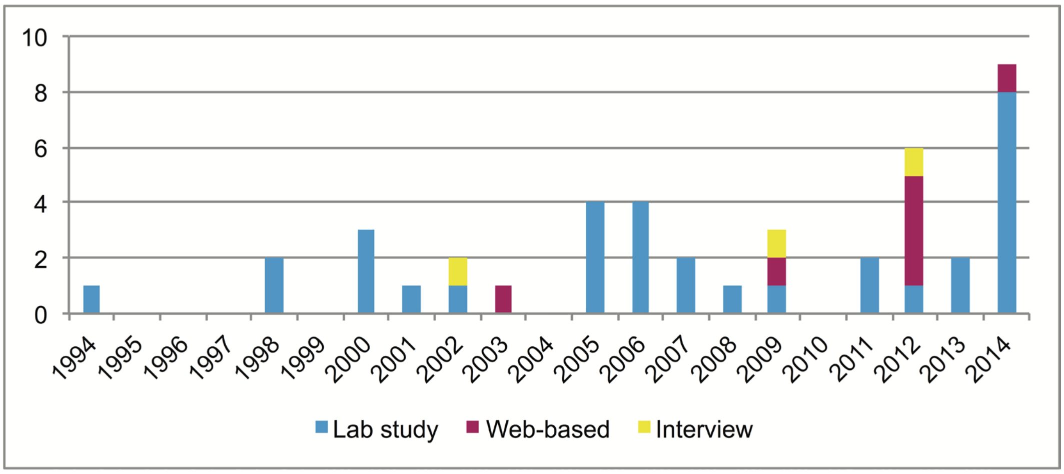

Figure 4.2. Number of user studies assessing the effect of uncertainty visualization (over time, separated by type).

4.3.1 Uncertainty categories

Traditionally, in geodata uncertainty research a distinction between three categories of uncertainty is made: attribute/thematic (what), positional/geometric (where), and temporal (when) uncertainty (MacEachren et al. 2005). From these categories, most studies we analyze here

(26 out of 43) cover attribute uncertainty, e.g., uncertainty in land suitability values. However, a number of studies (11 out of 43) deal with the effect of positional uncertainty, for instance, the uncertain location or boundaries of a geographic object (Andre and Cutler 1998; Ash, Schumann, and Bowser 2014; Brolese and Huf 2006; Cox, House, and Lindell 2013; Hope and Hunter 2007a; Kirschenbaum and Arruda 1994; Kirschenbaum et al. 2014; Pyysalo and Oksanen 2014; Simpson et al. 2006; St. John, Callan, and Proctor 2000). While none of the studies involve temporal uncertainty, a small number of studies include multiple types of uncertainty (Riveiro, Helldin, and Falkman 2014a; Riveiro et al. 2014b) or types that do not fit in one of the categories, e.g., uncertainty caused by lack of data agreement (Roth 2009b) or ambiguity in textual descriptions (Kobus, Proctor, and Holste 2001; Shattuck, Miller, and Kemmerer 2009).

4.3.2 Visualization techniques

A variety of visualization techniques were used in the studies presented here. As in the first part of the review, the majority of studies include coincident, static displays and use intrinsic techniques to visualize uncertainty. A minority involve extrinsic techniques (Brolese and Huf 2006; Ferreira, Fisher, and König 2014; Fisher et al. 2012; Hope and Hunter 2007b; Kirschenbaum and Arruda 1994; Kirschenbaum et al. 2014; Riveiro, Helldin, and Falkman 2014a; Riveiro et al. 2014b; Simpson et al. 2006; St. John, Callan, and Proctor 2000) or compare intrinsic and extrinsic (Andre and Cutler 1998; Cliburn et al. 2002; Hope and Hunter 2007a; Pyysalo and Oksanen 2014; Senaratne et al. 2012). Regarding display organization, approaches utilizing adjacent (i.e., side by side) views are rarely included – a few studies assess the effects when adjacent views are used and compare it to the use of coincident views (Aerts, Clarke, and Keuper 2003; Gerharz and Pebesma 2009; Retchless 2012; Senaratne et al. 2012; Viard, Caumon, and Levy 2011). With respect to dynamic uncertainty display, a small number of studies use animated and/or interactive approaches to depict uncertainty, comparing them to static ones (Aerts, Clarke, and Keuper 2003; Bisantz et al. 2011; Ferreira, Fisher, and König 2014; Gerharz and Pebesma 2009; Senaratne et al. 2012).

4.3.3 Application domains

Most studies included in this review (and more than in the complementary review, Kinkeldey, MacEachren, and Schiewe 2014) use data and tasks from specific domains. We identified 11 domain categories covered by the studies under review; see Table 4.1 for an overview (in which some publications appear more than once because they contain multiple studies covering more than one domain).

Table 4.1. Domains of data and tasks used in the reviewed studies.

Domain	Studies
Air defense, aviation, marine, military	Andre and Cutler (1998), Bisantz et al. (2011), Brolese and Huf (2006), Hope and Hunter (2007a), Kirschenbaum and Arruda (1994), Kirschenbaum et al. (2014), Kobus, Proctor, and Holste (2001), Riveiro, Helldin, and Falkman (2014a), Riveiro et al. (2014b), Shattuck, Miller, and Kemmerer (2009), St. John, Callan, and Proctor (2000)
Land-use planning, spatial planning, urban planning	Aerts, Clarke, and Keuper (2003), Hope and Hunter (2007b), Leitner and Buttenfield (2000), Senaratne et al. (2012)
Remote sensing, land cover/land use	Senaratne et al. (2012)
Health research, medical imaging, disease reporting	Severtson and Myers (2013), Simpson et al. (2006)
Environmental modeling, water management, hydrology, floodplain mapping, climate change, soil mapping, geology	Cliburn et al. (2002), Deitrick and Edsall (2006), Deitrick (2012), Gerharz and Pebesma (2009), Retchless (2012), Roth (2009b), Pyysalo and Oksanen (2014), Senaratne et al. (2012), Viard, Caumon, and Levy (2011)
Natural hazard management	Ash, Schumann, and Bowser (2014), Cox, House, and Lindell (2013)
Meteorology, weather forecasting	Nadav-Greenberg, Joslyn, and Taing (2008), Scholz and Lu (2014)
Stock trading	Bisantz, Marsiglio, and Munch (2005)
Autonomous driving	Riveiro, Helldin, and Falkman (2014a)
Engineering design	Dong and Hayes (2012)
Server operations, online game reporting, twitter analytics	Fisher et al. (2012)
No specific domain	Ferreira, Fisher, and König (2014), Finger and Bisantz (2002)

4.3.4 Participants

The studies regarded in this review show great variation in the numbers of participants (Figure 4.3). Reports on two lab studies and one web-based study did not contain the number and do not appear in the chart. The first finding is that lab studies and those conducted via the web mainly include less than 40 participants but that there are also a number of studies (10 out of 43) including more than 80 participants. The highest numbers of participants were 826 in web-based studies (Severtson and Myers 2013) and 501 in lab studies (Ash, Schumann, and Bowser 2014). Opposed to this, the three interview-based studies naturally include a low number of participants (3, 6, and 10).

Most studies recruit people from the university context, typically students or faculty. The participants' type and level of expertise is usually only vaguely defined (e.g., undergrads are taken as non-experts and grad students as experts) or even not reported at all. One of the exceptions is the study by Roth (2009b), measuring the effect of expertise on risk assessment. The author distinguished between domain and map-use expertise, both characterized by three realms of expertise: education/training, work experience, and self-reporting. Refer to the "Expertise" subsection in "Findings" for a discussion of the role of expertise in the studies reviewed here.

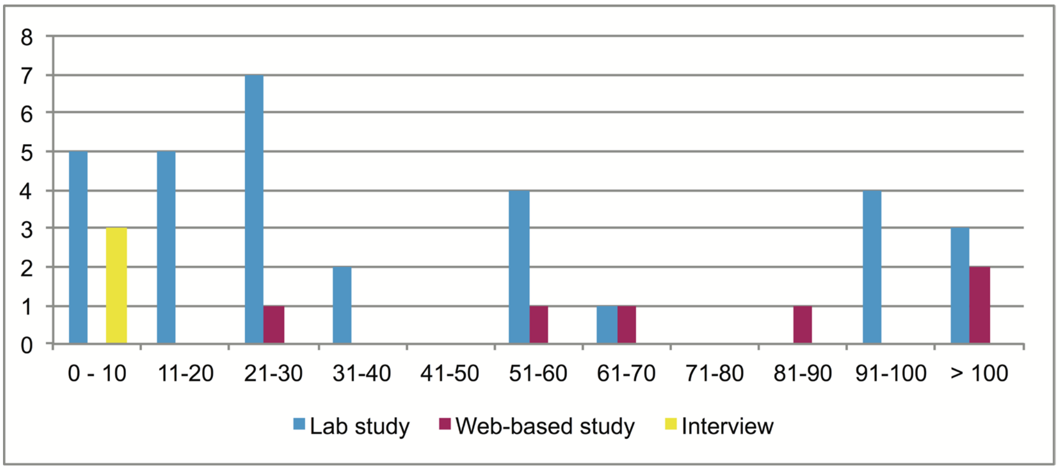

Figure 4.3. Number of participants by study type.

4.3.5 Tasks

The studies in the first review paper focusing on communication of uncertainty (Kinkeldey, MacEachren, and Schiewe 2014) included rather basic tasks such as value retrieval, value comparison, or ranking of data and/or uncertainty values. Compared to this, many studies assessing the effect of uncertainty utilize more complex tasks, i.e., tasks that relate to the analysis of data and associated uncertainty. We identified two main categories: the first one includes tasks that focus on whether or not communicating uncertainty results in more accurate and/or faster decisions, or otherwise different decision outcomes (*objective assessment*). The second category contains tasks to assess if the participants' perceptions (e.g., of risk) are different when uncertainty is incorporated or if user confidence with a decision changes, typically measured using self-report methods (*subjective assessment*). In the following subsections, we summarize tasks from the studies under review, separated by category.

Objective assessment

This category includes tasks to either assess user performance with the help of a correct solution or compare the answers of participants under different conditions (typically with and without uncertainty or comparing different techniques to portray uncertainty). The intention is to assess effects on decision-making in an objective way, e.g., if decision performance or the outcome of decisions are affected by information about uncertainty.

A commonly used decision task is *selection* from a number of options. Based on data and associated uncertainty, study participants have to decide which alternative is the optimal one. A typical example is location selection, a prototypical spatial decision that maps are used for. In this task, map users interpret information relevant to a decision (e.g., land parcel suitability for a facility) and select the "best" location. For instance, Leitner and Buttenfield (2000) asked their participants to decide where a new park or airport should be sited. This decision is made based on a map showing wetland areas and related uncertainty. Several studies include similar tasks (Aerts, Clarke, and Keuper 2003; Gerharz and Pebesma 2009; Scholz and Lu 2014; Viard, Caumon, and Levy 2011). However, although these tasks incorporate maps, location plays no or a minor role – decisions are solely or at least primarily based on attribute values and their uncertainty. An example is the study by Scholz and Lu (2014): a decision had to be made where to plant a crop (with a defined annual precipitation need). This was based on two uncertain precipitation values while their location was not important for the task (Figure 4.4).

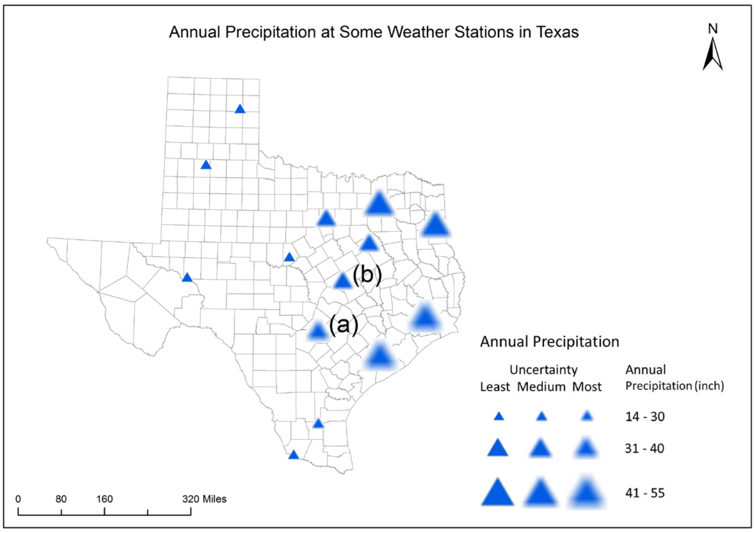

Figure 4.4. Selection task between two locations (a) and (b) by precipitation and related uncertainty [reprinted with permission from Scholz and Lu (2014)].

Only a few studies prompted participants to consider the geographical context in terms of how suitability of specific locations relates to their surroundings. An example is the above mentioned Leitner and Buttenfield (2000) study in which a number of planning criteria had to be evaluated for different locations in the map. Other selection tasks have a stronger spatial component. For instance, in a study by Hope and Hunter (2007a), one of the tasks was to select the estimated position of a boat relative to an uncertain boundary (from a predefined list of verbal descriptions). Other tasks of this kind included selection of the area in which a point lies based on uncertain boundaries (Pyysalo and Oksanen 2014) or selection of a target with uncertain location that seems closest to a marked position (Brolese and Huf 2006). However, in the studies reviewed here, selection tasks with a spatial component are rare.

An extension of selection tasks are *ranking* tasks, i.e., sorting a number of items by uncertain attributes. For instance, in the study by Deitrick and Edsall (2006), participants ranked different regions according to predicted water consumption and connected uncertainty (after aggregation of values over an area). In the study by Retchless (2012), one of the tasks was ranking of regions from most to least extensive harm and by suitability for a natural reserve (Figure 4.5). Similar examples can be found in Deitrick (2012), Hope and Hunter (2007b), and Viard, Caumon, and Levy (2011). Ranking tasks are more complex than those from the "selection" category because not only must the best option be found, but a number of options have to be prioritized.

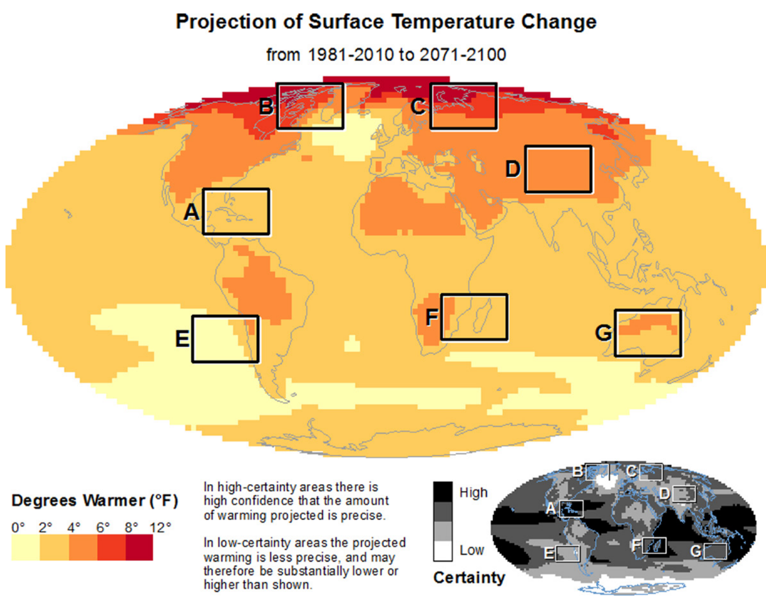

Figure 4.5. Ranking of regions A to G by environmental harm and by suitability for a natural reserve based on uncertain temperature predictions [reprinted with permission from Retchless (2012)].

Some studies use *game-like* tasks that are inspired by actions occurring in games, e.g., stock trading with virtual money or shooting of targets, often in a dynamic environment and under time pressure. For instance, in the second study presented by Finger and Bisantz (2002), participants conducted an identification task in which the status of moving objects with uncertain identity had to be determined. The objects moved on the screen and their identity changed dynamically; this and the fact that the goal was a high score in a certain time as well as the limited number of decisions (with a penalty for wrong guesses) make the study resemble a game. Similar tasks are used in several studies including dynamic environments (Andre and Cutler 1998; Bisantz, Marsiglio, and Munch 2005; Bisantz et al. 2011; Hope and Hunter 2007a; Riveiro, Helldin, and Falkman 2014a) and in one study using a static setup (Cox, House, and Lindell 2013).

A number of tasks of this category have a spatial component, e.g., Cox, House, and Lindell (2013) included a decision about the likely path of a storm (Figure 4.6). Further examples can be found in Andre and Cutler (1998) and Bisantz et al. (2011).

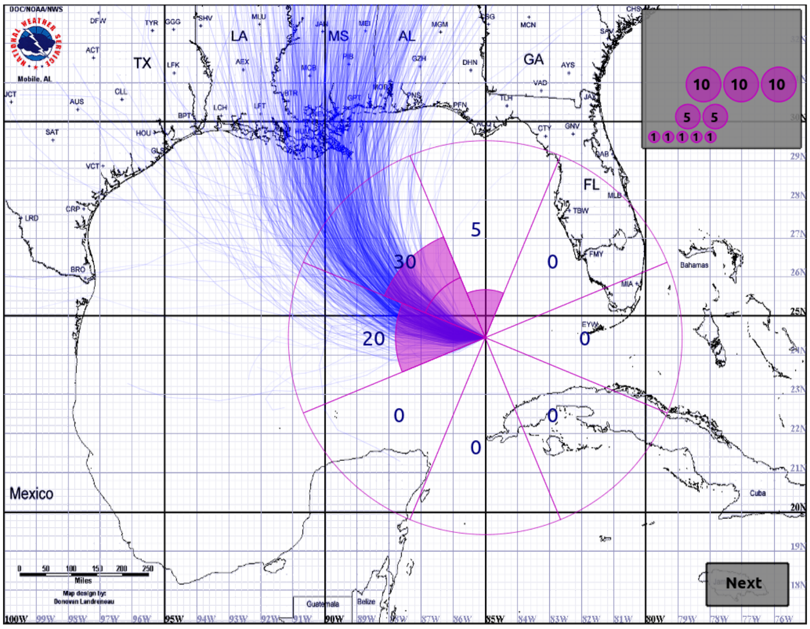

Figure 4.6. Interface to assess user confidence using a game-like task. Participants bet virtual money on directions of the storm they see as most likely [reprinted with permission from Cox, House, and Lindell (2013)].

Game-like tasks are also used to assess user confidence "indirectly" (instead of directly asking for the participants' confidence, see "subjective assessment" below), e.g., by the amount of stock purchases (Bisantz, Marsiglio, and Munch 2005) or of virtual money participants bet on a decision (Cox, House, and Lindell 2013, Figure 4.6).

Compared to the tasks we mentioned so far that remained relatively simple, some studies incorporate more complex *real-world* tasks from different domains. Although many of them resemble game-like tasks, the difference is that they occur (at least in a similar fashion) during practical work in the respective domain. Some of the real-world tasks are aspatial, e.g., in the second study described in Riveiro, Helldin, and Falkman (2014a), in a scenario dealing with autonomous driving, the task was to decide when to take over control of a car. This decision was based on a graphical representation of the car's ability to drive autonomously, including related uncertainty. Other tasks belonging to this category include decisions whether or not to follow a particular query sequence to dig for more information in the data (Fisher et al. 2012) or decisions between different design alternatives based on uncertain criteria (Dong and Hayes 2012).

Other studies utilize real-world tasks with spatial components. In the study by Nadav-Greenberg, Joslyn, and Taing (2008), participants from the weather forecasting domain were asked to make a decision about whether to issue a high wind weather warning based on uncertain wind speed values from a map and/or a box plot. Another example for a spatial decision task, which is taken from the medical domain, includes simulated drilling of a bone to remove a tumor with the help of a display depicting the uncertain path of the instrument (Simpson et al. 2006). A number of tasks with strong spatial components occur in military studies, e.g., selecting tactical leverage points from a list of locations and deciding between different courses of action (Kobus, Proctor, and Holste 2001). Further examples for this task type can be found in Riveiro et al. (2014b), Kirschenbaum and Arruda (1994), Kirschenbaum et al. (2014), Kobus, Proctor, and Holste (2001), Shattuck, Miller, and Kemmerer (2009), and St. John, Callan, and Proctor (2000).

Subjective assessment

As mentioned above, this category includes tasks that measure the impact of visually depicted uncertainty on subjective assessment, such as tasks to assess risk perception or user confidence. In the reviewed studies, these tasks were rarely used compared to tasks for objective assessment.

An example for this category is Roth's (2009b) study in which participants expressed perceived flood risk for different sites in a map showing different possible delineations of a floodplain (implicitly depicting uncertainty). Similar tasks in other studies were ranking of a number of areas on a map by the likelihood of harm to regional ecosystems (Retchless 2012), expressing perceived fear when participants imagine living at a certain location and the likelihood that they would take protective actions (Ash, Schumann, and Bowser 2014), or judging risk beliefs, emotion, and perceived ambiguity associated to locations on cancer risk maps (Severtson and Myers 2013).

Using subjective assessment, user confidence can be assessed "directly" using a Likert scale or percentage (Brolese and Huf 2006; Deitrick and Edsall 2006; Dong and Hayes 2012; Ferreira, Fisher, and König 2014; Kirschenbaum and Arruda 1994; Riveiro et al. 2014b; Roth 2009b). This is much more common than the "indirect" way to measure confidence mentioned above with respect to objective assessment.

All in all, the complexity of tasks to assess decision-making under uncertainty varies considerably, from a simple decision between two attributes and their uncertainty, to more complex decisions related to identifying moving targets with uncertain identity in a dynamic environment, to tasks from real-world applications. Furthermore, we can distinguish between aspatial decision tasks and those that have a spatial component. Most tasks are utilized for objective assessment of effects, whereas a lower number of tasks assess impacts in a subjective way. Refer to the "Tasks" subsection in "Lessons learned" for a further discussion of the role of tasks for uncertainty visualization assessment.

4.4 Findings

This section synthesizes findings from the studies under review, divided into subsections addressing the following main topics: acceptance of uncertainty, effects of visual communication of uncertainty, the role of uncertainty visualization techniques, and impacts of user expertise.

4.4.1 Acceptance

An important question addressed in multiple studies is whether users are able and willing to use uncertainty information for their decisions if it is available. First, adding depictions of uncertainty means that more information has to be processed when making a decision. Second, uncertainty is often quantified by measures that are difficult to understand, e.g., statistical measures of probability (Micallef, Dragicevic, and Fekete 2012). Even if information about uncertainty may be advantageous for decision makers, they have to accept the additional effort of incorporating it into decision-making.

Several studies provide evidence that users (in some contexts) find visual depiction of uncertainty useful. Aerts, Clarke, and Keuper (2003), for example, report that the majority of their participants (planners and decision makers) agreed that "[u]ncertainty visualization improves the decision-makers' view, analyses and model simulations" (258). Other studies also report positive reactions from participants towards visual depiction of uncertainty (Gerharz and Pebesma 2009; Pyysalo and Oksanen 2014; Scholz and Lu 2014). In contrast, there are studies reporting less positive results with respect to acceptance of uncertainty visualization. For example, participants in the Deitrick and Edsall (2006) study remarked that uncertainty information had an influence on their decisions but that "they would feel more confident if they had other data sources in addition to the uncertainty/certainty maps" (735). Our interpretation of this statement is that they did not see the information provided on uncertainty as sufficient for making confident decisions. Similarly, Cliburn et al. (2002) found that decision makers were able to work with the uncertainty visualization offered by a decision-support system but that they were skeptical of its value. They raised basic concerns about advisability of communicating uncertainty, e.g., that model outputs could be negated by this information. Further concerns were that decision makers needed additional information for interpreting and coping with uncertainty (e.g., when a high degree of uncertainty is a problem and when not) and that decision makers often have little time to explore uncertainty in the data. Based on these findings, the authors contend that well-crafted visualization methods alone may not be sufficient to support decision makers and that suitable strategies may be needed to utilize uncertainty information ("they need to know what choice is best", 948).

The findings show that not all participants in the reviewed studies found uncertainty visualization helpful. However, the question of how to establish sufficient acceptance cannot be fully answered without further studies assessing the utility of uncertainty depictions in practical work, with a special focus on the question of how to support users with different roles and expertise while making decisions based on uncertain data.

4.4.2 Effects of uncertainty visualization

When assessing effects of uncertainty visualization, most reviewed studies focus on decision performance, i.e., they measure decision accuracy and/or speed, often supplemented by measurement of decision confidence of participants. Only a small number of studies assess other aspects such as effects on the outcome of decisions or on how decisions are made. In the remainder of this subsection, we provide a summary of findings concerning the three main effect types: effects on decision performance, decision outcome, and decision confidence.

Effects on decision performance

In the following, we summarize results regarding decision performance, i.e., if decision accuracy or speed is different with and without uncertainty displayed. For this, the common strategy is to define a set of "correct" decisions (for assessing decision accuracy) and to measure the time needed to make the decision (for decision speed).

The general finding is that *most studies suggest evidence for effects on decision performance caused by visual uncertainty depiction*. An example for positive effects with respect to decision-making accuracy is the study by Leitner and Buttenfield (2000), since the availability of uncertainty information increased the number of correct siting decisions. This goes in line with findings by Simpson et al. (2006) who found that the visualization of uncertainty in a simulated surgical task significantly lowered the number of attempts needed to successfully locate a tumor. In addition, it decreased the number of incomplete tasks, i.e., fewer participants gave up on the task. Supporting results were also reported by Riveiro et al. (2014b) in a target identification scenario, where the participants aided by uncertainty representations needed significantly fewer attempts to make a final identification, even though the performance of both groups (with regard to correct identifications) was very similar. Another example of studies suggesting positive effects of uncertainty visualization was reported by Andre and Cutler (1998). In a flight task simulation, participants showed improved decision accuracy with different representations of positional uncertainty, compared to the "no uncertainty" baseline.

In contrast to the studies above, there are also findings suggesting negative effects of uncertainty representations on decision accuracy. In one such study by Hope and Hunter (2007b), participants were to select the most suitable site for a new airport. In cases when the optimal choice would have been the zone with the higher degree of uncertainty (e.g., when both zones have the lowest suitability), they still tended to decide for the option with low uncertainty, thus making an unreasonable decision. In addition, they showed a tendency to opt for "no preference" in the high uncertainty condition. The authors explained this outcome by loss aversion that may have caused a bias towards less uncertain information and state that "the inclusion of uncertainty information may therefore lead to irrational decisions" (212).

Besides decision accuracy, a question addressed in several studies is if there is an impact of uncertainty visualization on the time spent to make a decision, i.e., *decision-making speed.* Leitner and Buttenfield (2000) expected that an increase of map complexity resulting from the addition of uncertainty to the display would make decisions slower, but they found no significant differences in speed. This result is in line with findings by Riveiro et al. (2014b) who found no significant differences regarding response times between groups with or without uncertainty depictions. Shattuck, Miller, and Kemmerer (2009) observed that ambiguous and missing information led to slower decisions than completely certain information but not when uncertainty in terms of conflicting information was introduced; thus, only certain types of uncertainty slowed down decision-making. The above cited study by Andre and Cutler (1998) even reports upon reduced decision times through uncertainty visualization: they found that under highly uncertain conditions, participants with graphical depiction of uncertainty made more correct decisions in the same time than the group without uncertainty representation.

Overall, based on studies reviewed, uncertainty visualization has tended to result in a positive effect on decision accuracy. The evidence is less clear for decision speed, but it could be observed that usually, uncertainty visualization does not slow down decision-making and in one case it even decreased the decision time. The above mentioned findings on irrational decisions under uncertainty suggest that in order to use data and related uncertainty effectively, users must know how to interpret data together with related uncertainty, an issue we address in the "Expertise" subsection.

Effects on decision outcome

A small proportion of the studies under review assessed relative differences with respect to an outcome without defining correct decision results. Some show that adding uncertainty visualization can lead to different decision results than in the deterministic case when only the data are depicted. This was the case in the study by Deitrick and Edsall (2006) about the impact of uncertainty visualization on ranking of regions by predicted water consumption (based on a map including uncertainty). Rankings were significantly different based on maps with and without uncertainty but the kind of difference was not interpreted further. In a later study reinforcing the initial findings, Deitrick (2012) again found a significant difference in rankings of policy options from most to least robust for three options: a display without uncertainty, one depicting uncertainty implicitly (by showing all possible outcomes) and one signifying the degree of uncertainty explicitly (using variation of transparency). Once more, there was no further description of the differences. A further study reporting on effects on decision outcomes deals with uncertainty-aware drainage divides (Pyysalo and Oksanen 2014). Nearly 60% of the participants changed their answers after they were provided with information on uncertainty. Viard, Caumon, and Levy (2011) also report differences in decisions with and without uncertainty, but just for one of the tasks, a comparison of two uncertain locations.

The findings summarized here suggest that communicating uncertainty can have different kinds of effects on decision outcomes. However, more research is needed on how and why decisions change when uncertainty comes into play, an issue we discuss further in the "Lessons learned" section.

Effects on decision confidence

Multiple studies assess whether confidence in decision-making is influenced by uncertainty portrayal and how this relates to user performance. A frequent hypothesis is that users make decisions with higher confidence when they are informed about uncertainty. In contrast to this assumption, most studies investigating the impact of uncertainty signification on confidence did not measure a significant difference in confidence levels when uncertainty was added to the display (Deitrick and Edsall 2006; Dong and Hayes 2012; Kirschenbaum and Arruda 1994; Leitner and Buttenfield 2000; Riveiro et al. 2014b). For instance, in the Leitner and Buttenfield (2000) study, the self-assessed confidence of participants was uniformly high whether or not uncertainty was displayed. In the same study, decision accuracy was significantly higher with uncertainty depicted, meaning that user performance and confidence did not correspond.

Opposed to this, Fisher et al. (2012) reported that, "even relatively simple representations of uncertainty using error bars progressively updating over time, allowed analysts to trust their decision points" (1681f). In other words, there was a positive effect of uncertainty on user confidence when they made decisions based on uncertain (aspatial) data. Ferreira, Fisher, and König (2014) partly supported this observation: they compared decisions that were made based on bar charts without the use of uncertainty visualization to those based on charts with static error bars and “interactive annotations” (with uncertainty provided on demand). For the latter technique, they not only measured higher confidence in the answers than for the other two conditions, but answers with high confidence were also more likely to be right.

All in all, there is little evidence that uncertainty portrayal can have a positive impact on user confidence when decisions are made – most studies found no effects and effects tended to be small when found at all. This goes in line with findings from the complementary review (Kinkeldey, MacEachren, and Schiewe 2014) suggesting that, related to map reading incorporating uncertainty, user confidence can only be increased when well-crafted visualization techniques are involved that are tailored to the tasks they are used for. The findings presented here suggest that this is also true, and maybe even more important, for decision-making tasks.

4.4.3 Uncertainty visualization techniques

As outlined above, a basic question in uncertainty evaluation is whether displaying uncertainty has any effect on reasoning or decision-making based on uncertain data. When the answer is yes, the follow-on question is if different visualization techniques invoke different kinds or magnitudes of effects. A number of studies have investigated this aspect by varying display techniques for uncertainty and assessing the effect. Most studies did not only compare graphical representations of uncertainty to each other, but included numeric and/or textual depictions as well. The two subsections below focus on findings regarding decision-making and risk assessment.

Decision-making

Some studies did not suggest significant differences between representation types regarding the impact on decision-making. For instance, in Bisantz, Marsiglio, and Munch (2005) using a dynamic stock trading task the authors did not observe a significant influence of the type of uncertainty representation on user performance. Two different graphical representations (resolution, colored icons) yielded comparable results as numeric and linguistic expressions.

However, some studies support the hypothesis that the effect on decision-making is influenced by different uncertainty depictions. In the Hope and Hunter (2007a) study, a task was to stop a ship before it crosses an uncertain boundary defining a restricted territory. The boundary was represented in four ways: graphically, either as a 99% probability range (“limits”) or as graduated color, as well as non-graphically, as textual or numeric description. When the “limits” representation was used, the participants stopped the ship significantly earlier than with the other techniques, i.e., decision speed was higher. Finger and Bisantz (2002) also suggested differences while comparing different representations of uncertain identity of objects (degraded/non-degraded icons with and without uncertainty as numeric value). They found no significant impact of the display condition on the accuracy of decisions, but decision speed was higher for the condition with uncertainty depicted graphically. Participants that additionally had access to numeric values were more conservative and tended to wait for more certain information before making the decision. Andre and Cutler (1998), cited above, support the observation that graphical depictions can lead to higher decision performance than textual representations. The authors used two graphical representations of positional uncertainty (a color scheme and a circle glyph depicting the uncertainty of its position) and a numeric depiction. Compared to the "no uncertainty" baseline representation, they found improved decision performance (accuracy and time) with all three representations, while the highest increase was observed with the extrinsic technique using circle glyphs. In a similar study, Kirschenbaum and Arruda (1994) observed higher accuracy in range estimates using a graphic display of uncertain target movements (error ellipse) compared to a verbal description of uncertainty. However, they found that the difference between the two depictions was only significant under high uncertainty conditions.

In one study, a dynamic approach was more successful than a static one: Bisantz et al. (2011) compared different representations of the uncertain identity of planes, birds, and missiles as icons. Task performance was highest when the participants were able to switch among different display conditions ("toggling"). Apart from that, the additional display of numeric values did not improve decision performance – supporting the findings that graphical displays of uncertainty can be advantageous compared to numerical representations.

Another interesting finding provided by Viard, Caumon, and Levy (2011) suggested differences in decision performance between coincident and adjacent displays, but not for all the tasks they included. They found that with a more complex ranking task, the adjacent display resulted in less accurate judgments than the coincident representation; but with a simple comparison task, the differences were non-significant. This suggests that the strengths of well-designed uncertainty visualizations may only matter with decisions that are not too simplistic, an aspect further discussed the "Lessons learned" section.

Two studies suggest effects towards bias in participants' decisions: Nadav-Greenberg, Joslyn, and Taing (2008) observed that different visualization approaches resulted in different forecasts in terms of wind speed. The most important finding was that a "worst case" map that showed the upper boundary of projected wind speeds yielded biased forecasts with higher speeds than those based on the other displays. Thus, the authors give the warning that in a real-world setting providing worst-case maps could lead to more false alarms. They explain this effect with evidence from past research that anchors (in this case the depiction of the worst case) unconsciously influence people's judgments (Chapman and Johnson 2002). Similar results were provided by Riveiro et al. (2014b) in a target identification experiment, where an expert group aided by uncertainty visualization selected higher priority values and more hostile and suspect identities. This suggests that, when safety was an issue, the experts put themselves in the "worst-case scenario" in the presence of uncertainty.

Risk assessment

Regarding the effect of different depictions of uncertainty on risk assessment, two studies suggested a measurable impact. Severtson and Myers (2013) assessed the effect of uncertainty on risk beliefs and perceived ambiguity, utilizing maps that showed distributions of cancer risk from air pollution. Unclassed contours and verbal risk expressions generated higher perceived ambiguity while classed contours generated stronger risk beliefs for high risk levels and weaker beliefs for low risk. Similarly, Ash, Schumann, and Bowser (2014) reported on impacts of the display type on perceived fear. They compared the effect of three different designs of tornado warning: the original "cone of uncertainty", i.e., the boundary representing a zone of elevated probability of occurrence of a tornado (deterministic design), the same shape with a red gradient fill depicting different zones of probability (probabilistic design) and a third one with a spectral color scheme (Figure 4.7). They observed the highest fear and intent of protective action with the original design, the second highest levels with the red gradient probabilistic design, and the lowest levels with the gradient using spectral colors. Based upon application of a Mann–Whitney U test of differences between pairs of designs, the authors even report a significant difference in the perceived fear between the two gradient designs that only differed in their color scheme.

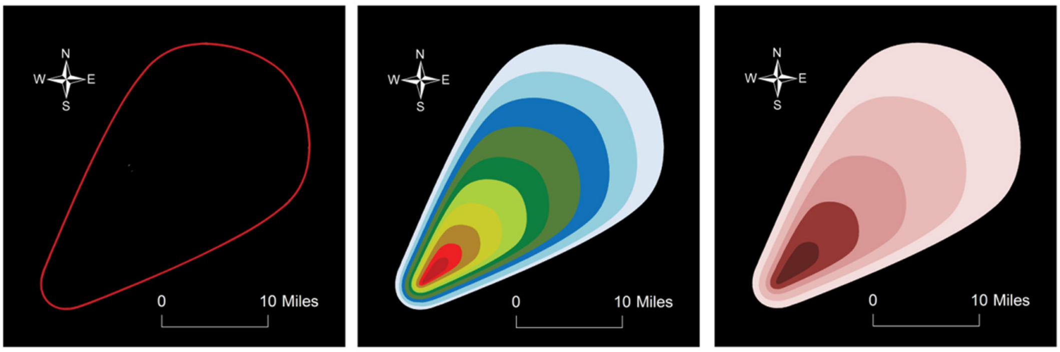

Figure 4.7. Three different visual depictions of uncertainty in Ash, Schumann, and Bowser's (2013) study: original (left), spectral (center) and red gradient (right) [reprinted with permission from Ash, Schumann, and Bowser (2014), © American Meteorological Society].

Overall, the findings summarized in this subsection suggest that the choice of uncertainty representations can make a difference in terms of decision accuracy and speed, as well as bias and risk perception. Moreover, there is evidence that with respect to decision performance, graphical representations can be advantageous over numeric and textual representations. For a discussion of these findings, refer to the "Lessons learned" section.

4.4.4 Expertise

From cognitive research we know that expertise can play an important role when decisions are made based on uncertain data (Klein 1999). For this reason, many studies we reviewed involve groups of participants with varying type and/or level of expertise. There are several studies that found no or non-significant differences with respect to user performance between groups with different expertise levels (Aerts, Clarke, and Keuper 2003; Hope and Hunter 2007a; Nadav-Greenberg, Joslyn, and Taing 2008; Scholz and Lu 2014). However, a number of studies evaluating this aspect measured different kinds of effects. The study by Roth (2009b), for instance, focused on the effect of user expertise on the aspects of risk perception, perceived assessment difficulty, and assessment confidence in a floodplain mapping scenario. The most important result was that expertise had effects regarding all three aspects. Generally, novices tended to underestimate flooding risk, accompanied by high perceived difficulty and low confidence in the assessment. Experts showed effects depending on the type of expertise, for instance, experienced map users found the task less difficult and were more confident with their decisions, but they still underestimated the risk because of the lack of domain expertise. The best results were achieved by participants with domain expertise as well as high degree of experience with map use, thus, both types of expertise played a role in this context. Another study worth mentioning, conducted by Hope and Hunter (2007b), showed that participants with a high level of experience had the strongest bias towards selecting areas of low uncertainty, regardless of the degree of land suitability. Their conclusion was that "many people, including those who consider themselves to be experienced in working with spatial information, do not have an intuitive understanding of how to handle uncertainty in decision-making" (199), which confirms the suggestion made above that it is important to provide users with strategies to incorporate uncertainty into their work.

Other studies showed that the participants' expertise can also affect the way decisions are made, as observed by Kobus, Proctor, and Holste (2001) in a realistic scenario from the military context. Under high uncertainty conditions, experienced officers took more time than less experienced to become aware of the situation, but after that, their decisions were more accurate and faster than those of a group with less experience. This confirms the findings by St. John, Callan, and Proctor (2000) who report that under high uncertainty, the less-experienced officers (with respect to command-post experience) were more likely to choose a "wait-and-see" response than the more experienced officers who were not affected by uncertainty.

Two studies from our selection suggested differences in the influence of the display type depending on the level of expertise. The first one was the weather forecasting study by Nadav-Greenberg, Joslyn, and Taing (2008) mentioned above. Although there were no significant differences between the low- and high-expertise groups, the authors suggested that experts were less influenced by different visual depictions of uncertainty than non-experts (they "did not appear to be influenced by visualization", 39). In the other study, reported by Kirschenbaum et al. (2014), the expert group was very accurate in their decisions regardless of the display type, whereas the non-expert group was significantly more accurate using the geographic display than with the other display types. The authors' hypothesis was that the experts were so well trained that they were capable of conducting the task with the information given, no matter how it is communicated.

Expertise has also been shown to play a role for user preference towards uncertainty visualization techniques. For example, Cliburn et al. (2002), in the usability engineering study cited above, observed differences in preference between the interviewees with domain expertise (who preferred the extrinsic glyph-based approach) and decision makers with less experience in the sciences (who preferred the intrinsic RGB coloring technique). This observation may be explained by different tasks the two groups tried to accomplish and provides evidence for the importance to consider tasks for the development of suitable uncertainty visualization techniques (a discussion included in the complementary review paper, Kinkeldey, MacEachren, and Schiewe 2014).

4.5 Lessons Learned

In the following section, we derive "lessons learned" and recommendations from the summary of findings, related to study focus, study design, evaluation methodology, effects, choice of tasks, expertise, and decision-making research. Based on the findings, we developed a categorization of research foci directed to the effect of uncertainty visualization on decision-making (Table 4.2). It includes the main categories of research foci we identified: uncertainty, visualization, methodology, participants, tasks, and effects, further divided into subcategories. In order to give an overview on what topics have been covered so far, we list the studies covering each subcategory. Furthermore, we suggest open research topics for each main category to highlight possible future work in the field.

Table 4.2. Categorical overview: evaluating the effect of uncertainty visualization on decision-making. Several studies contain more than one study and may appear in multiple subcategories.

Categories			**Studies**	**Open evaluation research topics**
Uncertainty	Type	Attribute / thematic	Majority of studies involve attribute uncertainty	Visualization techniques to display temporal uncertainty Visualization techniques to display multiple types of uncertainty Role of uncertainty classification (continuous / classed, number of classes) for different tasks
		Positional / geometric	Andre and Cutler (1998), Ash, Schumann, and Bowser (2014), Brolese and Huf (2006), Cox, House, and Lindell (2013), Hope and Hunter (2007a), Kirschenbaum and Arruda (1994), Kirschenbaum et al. (2014), Pyysalo and Oksanen (2014), Simpson et al. (2006), St. John, Callan, and Proctor (2000)	
		Temporal	None	
		Multiple	Kobus, Proctor, and Holste (2001), Riveiro, Helldin, and Falkman (2014a), Riveiro et al. (2014b), Roth (2009b), Shattuck, Miller, and Kemmerer (2009)	
Visualization	Display strategy	Intrinsic	Majority of studies involve intrinsic visualizations	Development of task-based typologies for systematic selection of visualization techniques Use of intrinsic and extrinsic depictions for different tasks Use of adjacent views for different tasks Use of dynamic (animated and/or interactive) views for different tasks Impact of visualization technique on decisions (bias, outcome etc.) Interaction between graphical, numerical & textual representation of uncertainty
		Extrinsic	Brolese and Huf (2006), Ferreira, Fisher, and König (2014), Fisher et al. (2012), Hope and Hunter (2007b), Kirschenbaum and Arruda (1994), Kirschenbaum et al. (2014), Riveiro, Helldin, and Falkman (2014a), Riveiro et al. (2014b), Simpson et al. (2006), St. John, Callan, and Proctor (2000)	
		Intrinsic vs. extrinsic	Andre and Cutler (1998), Cliburn et al. (2002), Hope and Hunter (2007a), Pyysalo and Oksanen (2014), Senaratne et al. (2012)	
	Display organization	Adjacent	None	
		Coincident	Majority of studies involve coincident visualizations	
		Adjacent vs. coincident	Aerts, Clarke, and Keuper (2003), Gerharz and Pebesma (2009), Retchless (2012), Senaratne et al. (2012), Viard, Caumon, and Levy (2011)	
	Display dynamics	Static	Majority of studies involve static visualizations	
		Dynamic	None	
		Static vs. dynamic	Aerts, Clarke, and Keuper (2003), Bisantz et al. (2011), Ferreira, Fisher, and König (2014), Gerharz and Pebesma (2009), Senaratne et al. (2012)	

Methodology	Research methods	Quantitative methods	Most studies apply quantitative methods	Studies using qualitative and mixed methods to better understand how and why impacts of uncertainty visualization occur
		Qualitative methods	Fisher et al. (2012)	Qualitative studies to identify the most important questions to focus on in (quantitative) user studies
		Mixed Methods	Cliburn et al. (2002), Deitrick and Edsall (2006), Gerharz and Pebesma (2009)	Field research involving observation of the role of uncertainty visualization in real-world tasks
Participants	Type of expertise	Map use	Roth (2009b), Senaratne et al. (2012)	Systematic definition of type and level of expertise Expertise in uncertainty measures and statistics Role of expertise in visually-supported reasoning and decision making under uncertainty
		Statistics / uncertainty measures	Gerharz and Pebesma (2009), Senaratne et al. (2012)	
		Domain expertise	Cliburn et al. (2002), Kobus, Proctor, and Holste (2001), Pyysalo and Oksanen (2014), Roth (2009b), Senaratne et al. (2012), Shattuck, Miller, and Kemmerer (2009), St. John, Callan, and Proctor (2000)	
	Level of expertise	Novices	Bisantz, Marsiglio, and Munch (2005), Brolese and Huf (2006), Leitner and Buttenfield (2000), Severtson and Myers (2013), Simpson et al. (2006)	
		Experts	Ash, Schumann, and Bowser (2014), Deitrick (2012), Ferreira, Fisher, and König (2014), Fisher et al. (2012), Gerharz and Pebesma (2009), Kirschenbaum and Arruda (1994), Pyysalo and Oksanen (2014), Riveiro, Helldin, and Falkman (2014a), Riveiro et al. (2014b), Shattuck, Miller, and Kemmerer (2009)	
		Novices vs. experts	Aerts, Clarke, and Keuper (2003), Cliburn et al. (2002), Deitrick and Edsall (2006), Dong and Hayes (2012), Hope and Hunter (2007a), Hope and Hunter (2007b), Kirschenbaum et al. (2014), Kobus, Proctor, and Holste (2001), Nadav-Greenberg, Joslyn, and Taing (2008), Roth (2009b), Scholz and Lu (2014), Senaratne et al. (2012), St. John, Callan, and Proctor (2000), Viard, Caumon, and Levy (2011)	
Tasks	Type	Selection	Aerts, Clarke, and Keuper (2003), Brolese and Huf (2006), Gerharz and Pebesma (2009), Hope and Hunter (2007a), Leitner and Buttenfield (2000), Pyysalo and Oksanen (2014), Scholz and Lu (2014), Viard, Caumon, and Levy (2011)	More complex decision tasks, e.g., multi-criteria decisions requiring external knowledge Tasks with spatial actions, e.g.,

		Ranking	Deitrick (2012), Deitrick and Edsall (2006), Hope and Hunter (2007b), Retchless (2012), Viard, Caumon, and Levy (2011)	drawing boundaries Further development of game-like tasks to motivate participants and create time pressure Integration of decision theory into evaluation of geodata uncertainty visualization
		Game-like	Andre and Cutler (1998), Bisantz, Marsiglio, and Munch (2005), Bisantz et al. (2011), Cox, House, and Lindell (2013), Finger and Bisantz (2002), Hope and Hunter (2007a), Riveiro, Helldin, and Falkman (2014a)	
		Real-world	Dong and Hayes (2012), Fisher et al. (2012), Kirschenbaum and Arruda (1994), Kirschenbaum et al. (2014), Kobus, Proctor, and Holste (2001), Nadav-Greenberg, Joslyn, and Taing (2008), Riveiro, Helldin, and Falkman (2014a), Riveiro et al. (2014b), Shattuck, Miller, and Kemmerer (2009), Simpson et al. (2006), St. John, Callan, and Proctor (2000)	
Effect	Type (effect on…)	Decision performance	Most studies focus on decision performance	Effects of uncertainty visualization on reasoning with geodata Studies observing other effect types, e.g., how and why decisions change when uncertainty comes into play Impact of uncertainty visualization on decision confidence
		Decision outcome	Deitrick (2012), Deitrick and Edsall (2006), Pyysalo and Oksanen (2014), Viard, Caumon, and Levy (2011)	
		How a decision is made	Bisantz, Marsiglio, and Munch (2005), Bisantz et al. (2011), Dong and Hayes (2012), Finger and Bisantz (2002), Shattuck, Miller, and Kemmerer (2009)	
		Decision confidence	Deitrick and Edsall (2006), Dong and Hayes (2012), Ferreira, Fisher, and König (2014), Fisher et al. (2012), Kirschenbaum and Arruda (1994), Leitner and Buttenfield (2000), Riveiro et al. (2014b)	

4.5.1 Study focus

The first lesson learned is that most studies from this review on the effect of uncertainty depiction focus on decision-making, measuring impacts on decision performance, decision outcome, etc. Only a few studies assess the impact of uncertainty depiction on reasoning, i.e., on how users construct an understanding of the information in a map that is then used to arrive at an interpretation or decision. Thus, for most studies that suggest an impact of uncertainty depiction, we lack the evidence to determine why and how the impact happens, because we do not yet fully understand how uncertainty affects the steps before a decision is made. That is why we see a need for studies focusing on the role of uncertainty visualization during reasoning with geodata, in addition to those dealing with decision-making.

4.5.2 Study design

There are a number of issues regarding the design of the studies under review. The first one we address here is that most authors were not systematic when determining the "universe" of phenomena that the uncertainty signification considered by the study relates to. Thus, the data sets used are often very ad hoc and difficult to compare. Similarly, the selection of visualization techniques generally seems arbitrary and not well-founded (in relation to building a comprehensive understanding of uncertainty visualization effects), e.g., when two intrinsic techniques are compared to each other and not to others.

Another limitation of several studies, often mentioned by the authors themselves, is that the numbers of participants were relatively small, perhaps too small for a valid evaluation, e.g., 14 students in the second laboratory experiment reported in Brolese and Huf (2006). An additional issue is that statistical power was never reported, thus, the published reports provide no way to assess specifically whether the sample size was sufficient.

Some studies had specific experiment design flaws. For instance, Leitner and Buttenfield (2000) identified the lack of a training phase as the reason for a strong learning bias. Another issue they mention is that map complexity was not properly balanced in the four parts of their experiment. Both flaws may have had a negative impact on the validity of the results. Simpson et al. (2006) used the same tasks in the same order with all participants (10 tasks without uncertainty, same 10 tasks with uncertainty) and thus believed that "the major source of error in the study is that all participants were shown the two tasks in the same order, which may have favored the second task" (403). This issue can be avoided by either using a between-subjects design (one group with and one without uncertainty signification) or through balancing of order in a within-subjects design, either covering all possible orders across the participants or using a Latin Square method (Winer 1962) to systematically balance order when this is impractical.

4.5.3 Methodology

For evaluating the basic usability of uncertainty signification techniques ("Can map readers retrieve data and uncertainty values?"), both laboratory studies and web-based studies have been found to be appropriate (Kinkeldey, MacEachren, and Schiewe 2014a). But for the assessment of the effect ("Does uncertainty facilitate better informed decisions?") a combination of laboratory studies and a range of qualitative methods applied in the field may be required (mixed methods). For instance, studies by Cliburn et al. (2002), Deitrick and Edsall (2006), Fisher et al. (2012), or Gerharz and Pebesma (2009) have successfully used qualitative methods (interview and verbal protocol analysis/"think aloud"), mostly in combination with quantitative approaches. We see great potential in qualitative and mixed methods to address deeper questions of how and why particular signification methods result in positive or negative outcomes. Particularly, to investigate the effect of uncertainty signification on reasoning and decision-making, quantitative web-based studies need to be complemented with in-depth and longitudinal studies that may incorporate ethnographic, participant observation, or other methods. Those designing such studies should consider approaches developed for investigating the process of work and reasoning with information and visual analytics methods such as multidimensional in-depth long-term case studies (MILC), a method to assess information evaluation tools by documenting the usage as well as success over a long time (Shneiderman and Plaisant 2006). Another approach with the potential to provide insights about how reasoning is affected by uncertainty depictions is pair analytics, a new method developed to study reasoning with visual analytics methods and tools by including a visualization expert and a domain expert in the study (Arias-Hernandez et al. 2011).

4.5.4 Effects

The majority of studies included in this review deal with effects that visually depicted uncertainty may have on decision-making, and most of them are limited to measuring if uncertainty visualization results in an increased numbers of correct decisions in a shorter time. Thus, past research has covered just a part of the possible effects that may occur, which are at least the following differences:

1. in decision outcomes,
2. in correctness of decisions,
3. in kinds of errors made,
4. in decision time,
5. in confidence in a decision,
6. in willingness to make a decision,
7. in how much workload decision-making causes, or,
8. in how a decision is made.

Some of the effects listed, such as the difference in kinds of errors made, have not been covered in the studies reviewed here. For further studies we recommend more systematic assessment of effects, for instance, conducting interviews after quantitative studies or using “think aloud” during such studies to learn more about why there was a difference in decision-making with and without representation of uncertainty. Without this information, it is difficult to interpret the effects that were measured due to the high number of possible factors that could have caused them.

A number of studies provided evidence for differences in effects depending on the representation of uncertainty that was used, for instance, regarding decision performance or risk assessment. This suggests the important role of uncertainty representation when the goal is to improve reasoning or decision-making and avoid bias introduced by the representation (as observed with "worst case" maps in Nadav-Greenberg, Joslyn, and Taing 2008). Another important finding from studies such as Bisantz et al. (2011), Finger and Bisantz (2002), or Hope and Hunter (2007a) is that graphical representations of uncertainty generally have advantages over numerical depictions and that adding a numeric value to an icon representing uncertainty can even lead to decreased decision performance. This interaction between graphical and numerical signification of uncertainty is important to understand more fully. Studies assessing the differences between and interactions among graphical and numeric or textual uncertainty depictions may provide insight about what characteristics a well-crafted uncertainty visualization technique must have.

Another crucial aspect is the number of uncertainty classes that was deemed to play a role (Bisantz, Marsiglio, and Munch 2005; Severtson and Myers 2013). This is not only a topic related to visualization but reveals the importance of considering what number of uncertainty classes is needed for different tasks. There are at least two ways through which a high number of classes or a continuous scale may overwhelm the user: on the one hand, the resulting visual complexity of the display, on the other hand, the higher cognitive complexity when interpreting the information on uncertainty. The findings here reinforce the recommendation to assess the impact of number of uncertainty classes that we give in the complementary review paper on communicating uncertainty (Kinkeldey, MacEachren, and Schiewe 2014).

4.5.5 Tasks

A central methodological decision when assessing effects of data uncertainty in user studies is the choice of appropriate tasks. As discussed in the section on "Analysis of the literature", in the studies dealing with decision-making, the complexity of decision tasks varies substantially; examples include:

1. the decision to select one of two regions for some use or other outcome on the basis of suitability and uncertainty values, which is essentially a value comparison task incorporating uncertainty,
2. the ranking of site suitability, with suitability given as a value (+ uncertainty), which is a task focused on judging the relative value of multiple entities with the decision being simply whether or not one is "more" on some scale than another, or,
3. the dynamic decision about object identity and uncertainty (under time pressure), which is essentially an entity recognition task in which the decision is an either-or (is it an X or Y).

Some decision tasks include an explicit spatial component such as in the Hope and Hunter (2007a) study in which participants had to decide when to stop a boat before it enters a restricted zone with a fuzzy boundary. However, these tasks are limited to spatial judgments and do not require any spatial actions, e.g., drawing a route or boundary. Thus, they only cover parts of what is needed to assess spatial decisions and there is a lack of studies including spatial action tasks.

As already alluded to with examples above of tasks that have been used, few studies assess decisions that have much complexity. Thus, most decisions in existing studies have not reached the level of "decision-making" as defined in the Oxford Dictionary,[4] "the action or process of making decisions, especially important ones." The Bussinessdictionary.com adds that, "[f]or effective decision making, a person must be able to forecast the outcome of each option as well, and based on all these items, determine which option is the best for that particular situation."[5] This emphasizes that real-world decision-making is embedded in a problem context with multiple goals and sub-goals; and problem context has typically been included superficially if at all in research on decision effects. Instead, the tasks labeled as decisions in the studies reviewed can typically be made by "advanced map reading" for which no external knowledge is necessary. We see the need for studies involving more complex tasks, for instance, a realistic site location task that multi-criteria decision-making might be used for, a decision about whether and when to evacuate a region due to a pending risk such as a hurricane (and if yes by which route), or a decision about a disaster declaration in which choices must be made about places in and out of the declared region, etc.

Game-like tasks (as described in the section on "Analysis of the literature") can motivate participants and put time pressure on them (Deterding et al. 2011). They are more holistic than the tasks mentioned above and often require sub-tasks to be accomplished. Such tasks can potentially help reveal effects regarding decision speed that cannot be observed when there is no benefit for quick task completion. Moreover, the implicit assessment of decision confidence by for instance using virtual money that is bet on a decision is an interesting alternative approach to self-assessment on a Likert scale or similar methods. So far game-like tasks have been used in cognitive science studies mainly, but we see great potential for their use in studies in the field of geovisualization (as one recent example of the potential – focused on decision-making but not uncertainty specifically – see Tomaszewski, Szarzynski, and Schwartz 2014).

4.5.6 Expertise

When assessing the effect of uncertainty visualization, it is important to define expertise in detail and pick appropriate groups of participants. The first criterion is the type of expertise that is often

[4] http://www.oxforddictionaries.com/us/definition/american_english/decision-making [Accessed 20 June 2015].

[5] http://www.businessdictionary.com/definition/decision-making.html [Accessed 20 June 2015].

not clearly expressed in the studies under review. This can be expertise in visualization, map reading, map production, domain expertise in theory and practice, or experience with statistics or uncertainty measures, just to name a few. The second criterion is the level of expertise that most studies determine via self-assessment or self-reporting of metrics, i.e., years in a profession, number of courses taken, frequency of encountering a map type, etc.

Regarding the role of expertise in uncertainty visualization, we partly agree with Roth (2009b) that "[e]xpertise does not matter when the task is simply to retrieve facts from maps displaying uncertainty (e.g., Evans 1997) or to make qualitative judgments about which representation may work best (e.g., Aerts, Clarke, and Keuper 2003)" and that "when the task involves a complex and realistic risk assessment or decision that requires sophisticated human reasoning (e.g., Kobus, Proctor, and Holste 2001; Hope and Hunter 2007b), expertise may make a profound difference" (42). In light of the insights gained in our complementary review (Kinkeldey, MacEachren, and Schiewe 2014), we agree that domain expertise may not matter much when values are retrieved from a map but map reading expertise is likely to matter. Thus, we would support the statement but restrict it to certain types of expertise. Here, we propose that there are at least two main categories of expertise that have to be distinguished:

1. Expertise relevant to map reading
    a. Spatial ability (assessed with formal tests, e.g., see the "Tests of Spatial Skills"[6] from Spatial Intelligence and Learning Center)
    b. Explicit map reading expertise (subdivided into kinds of map; measured by either amount of education of through task-based metrics)
2. Expertise relevant to decision-making under uncertainty
    a. Expertise in probability, formal logic, etc. that supports understanding of evidence-based reasoning (e.g., as discussed in Micallef, Dragicevic, and Fekete 2012)
    b. Expertise in the specific knowledge domain necessary to judge evidence, understand processes, etc. (measured by either level of education or through task-based metrics)

For the future, we see a more careful definition of expertise as a crucial step to improve the quality of uncertainty visualization "effect" evaluation and the generalizability of the findings.

4.5.7 Decision-making research

While there is a fairly large literature on decision-making under uncertainty (e.g., see Snyder 2006

[6] http://spatiallearning.org/index.php/testsainstruments [Accessed 20 June 2015].

for a review related to facility location), little attention has been given to insights from this domain in studies of geodata uncertainty visualization for decision-making. Research grounded in the cognitive aspects of decision-making (e.g., Busemeyer and Townsend 1993; Bagassi and Macchi 2006; Klein 1999) may be particularly relevant for understanding the effect of visual depiction of uncertainty on decisions. One example of research on integration of decision theory into geographic scale decisions (of the sort where geodata uncertainty visualization might be relevant) is a study by Regan et al. (2005) in the context of conservation management. They explore the potential of using *information gap theory* to understand the level of uncertainty that decision makers can tolerate before their decision would change. With this theory, the range of utility values that can be achieved reliably with a selected action can be determined, based on initial input from the analyst of a utility measure (in the context of conservation management, such measures might include metrics for biodiversity, for minimizing economic impact of conservation decisions, etc.). Their research, however, does not incorporate any visual depictions. A promising strategy would be to select scenarios like this one from problem domains in which decision-making processes are well described (by a theory) and use them for assessing the role of uncertainty visualization in decision-making. This suggests an opportunity to investigate the effect of such depictions within problem domains where knowledge of the decision-making process exists.

4.6 Conclusion

In this paper, we presented a systematic review of studies assessing the impact of uncertainty visualization on decision-making and risk assessment. We identified 43 studies of this type in 34 publications from the fields of geovisualization, information visualization, scientific visualization, and cognitive science (1994–2014). We focused on user studies including geographical visualization, only a small number of studies with aspatial displays (charts) were added that were deemed to be particularly relevant to the discussion. A complementary review article (Kinkeldey, MacEachren, and Schiewe 2014a) covers aspects related to evaluating communication (rather than effects of) visually depicted uncertainty.

We analyzed the findings from the studies under review, trying to summarize commonalities and differences. However, due to a lack of systematic evaluation strategies in the studies reviewed, reported findings were difficult to compare and meaningful overarching conclusions were often not possible. A variety of effects have been measured using diverse kinds of data in scenarios from different domains, from uncertain land suitability, to object identities, to vague boundaries. Still we were able to derive a number of lessons learned and recommendations regarding study focus and design, evaluation methodology, effects under evaluation, the choice of appropriate tasks, the role of expertise, as well as decision-making theory. From this, we identified the main categories of research foci: uncertainty, visualization, methodology, participants, tasks, and effects, compiled in Table 4.2. Furthermore, we list the publications that cover each category and suggest open research topics to fill the gaps we identified in the literature. The table provides a basis for future evaluation research in the field of visually supported decision-making under uncertainty.

Given the complexity of the topic and the lack of a common framework for integration of findings, one of our main conclusions is that we need research efforts to better understand the process of working with information on uncertainty as a basis for subsequent studies of its visualization. This will require in-depth qualitative studies (e.g., using "think aloud" protocols or pair analytics) in order to gain more knowledge about the process (with respect to reasoning as well as decision-making) and to identify the most important questions to evaluate in subsequent quantitative studies.

Reflecting the findings from this review about the effect of uncertainty, we still support the statement by MacEachren et al. (2005) that "we cannot yet say definitively whether decisions are better if uncertainty is visualized or suppressed, or under what conditions they are better; nor do we understand the impact of uncertainty visualization on the process of analysis or decision making" (155). The main reason is that there is evidence for positive as well as negative effects of visually depicted uncertainty, the results are difficult to generalize (for the reasons mentioned above), and there has been limited attention to the processes of analysis or decision-making in research thus far. A systematic effort is needed to create comparable results in experiments that focus on the impact of uncertainty on reasoning and decision-making. To achieve this, we need not only a typology of uncertainty representations but also one that defines categories of reasoning tasks or decisions that uncertainty could make a difference for. External validity of evaluation results could be increased if data, tasks, and effects to be evaluated were chosen in a more systematic way; the categorization in Table 4.2 may serve as a first guideline. This strategy would lead to more reliable insights about the role of uncertainty, insights that are needed to come closer to the overall goal to successfully incorporate visually depicted uncertainty in reasoning and decision-making.

Acknowledgments

Some of the issues discussed in this article came up during meetings of the uncertainty reading group at GeoVISTA Center, Penn State University – special thanks to Jennifer Mason, David Retchless, and Alexander Savelyev for inspirational discussions. Further we would like to thank the anonymous reviewers for their constructive comments and suggestions that helped improve the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

5 Evaluation of Noise Annotation Lines: Using Noise to Represent Thematic Uncertainty in Maps

This chapter was previously published as peer-reviewed journal paper:

Kinkeldey, C., Mason, J., Klippel, A., and Schiewe, J., 2014. "Evaluation of Noise Annotation Lines: Using Noise to Represent Thematic Uncertainty in Maps." *Cartography and Geographic Information Science* 41(5), 430–439.

The candidate is the primary author and responsible for developing the study, conducting the study, and the major part of the analysis of the results (70% of the overall work).

Abstract

Noise annotation lines are a promising technique to visualize thematic uncertainty in maps. However, their potential has not yet been evaluated in user studies. In two experiments, we assessed the usability of this technique with respect to a different number of uncertainty levels as well as the influence of two design aspects of noise annotation lines: the grain and the width of the noise grid. We conducted a web-based study utilizing a qualitative comparison of 2 areas in 150 different maps. We recruited participants from Amazon Mechanical Turk with the majority being nonexperts with respect to the use of maps. Our findings suggest that for qualitative comparisons of "constant uncertainty" (i.e., constant uncertainty per area) in thematic maps, noise annotation lines can be used for up to six uncertainty levels. During comparison of four, six, and eight levels, the different designs of the technique yielded significantly different accuracies. We propose to use the "large noise width, coarse grain" design that was most successful. For "mixed uncertainty" (i.e., uncertainty is not constant per area) we observed a significant decrease in accuracy, but for four levels the technique can still be recommended. This article is a follow-up to our conference paper reporting on preliminary results of the first of the two experiments.

Keywords

uncertainty; geovisualization; user evaluation; usability; AMT

5.1 Introduction

Uncertainty is inherent in all geospatial data arising from various sources such as measurement errors and inaccuracies, model ambiguity, and vagueness, or loss of quality during processing of the data (Atkinson and Foody 2002; Heuvelink and Brown 2008; Shi 2010; Zhang and Goodchild 2002). With many applications, ignoring uncertainty can result in misleading or unusable results: "Error-laden data, used without consideration of their intrinsic uncertainty, are highly likely to lead to information of dubious value" (Zhang and Goodchild 2002, 3). Past research suggested that communicating uncertainty through visualization can support analyses and decision making (Deitrick and Edsall 2006; Hope and Hunter 2007a) and can increase analysts' trust in their results (Fisher et al. 2012). A variety of methods for visualizing uncertainty exist, especially in the area of GIScience and scientific visualization (MacEachren et al. 2005; Pang 2008; Brodlie, Allendes Osorio, and Lopes 2012). Typically, uncertainty is categorized by type, that is, thematic (also: attribute), geometric (positional), or temporal uncertainty. To display different types of uncertainty, different approaches can be combined, for example, integrated or adjacent views, static or dynamic approaches, and the use of interaction, etc. This results in a high number of possible approaches, and it can be difficult to choose a suitable technique for a specific application. Different typologies for uncertainty visualization exist (Thomson et al. 2005; Senaratne et al. 2012) but mainly focus on data characteristics (dimensionality, type, etc.) and do not account for other aspects such as the tasks involved. Thus, they can only offer limited support for the selection of suitable techniques. Besides typologies, other categorizations of the techniques can also be helpful, for instance, the distinction between intrinsic and extrinsic approaches (Gershon 1998). Intrinsic approaches utilize visual variables from existing objects in the visualization to represent uncertainty, mostly including visual variables from cartography. In addition to the seven visual variables described by Bertin (1983), variables including symbol focus and clarity are used for intrinsic displays (MacEachren 1992; MacGranaghan 1993). Extrinsic approaches, on the other hand, incorporate additional graphical objects to represent uncertainty, for example, glyphs (Pang 2001) or other objects such as bars or dials that are added to the display. Unlike most intrinsic variables, they can be visually separated from the other content.

In this research, we evaluate an extrinsic method that we term *noise annotation lines*, a method first described as procedural annotations by Cedilnik and Rheingans (2000). The technique is a promising way to display thematic uncertainty in maps that involve heterogeneous geometries, for example, land cover maps (Figure 5.1). This work contributes to the evaluation of extrinsic uncertainty visualization methods by testing usability aspects of noise annotation lines, with a focus on the impact of design factors on the usability of the method. After discussing related work (the "Related Work" section) we report the results from two web-based experiments (the "Evaluation of Noise Annotation Lines" section) and discuss the implications of the experiments in the "Discussion" section. In the last section, we conclude our findings and the limitations of the experiment and provide suggestions for future evaluation in this field.

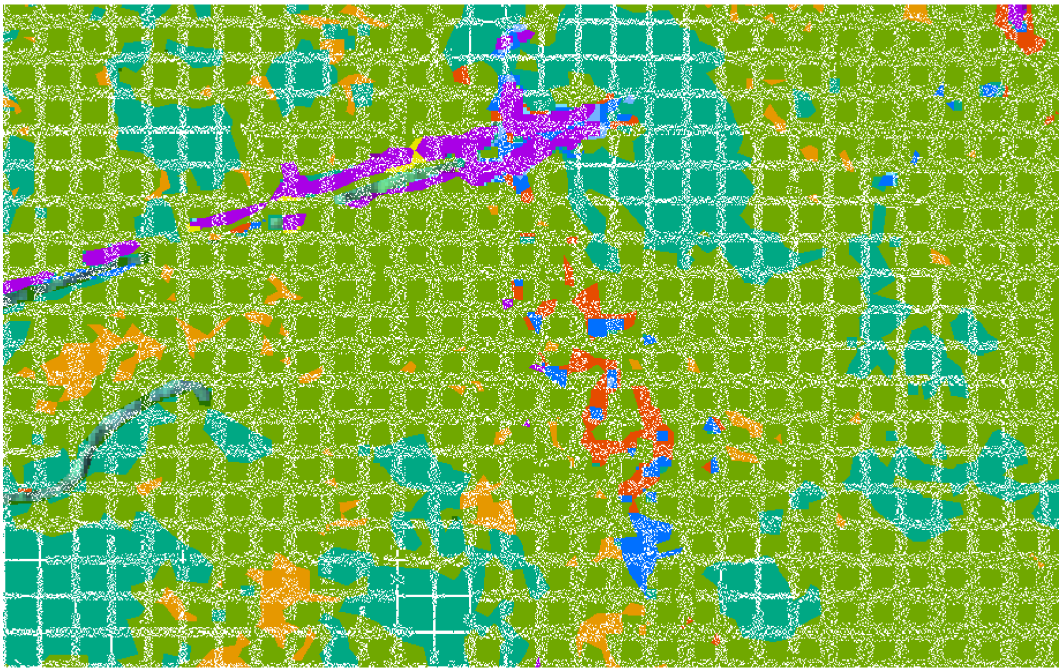

Figure 5.1. Noise annotation lines representing classification uncertainty of a vegetation land cover map.

5.2 Related work

So far there has been very little research to investigate the usability of the noise annotation lines technique. An exception is a qualitative assessment by Zuk and Carpendale (2006) who evaluated noise annotations along with the three other types suggested by Cedilnik and Rheingans (2000): width, sharpness, and amplitude. This was done from a theoretical standpoint using heuristics from theories introduced by Bertin, Tufte, and Ware. Zuk and Carpendale point out that the data-ink ratio of a noise grid is relatively small, compared with other annotation types (e.g., the amplitude grid). Thus, they hypothesize that the noise grid may not be able to represent as many uncertainty levels as the other types but they do not provide further evidence for this assumption. They remark that more formal testing of procedural annotations, especially concerning perceptual aspects, is needed.

Similar to the work we present here, a number of studies by Kardos and colleagues evaluate a technique called "trustree" that depicts uncertainty in census maps by varying the level of detail locally (Kardos, Benwell, and Moore 2007; Kardos, Moore, and Benwell 2008). They found that the visual metaphor of “detail”, that is, a coarser grid in uncertain areas is more usable than a metaphor of “clutter” that represents uncertain areas with a finer grid.

Despite the variety of methods for representing uncertainty, a systematic evaluation of their usability is still needed (MacEachren et al. 2005). Studies on extrinsic approaches deal with various data and display types, including multivariate vector glyphs (Wittenbrink, Pang, and Lodha 1996) or 3D displays (Newman and Lee 2004). Thus, the results do not always help in choosing a suitable technique for 2D maps since the requirements for displaying uncertainty differ. For instance, in 3D environments the question of clutter through occlusion in different perspectives is important but does not exist with 2D maps.

5.3 Evaluation of noise annotation lines

Many thematic maps such as land cover maps contain objects of high geometric variability, that is, objects that differ considerably in size and shape. Representing uncertainty integrated into such maps is challenging compared to maps with more homogeneous objects such as choropleth maps. For maps with geometrically diverse areas, extrinsic methods, especially those based on uniform grids, seem promising because they are independent of the underlying geometry. *Noise annotation lines* (Kinkeldey et al. 2013) are a grid-based extrinsic method. A regular grid is placed onto the map and is altered locally to represent the degree of uncertainty (Figure 5.2). Cedilnik and Rheingans (2000) proposed four different versions of annotations: variation of width, sharpness, noise, and amplitude. We focus here on the noise grid because we expected noise to be a particularly suitable metaphor for uncertainty. This was substantiated in a qualitative pretest where people found the noise display particularly intuitive (Kinkeldey and Schiewe 2012). Additionally, recent research suggests “noise was seen as a promising graphic variable, worth further investigation” (Vullings et al. 2013). The noise grid is varied in size locally to represent the level of uncertainty; the more uncertain the underlying content, the more scattered the noise line. Regardless of the level of uncertainty, the number of noise particles and their size (grain) remain constant. An important characteristic of this approach is that it only represents the values beneath the lines, where the values in the cells of the grid are not depicted thus showing a generalization of the uncertainty data. However, since the size of the grid cells can be varied according to the scale of the map, a compromise can be made between maximum coverage of uncertainty data and minimum occlusion of the underlying content. Related research suggested that

> [f]rom a design standpoint, selection of an appropriate number of levels should be guided first by task demands, such as the level of detail necessary for people to differentiate among potential actions. Information should not be displayed at a greater level of detail than is required by the task. (Bisantz et al. 2009, 78)

This could be one of the potential advantages of noise annotation lines and makes this method promising for use in maps.

We plan to use the noise annotation lines approach for exploratory analyses of change in land cover maps. Understanding general changes of land cover does not require precise values of uncertainty, rather a qualitative estimation to compare general differences. Thus, we tested the qualitative comparison of uncertainty between different areas, that is, no specific uncertainty values had to be retrieved from the display and no legend was provided.

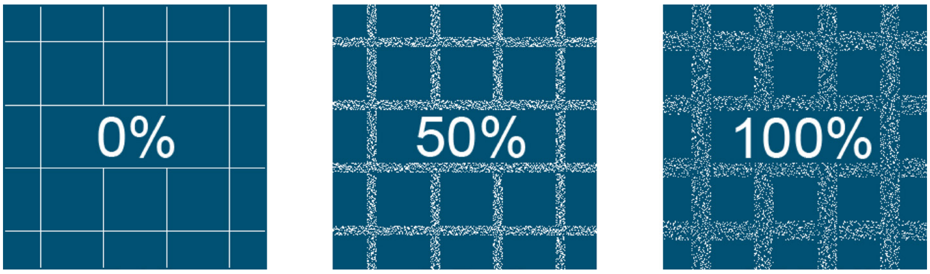

Figure 5.2. Variation of noise width to represent uncertainty: the higher the uncertainty, the larger the width of the noise grid.

5.4 Experiment 1

5.4.1 Research questions

The goal of this experiment was to make a first estimation about the usability of noise annotation lines to represent thematic uncertainty in maps. As there are different ways to display the noise grid, the first question that emerged was whether changes in the design affect user performance when areas are compared in terms of their uncertainty. The second question was how many levels of uncertainty could be easily distinguished, and third, if and how user performance changes when uncertainty does not remain constant over the areas in the map (mixed uncertainty). These broad questions led to the following specific research questions:

1. How do different design parameters impact the usability of noise annotation lines as a representation of thematic uncertainty in a land cover map?
2. How does the number of uncertainty levels affect user performance?
3. Can users accurately compare the overall degree of uncertainty between two defined areas when the values vary within the areas (mixed uncertainty)?

5.4.2 Variables

The appearance of noise annotation lines can be changed by altering different design parameters. Our hypothesis was that these variations will impact the effectiveness and efficiency of the uncertainty display. The following three major design parameters were chosen for evaluation: the number of uncertainty levels, the width of the noise grid, and its grain (Table 5.1).

Table 5.1. Factors used in experiment 1.

Factor	Number of levels	Levels
Noise width	2	Small (40%), large (50%)
Noise grain	2	Fine (1 × 1), Coarse (2 × 2)
Uncertainty levels	3	4, 5, 6 levels

The width of the noise grid was defined with respect to the size of the grid cells (Figure 5.3). With a smaller noise width the grid covers less area, and there is less space to represent different levels of uncertainty. Consequently, the choice of this parameter was a compromise between the visual interference of the grid with the underlying content and the number of levels that are discernible. If the grid width is too large, the grid occludes more of the underlying content and the structure of the noise grid is not preserved. This effect already occurs with a noise width of 60% of the grid cell size. On the other hand, if the width is too small, it limits the number of uncertainty levels to be discerned. Thus, we chose 40% and 50% of the grid cell size as levels to assess for this factor.

Figure 5.3. Both grids represent the same degree of uncertainty (100%), but with different widths: 40% (left) and 50% (right) of the grid cell size.

The grain of the noise particles was the second design parameter we manipulated (Figure 5.4). A finer grid consists of a higher number of small particles, while a coarse grid contains fewer, but larger particles. Since we kept a constant pixel resolution across all maps in this study we implemented 1 × 1 pixels and 2 × 2 pixels for the factor "grain."

Figure 5.4. Design parameter “noise grain.” Both grids represent the same degree of uncertainty (100%), but with different grain sizes: “fine” (left) and “coarse” (right).

For the third factor, we varied the number of uncertainty levels. We chose a minimum of four levels because a pretest revealed that three levels (0%, 50%, and 100% uncertainty) are straightforward to discern in contrast to four levels that already led to errors in uncertainty value retrieval. We hypothesized that a variation up to six levels would be appropriate to determine the limit of levels that people are able to compare. This resulted in four, five, and six uncertainty levels so subjects had to discern intervals of 33%, 25%, and 20% uncertainty (Table 5.2).

As dependent variables, response accuracy and the time the subjects needed for each question were measured during the experiment.

Table 5.2. Factor “uncertainty levels” in experiment 1.

Number of levels	Interval (%)	Levels
Four levels	33	0%, 33%, 66%, 100%
Five levels	25	0%, 25%, 50%, 75%, 100%
Six levels	20	0%, 20%, 40%, 60%, 80%, 100%

5.4.3 Task

For the main part of the study we chose a uniform task for all maps: a comparison of uncertainty between two areas. This was done in a qualitative way (no specific values had to be retrieved) because during real world analyses it is rarely the case that the user needs exact uncertainty values. Instead, with most of the tasks it is more important to know how uncertainty of one area relates to uncertainty in a different area. For example, if an area is classified as water body, high uncertainty can be an evidence for a misclassification. But this is just the case if uncertainty is relatively higher than with other areas of this type – the specific amount is not of interest. Thus, we asked participants to compare uncertainty between two marked areas: “A” and “B.” The question and possible answers remained the same for all maps: “Which area is more uncertain?” Potential answers included “A is more uncertain,” “B is more uncertain,” “A and B are equal”, and

"I can't tell." Since the questions were mandatory the latter answer was included to minimize nonsense answers when participants could not read the map or had little confidence in their answers.

5.4.4 Stimuli

We created 10 different maps per factor combination to establish 10 repetitions. All maps were taken from the same vegetation land cover map representing equal-sized areas (100 m × 100 m) at the same scale. Furthermore, the size of the noise grid cells in all maps was the same (4 m). We varied the background colors according to a qualitative color scheme recommended by ColorBrewer (Brewer, Hatchard, and Harrower 2003). The utilized color scheme ("Paired") is indicated to be colorblind safe and laptop / LCD friendly. In each map, two square areas in the size of 3 × 3 noise grid cells were drawn on areas of the same color and labeled "A" and "B" (Figure 5.5). The value within the areas was either equal or differed by one level, for example, 66% vs. 100% with four levels or 60% vs. 40% with six levels. We placed the squares on areas with the same background color, either light blue or light green. These two colors have a very similar contrast distance from the white color of the grid. Hence, we varied the color but not the contrast between grid lines and the background. Additionally, the low contrast between grid and background assures we evaluate the more critical point to find the limits of this technique.

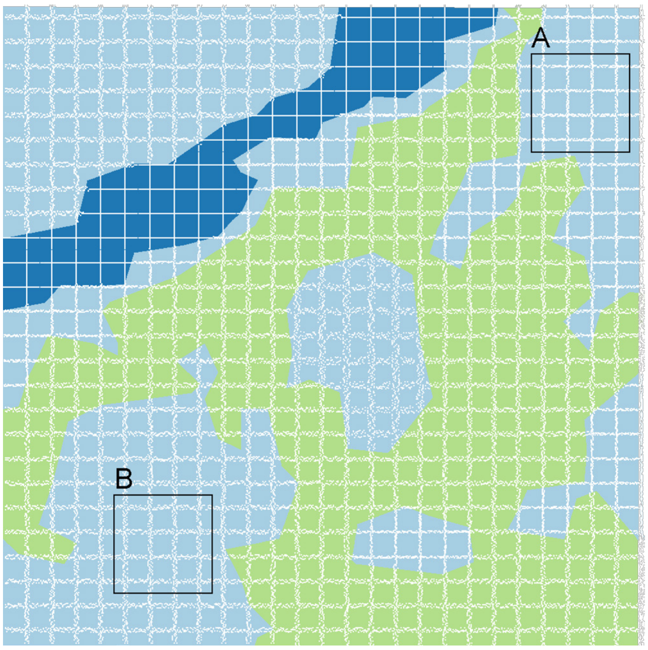

Figure 5.5. Example map representing constant uncertainty (uncertainty remains constant per map region).

In order to answer the third research question (constant vs. mixed uncertainty), we included maps showing a mixed representation of uncertainty. This means, in contrast to the constant case, uncertainty values do not remain constant within each area (Figure 5.6). Thus, uncertainty is not constant in the marked regions A and B. For the mixed uncertainty case, we did not involve all combinations as with the constant case. In order to keep the number of combinations low we only varied the number of uncertainty levels and not noise width and grain. The three levels (four, five, and six uncertainty levels) were repeated ten times, resulting in thirty maps. Overall, each participant answered 150 (120 constant + 30 mixed uncertainty) questions in the main portion. In the survey, the maps with constant uncertainty and those with mixed uncertainty were shown in a randomized order.

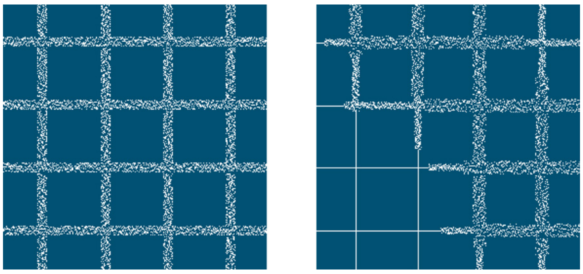

Figure 5.6. Constant (left) vs. mixed uncertainty distribution (right).

We conducted the experiment as a web-based survey, making it possible to recruit participants via Amazon Mechanical Turk (AMT, http://www.mturk.com). Web-based experiments have proven to be suitable for evaluating usability aspects of uncertainty visualization (Aerts, Clarke, and Keuper 2003). There have been substantial review articles by Rand (2012) and Crump et al. (2013) showing that behavioral experiments via AMT yield essentially the same results as those in controlled environments. Crump et al. (2013) were able to replicate findings of popular reaction times tasks requiring participant attention in AMT to that of traditional laboratory settings. In an online cognitive and perceptual experiment by Germine et al. (2012), they found that their data from challenging timed tasks were similar in quality to performance data that can be gathered in a laboratory setting despite being anonymous, uncompensated, and unsupervised. Additionally, guidelines exist that describe how experiments based on AMT can be designed to yield valid results (Mason and Suri 2012). In our case, we were aware we could not control the display type, color calibration, or distractions that potentially influence the participant; that is why our experiments are similar to field experiments rather than controlled lab experiments. But this fact could even make the results more valid for the use in real world applications.

5.4.5 Survey

The survey comprised of the following parts:

1. Introduction: The participants were provided with an introductory explanation of uncertainty and noise annotation lines. Three figures of noise annotation lines were shown (no, medium and high uncertainty) to clarify the method. We also included a note to not use a smartphone or similar device and when using a tablet, to not zoom in and out to make sure that all subjects see each map in its entirety when answering the questions.
2. Personal information: We asked for gender, age, and a self-assessment in terms of experience with uncertainty visualization in maps.
3. Maps: The main section of the study showed the 150 maps in succession including its uncertainty. In each map, the two areas A and B were compared. In order to avoid bias and learning effects, we randomized the order of the questions.
4. Comments: An opportunity to provide feedback on the survey. All questions except the comments at the end were mandatory.

We used LimeSurvey (http://www.limesurvey.org), survey software freely available under an open source license. We decided to use version 1.92+ since we noticed problems with the randomization that occurred in the latest version (2.05).

5.4.6 Participants

We recruited participants using the online crowdsourcing service AMT. The reasons for utilizing this service are threefold: First, it is efficient to recruit subjects, second, we aimed to obtain participants with different backgrounds and expertise (not only from our domain) and third, paid participants were likely to be motivated to finish the survey even though it took 20 to 30 minutes. Participants were reimbursed with $0.50 for their participation. Among the 32 participants, 17 declared themselves as female and 15 as male. Regarding age, most of the participants classified themselves between 20 and 29 years old (16/32), followed by 50 to 59 years (8/32) and 30 to 39 (5/32). Very young people and the group 40 to 49 years were barely represented. People who are 60 years and older did not participate at all.

Concerning the subjects' experience with uncertainty maps, we asked three questions: If they had known about the concept of uncertainty before, how often they used maps, and if they had seen a map including uncertainty information before. From the three answers we determined a level of experience per participant (little, average, extensive experience). More than half of the participants (18 out of 32) had little experience while roughly one-quarter had average experience

(7/32) or extensive experience (7/32) with uncertainty maps. This is not surprising as one can expect that participants acquired via AMT will be primarily lay people.

5.4.7 Results

We obtained the results through rounds of 10 participants and checked the integrity and validity of the data after each step. In the end the turnout was 32 because of incomplete replies that we ignored. Since there were 10 maps for each combination in our $3 \times 2 \times 2$ factorial design we had 150 answers from each subject, totaling 4800 single answers in the main section. In case a participant responded that he or she was not able to provide an answer, we treated this as a missing value. Since there were only 96 missing values out of 4800 (2%) we made the assumption that sufficient responses were collected for each participant and map. We computed accuracy for each participant and factor combination as the percentage of correct answers.

Figure 5.7 shows the mean accuracy and the standard error for the maps with constant uncertainty. The charts are grouped by combination of the factors "noise width" and "noise grain" and each chart shows the accuracies for four, five, and six uncertainty levels. Generally, for all factor combinations, the mean values are higher than 76%. The "small noise width" conditions (two charts on the left) show a stronger trend of decreasing accuracy with an increase of uncertainty levels from 4 to 5. In the "large noise width" conditions (two charts on the right), the accuracy values are, compared to the other conditions, lower for four levels and roughly equal for five levels, but increasing for six levels again, especially for the coarse grain grid (increase from 77% for five levels to 83% for six levels).

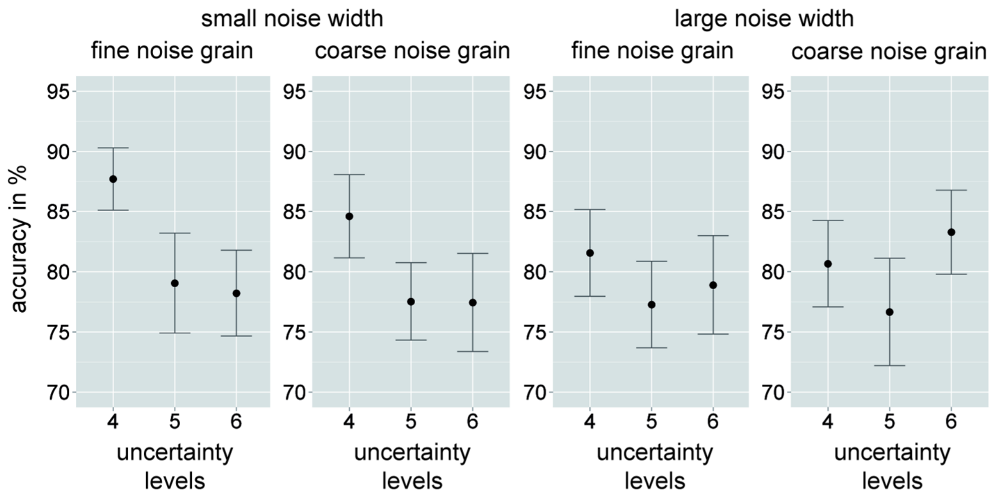

Figure 5.7. Experiment 1: Accuracy for constant uncertainty (mean and standard error).

Repeated measures ANOVA with the three factors revealed the following: Mauchly's test of sphericity showed that the assumption of sphericity is violated for the factor "uncertainty levels" ($\chi^2(2) = 7.23$, $p = .27$) and the interaction of "uncertainty levels" and "noise grain" ($\chi^2(2) = 8.22$, $p = .16$). Hence, we used the Greenhouse– Geisser correction as suggested by Tabachnick and Fidell (2013). Of the three main effects, only "uncertainty levels" is statistically significant ($F(1.647,51.065) = 7.024$, $p = .004$, $\eta^2 = .185$): the accuracy decreases with a higher number of levels. The factors "noise width" and "noise grain" do not significantly change user performance; however, there is a statistically significant interaction effect of "uncertainty levels" and "noise width" ($F(1.960,60.747) = 7.295$, $p = .002$, $\eta^2 = .19$).

Our third research question addressed the change from constant to mixed uncertainty data (Figure 5.6) and its effect on user performance.We did not vary noise grain and width for this comparison and selected a salient combination ("fine grain, large width") visualizing four, five, or six uncertainty levels. A graphical comparison of accuracy between constant and mixed case can be found in Figure 5.8. The most obvious difference is that with mixed data, accuracy is generally lower than in the constant case. In contrast to constant uncertainty the results for the mixed case show a consistent decrease of accuracy with growing number of uncertainty levels.

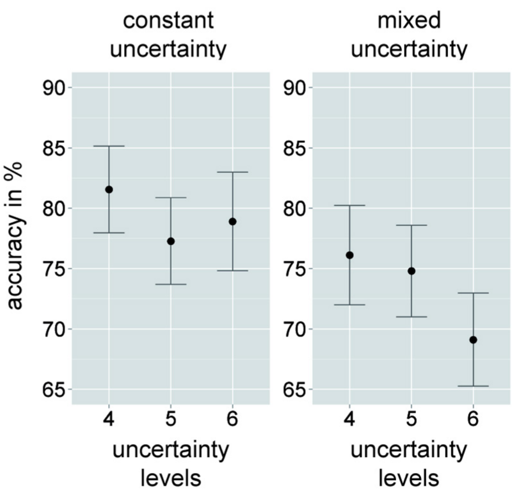

Figure 5.8. Experiment 1: Accuracy for constant vs. mixed uncertainty (mean and standard error) for the factor combination "large noise width", "fine noise grain".

Repeated measures ANOVA provided further insight. Mauchly's test indicated that only the main effect of uncertainty levels is statistically significant ($\chi^2(2) = 8.56$, $p = .014$). Therefore degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity. The main effect "constant vs. mixed" is statistically significant ($F(1,31) = 10.59$, $p = .003$, $\eta^2 = .255$). Likewise, the main factor “uncertainty levels” is statistically significant ($F(1.60,49.68) = 3.912$, $p = .035$, $\eta^2 = .112$) but there is no significant interaction between “constant vs. mixed” and “uncertainty levels” ($F(2,62) = 1.74$, $p = .183$). Hence, the results for the number of uncertainty levels (the more uncertainty levels the lower the accuracy) is present in this comparison, too, and making judgments in which multiple levels of uncertainty (mixed case) have to be taken into consideration as further decreasing the accuracy of user responses.

5.5 Experiment 2

In the first experiment, the use of four, five, and six uncertainty levels yielded relatively high accuracy, even with six levels (>76% mean accuracy with a maximum standard error of 4.5%). In order to determine the limitations of noise annotation lines with respect to the number of uncertainty levels, we repeated the experiment with up to eight levels. The setup was the same as with experiment 1, that is, the structure and the task were not changed. However, we removed the maps with five levels and added maps using eight uncertainty levels. Thus, the uncertainty levels changed to 4, 6 and 8 (Table 5.3). The other two factors “noise width” and “noise grain” remained the same as in experiment 1.

Table 5.3. Levels for factor “uncertainty levels” in experiment 2.

Uncertainty levels	**Step (%)**	**Levels**
4 levels	33	0%, 33%, 66%, 100%
6 levels	20	0%, 20%, 40%, 60%, 80%, 100%
8 levels	14	0%, 14%, 28%, 43%, 57%, 71%, 86%, 100%

5.5.1 Participants

We recruited participants in the same way as in experiment 1, that is, via AMT offering the same amount for a complete set of answers. We had the same number of responses as in experiment 1 (32 full responses) after sorting out incomplete datasets.

There were more female participants (18) than male (14). The distribution of age was similar to experiment 1; most of the participants classified themselves to be between 20 and 29 years old (14/32), followed by 30 to 39 (8/32) and 40 to 49 years (6/32). Very young (below 20) and older respondents (over 60) were not represented. Self-assessment of experience resulted in 14 subjects with little, 11 with average, and 7 with extensive experience interacting with uncertainty in maps. Hence, the overall experience was higher than with experiment 1 where half of the subjects had little experience and fewer subjects in the average experience group.

5.5.2 Results

Figure 5.9 shows the mean accuracies for all factor combinations. Generally, accuracy is lower than in experiment 1 (between 4% and 12% lower for four levels and 2.5% and 9.5% for six levels). While most results are similar to experiment 1 there were a number of participants in experiment 2 with constantly low accuracy. Since we could not determine the reason for their low performance we did not treat them as outliers.

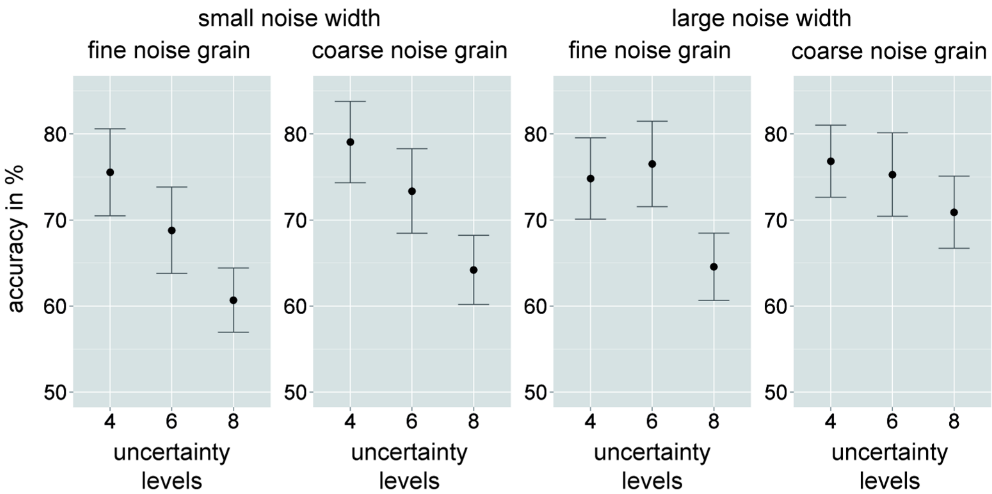

Figure 5.9. Experiment 2: accuracy for constant uncertainty (mean and standard error).

As expected, the decrease in accuracy with more than six uncertainty levels is higher: for eight levels, accuracies are generally lower than with the same design using four or six levels. For example, in the "small noise width," "fine noise grain" condition, accuracy decreases from 75.6% (four levels) to 60.7% (eight levels). Again, the magnitude of change depends on the design of the grid; the "coarse grain, large width" condition results in higher accuracy values (76.9% for four levels to 70.9% for eight levels).

Repeated measures ANOVA showed that Mauchly's test of sphericity is not significant for experiment 2, so we assumed sphericity. In contrast to experiment 1, results suggest that all three factors have a statistically significant impact on accuracy: "uncertainty levels" ($F(1.993,61.777) = 15.337$, $p < .001$, $\eta^2 = .331$), "noise width" ($F(1,31) = 5.958$, $p = .021$, $\eta^2 = .161$), and "noise grain" ($F(1,31) = 6.039$, $p = .02$, $\eta^2 = .163$). Additionally, we found a statistically significant interaction effect between "uncertainty levels" and "noise width" ($F(2,62) = 3.875$, $p = .026$, $\eta^2 = .111$). Contrast revealed that this interaction effect is specific to the change from four to six uncertainty levels but does not apply to eight uncertainty levels. These results largely reflect experiment 1 showing that more uncertainty levels lead to lower accuracy. The expected additional challenge using four, six, and eight uncertainty levels is nicely reflected in the increased η^2 for the main effect "uncertainty levels." It becomes obvious that in more challenging situations, design characteristics become more important and can support reasoning with uncertainty: salient visual characteristics can offset the negative effect of higher numbers of uncertainty levels.

Comparing constant and mixed uncertainty representations, accuracy values for the mixed case were found to be lower than that for the maps that depicted constant uncertainty (Figure 5.10). Mauchly's test was not significant, thus we assumed sphericity. ANOVA revealed that uncertainty levels were again significant ($F(2,62) = 7.444$, $p = .001$, $\eta^2 = .194$): more levels result in lower accuracy. Additionally, there is a significant difference between constant and mixed uncertainty ($F(1,31) = 15.255$, $p < .001$, $\eta^2 = .330$), indicating the difficulties participants have when uncertainty is not constant per area (mixed). There are no significant interaction effects.

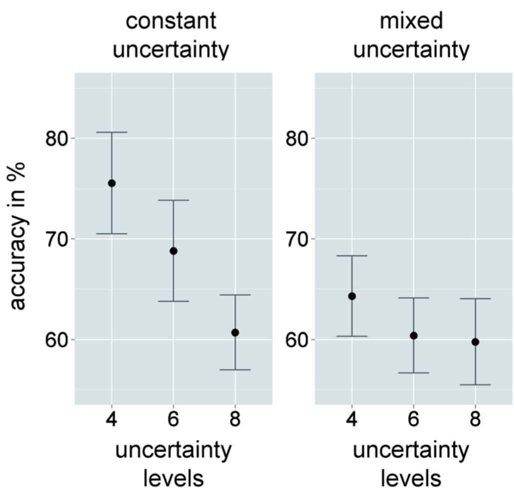

Figure 5.10. Experiment 2: accuracy of constant vs. mixed uncertainty (mean and standard error) for the factor combination "large noise width", "fine noise grain."

5.6 Discussion

The results from the two experiments offer insights into various aspects of uncertainty visualization. The first thing to note is that, as expected, the number of uncertainty levels has an influence on people's abilities to make judgments about uncertainty. Simply put, the fewer levels participants have to distinguish, the better their performance was in terms of correct answers. This result can be explained by numerous studies in both perception and cognition literature that indicate that the more information that has to be distinguished and kept in working memory, the more difficult it is to reason with this information (Lloyd and Bunch 2005).

The downward trend of user performance can be halted using appropriate visualizations, which is an important finding for research on visualizing uncertainty. Experiment 2 revealed that when comparing four, six, and eight uncertainty levels, the two design parameters of the grid had a significant impact on the decrease of user performance. The most visually salient design of the grid (large noise width and coarse grain) was able to leverage the negative effect of an increased number of uncertainty levels (as revealed by the significant interaction effects and the graph in Figure 5.9). In this combination, accuracies in comparing the two areas were higher than expected: even for eight uncertainty levels we observed a mean accuracy of 71% compared to 61% with the small width and fine grain design. The more levels to distinguish, the more important it is to use a salient design for the grid. For six levels, we observed a mean accuracy of more than 75%. Taking into account that we evaluated the case of very low contrast to the background, we see this as recommendation for the use of noise annotation lines with up to six uncertainty levels.

The comparison of constant and mixed visualizations of uncertainty shows that accuracy with mixed uncertainty is generally lower. This confirms the assumption that the more complex the information is that is offered to participants, the more errors they make. We only compared one factor combination (fine grain and large width) for constant and mixed uncertainty visualization and expected that the increase in the number of levels would lead to lower user performance. Regarding the dependency on the number of uncertainty levels the two experiments show ambiguous results. In experiment 1 (four, five, six levels), only the accuracy values for mixed uncertainty decreased with more uncertainty levels. In experiment 2 (four, six, eight levels) accuracy decreased in both cases (constant and mixed) but the extent was higher with constant uncertainty. Hence, mixed uncertainty generally leads to lower accuracy but with up to eight levels the effect of decreasing accuracy is weaker. This can be explained by the implicit difference between the way of comparing constant and mixed uncertainty. In the constant case, the comparison task will be successful if the noise grid in the two areas can be visually distinguished. With mixed data subjects have to estimate the overall uncertainty in both areas first and subsequently compare them. The estimation part of the task could be less influenced by the number of uncertainty levels than the direct comparison lessening the effect of decreased accuracy with more uncertainty levels.

Given the difficulty of controlling time in web-based studies (compared to lab experiments) it does not come as a surprise that no significant effects related to response times could be observed. For instance, we recognized that the loading times of the maps varied a lot depending on the Internet connection. We tried to minimize these effects by compressing the map images but there were still differences.

5.7 Conclusions and outlook

We have presented a study to evaluate noise annotation lines for visualizing thematic uncertainty. In a web-based survey, subjects compared the uncertainty of two equal-sized areas A and B. This was done for constant and mixed uncertainty representations (see Section "Evaluation of Noise Annotation Lines"). From the experiments, the following results can be concluded:

- the number of uncertainty levels has a significant influence on the participants' judgment (generally, with more levels user performance decreases),
- the variation of the design parameters "noise width" and "noise grain" has a significant impact on user performance with a higher number of (up to 8) levels, meaning that
- the decrease of user performance when more uncertainty levels need to be distinguished can be counterbalanced with changes in the design of the noise grid,
- more complex uncertainty information ("mixed" uncertainty) leads to a significant decrease in user performance, and,
- there are no significant effects with respect to response time.

All in all, we could show the potential of noise annotation lines for the representation of thematic uncertainty in qualitative comparison tasks. We can recommend their use in maps that are geometrically diverse or already make extensive use of color so that intrinsic techniques are hard to apply. Our findings show that the technique can successfully be used in qualitative analyses of up to six uncertainty levels when more salient grid designs are used.

There are several *limitations* worth mentioning with respect to the two experiments. First, we did not account for the impact of noise annotation lines on the readability of the map, that is, we did not evaluate if the background could still be read despite the noise grid. Second, we did not evaluate the influence of background colors with different contrast levels. Third, the fact that response times did not show any effect could be due to our setup. There was no time pressure or incentive for quick replies.

Based on the limitations, we recommend the following aspects as *future work*. The first one is the evaluation of response time. As already discussed above, our experiments did not show any effects regarding time. For applications in which quick comparisons are of importance this could be of high interest, for instance, in dynamic decision making (Kobus, Proctor, and Holste 2001). We are convinced that there are alternative experimental designs that could help reveal time effects, for example, the use of incentives for quick answers combined with a reward for the most accurate responses. A second aspect worth investigating is the intuitiveness of noise annotation lines. So far we experienced that when people see noise annotation lines for the first time they seem to understand the concept very quickly. This could be a potential strength of the technique and has not been evaluated systematically so far. A third aspect worth looking at is the comparison with other types of annotations. Zuk and Carpendale (2006) assumed that with a different type of annotation, a grid representing uncertainty by amplitude of a sine curve, more levels of uncertainty may be discerned than with a noise grid (because of the higher data-to-ink ratio). It would be interesting to evaluate if this assumption is true and how much difference there is. Finally, making the grid size interactively adaptable by the user is a promising approach for the use of noise annotation lines in exploratory analyses. When the data are unknown the user can reveal patterns in the data by altering the grid width. This may open this approach for other usages than the representation of uncertainty – for explorative analysis of other data, for example, pollution or population density, it seems promising as well.

6 Development of a prototype for uncertainty-aware geovisual analytics of land cover change

This chapter was previously published as peer-reviewed journal paper:

Kinkeldey, C., 2014. "Development of a prototype for uncertainty-aware geovisual analytics of land cover change." *International Journal of Geographical Information Science* 28(10), 2076–2089.

The candidate is the sole author of this work.

Abstract

Analyses of spatiotemporal data are affected by different kinds of uncertainty, and various studies have shown that their impact can be severe. This is especially true for land cover change analysis based on remotely sensed data – it has been shown that ignoring uncertainty can lead to unreliable results. Approaches are needed that incorporate information on uncertainty into the analysis. However, usable models and methods are still rare and the supposed positive effect of uncertainty information has not been extensively studied. In this contribution, we describe the development of *ICchange*, an interactive visual prototype for the exploratory analysis of land cover change following a geovisual analytics approach. Apart from serving as proof of concept, the prototype will be used to evaluate the role of uncertainty information during change analyses. We conducted a qualitative evaluation of the prototype using low-level tasks to test basic usability. The input from the study is used to improve the design. A special focus is placed on the visual representation of uncertainty in one of the views that did not perform satisfyingly. The prototype presented here will be used in future studies to evaluate the role of uncertainty in real-world change analysis.

Keywords

geovisual analytics; change analysis; uncertainty; land cover; categorical change

6.1 Introduction

Uncertainty is inherent in all kinds of spatiotemporal data and is caused by uncertainty in the real world, limitation of human knowledge, limitations of measurement technologies, and the potential to generate and propagate uncertainty in processing and analysis (Shi 2010). Ignoring the uncertain nature of data is 'highly likely to lead to information of dubious value' (Zhang and Goodchild 2002) and can cause ill-informed decisions, for instance in public policy decision support (Deitrick 2012).

In the field of remote sensing (RS) the degree of uncertainty in the data generally tends to be high because images of the Earth's surface are taken under varying conditions. For example, different sun angles, haze or clouds can affect the quality of the imagery. This holds especially true when changes are detected and analyzed based on RS data, e.g., in environmental monitoring, natural resource management or urban development. When two or more datasets are involved uncertainty accumulates and can lead to highly uncertain results. For instance, Pontius and Lippitt (2006) showed that even with highly accurate classified datasets, about half of the observed change could be explained by error. That is why researchers from the field of RS recommend the use of uncertainty information to support change analysis (van der Wel 2000, p. 56).

We incorporate uncertainty into change analysis by following a geovisual analytics (GVA) approach. The discipline of GVA is derived from visual analytics (VA) which can be defined as 'the science of analytical reasoning facilitated by interactive visual interfaces' (Thomas and Cook 2005, p. 4). Keim et al. (2010) extend upon this definition by stressing the aspect of 'combining automatic and visual analysis methods with a tight coupling through human interaction'. GVA can be seen as a subarea focused on spatiotemporal phenomena that 'integrates perspectives from Visual Analytics […] and Geographic Information Science' (Tomaszewski et al. 2007).

In this article we present a concept to incorporate uncertainty into change analyses using GVA. After discussing related work (Section 6.2) we present the approach in detail (Section 6.3). The core of the work described here is the development and evaluation of an interactive software prototype (Section 0). Section 6.5 concludes this article and provides an outlook on future work.

6.2 Related work

There are a number of approaches to visually explore and analyze uncertainty in remotely sensed imagery and derived datasets in visual environments. An example is the visual environment employed by Lucieer (2004) to explore uncertainty in segmentation and classification of remotely sensed imagery. This tool supports the uncertainty-aware, iterative choice of classification parameters that has the potential to increase the quality of the results as shown in his study. Related to this, Ahlqvist (2008) handles incompatibilities of class definitions for land cover and land use by visualizing semantic similarity and overlap between class definitions. Other examples include tools that help measure and model uncertainty in RS data. van der Wel et al. (1998) use a visual tool from the *CAMOTIUS* package to compare different uncertainty measures for RS data. This approach helps clarify how different uncertainty measures reveal different information and helps choose an appropriate measure for a specific application. Bastin et al. (2002) present the *VTBeans* toolkit to visualize and explore fuzzy classifications of RS imagery. It uses interactive, linked views to enable the visualization of data uncertainty by a variety of means. Pebesma et al. (2007) and Gerharz et al. (2010) use an interactive visualization tool named *AGUILA* to handle and visualize uncertainty in spatiotemporal data using probability density functions. With the help of this tool, different approaches to interpolate air quality measurement can be compared regarding their uncertainty.

Despite this development, information on uncertainty is still rarely used in practical analyses. The main reason for this can be seen in a lack of support by most GIS, where standard models and procedures for the use of uncertainty are largely missing. Past research in the areas of uncertainty modeling, quantification, and visualization has laid a broad foundation for leveraging information about uncertainty. However, there is still a gap when it comes to integrating this information into analyses in a way that analysts can effectively use it: 'They [the users] will, however, only accept the assumed extra value if they have the disposal of easy-to-use tools and methodologies to derive and handle this particular meta-information' (van der Wel 2000, p. 62). The intention of this work is to help close this gap by following the approach described in the next section.

6.3 Approach and background

This research is based on the hypothesis that the use of GVA is the key to enable the use of uncertainty in spatiotemporal analyses. Past research has shown that GVA can be successfully utilized to enhance the analysis of changes in spatiotemporal data. Andrienko and Andrienko (2005, 2006) have contributed a lot of relevant research in this field, including fundamental theoretical work as well as practical tool development. Related to the analysis of categorical change, Landesberger et al. (2012) present a GVA approach to combine visualizations of categorical changes with computational techniques that help gain new insight. Two applications from the domains of movement tracking and weather data analysis demonstrate the strength of the approach. For the exploratory analysis of remotely sensed time series, Zurita-Milla et al. (2012) extend the GIS software package *ILWIS (Integrated Land and Water Information System)* with a toolbox called *SITS (Satellite Image Time Series)*. It facilitates animation and interaction to analyze changes directly from the imagery and provides filtering and aggregation functionality to conduct analyses at different levels of granularity. Green (2011) follows a different approach:

Change Matters is a web-based application for interactive change analyses of an extensive library of Landsat satellite imagery (http://changematters.esri.com/compare). A change overlay is displayed on the imagery, and the user can modify thresholds, e.g., for vegetation gain and loss. Based on this, appropriate change maps can be created interactively and distributed over the Internet. Another interesting approach is *GTDiff (Geo-Temporal Differences)* by Hoeber et al. (2010), a tool that allows for the interactive exploration and analysis of multi-temporal changes based on the visualization of spatiotemporal difference graphs.

However, related work on exploratory change analysis does not usually incorporate uncertainty information although existing studies indicate positive effects, e.g., on decision making (Leitner and Buttenfield 2000, Deitrick and Edsall 2006, Hope and Hunter 2007a). We see great potential in GVA for enhancing uncertainty-aware analysis, because the use of interactive visual environments allows uncertainty information to be directly communicated to the user during the process. Hence, we see the need for a concept to evaluate if and how knowledge about uncertainty affects exploratory analyses in GVA environments.

We focus on the domain of RS and a defined target user group because studies in the field of uncertainty visualization indicate that the usability of visual communication of uncertainty is highly domain-, user-, and task-dependent (Nadav-Greenberg et al. 2008, Sanyal et al. 2009). Thus, our approach is to gain new insight by focusing on one domain and making generalizations about the results afterwards. As a basis for future evaluations we developed an interactive prototype tailored to the exploratory visual analysis of land cover change. We plan to conduct studies including expert analysts because it 'does appear that experts and novices may incorporate uncertainty into their decision-making processes differently' (MacEachren et al. 2005, p. 155). We intend to follow qualitative approaches rather than controlled experiments because the exploratory nature of the VA approach involves tasks that are 'nearly impossible to reconstruct in a controlled experiment for a variety of reasons' (Perer and Shneiderman 2009). The main goal is to assess if analysts use the information and if it helps them find new or better hypotheses and insights. In addition, the influence on user confidence will be measured.

The results from future studies will eventually be compiled in a framework to provide guidelines for the development of tools for uncertainty-aware GVA. This has been identified as one of the major challenges in VA research: 'In general, we have a host of technology, but for a given task, the challenge is to provide guidance on what to use (e.g., method of analysis, type of visualisation), how to use it, and how to decide if it was a good choice' (van Wijk et al. 2010, p. 147).

6.4 Prototype *ICchange*

We have developed a software prototype named *ICchange* ('*I see change*') for the interactive exploratory analysis of land cover change. It is intended to be a proof of concept for the approach and to provide the basis for future user studies on the role that uncertainty plays in GVA. The development was divided into three steps: The first step was to create a typology of tasks that occur during the exploratory analysis of land cover change. From the comprehensive typology we selected a subset of tasks to be supported by the prototype (Section 6.4.1). The second step was the design of the prototype and its implementation. We designed two linked views to provide different perspectives of the change data: The *map view* shows the spatial attributes of changes, and the *change info view* provides an overview of further characteristics. In combination, the two different views facilitate an exploration of the changes through interactive filtering (Section 6.4.2). The third step in the development was a user study which revealed the strengths and weaknesses of the design (Section 6.4.3).

6.4.1 Tasks

The first step in the development process was to identify typical recurring tasks in exploratory change analysis from literature, personal experience, and discussion with domain experts. Based on the task typology by Andrienko and Andrienko (2006), we systematically described possible lookup, comparison, and relation-seeking tasks in the dimensions of space, time, and change attributes. The typology was sent to three different experts from the RS domain to receive feedback. The initial goal was to let them rate the tasks in order of importance for their work, but this turned out to be difficult. Although concrete examples were given (e.g., 'Find out where and when changes of the type "agriculture to built-up land" have occurred') the more complex tasks were hard to understand and a reliable rating was not possible. Concerning completeness, all three experts stated they were able to identify the main tasks they could think of, but they could not tell if the list was complete. Although we did not get a task rating in order of importance we were able to use the feedback to compile a reasonable selection of tasks which the prototype had to support. The challenge was to find a subset that limits the complexity of the tool but still makes exploratory analysis possible. We decided to concentrate on the most common tasks, i.e., lookup tasks as well as basic comparison and relation-seeking tasks.

Table 6.1 contains the list that served as the basis for the development of the prototype. It shows the tasks per category and how each task is supported in the prototype (refer to the next section for a description of the map view, the change info view and their functionality).

6.4.2 Design and implementation

The prototype was intentionally kept simple and was not supposed to be a full-featured change analysis tool. Therefore it only supports selected tasks of exploratory change analysis that we defined in advance. We decided to use two linked views to facilitate the analysis of changes in space and time. The first view is an interactive map showing changes that have occurred in the area of interest. Generally, it is not trivial to represent temporal characteristics on a map. Animation can be used, but it increases the complexity of the view. In addition, there is evidence that representing change through animation can lead to the effect of 'change blindness' (Blok et al. 2011, Zurita-Milla et al. 2012). For this reason we decided to add a nonspatial view called the *change info view* in order to show changes along with information on time, amount of change and uncertainty. We created different drafts and discussed them at conferences, workshops, and internal meetings to iteratively improve the prototype design. The implementation of the entire software was done using Java and geotools, an open source library providing GIS functionality (http://geotools.org). In the following, the design is described in detail with a focus on uncertainty visualization. For the complete design of the prototype refer to Figure 6.5.

Map view

The map view was created to provide a view of spatial attributes showing changes in the area of interest, mainly concerning lookup, comparison, and relation-seeking tasks related to space. A number of layers are available on the interactive map: Layers showing the RS images and the land cover classifications derived from them. Another layer serves as a mask to exclusively show changed areas. The map view provides standard navigation (panning and zooming) and by clicking on the map the user gets information about the type and uncertainty of the change at a selected location.

Representing the different types of change ('What has changed to what?') was challenging because the number of types grows significantly with an increasing number of land cover classes and number of datasets. For instance, five classes and three datasets already result in 125 possible different change types. Although not all variations occur in real analyses the number of different change types tends to be very high and thus difficult to display. Hence, we did not try to encode the change type graphically. Nevertheless, the user can toggle between the land cover layers on the map and by using filtering and tooltips it is possible to identify the change types (see Section 4.2.2).

Table 6.1. Selected tasks supported by the prototype.

Category	Task	Support in Prototype
Direct Lookup	Determine the type/uncertainty of the change in a defined location	Map view: tooltip
	Determine the amount of changes of a defined type	Change info view: left column
	Determine the mean uncertainty of changes of a defined type	Change info view: right column/tooltip
Inverse Lookup	Find out *where* changes with a defined type/uncertainty have occurred	Change info view: filter, map view
	Find out *when* changes with a defined type/uncertainty have occurred.	Change info view: filter, center column
Direct Comparison	Compare the type/uncertainty of change A and change B	Change info view: sort, center column
Inverse Comparison	Find changes of the same type/uncertainty as change A in a defined time period	Change info view: sort, center column
Relation-Seeking	Find changes of the same type/uncertainty that occurred during the same time period	Change info view: sort, center column
	Find changes that have occurred in the same time period as change A	Change info view: sort, center column
Pattern Definition	Find out how changes are distributed in *space*	Map view
	Find out how changes are distributed over *time*.	Change info view: center column
	Determine how often different change types occur	Change info view: left column
	Determine the distribution of uncertainty in a group of changes	Change info view: right column
Pattern Search	Find regions in which changes form a spatial cluster	Map view
	Find time periods in which changes accumulate	Change info view: left/center column
	Find out which types of change occur most often	Change info view: sort, center column
Direct comparison (pattern)	Find out if changes of a defined type form a spatial cluster	Change info view: filter, map view
	Find out if changes of a defined type have occurred most often in a defined time period	Change info view: sort, left/center column
Inverse comparison (pattern)	Compare the position and extent of the clusters of large-area changes and small-area changes	Map view
Relation-seeking (pattern)	Find out if there are regions where changes of the same type are accumulated in different time periods	Change info view: filter, map view
	Find changes of the same type that have occurred with a similar frequency in different time periods	Change info view: sort, center column
	Find out if the major change types in different time periods are the same	Change info view: sort, left column

Another challenge was to create an appropriate visual representation of change uncertainty. A variety of uncertainty visualization techniques exist, and it is not trivial to find a suitable one for a certain application. A first guideline was the distinction between intrinsic and extrinsic approaches (Gershon 1998). Intrinsic techniques use existing graphical variables, e.g. color hue, saturation or transparency to represent uncertainty. Extrinsic approaches incorporate additional graphical objects to signify uncertainty, e.g., glyphs or other objects that are added to the display. Unlike most intrinsic variables, they can be visually separated from the content. As existing studies indicate that visually separable representations are preferable for exploratory use (MacEachren et al. 1998, Blenkinsop et al. 2000) we focused on extrinsic approaches. Our choice was to implement *annotation lines*, an extrinsic technique originally described by Cedilnik and Rheingans (2000) under the term *procedural annotations*: A regular grid is placed on the map and is distorted locally to represent the degree of uncertainty. It follows the suggestion to use metaphors to make uncertainty visualization more understandable (Gershon 1998). The authors of the original paper proposed four different versions of annotations (variation of width, sharpness, noise, and amplitude) from which we decided to implement noise. On the one hand, from the four alternatives, noise, and sharpness seemed to be the most suitable metaphors for uncertainty while on the other hand, sharpness (often referred to as its opposite, blur) has been shown to be hard to read (Boukhelifa et al. 2012). *Noise annotation lines* display uncertainty as noise ranging from a solid line (very low uncertainty) to a diffuse line (very high uncertainty) (Figure 6.1). They cause less occlusion of the underlying map than, for instance, areal glyphs. We conducted an experiment to evaluate how many levels of uncertainty can be distinguished using different design variations of noise annotation lines (Kinkeldey et al. 2013). We found that for a comparison task, user performance was relatively high, even with six levels of uncertainty to distinguish, but the design of the grid influences this. For use in the prototype we extended the technique by interactive means: The user can vary the cell size of the grid by using the mouse wheel and change the position of the grid by dragging the mouse. This allows the user to be able to explore the spatial distribution of uncertainty.

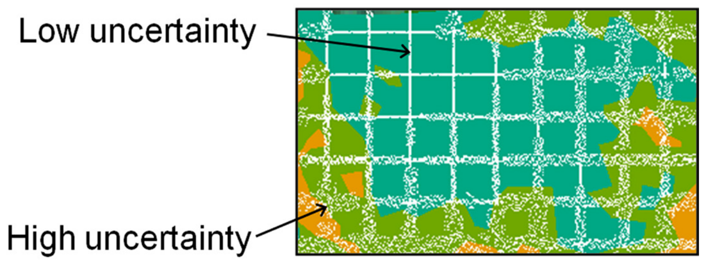

Figure 6.1. Principle of noise annotation lines. Refer to Kinkeldey et al. (2013) for a detailed description of the technique.

Change info view

A second view was employed as a nonspatial overview visualizing changes and their attributes over time. We decided to display the changes, grouped by type, in a scrollable list with each row representing all changes of a certain type. The horizontal axis represents time, and color-coded circles indicate the land cover class at a certain date. Figure 6.2 shows the principle: changes over three dates are displayed (May, August, and October 2010). The temporal distance between the dates is represented graphically, i.e., the gap between May and August is larger than that between August and October. In the left column, white bars show the amount of each change, i.e., the number of pixels. By hovering over one of the rows, a tooltip provides additional information such as the exact amount of changed area and uncertainty.

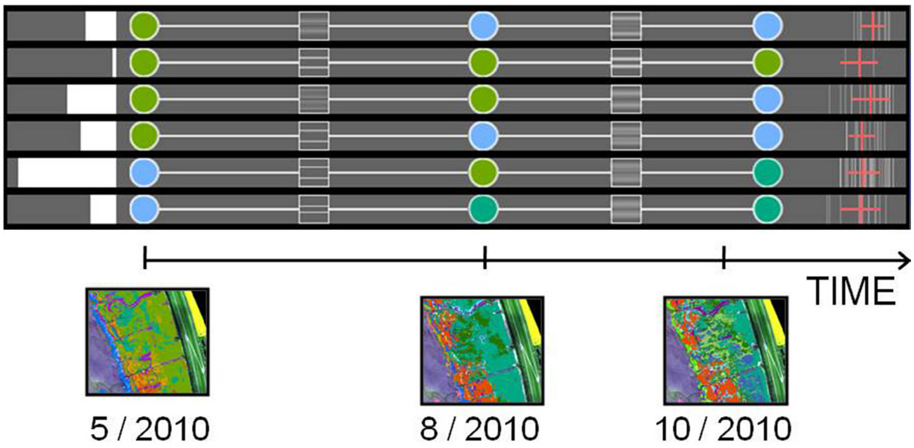

Figure 6.2. Change info view.

Two kinds of uncertainty distributions had to be displayed in each row of the change info view: Uncertainty of changes in specific time periods (e.g., between May and August) and over the whole time period (overall uncertainty). In the first designs we used histograms for both, but it was hard to compare several rows due to their visual complexity. Additionally, we were looking for a type of visualization that represents uncertainty more intuitively. So we decided to design novel symbolizations, one for single period change uncertainty (between each pair of circles) and a different one for overall change uncertainty (in the right column). The reason why we used different representations is that they serve different purposes: for the first one (single change uncertainty) compact size was most important. For the second one (overall uncertainty per row) the primary goal was to make the comparison between the rows possible.

For depicting the uncertainty of specific time periods we have designed a symbol called *barcode* that is based on the following idea: when all changes of the group are absolutely certain it appears as a straight line. The higher the uncertainty, the discontinuity of the line becomes larger. Furthermore, not only a single uncertainty value but the distribution of a set of values can be represented: several lines representing uncertainty values are drawn on top of one another as one symbol using transparency. Thus, if a symbol shows clear lines, all uncertainty values are equal or similar, whereas a blurry symbol indicates that the distribution contains many different values. Figure 6.3 contains four examples to clarify the idea: The first three barcodes on the left represent changes of a uniform uncertainty value (0%, 50%, and 100%). For 0% uncertainty the connection is a straight line. With higher uncertainty (50% and 100%) the connection appears 'broken'. The fourth barcode on the right shows how a whole set of uncertainty values can be represented. Here, the values are distributed equally from 0% to 100% which results in a blurry appearance.

For visualizing the uncertainty over the whole time (overall uncertainty) we did not use the barcode display, instead we decided to create a line display showing all uncertainty values as gray lines on a horizontal scale from 0% to 100% uncertainty (Figure 6.4). With the use of a red cross the mean value (vertical) and standard deviation (horizontal) are displayed. The position of the vertical line on the scale signifies the mean value, and the length of the horizontal line shows the standard deviation. The intent of the prominent red crosses was to make comparisons between different rows easier.

The change info view offers sorting and filtering functionality. The list of changes can be sorted by different attributes, i.e., by land cover type for each date, by amount of changed area, and by mean uncertainty. For filtering the changes, one or more rows in the list can be selected so that they appear highlighted. This also affects the display of changes in the map view and makes it possible to compile subsets of changes of interest and view their spatial distribution in the map.

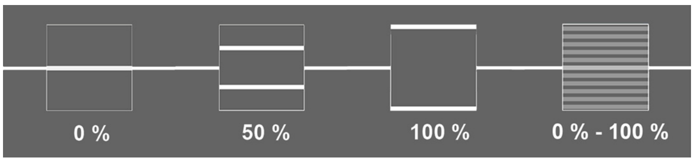

Figure 6.3. Barcode symbols which represent the change uncertainty given a specific time period.

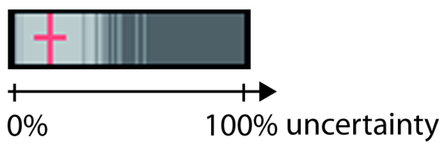

Figure 6.4. Line symbol which represents the change uncertainty over the whole time period.

Limitations

This prototype is tailored to facilitate simple but yet realistic exploratory change analyses. For use with arbitrary change data it has several limitations, the first one being scalability in terms of the number of land cover classes. On the one hand, classes are coded as color (hue) so there is a natural restriction regarding the use of easily separable colors. This can be solved by the use of additional shapes other than circles to symbolize land cover classes. But apart from that, the number of rows in the info view grows exponentially with the number of land cover classes, e.g., with five classes and three datasets there are already more than 100 rows to display. Although this is a theoretical value as not all combinations will occur in practice, a high number of rows is likely when using more classes. Scrolling the rows can function for hundreds of rows, but not with thousands or more. In this case, data reduction techniques such as clustering algorithms could help tackle this problem.

Another limitation refers to temporal scalability, i.e., a high number of dates would be difficult to display because the amount of space in the horizontal axis is limited. It is possible to scroll the view horizontally, but this slows down the tasks for which the information of an entire row is needed. However, this should not be seen as a severe limitation in our case as multi-temporal change analyses from satellite imagery typically deal with a low number of (usually three) datasets.

6.4.3 User evaluation

In order to assess the usability of the tool we conducted a low-level user evaluation based on a scenario of multi-temporal land cover change. The main question was if the tool is usable and understandable for expert users. An important question was whether or not users would be able to understand the uncertainty visualizations in the two views. The following sections describe the study setup, the group of participants and a discussion of the results.

Setup

We created a land cover change scenario based on real data but intentionally kept simple. We classified three SPOT satellite images from 2003 and 2004 that cover an agricultural area in the north of Spain. Figure 6.5 shows the prototype and the data used in this study. For depicting uncertainty in the changes, we used a straightforward measure based on the maximum ambiguity of each pixel during the classification of the three images. In this low-level study, the participants were not supposed to use contextual information. Hence, we did not use the real names of the land cover classes but named them by their color: green, blue, and orange. This made it easier to describe the tasks and for the subjects to make comments.

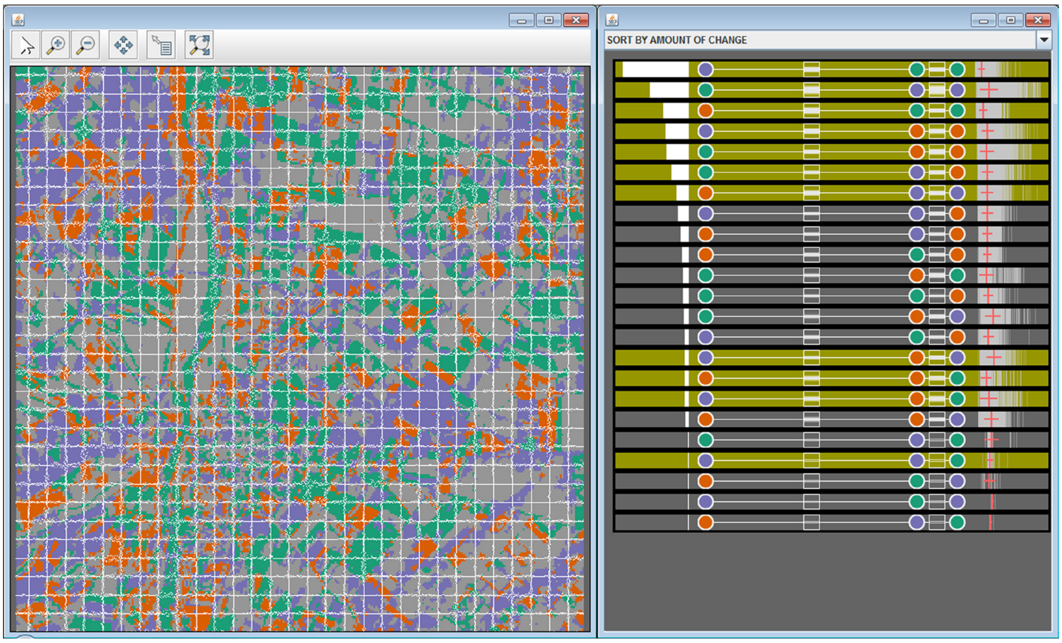

Figure 6.5. Prototype showing the scenario from the user study including a *map view* (left) and a *change info view* (right).

The study was conducted as verbal protocol analysis (VPA) that is described as 'especially valuable for understanding both the critical needs of an application, as well as its expected behavior from the perspective of the end-user' (Robinson et al. 2005, p. 246). The participants were provided with the tool and a number of tasks. They were asked to 'think aloud' while using the tool and we wrote down their comments during the procedure. The study comprised of three parts: First, the subjects were introduced to the purpose of the tool, its visualizations and the user interface. Special attention was given to the explanation of uncertainty and how it is being displayed in the two views. In the second part, a set of defined tasks were conducted using the tool. We used information retrieval tasks with a maximum of two steps, e.g.,

- Lookup (amount): 'Determine the amount of changes of the type "orange-orange-green"',
- Comparison (change type): 'Determine the type of change that occurred most over the whole time period' or
- Relation-seeking (uncertainty): 'Find out if the distribution of uncertainty in the north of the area is the same as in the south'.

We read the tasks out loud, let the subjects use the tool and wrote down how they did it and what their comments were during the analysis. The sessions lasted between 25 and 60 minutes depending on the speed and the interest of the participant. In the third part, we asked for their overall impression of the tool and if it lacked certain functionality that they needed. We also discussed their opinion about the uncertainty displays in both views. Finally, we asked them if they had any ideas for enhancing the tool or if they had general comments.

Participants

We recruited five male participants between 25 and 35 years old who were not paid for taking part in the study. Two of them were graduate students from the Geomatics master programme at HafenCity University. The other three consisted of a faculty member from Lab for Geoinformatics and Geovisualization, a GIS expert working for a telecommunication company, and a software developer from a company for RS applications. All of them had high expertise in GIS and RS gained from university education, whereas the latter three had additional work experience in these fields. None of them had everyday experience with change analyses, but we did not believe this to be important for a lowlevel test. Although the concept of uncertainty was not new to three of the five participants, none of them had ever used uncertainty in an analysis.

Results and discussion

In general, the tool was perceived as clear and understandable. Two participants said they liked the way time is displayed explicitly and that they could easily see what changes occurred during what time periods. Although the tool was supposed to focus on the analysis of changes only, three participants were missing information on the no-change areas for comparison. All of them liked that the amount and uncertainty of all changes is shown qualitatively in graphical form and that the exact values are provided by a tooltip. Two subjects remarked that the list was a feasible approach even with more rows than in this example. One participant suggested that a vertical line resembling a ruler could help make comparisons over a high number of rows. Related to sorting the list, three subjects said they would like to manually arrange the rows so that a set of changes could be positioned next to each other for comparison. There was one task that was hard to accomplish for all subjects: Determining how many changes had occurred in the first, in the second, or in both time periods. This is because it was not possible to sort the changes according to time periods when the change had happened. One of the participants suggested an additional function to sort by similarity of the rows.

Regarding uncertainty visualization in the change info view, all subjects had problems with the barcode display. Although they understood the idea of a straight connection when certain and a 'broken' connection when uncertain they could not read the symbols well. Two persons especially had problems when comparing them: They found the boxes too small and found it difficult to visually 'jump' from one row to the other. The line symbol used for overall uncertainty in the right column of the view was easier to understand for all subjects, because the prominent display of the mean value and standard deviation made it easier to compare the rows visually. Most participants did not understand the role of the gray lines that represented the uncertainty values. Although it simply showed the values on a scale from 0% to 100% of uncertainty they got confused. Three subjects proposed more standardized methods, e.g., box plots, to show the distribution of uncertainty.

A further intention of the study was to have a discussion on the noise annotation lines technique used to represent uncertainty in the map view. All participants quickly understood the way it worked although some were unsure if the values in the cell or directly below the grid lines were represented. They found the method useful for bigger clusters but remarked that with small clusters or single pixels the degree of uncertainty could no longer be identified without zooming in. The modification of the grid using the mouse wheel seemed to be straightforward, but two subjects suggested to additionally display the grid size in meters to get an impression on its relative dimension. One person tended to use cell sizes that were so small that the grid degenerated to a texture. This revealed very fine details but caused the technique to lose the advantage of covering a relatively small portion of the background only. Another participant proposed to additionally alter the size of the noise particles to highlight regions with high uncertainty. When we asked for the usability of the technique everyone agreed that they were able to estimate the distribution of uncertainty over the whole area and within clusters of change, that they could identify uncertainty hotspots, and that they were able to compare the degree of uncertainty in different clusters of change.

All in all the subjects agreed that the tool had potential to provide information and functionality that common GIS systems do not offer. They stated that the change info view clearly shows what changes have occurred over time and that it successfully complements the information in the map view. The reactions towards the uncertainty displays were, aside from the aforementioned criticism, predominantly positive although we should point out that a direct comparison to other techniques was not possible. Concerning the general reaction towards the uncertainty displays we can confirm the insight from other studies that '[n]one of the experts found the offered uncertainty visualizations overwhelming, confusing, or useless' (Kunz et al. 2011).

As expected, practical expertise (that was higher with three of the five participants) did not seem to play a relevant role in this study. There were differences in how the participants approached the tasks, but the issues they mentioned were similar regardless of practical experience in the field. This can be explained by the fact that the GIS and RS expertise was at a comparable level and that practical experience with change analysis was not necessarily needed in this study.

6.5 Conclusions and outlook

This contribution described the development of *ICchange*, an interactive visual prototype for exploratory land cover change analyses. This prototype was intended to provide a basis for user studies to evaluate the role of uncertainty in change analyses using GVA environments. The development process included the definition of a task typology for exploratory change analysis, a design phase using an iterative design approach, and a user evaluation.

In the first step of the development, a typology was created that described the tasks during the exploratory analysis of categorical change. It included tasks related to space, time, change type, and uncertainty. From this, only a subset of tasks was selected for the prototype, on the one hand to decrease the complexity but on the other hand to preserve the exploratory nature of the tool. The design process itself was based on sketches that were discussed with visualization experts to iteratively improve the design. As we intended to develop a tool with low complexity we decided to restrict the basic design to two linked views, the *map view* and the *change info view*. The *map view* provided an interactive map showing the distribution of changes over the area. The choice of an appropriate uncertainty visualization method was subject to extensive discussion. We ended up using what we call *noise annotation lines*, an extrinsic uncertainty visualization technique that has rarely been used in the past but met the requirements to represent uncertainty in the map view. Alongside this, we proved its basic usability in a user study. The *change info* view showed a list where each row represented a group of changes of a certain type over time while including additional information such as their amount and their uncertainty. For representing uncertainty we created a symbol we call *barcode* to represent uncertainty of change during a specific time period. In addition, another display showed the distribution of overall uncertainty and basic statistical information as line glyphs. After the description of the design process, limitations of the design have been discussed.

A qualitative study to evaluate the basic usability of the prototype was conducted as VPA. Five participants with a high level of expertise in GIS and RS individually conducted predefined low-level tasks with the help of the prototype. All in all, the study indicated that the design and implementation of the prototype had reached a usable level. However, the results showed that specifically the uncertainty displays in the change info view were not easily understood. The suggestions made by the participants will help improve this aspect of the design. A similar study after the optimization step will show whether or not the issues with the current design continue to persist. In retrospect, it would have been advantageous to conduct a study with concrete tasks earlier in the process. Previous discussions, based on sketches, were helpful but did not reveal difficulties people had when they were actually conducting a task. For instance, no one noticed difficulties with the barcode display before the actual study. For similar projects we recommend that people conduct concrete tasks as early as possible even if based on non-interactive, static images.

The following step of this research will be to develop a concept for user studies based on the prototype. The goal is to assess the role of uncertainty information in exploratory change analysis. For this, different existing case studies from change analysis will be selected and, with the help from domain experts, enriched by uncertainty information for the use in the prototype. Analyses with and without uncertainty representation will show whether or not this information leads to new hypotheses and improved insights. The results gained in the study and the experience from the development of the prototype will be compiled in a framework to provide guidelines to support the development of GVA tools for uncertainty-aware change analysis and help get closer to the overall goal of making uncertainty information more usable in this field.

Acknowledgments

The author would like to thank Jörg Münchow, Tonio Fincke, Jan Dawidzinski, Johannes Kröger, and Julian Timm for voluntarily taking part in the study and for their ideas and suggestions regarding the prototype. Helpful comments on the article provided by Jochen Schiewe (HafenCity University), Till Nagel (FH Potsdam), Mark Iliffe (University of Nottingham), Jennifer Mason (The Pennsylvania State University), and Maya Donelson (HafenCity University) were also greatly appreciated.

7 Evaluating the use of uncertainty visualization for exploratory analysis of land cover change: A qualitative expert user study

This chapter was previously published as peer-reviewed journal paper:

Kinkeldey, C., Schiewe, J., Gerstmann, H., Götze, C., Kit, O., Lüdeke, M., Taubenböck, H., and Wurm, M., 2015. Evaluating the use of uncertainty visualization for exploratory analysis of land cover change: A qualitative expert user study. Computers & Geosciences 84, 46–53.

The candidate was the primary author and responsible for the interviews incl. data preparation, the summary of the findings, and contributed to the desription of the change scenarios and the discussion (70% of the overall work).

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 fields 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 expertize. 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.

Keywords

Uncertainty visualization; Remote sensing; Land cover; Change analysis; User study

7.1 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., 2014b), 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 recommendations for future work.

7.2 Method

The goal of this research was to assess the concept for uncertainty-aware land cover change analysis described in 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., 2014b). 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 depict 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.

7.2.1 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 *ICchange* software prototype and on *noise annotation lines*, the technique we used to display uncertainty in the map (Kinkeldey et al., 2014b). 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’.

7.2.2 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 μ_0 and μ_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 (Figure 7.1, ‘u classed’). Please refer to Kinkeldey (2014a) for a detailed description and discussion of the measure.

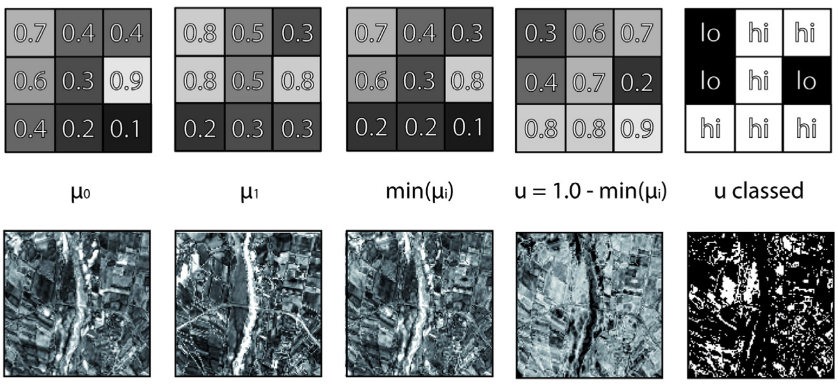

Figure 7.1. Change uncertainty measure derived from class membership values μ_i [reprinted from Kinkeldey (2014a)].

7.2.3 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 (Figure 7.2). 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).

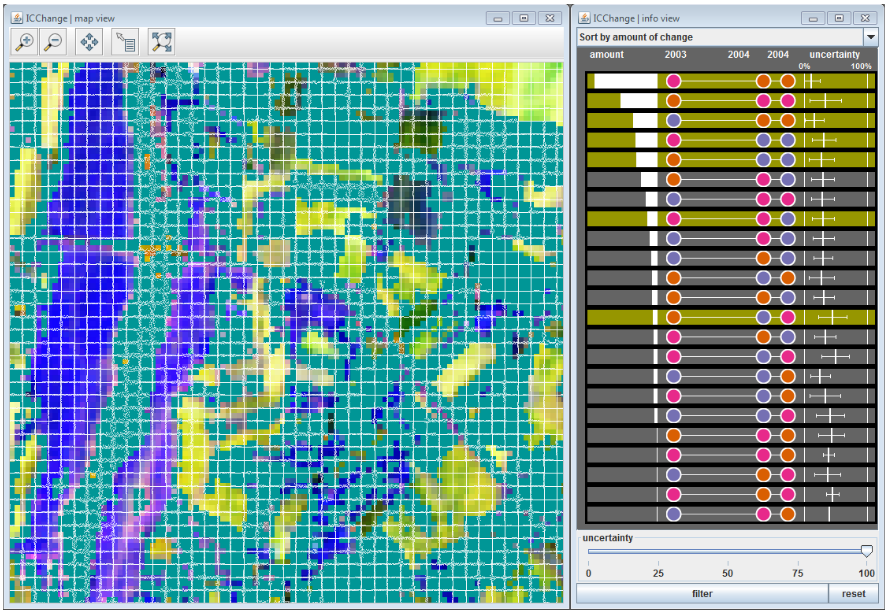

Figure 7.2. Software prototype ICchange. In the map view (left) the green layer on top of a satellite image represents land cover change and noise annotation lines depict connected uncertainty. The info view (right) shows supplementary information about occurring changes and provides a slider to filter by level of uncertainty.

7.3 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.

7.3.1 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 (Figure 7.3). 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 (Figure 7.4). 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.

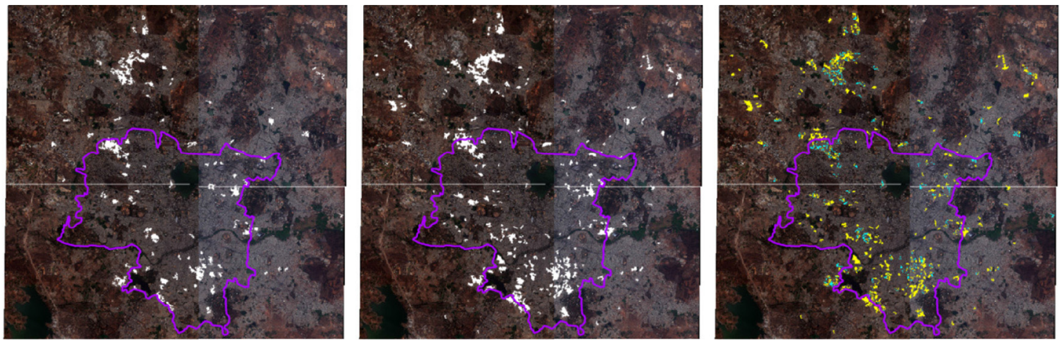

Figure 7.3. Informal settlements of Hyderabad in 2003, 2010 (white: detected informal settlements), and the change map (yellow: growth, blue: reduction) [adapted from Kit and Lüdeke (2013)].

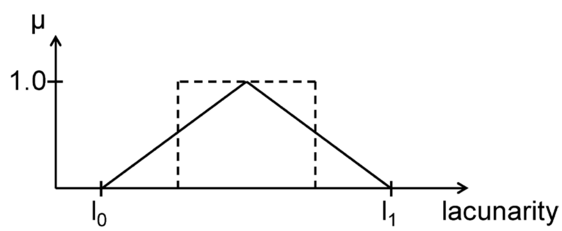

Figure 7.4. Fuzzy function defining class membership values μ based on the lacunarity interval for the 'informal settlements' class. Instead of the crisp definition (dashed line) we used a fuzzy interval (solid line) to define the membership values for the change uncertainty measure.

7.3.2 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[7]. 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 Figure 7.5, 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. Figure 7.6 shows the mapview including the change map and related uncertainty represented by noise annotationlines.

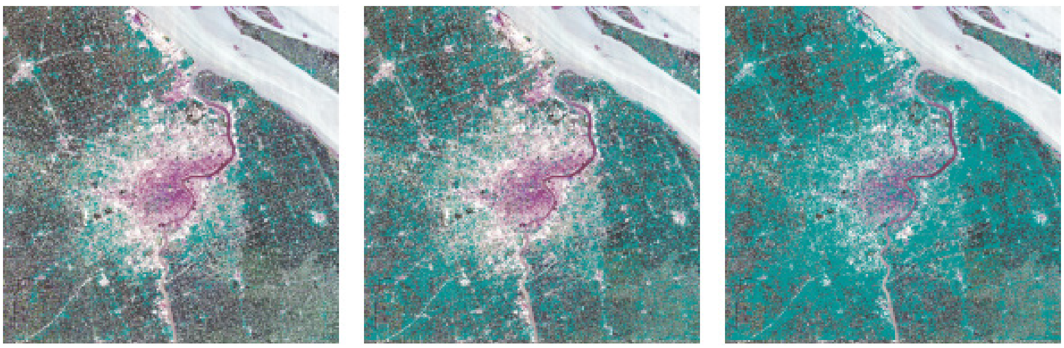

Figure 7.5. Change in urban areas in Shanghai, China, shown in green: between 1987 and 1995 (left), between 1995 and 2004 (center) and between 1987 and 2004 (right). Change is displayed over the Landsat image for the year 1987.

[7] http://www.ecognition.com/

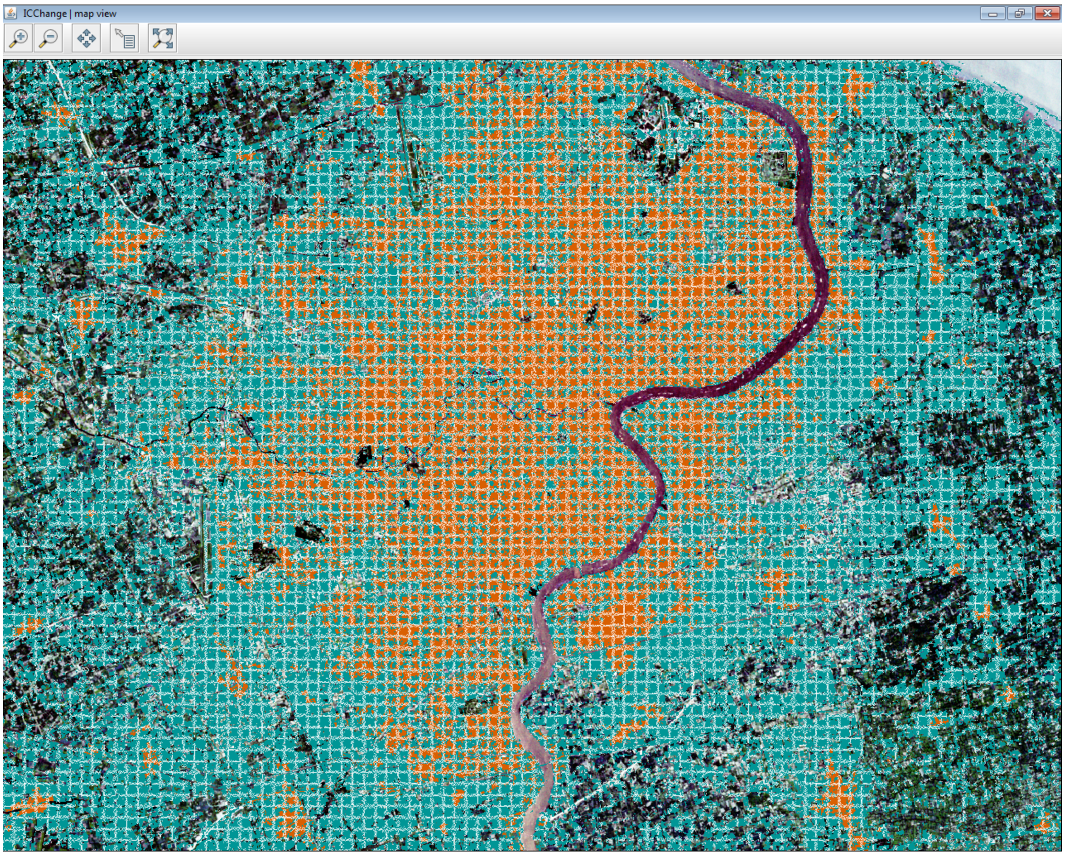

Figure 7.6. Map view showing (unchanged) urban areas in Shanghai, 1987 (orange) and urbanization between 1987 and 2004 (green). Uncertainty related to changes and non-changes is depicted by *noise annotation lines* (Kinkeldey et al., 2014b).

7.3.3 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 data sets 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 and Gerstmann et al., 2014). The separation of the different grassland types turned out to be the most uncertain class separation. Refer to Figure 7.7 for the land cover datasets for 2000, 2003, and 2009.

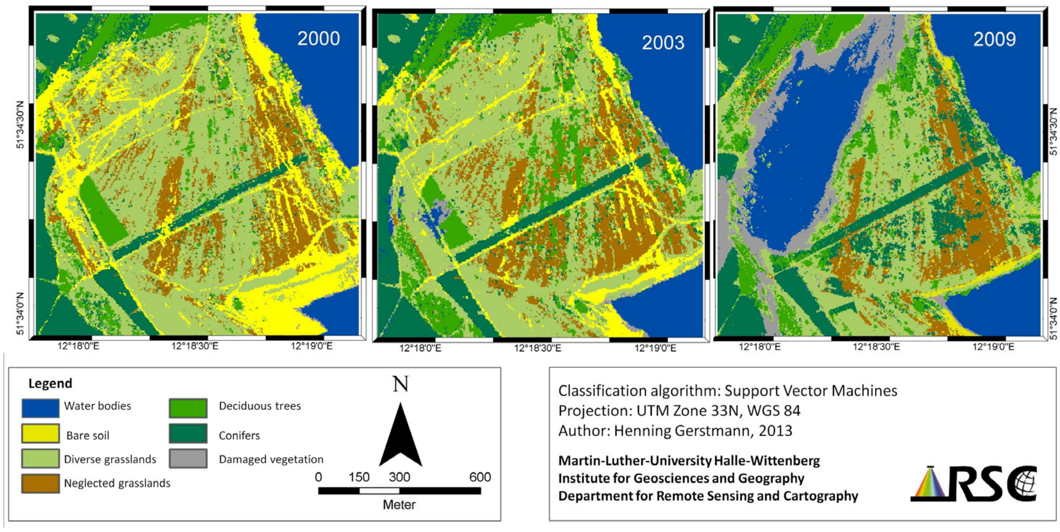

Figure 7.7. Land cover maps of Petersroda, Germany for 2000, 2003, and 2009 [reprinted from Gerstmann (2013)].

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.

7.4 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 prototype and uncertainty visualization technique.

7.4.1 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 7.3). 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 me 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.

7.4.2 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.

7.4.3 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 expertize. 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.

7.4.4 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*[8], *QGIS*[9]) 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., 2014b). Some participants recognized limitations, e.g., that readability may be low with a bright or visually complex background.

[8] http://www.arcgis.com

[9] http://qgis.org

7.5 Discussion

In this section we discuss selected findings from the summary in Section 7.4. In addition we highlight the benefits and limitations of the method used for the study.

7.5.1 Findings

After summarizing the findings from 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 expertize, confirming results from past uncertainty visualization studies (Cliburn et al., 2002, Hope and Hunter, 2007b and Roth, 2009b). All this stresses the importance of carefully defining types and levels of expertize 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.

7.5.2 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.

7.6 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 ArcGIS 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 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.

8 Conclusion

Research about geodata uncertainty has been conducted for more than two decades, focusing on models to describe and quantify uncertainty (Foody and Atkinson 2002, Shi 2010, Zhang and Goodchild 2002) and ways to visually communicate this to users (Buttenfield 1993, Evans 1997, Griethe and Schumann 2006, Pang 2001). Over the years, visualizing uncertainty during spatial analysis has been recognized as beneficial (Caers 2011, Deitrick and Edsall 2008) and has been shown to impact decision making (Deitrick and Edsall 2006, Leitner and Buttenfield 2000) or risk assessment (Roth 2009, Severtson 2012).

While most evaluation efforts regarding the effects of uncertainty visualization are related to decision making, empirical research on the role of uncertainty in exploratory analysis of geodata is still underrepresented (MacEachren et al. 2005). This doctoral dissertation contributed to filling this gap with a concept to model, quantify, and communicate uncertainty connected to land cover change in remote sensing (RS). As a practical implementation of the concept, an interactive visual prototype for change analysis was developed, including a technique to visualize uncertainty connected to change data. Based on this, an expert study assessed the applicability of the concept in practice.

8.1 Summary of results

RQ1. How can information about uncertainty be incorporated into land cover change analysis?

The core of this research was to develop a concept to utilize uncertainty for exploratory land cover change analysis (Chapter 2). The main goals involved closer coupling of change detection and analysis steps, as well as communication of uncertainty during analysis. The concept aimed to fulfill both goals by following a geovisual analytics (GVA) approach which facilitates interactive visual interfaces to integrate visual analysis with automated algorithms. An important part of the concept was the definition of an uncertainty measure for land cover change analysis, taking into account classification confidence in terms of class membership values and providing a summative estimation of change uncertainty. The paper discussed three potential applications of uncertainty information: better informed analysis of change, optimization of change detection parameters, and removal of erroneous change using filtering by uncertainty. A case study utilizing RapidEye satellite data of a rural area near Hamburg, Germany, showed how erroneous change can be identified visually and filtered out based on change uncertainty using a simple prototype. The change uncertainty measure was shown to be a viable indicator of erroneous change, but limited by the fact that filter thresholds determined for a part of the dataset are not necessarily transferable to the whole dataset. Supplementary filter criteria can be used to counter this limitation, e.g., the size or spectral attributes of changed areas. The case study showed how the concept can help decrease error in land cover change data by incorporating uncertainty into the analysis.

RQ2. What is a suitable method for the visual representation of uncertainty in land cover change maps?

To answer this research question, two steps were taken: a literature review of uncertainty visualization user studies, and the selection and evaluation of a technique to display uncertainty in land cover change maps. The first step, a systematic literature review of user studies in the field of uncertainty visualization, consisted of two articles (Chapter 3 and 4). The first article covered studies that assess communication aspects ('how well do different techniques communicate uncertainty?') and the other considered studies that deal with the effect that visualized uncertainty can have ('if and how do decisions change when uncertainty is displayed?').

In the first part of the review, the main findings were the following: Generally, a number of studies have shown that users accept the additional effort of reading and processing uncertainty visualizations, but only when the visual depiction is well chosen and understandable – an observation that stresses the importance of user evaluation in the field of uncertainty visualization. In terms of user performance (accuracy, speed) of different uncertainty depictions, intrinsic[10] techniques such as color hue, color value, and transparency provided good results, as well as extrinsic[10] approaches such as glyphs. Regarding the question of intuitiveness, techniques utilizing fuzziness, position, and color value were found to be successful. From what we know today, color saturation, which is often used to visually represent uncertainty ('fading out' metaphor), cannot be recommended as a first choice. First, it has been shown that it is not as intuitive as other techniques, especially regarding the question of 'which end is up', i.e., if lower or higher saturation represents higher levels of uncertainty. Second, when saturation is low and color value is used to express the data, different values can be hard to separate from each other.

However, these findings have to be judged carefully. Results were difficult to summarize and generalize because results tended to be ambiguous and hard to comapre between most studies. For this reason, general guidelines were difficult to derive. Another 'lesson learned' was that special techniques may be needed to effectively visualize uncertainty, in comparison to other map attributes (such as land cover classes). For instance, the use of metaphors such as fog, blur, or

[10] For a brief introduction into the categories used here, refer to Chapter 1. A detailed discussion can be found in Chapter 3.

noise was deemed to increase the intuitiveness of the uncertainty display, but research on this aspect remains rare. For future work, a more systematic selection of techniques was recommended, based on typologies of uncertainty visualization. Evidence in the literature suggests that beside types of data and uncertainty which common typologies are based upon, the task also plays an important role. Thus, the recommendation was to develop novel typologies that consider tasks to support the development, evaluation, and selection of uncertainty visualization techniques in the future. For a further discussion of this issue, refer to Subchapter 8.2.

The second part of the review analyzed studies that assess the potential effects of uncertainty, most of them focusing on decision making. A number of different effects were observed, e.g., on decision performance (accuracy, speed), on decision outcome, or on user confidence. The majority of studies suggested that uncertainty visualization affects decision making performance, often in terms of higher decision accuracy. In addition, visual depiction of uncertainty did not overwhelm the participants in most cases, and did not slow down decision making although more information had to be processed. Another important finding was that expertise was deemed to play an crucial role on the impact uncertainty can have. However, as with the first part of the review, it was difficult to compare and generalize the results from different studies because of the unsystematic selection of tasks, data, and participants. That is why for the future, more systematic study setups are needed, e.g., more careful definition of the type and the level of the participants' expertise, to produce results that can be analyzed and compared in a systematic way, with the aim of generating guidelines.

Based on the findings from the review of user studies, the second step needed to address this research question was to choose a suitable technique for uncertainty visualization in land cover change maps. Since maps of this type typically contain areas that are very irregular in shape and size, we disregarded intrinsic techniques (such as manipulation of color attributes to represent uncertainty), which were likely to cause clutter. In addition, there was the assumption from the earlier review that the use of metaphors could be a means to increase the intuitiveness of uncertainty visualizations. The *noise annotation lines* technique, introduced by Cedilnik and Rheingans (2000) under the term *procedural annotations*, fulfilled these requirements as an extrinsic, grid-based method that signifies uncertainty using a noise metaphor. This approach had not been evaluated in user studies before, thus, two web-based user studies were conducted to assess the method's usability with different grid line designs and varying number of uncertainty levels. Using a comparison task between the levels of uncertainty in two areas of a map, high user accuracy was observed for up to six levels with the most salient design of the grid. The results from the user study confirmed the hypothesis that noise annotation lines are an appropriate technique to represent attribute uncertainty in land cover change maps.

In the final study of this dissertation (Chapter 7), experts conducted change analyses with the *ICchange* software prototype (see RQ3), which uses noise annotation lines to display change uncertainty in a map. The experts' opinions about the technique were predominantly positive. Most participants especially liked the following three characteristics: First, the fact that the change map does not have to be modified, second, that it decreases the 'salt-and-pepper-effect' in noisy uncertainty layers, and third, that the level of detail can be varied by changing the cell size of the grid. They also identified potential drawbacks such as the low contrast between the white grid and bright map content.

All in all, it can be stated that noise annotation lines have been proven to be an appropriate technique for visualization of uncertainty during exploratory change analysis. Their use could as well be beneficial for other applications (see Subchapter 8.2.1).

RQ3. How can a software tool support uncertainty-aware land cover change analysis?

As mentioned above, an important part of this research was the development of a software prototype for uncertainty-aware change analysis (*ICchange*). It served as proof-of-concept and as a vehicle for the expert study (see RQ4). The paper in Chapter 6 provided an in-depth discussion of the prototype and its development. A description of low-level tasks occurring during exploratory change analysis served as the basis for the development of the prototype. It was designed to support these tasks by facilitating two linked views: the 'map view' and the 'info view'. The map view showed the change map and additional data (satellite scenes, land cover data, etc.) as well as the change uncertainty layer visualized using noise annotation lines (see RQ2). The info view provided an abstract overview on the occurring changes, including their type, amount, and uncertainty (Figure 8.1).

To assess the usability of the prototype, a verbal protocol analysis (VPA) user study was conducted, involving five participants with a high level of expertise in GIS and RS. The participants were asked to conduct a number of analysis tasks from the list mentioned above and to 'think aloud' during the analysis. The findings suggested that generally, all analysts could understand and successfully use the tool. A critical issue was the way uncertainty was displayed in the first version of the info view ('barcode display'), which was not understandable for all participants (Chapter 6). This display was replaced by a simple glyph depicting mean and standard deviation of the uncertainty of all changes of a type. Refer to Chapter 7 for a description of the updated prototype.

Further findings regarding this research question were provided by the expert study described in Chapter 7. In the interviews, several participants mentioned the lack of tool support as the main reason why they had not used uncertainty in practice. Therefore, it was interesting to find out how they would judge the *ICchange* prototype, and if they could imagine using it in their work. Generally, the prototype received positive reactions with all groups. The experts quickly understood its operation, wanted to use it with their data and came up with possible applications. This outcome was not expected because the prototype was intentionally kept simple and it was presumed that experts wanted richer functionality. However, the positive results from the expert study, supported by the outcome of the above mentioned usability study, suggest that the prototype has met its purpose of being a simple, yet powerful implementation of the concept presented in this dissertation.

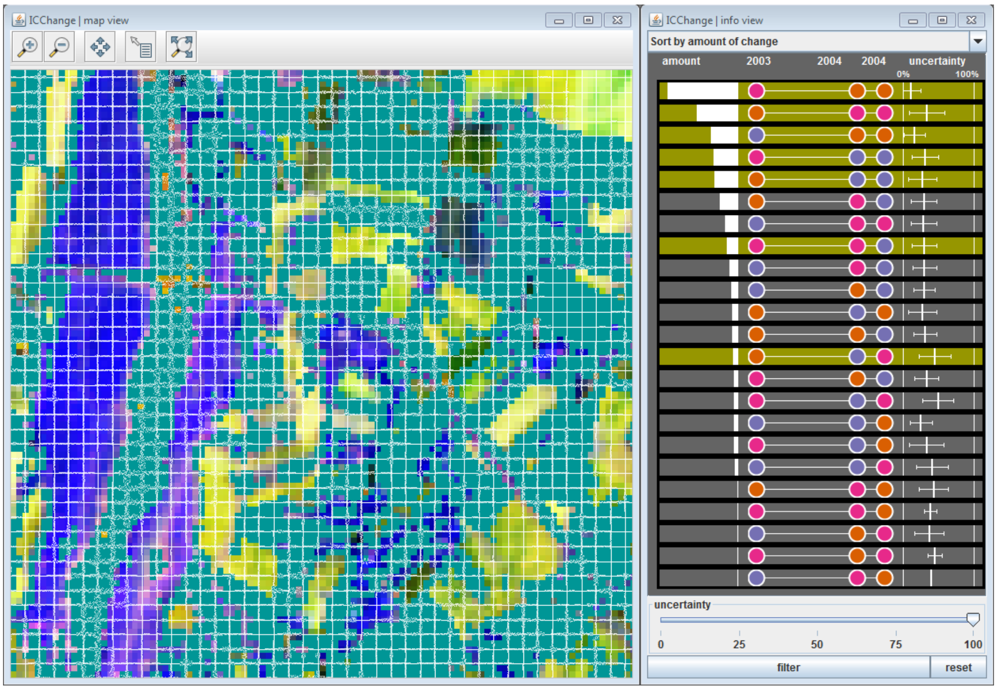

Figure 8.1. Software prototype ICchange including 'map view' (left) and 'info view' (right).

RQ4. Does the developed concept help experts with land cover change analysis?

An important part of this research was an expert study to learn more about the question of if and how analysts could use the developed concept to incorporate uncertainty information into the analysis of land cover change. For this, semi-structured interviews were conducted with three expert groups who deal with change analysis in their work. The change scenarios were provided by the groups, related to informal settlements (Hyderabad, India), growth of urban areas (Shanghai, China), and vegetation change in a former mining landscape (Petersroda, Germany). During the interviews, each group was confronted with their own change scenario in the *ICchange* prototype, thus, the experts were already familiar with the respective dataset. In addition to the change data, the change uncertainty layer was computed for each dataset and displayed in the prototype. Since the interviewer operated the prototype on behalf of the experts, they could focus on analyzing the change data and associated uncertainty.

The interviews were divided into four parts: The first part was an introduction to explain the prototype and the uncertainty measure. The second part was concerned with the use of uncertainty for change analysis: we first posed questions regarding the specific change scenario and, second, continued with a general inquiry about the potential of uncertainty information regarding their work. Questions about the ICchange software prototype and the noise annotation lines technique were asked in the third part of the interviews. In the fourth and last part, the experts could make free comments. The recorded interviews were transcribed and categorized related to 'change detection and analysis', 'reasoning with uncertainty', 'communication of uncertainty', and 'tool and visualization'.

The first general finding from the interviews was that all groups were interested in using change uncertainty. Geographically varied information on uncertainty (opposed to one quality value for a whole dataset) was novel for them and they came up with ideas how to utilize it to enhance change detection outcomes, e.g., optimizing detection parameters or by filtering out erroneous change with the help of uncertainty. Regarding reasoning, the analysts were interested in using change uncertainty for the interpretation of change maps. They stated that the information could be useful to confirm or revise hypotheses about change, but in the end, most of them wanted to derive hard facts ("the city has grown by 20%"). Uncertainty was seen as important for internal use and for communication within academia, e.g., when characteristics of change detection algorithms are described. However, communication of uncertainty to users of change data, such as decision makers, was seen as critical by the majority of experts. On the one hand, they were concerned that their data may be seen as inferior because of the communicated uncertainty. On the other hand, they were not convinced that users would be able to use data annotated with uncertainty information in a beneficial way (an issue revisited in Chapter 8.2).

Regarding the methodology, the choice of semi-structured interviews and the decision to utilize real data in a software prototype were proven to be successful for the purpose of this work. The participants seemed inspired by seeing their data and related uncertainty, and it is likely that the discussions would not have been as deep if they would have been based on a theoretical description of the concept. All in all, the study showed how qualitative methods can successfully contribute to improve the understanding of uncertainty-aware geodata analysis.

8.2 Implications and future work

The findings summarized above made it possible to generate several implications related to the fields of uncertainty visualization and uncertainty-aware change analysis. They will be addressed in the following two subsections, including generalizations of the findings and recommendations for future work.

8.2.1 Uncertainty visualization

As discussed above, a large variety of uncertainty visualization techniques for geographic displays exist. To systematize the approaches, three main categories were identified: ‘intrinsic / extrinsic’, ‘coincident / adjacent’, and ‘static / dynamic’, represented as axes of an ‘Uncertainty Visualization cube’ (*UVis*^{3}) (Figure 8.2). Apart from the use to structure the literature review in Chapters 3 and 4, it can be a means to systematize the selection of techniques for uncertainty displays during tool development.

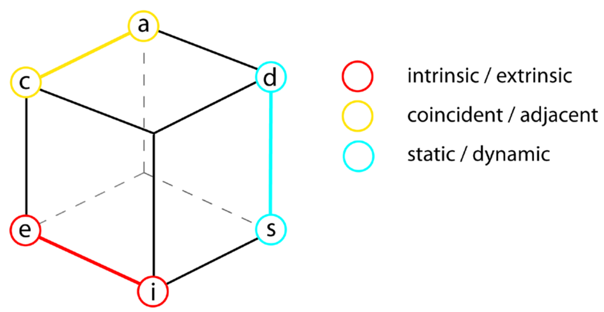

Figure 8.2. The UVis³ systematization for uncertainty visualization techniques (refer to Chapter 3).

One important part of this work was to find a suitable uncertainty visualization method for exploratory analysis of land cover change maps. The requirements were set by the characteristics of the map type, but also by the exploratory nature of change analysis. When precise analytic questions cannot be defined in advance, the use of adaptive (or adaptable) techniques is recommended. In this case, *noise annotation lines* were found suitable, which use a regular grid and a noise metaphor for representing uncertainty (Figure 8.3): the width of the noise grid represents the level of uncertainty below the grid line. They fulfill two requirements: first, they are relatively independent from the content of the map and do not need regularly shaped map objects to be readable. Second, the grid cell size can be adapted to obtain different levels of detail needed for the task, as shown in Figure 8.3. The software implementation of the technique is available on the web[11] as free and open source Java code under the Apache License 2.0.

The noise annotation lines technique has been shown to be a suitable technique to display change uncertainty (as defined in this work). However, it is not restricted to the representation of uncertainty in land cover change and can be recommended for various types of attribute uncertainty (e.g., in elevation) as well as temporal uncertainty (e.g., in measurement dates), especially with data depicting irregularly shaped areas. It may be useful to depict positional (geometric) uncertainty, but we would not see it as first choice for representation of vague objects with uncertain locations and / or boundaries.

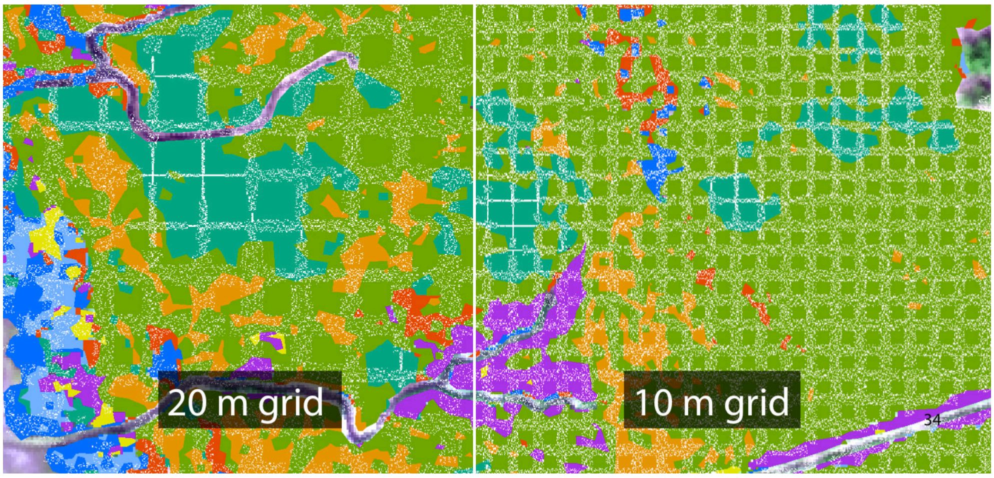

Figure 8.3. Noise Annotation Lines with 20 m and 10 m grid cell size.

A reasonable step in future work will be to convert the code into other programming languages. The implementation provided here uses Java which is a widespread language and offers a multitude of libraries for geodata processing. However, nowadays, Java is rarely run in web browsers, mainly because of security issues with the Java browser plug-in[12], which can be seen as a major drawback. A JavaScript version of the code would be meaningful, based on libraries such as D3.js, a powerful library for dynamic visualization in web browsers[13]. This would make noise annotation lines usable in browser-based web applications such as map clients. It is likely that this step would increase the usage of the technique, not only in geographical visualization, but also in related fields such as information visualization, and arouse interest in further development of this promising technique.

[11] https://github.com/ckinkeldey/nalines

[12] http://www.zdnet.com/article/how-big-a-security-risk-is-java-can-you-really-quit-using-it

[13] http://d3js.org/

The example of noise annotation lines being a suitable approach for exploratory analysis shows that it is not only the characteristics of data and uncertainty that determine if a technique can be recommended, but also the tasks it is used for. Evidence for this assumption can be found in uncertainty visualization literature (Pang 2008, Sanyal et al. 2009). One of the conclusions from our literature review on uncertainty visualization user studies (Chapters 3 and 4) was that a task-oriented systematization of uncertainty visualization can be beneficial, i.e., typologies using the task type as a characteristic (in addition to data characteristics). Task categories proposed in the conclusion of the review article could be a basis for this new kind of typologies (Table 8.1). This recommendation is not limited to geographic visualization of uncertainty and is also valid for related fields such as information visualization or scientific visualization; the development of task-based typologies could be a substantial step forward for the research field of uncertainty visualization.

Table 8.1. Task categories as starting point for task-oriented typologies for uncertainty visualization.

Task type	**Description**	**Example**	**Strategy**
Communication tasks	Map reading tasks	'Determine the level of uncertainty in area A'	Traditional rules from cartography apply
Analytical tasks	Predefined analytical questions are answered	'What is the best location for drilling a new well?'	Uncertainty visualization must be tailored to the task
Exploratory tasks	Questions occur during data exploration	'Is the high uncertainty in the vicinity of the airport due to misclassification?'	Uncertainty visualization must be versatile (adaptable and / or adaptive)

8.2.2 Uncertainty-aware change analysis

This research showed how a concept based on a GVA approach can facilitate the use of uncertainty information in exploratory change analysis. As proof-of-concept, the *ICchange* prototype was developed for uncertainty-aware land cover change analysis, using a task-based design approach. Findings from semi-structured interviews with three expert groups based on the prototype showed a high potential of the concept for practical use. Further findings highlighted applications and revealed barriers regarding the use of uncertainty information for real-world change analysis. As another contribution of this dissertation, the Java code of the prototype implementation, is available[14] under free and open source license (Apache License 2.0).

[14] https://github.com/ckinkeldey/icchange

A number of applications incorporating uncertainty in change analysis have been discussed in Chapter 2: optimization of change detection parameters, removal of erroneous change through filtering by uncertainty, and support of reasoning with change data. These applications, although related to change analysis in this work, can be transferred to other fields that involve spatial analysis. Imagine a map with projected future population density, including uncertainty from probabilities related to the model output. The choice of suitable prediction model parameters could be supported by uncertainty information, analogous to the optimization of change detection parameters discussed in this work. Through filtering by uncertainty, it may be possible to detect errors or inconsistencies in the projected population density, with high uncertainty indicating less reliable results. With respect to the support of reasoning, incorporating uncertainty related to the predicted population density could lead to better informed interpretations of the data, similar to supporting interpretation of change. Thus, the concept developed here is not limited to change analysis from RS data and can be transferred to other applications of spatial analysis.

However, related to applicability of the concept, the general question remains as to how the use of uncertainty can be established in practice. In the interviews reported in Chapter 7, experts identified the lack of tool support as one of the main reasons for the rare use of uncertainty information. In this work, several contributions were made to support tool development. First, a task-based approach for prototype development was described. Second, recommendations were given on how to select suitable uncertainty visualization techniques for a tool under development. This provided a first orientation, but the need for new kinds of typologies was discussed as well, and additional effort is required to pursue this aspect further. Third, the *ICchange* prototype was made available, including the Java code[15]. At the time of writing, the prototype is intended for users with programming skills, but in the future it would be interesting to implement an end user tool based on the prototype as extension of a standard GIS, e.g., of ArcGIS[16], ERDAS Imagine[17], or ENVI[18]. An implementation in a free and open source GIS such as QGIS[19] could be a viable alternative to increase the availability of the software.

The development of well-crafted tools is only one side of the medal. This work and related research have shown the importance of appropriate expertise to successfully utilize information on geodata uncertainty. Beside domain expertise and experience with maps or GIS, expertise with uncertainty measures has been shown to be an important factor. The experts in our interviews

[15] https://github.com/ckinkeldey/icchange

[16] http://www.arcgis.com

[17] http://www.hexagongeospatial.com/products/remote-sensing/erdas-imagine/overview

[18] http://www.exelisvis.com/ProductsServices/ENVIProducts/ENVI.aspx

[19] http://www.qgis.org

were able to understand the uncertainty measure used here, because it was based on classification confidence, a concept they were already aware of. However, for users without this expertise, it could have been much more difficult to understand how to interpret the measure. Thus, in order to support analysts or decision makers, we do not only need to provide them with information about uncertainty, but "to maximize the effectiveness of visualization, uncertainty must be represented and ways to deal with it must be provided" (Cliburn et al. 2002, p.948). This topic has barely been addressed so far and will be important to consider for future work.

All in all, this dissertation research provided findings that represent a substantial step regarding utilization of uncertainty information for exploratory geodata analysis. A number of results have implications beyond land cover change analysis, the application this research was focused on. The insights generated in this research help bridge theory and practice and serve as a foundation for future research in the fields of uncertainty visualization and exploratory analysis of geodata.

Bibliography

Aerts, J. C., K. C. Clarke, and A. D. Keuper. 2003. "Testing Popular Visualization Techniques for Representing Model Uncertainty." *Cartography and Geographic Information Science* 30 (3): 249–261.

Ahlqvist, O. 2008. "Extending post-classification change detection using semantic similarity metrics to overcome class heterogeneity: A study of 1992 and 2001 U.S. National Land Cover Database changes." *Remote Sensing of Environment* 112 (3): 1226–1241.

Alberti, K. 2013. "Web-based visualization of uncertain spatiotemporal data." MSc thesis. Utrecht University. Utrecht, The Netherlands.

Allendes Osorio, R. S., and K.W. Brodlie. 2008. "Contouring with uncertainty." In *Proceedings of 6th Theory & Practice of Computer Graphics Conference (TP.CG.08)*, edited by I. S. Lim, and W. Tang, 59–66, Eurographics Association, Manchester.

Andrienko, G. and N. Andrienko. 2005. "Visual exploration of the spatial distribution of temporal behaviors." In *Proceedings of the ninth international conference on information visualisation (IV '05)*, July 6–8. Washington, DC: IEEE Computer Society, 799–806.

Andrienko, N. and G. Andrienko. 2006. *Exploratory analysis of spatial and temporal data: a systematic approach*. Secausus, NJ: Springer-Verlag New York.

Andre, A., and H. Cutler. 1998. "Displaying Uncertainty in Advanced Navigation Systems." In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, Santa Monica, CA, 31–35.

Arias-Hernandez, R., L. T. Kaastra, T. M. Green, and B. Fisher. 2011. "Pair Analytics: Capturing Reasoning Processes in Collaborative Visual Analytics." In *44th Hawaii International Conference on System Sciences (HICSS)*, January 4–7, 1–10.

Ash, K. D., R. L. Schumann, and G. C. Bowser. 2014. "Tornado Warning Trade-Offs: Evaluating Choices for Visually Communicating Risk." *Weather, Climate, and Society* 6 (1): 104–118.

Atkinson, P. M., and G. M. Foody. 2002. "Uncertainty in Remote Sensing and GIS: Fundamentals." In *Uncertainty in Remote Sensing and GIS*, edited by G. M. Foody and P. M. Atkinson. Hoboken, NJ: John Wiley.

Bagassi, M., and L. Macchi. 2006. "Pragmatic Approach to Decision Making under Uncertainty: The Case of the Disjunction Effect." *Thinking & Reasoning* 12 (3): 329–350.

Bastin, L.; Fisher, P.; Wood, J. "Visualizing uncertainty in multi-spectral remotely sensed imagery." *Computers and Geosciences* 28 (3): 337–350.

Bertin, J. 1983. *Semiology of Graphics*. Madison: University of Wisconsin Press.

Bisantz, A. M., D. Cao, M. Jenkins, P. R. Pennathur, M. Farry, E. Roth, S. S. Potter, and J. Pfautz. 2011. "Comparing Uncertainty Visualizations for a Dynamic Decision-Making Task." *Journal of Cognitive Engineering and Decision Making* 5 (3): 277–293.

Bisantz, A. M., R. Finger, Y. Seong, and J. Llinas. 1999. "Human Performance and Data Fusion Based Decision Aids." In *2nd International Conference on Information Fusion FUSION'99*, Sunnyvale, CA, July 6–8, 918–925.

Bisantz, A. M., S. S. Marsiglio, and J. Munch. 2005. "Displaying Uncertainty: Investigating the Effects of Display Format and Specificity." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 47 (4): 777–796.

Bisantz, A. M., R. T. Stone, J. Pfautz, A. Fouse, M. Farry, E. Roth, A. L. Nagy, and G. Thomas. 2009. "Visual representations of meta-information." *Journal of Cognitive Engineering and Decision Making* 3 (1): 67–91.

Blenkinsop, S., P. Fisher, L. Bastin, and J. Wood. 2000. "Evaluating the perception of uncertainty in alternative visualization strategies." *Cartographica: The International Journal for Geographic Information and Geovisualization* 37 (1): 1–14.

Blok, C.A., U. D. Turdukulov, R. Zurita-Milla, V. Retsios, M. Schouwenburg, and M. Metaferia. 2011. Development of an open-source toolbox for the analysis and visualization of remotely sensed time series. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 46 (4): 227–238.

Boukhelifa, N., A. Bezerianos, T. Isenberg, and J.-D. Fekete. 2012. "Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty." *IEEE Transactions on Visualization and Computer Graphics* 18 (12): 2769–2778.

Brewer, C. A., G. W. Hatchard, and M. A. Harrower. 2003. "ColorBrewer in Print: A Catalog of Color Schemes for Maps." *Cartography and Geographic Information Science* 30 (1): 5–32.

Broad, K., Leiserowitz, A., Weinkle, J., and M. Steketee. 2007. "Misinterpretations of the 'Cone of Uncertainty' in Florida during the 2004 Hurricane Season." *Bulletin of the American Meteorological Society* 88 (5): 651–667.

Brodlie, K., R. A. Osorio, and A. Lopes. 2012. "A Review of Uncertainty in Data Visualization." In *Expanding the Frontiers of Visual Analytics and Visualization*, edited by J. Dill, R. Earnshaw, D. Kasik, J. Vince, and P. C. Wong, 81–109. London: Springer.

Brolese, A., and S. Huf. 2006. "Visualizing Uncertainty to Improve Operators' Spatial Proximity Judgments in Uncertain Surroundings." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50 (3): 294–298.

Brown, K.M., G.M. Foody, and P.M. Atkinson. 2006. "Deriving thematic uncertainty measures in remote sensing using classification outputs." In *Proceedings of the 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Lisbon, Portugal, July 5–7, edited by M. Caetano, and M. Painho, Instituto Geográfico Português: Lisbon, Portugal.

Burnicki, A.C. 2011. "Modeling the probability of misclassification in a map of land cover change." *Photogrammetric Engineering and Remote Sensing* 77 (1): 39–49.

Busemeyer, J. R., and J. T. Townsend. 1993. "Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment." *Psychological Review* 100 (3): 432–459.

Buttenfield, B. P., and R. Weibel. 1988. "Visualizing the Quality of Cartographic Data." In *Third International Geographic Information Systems Symposium (GIS/LIS '88)*, San Antonio, TX, November 30–December 2.

Caers, J. 2011. *Modeling Uncertainty in the Earth Sciences*. Chichester, UK: Wiley-Blackwell.

Cedilnik, A., and P. Rheingans. 2000. "Procedural Annotation of Uncertain Information." In *IEEE Visualization 2000*, Salt Lake City, UT, October 8–13, 77–84.

Chapman, G. B., and E. J. Johnson. 2002. "Incorporating the Irrelevant: Anchors in Judgments of Belief and Value." In *Heuristics and Biases: The Psychology of Intuitive Judgment*, edited by T. Gilovich, D. Griffin & D. Kahneman, Cambridge: University Press, 120–138.

Cliburn, D. C., J. J. Feddema, J. R. Miller, and T. A. Slocum. 2002. "Design and Evaluation of a Decision Support System in a Water Balance Application." *Computers & Graphics* 26 (6): 931–949.

Congalton, R.G., and K. Green. 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed., CRC Press/Taylor & Francis: Boca Raton, FL, USA.

Coppin, P.; Jonckheere, I.; Nackaerts, K.; Muys, B., 2004. "Digital change detection methods in ecosystem monitoring: A review." *International Journal of Remote Sensing* 25 (9): 1565–1596.

Cox, J., D. House, and M. Lindell. 2013. "Visualizing Uncertainty in Predicted Hurricane Tracks." *International Journal for Uncertainty Quantification* 3 (2): 143–156.

Crump, M. J. C., J. V. McDonnell, T. M. Gureckis, and S. Gilbert. 2013. "Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research." *PLoS ONE* 8 (3): e57410.

De Chiara, D., 2012. "From GeoVisualization to Visual-Analytics: Methodologies and Techniques for Human-Information Discourse." Ph.D. Thesis, University of Salerno, Salerno, Italy.

Deitrick, S. 2012. "Evaluating Implicit Visualization of Geographic Uncertainty for Public Policy Decision Support." In *Proceedings of AutoCarto 2012*, September 16–18, Cartography and Geographic Information Society.

Deitrick, S., and R. Edsall. 2006. "The Influence of Uncertainty Visualization on Decision Making: An Empirical Evaluation." In *Progress in Spatial Data Handling*, edited by A. Riedl, W. Kainz, and G. A. Elmes, Berlin: Springer, 719–738.

Deterding, S., D. Dixon, R. Khaled, and L. Nacke. 2011. "From Game Design Elements to Gamefulness." In *Proceedings of the 15th International Academic MindTrek Conference*, Tampere, Finland, September 28–30, 9–15.

Dong, X., and C. C. Hayes. 2012. "Uncertainty Visualizations: Helping Decision Makers Become More Aware of Uncertainty and Its Implications." *Journal of Cognitive Engineering and Decision Making* 6 (1): 30–56.

Dooley, M. A. and S. J. Lavin. 2007. "Visualizing method-produced uncertainty in isometric mapping." *Cartographic Perspectives* 56: 17–36.

Drecki, I. 2002. "Visualization of uncertainty in geographical data." In *Spatial Data Quality*, edited by W. Shi, P. F. Fisher, and M.F. Goodchild. 140–159, Taylor & Francis, London/New York.

Edwards, L. D., and E. S. Nelson. 2001. "Visualizing data certainty: a case study using graduated circle maps." *Cartographic Perspectives* 38: 19–36.

Ehlschlaeger, C. R., A. M. Shortridge, and M. F. Goodchild. 1997. "Visualizing spatial data uncertainty using animation." *Computers & Geosciences* 23 (4), 387–395.

Evans, B. J. 1997. "Dynamic display of spatial data reliability: does it benefit the map user?" *Computers & Geosciences* 23 (4): 409–442.

Ferreira, N., D. Fisher, and A. C. König. 2014. "Sample-Oriented Task-Driven Visualizations: Allowing Users to Make Better, More Confident Decisions." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Paris, France, April 27–May 2, New York, NY, ACM, 571–580.

Finger, R., and A. M. Bisantz. 2002. "Utilizing Graphical Formats to Convey Uncertainty in a Decision-Making Task." *Theoretical Issues in Ergonomics Science* 3 (1): 1–25.

Fisher, D., I. Popov, S. Drucker, and M. C. Schraefel. 2012. "Trust Me, I'm Partially Right: Incremental Visualization Lets Analysts Explore Large Datasets Faster." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin, TX, May 5–10, New York: ACM, 1673–1682.

Fisher, P. 1993. "Visualising uncertainty in soil maps by animation." *Cartographica* 30: 20–27.

Fisher, P., C. Arnot, R. Wadsworth, and J. Wellens. 2006. "Detecting change in vague interpretations of landscapes." *Ecological Informatics* 1 (2): 163–178.

Fisher, P., Comber, A., and R. Wadsworth. 2006. "Approaches to Uncertainty in Spatial Data." In *Fundamentals of Spatial Data Quality*, edited by R. Devillers and R. Jeansoulin. London, UK: ISTE, 43–59.

Flannery, J. J. 1956. "The graduated circle: a description, analysis and evaluation of a quantitative map symbol." Unpublished PhD dissertation. University of Wisconsin.

Flannery, J. J. 1971. "The relative effectiveness of some common graduated point symbols in the presentation of quantitative data." *Cartographica* 8: 96–109.

Foody, G.M., and P.M. Atkinson. 2002. Eds. *Uncertainty in Remote Sensing and GIS*; Wiley: Hoboken, NJ, USA.

Gahegan, M. and M. Ehlers. 2000. "A framework for the modelling of uncertainty between remote sensing and geographic information systems." *ISPRS Journal of Photogrammetry and Remote Sensing* 55 (3): 176–188.

Gerharz, L. E., and E. J. Pebesma. 2009. "Usability of Interactive and Non-Interactive Visualization of Uncertain Geospatial Information." In *Proceedings of Geoinformatik 2009*, edited by W. Reinhardt, A. Krüger, and M. Ehlers, Osnabrück, Germany, March 31–April 2, 223–230.

Gerharz, L., E. Pebesma, and H. Hecking. 2010. "Visualizing uncertainty in spatio-temporal data." In *Proceedings of the ninth international symposium on spatial accuracy assessment in natural resources and environmental sciences*, edited by N.J. Tate and P.F. Fisher, July 20–23, Leicester.

Germine, L., K. Nakayama, B. C. Duchaine, C. F. Chabris, G. Chatterjee, and J. B. Wilmer. 2012. "Is the Web as Good as the Lab? Comparable Performance from Web and Lab in Cognitive/Perceptual Experiments." *Psychonomic Bulletin & Review* 19 (5): 847–857.

Gershon, N. 1998. "Visualization of an imperfect world." *IEEE Computer Graphics and Applications* 18, 43–45.

Green, K. 2011. "Change matters." *Photogrammetric Engineering and Remote Sensing* 77: 305–309.

Griethe, H., and H. Schumann. 2005. "Visualizing Uncertainty for Improved Decision Making." In *Proceedings of 4th International Conference on Business Informatics Research (BIR 2005)*, Skövde, Sweden.

Griethe, H., and H. Schumann. 2006. "The Visualization of Uncertain Data: Methods and Problems." In *Proceedings of the 17th Simulation and Visualization Conference (SimVis 2006)*, Magdeburg, Germany, March 2–3, 143–156.

Grigoryan, G., and P. Rheingans. 2004. "Point-based probabilistic surfaces to show surface uncertainty." *IEEE Transactions on Visualization and Computer Graphics* 10 (5): 564–573.

Goodchild, M. F. and T. J. Case. 2001. "Introduction." In *Spatial uncertainty in ecology: Implications for remote sensing and GIS applications*, edited by C.T. Hunsaker. New York, NY: Springer, 3–14.

Hengl, T. 2003. "Visualization of Uncertainty Using the HSI Colour Model: Computations with Colours." In *7th International Conference on GeoComputation*, Southampton, September 8–10, 8–17.

Heuvelink, G. B. M., and J. D. Brown. 2008. "Uncertain Environmental Variables in GIS." In *Encyclopedia of GIS*, edited by S. Shekhar and H. Xiong, New York: Springer, 1184–1189.

Hoeber, O., G. Wilson, S. Harding, R. Enguehard, and R. Devillers. 2010. "Visually representing geo-temporal differences." In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology 2010*, edited by MacEachren, A., Miksch, S., Salt Lake City, USA, October 25–26, IEEE Press: New York, NY, USA, 229–230.

Hope, S., and G. J. Hunter. 2007a. "Testing the Effects of Positional Uncertainty on Spatial Decision-Making." *International Journal of Geographical Information Science* 21 (6): 645–665.

Hope, S., and G. J. Hunter. 2007b. "Testing the Effects of Thematic Uncertainty on Spatial Decision-Making." *Cartography and Geographic Information Science* 34 (3): 199–214.

Howard, D., and A. M. MacEachren. 1996. "Interface design for geographic visualization: tools for representing reliability." *Cartography and Geographic Information Science* 23, 59–77.

Isenberg, T., P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. 2013. "A Systematic Review on the Practice of Evaluating Visualization." *IEEE Transactions on Visualization and Computer Graphics* 19 (12): 2818–2827.

ISO. 2006. International Organization for Standardization 2006. ISO 19138: geographic information: data quality measures. Geneva, Switzerland: ISO/TC211

Johnson, C. R., and A. R. Sanderson. 2003. "A Next Step: Visualizing Errors and Uncertainty." *IEEE Computer Graphics and Applications* 23 (5): 6–10.

Kardos, J. D., G. L. Benwell, and A. Moore. 2007. "Assessing different approaches to visualise spatial and attribute uncertainty in socioeconomic data using the hexagonal or rhombus (HoR) trustree." *Computers, Environment and Urban Systems* 31 (1): 91–106.

Kardos, J. D., A. Moore, and G. L. Benwell. 2008. "Exploring tessellation metaphors in the display of geographical uncertainty." In *Geospatial Vision – New Dimensions in Cartography*, edited by A. Moore, and I. Drecki, Springer, Berlin/Heidelberg, 113–140.

Kardos, J. D., A. Moore, and G. L. Benwell. 2003. "The visualization of uncertainty in spatially-referenced attribute data using TRUSTworthy data structures." In *7th International Conference on GeoComputation*, Southampton, September 8–10, 11–26.

Keim, D.A., J. Kohlhammer, G. Ellis, and F. Mannsmann. Eds. 2010. *Mastering the information age – solving problems with visual analytics*. Goslar: Eurographics Association.

Khorram, S. 1999. *Accuracy Assessment of Remote Sensing-Derived Change Detection.* American Society for Photogrammetry and Remote Sensing: Bethesda, MD, USA, 1999.

Kinkeldey, C., 2014a. "A Concept for Uncertainty-Aware Analysis of Land Cover Change Using Geovisual Analytics." *ISPRS International Journal of Geo-Information* 3 (3), 1122–1138.

Kinkeldey, C., 2014b. "Development of a prototype for uncertainty-aware geovisual analytics of land cover change." *International Journal of Geographical Information Science* 28 (10), 2076–2089.

Kinkeldey, C., A. M. MacEachren, and J. Schiewe. 2014a. "How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies." *The Cartographic Journal* 51 (4): 372–386.

Kinkeldey, C., J. Mason, A. Klippel, and J. Schiewe. 2013. "Assessing the Impact of Design Decisions on the Usability of Uncertainty Visualization: Noise Annotation Lines for the Visual Representation of Attribute Uncertainty." *Proceedings of 26th International Cartographic Conference (ICC 2013)*, Dresden, Germany, August 25–30.

Kinkeldey, C., J. Mason, A. Klippel, and J. Schiewe. 2014b. "Evaluation of Noise Annotation Lines: Using Noise to Represent Thematic Uncertainty in Maps." *Cartography and Geographic Information Science* 41 (5): 430–439.

Kinkeldey, C., and J. Schiewe. 2012. "Visualisierung Thematischer Unsicherheiten mit Noise Annotation Lines." Kartographische Nachrichten 62 (5): 241–249.

Kinkeldey, C., and J. Schiewe. 2014. "Expert interviews about the use of visually depicted uncertainty for analysis of remotely sensed land cover change." *Visually-Supported Reasoning with Uncertainty Workshop*, GIScience 2014, Vienna, Austria, September 23.

Kirschenbaum, S. S., and J. E. Arruda. 1994. "Effects of Graphic and Verbal Probability Information on Command Decision Making." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 36 (3): 406–418.

Kirschenbaum, S. S., J. G. Trafton, C. D. Schunn, and S. B. Trickett. 2014. "Visualizing Uncertainty: The Impact on Performance." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 56 (3): 509–520.

Kit, O., and M. Lüdeke. 2013. "Automated detection of slum area change in Hyderabad, India using multitemporal satellite imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 83, 130–137.

Klein, G. 1999. *Sources of Power: How People Make Decisions*. Cambridge, MA: The MIT Press.

Kobus, D. A., S. Proctor, and S. Holste. 2001. "Effects of Experience and Uncertainty during Dynamic Decision Making." *International Journal of Industrial Ergonomics* 28 (5): 275–290.

Kolbeinsson, A. 2013. "Visualising uncertainty in aircraft cockpits. Is icon degradation an appropriate visualization form." MSc thesis. University of Skövde, Skövde, Sweden.

Kubíček, P. and Č. Šašinka. 2011. "Thematic uncertainty visualization usability – comparison of basic methods." *Annals of GIS* 17, 253–263.

Kunz, M., A. Grêt-Regamey, and L. Hurni. 2011. "Visualization of uncertainty in natural hazards assessments using an interactive cartographic information system." *Natural Hazards* 59, 1735– 1751.

Larkin, J. H., and H. A. Simon. 1987. "Why a diagram is (sometimes) worth ten thousand words." *Cognitive Science* 11 (1): 65–100.

Leitner, M., and B. P. Buttenfield. 2000. "Guidelines for the Display of Attribute Certainty." *Cartography and Geographic Information Science* 27 (1): 3–14.

Leyk S., Boesch, R. and R. Weibel. 2005. "A conceptual framework for uncertainty investigation in map-based land cover change modelling." *Transactions in GIS* 9 (3): 291–322.

Lloyd, R., and R. Bunch. 2005. "Individual Differences in Map Reading Spatial Abilities Using Perceptual and Memory Processes." *Cartography and Geographic Information Science* 32 (1): 33–46.

Lowry, J.H.; R.D. Ramsey, L. Langs Stoner, J. Kirby, and K. Schulz. 2008. “An ecological framework for evaluating map errors using fuzzy sets.” *Photogrammetric Engineering and Remote Sensing* 74 (12): 1509–1519.

Lucieer, A. 2004. “Uncertainties in Segmentation and Their Visualization.” Ph.D. Thesis, Utrecht University, Enschede, The Netherlands.

MacEachren, A. M. 1992. “Visualizing uncertain information.” *Cartographic Perspectives* 13: 10–19.

MacEachren, A. M., C. A. Brewer, and L. W. Pickle. 1998. “Visualizing georeferenced data: representing reliability of health statistics.” *Environmental and Planning A* 30 (9): 1547–1561.

MacEachren, A. M., A. Robinson, S. Hopper, S. Gardner, R. Murray, M. Gahegan, and E. Hetzler. 2005. “Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know.” *Cartography and Geographic Information Science* 32 (3): 139–160.

MacEachren, A. M., R. E. Roth, J. O’Brien, B. Li, D. Swingley, and M. Gahegan. 2012. “Visual semiotics & uncertainty visualization: an empirical study.” *IEEE Transactions on Visualization and Computer Graphics* 18 (12): 2496–2505.

MacGranaghan, M. 1993. “A Cartographic View of Spatial Data Quality.” *Cartographica: The International Journal for Geographic Information and Geovisualization* 30 (2): 8– 19.

Mason, W., and S. Suri. 2012. “Conducting Behavioral Research on Amazon’s Mechanical Turk.” *Behavior Research Methods* 44 (1): 1–23.

Micallef, L., P. Dragicevic, and J. Fekete. 2012. “Assessing the Effect of Visualizations on Bayesian Reasoning through Crowdsourcing.” *IEEE Transactions on Visualization and Computer Graphics* 18 (12): 2536–2545.

Nadav-Greenberg, L., S. L. Joslyn, and M. U. Taing. 2008. “The Effect of Uncertainty Visualizations on Decision Making in Weather Forecasting.” *Journal of Cognitive Engineering and Decision Making* 2 (1): 24–47.

Newman, T. S., and W. Lee. 2004. “On visualizing uncertainty in volumetric data: techniques and their evaluation.” *Journal of Visual Languages & Computing* 15 (6): 463–491.

Pang, A. 2001. “Visualizing Uncertainty in Geo-Spatial Data.” In *Proceedings of the Workshop on the Intersections between Geospatial Information and Information Technology*, Arlington, TX, October 1–2.

Pang, A. 2008. “Visualizing Uncertainty in Natural Hazards. Risk Assessment, Modeling and Decision Support.” In *Risk Assessment, Modeling and Decision Support*, edited by A. Bostrom, S. French, and S. Gottlieb, Berlin: Springer, 261–294.

Pang, A. T., C. M. Wittenbrink, and S. K. Lodha. 1997. “Approaches to uncertainty visualization.” *The Visual Computer* 13: 370–390.

Pebesma, E.J., K. de Jong, and D. Briggs. 2007. “Interactive visualization of uncertain spatial and spatio-temporal data under different scenarios: an air quality example.” *International Journal of Geographical Information Science*, 21 (5), 515–527.

Perer, A. and B. Shneiderman. 2009. “Integrating statistics and visualization for exploratory power: from long-term case studies to design guidelines.” *IEEE Computer Graphics and Applications*, 29 (3), 39–51.

Peuquet, D. J. 1994. “It’s about time: a conceptual framework for the representation of temporal dynamics in geographic information systems.” *Annals of the Association of American Geographers* 84 (3): 441–461.

Pontius, R.G., and C.D. Lippitt. 2008. “Can error explain map differences over time?” *Cartography and Geographic Information Science* 33 (2), 159–171.

Potter, K., P. Rosen, and C. R. Johnson. 2012. “From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches.” In *Uncertainty Quantification in Scientific Computing*, edited by A. M. Dienstfrey and R. F. Boisvert, Berlin: Springer, 226–249.

Pyysalo, U., and J. Oksanen. 2014. “Visualization of Uncertain Catchment Boundaries and Its Influence on Decision Making.” *Poster, 17th AGILE International Conferences on Geographic Information Science*, Castellón, Spain, 3–6 June.

Rand, D. G. 2012. “The Promise of Mechanical Turk: How Online Labor Markets Can Help Theorists Run Behavioral Experiments: Evolution of Cooperation.” *Journal of Theoretical Biology* 299: 172–179.

Regan, H. M., Y. B. Haim, B. Langford, W. G. Wilson, P. Lundberg, S. J. Andelman, and M. A. Burgman. 2005. “Robust Decision-Making under Severe Uncertainty for Conservation Management.” *Ecological Applications* 15 (4): 1471–1477.

Retchless, D. 2012. “Mapping Climate Change Uncertainty: Effects on Risk Perceptions and Decision Making.” *Poster. AGU Fall Meeting*, San Francisco, USA.

Rhodes, P.J., R. S. Laramee, R. D. Bergeron, and T. M. Sparr. 2003. “Uncertainty Visualization Methods in Isosurface Volume Rendering.” In *Eurographics 2003, short papers*, Granada, Spain, September 1–5, 83–88.

Riveiro, M. 2007a. “Cognitive evaluation of uncertainty visualization methods for decision making.” In *4th Symposium on Applied Perception in Graphics and Visualization (AGPV 2007)*, Tübingen, July 26–27, 133.

Riveiro, M. 2007b. “Evaluation of uncertainty visualization techniques for information fusion.” In *10th International Conference on Information Fusion*, Quebec, Canada, July 9–12, 1–8.

Riveiro, M., T. Helldin, and G. Falkman. 2014a. “Influence of Meta-Information on Decision-Making: Lessons Learned from Four Case Studies.” In *Proceedings of 2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 3–6 March 2014, San Antontio, TX, 14–20.

Riveiro, M., T. Helldin, G. Falkman, and M. Lebram. 2014b. "Effects of Visualizing Uncertainty on Decision-Making in a Target Identification Scenario." *Computers & Graphics* 41: 84–98.

Robinson, A.C., J. Chen, E. J. Lengerich, H. G. Meyer, and A. M. MacEachren. 2005. "Combining usability techniques to design geovisualization tools for epidemiology." *Cartography and Geographic Information Science*, 32 (4), 243–255.

Roth, R.E. 2009a. "A qualitative approach for understanding the role of geographic information uncertainty during decision making." *Cartography and Geographic Information Science* 36 (4), 315–330.

Roth, R. E. 2009b. "The Impact of User Expertise on Geographic Risk Assessment under Uncertain Conditions." *Cartography and Geographic Information Science* 36 (1): 29–43.

Sanyal, J., S. Zhang, G. Bhattacharya, P. Amburn, and R. Moorhead. 2009. "A User Study to Compare Four Uncertainty Visualization Methods for 1D and 2D Datasets." *IEEE Transactions on Visualization and Computer Graphics* 15 (6): 1209–1218.

Sanyal, J., Song Z., Dyer, J., Mercer, A., Amburn, P. and Moorhead, R. J. (2010). 'Noodles: a tool for visualization of numerical weather model ensemble uncertainty', IEEE Transactions on Visualization and Computer Graphics, 16, pp. 1421–1430.

Schiewe, J. 2013. "Geovisualisation and geovisual analytics: The interdisciplinary perspective on cartography." In *Proceedings of 26th International Cartographic Conference ICC 2013*, Dresden, Germany, August 25–30, 122–126.

Schiewe, J., Ehlers, M., Kinkeldey, C., and D. Tomowski. 2009. "From fuzzy and object based classification to fuzzy and object based uncertainty evaluation." In *Proceedings of SPIE 7478, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IX*, edited by U. Michel and D. L. Civco, 74781L0-74781L0.

Scholz, R., and Y. Lu. 2014. "Uncertainty in Geographic Data on Bivariate Maps: An Examination of Visualization Preference and Decision Making." *ISPRS International Journal of Geo-Information* 3 (4): 1180–1197. doi:10.3390/ijgi3041180.

Schreck, T., T. Tekušová, J. Kohlhammer, and D. Fellner. 2007. "Trajectory-based visual analysis of large financial time series data." *ACM SIGKDD Explorations Newsletter* 9, 30–37.

Schuster, C.; M. Förster, and B. Kleinschmit. „Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data." *International Journal of Remote Sensing* 33, 5583–5599.

Schweizer, D. M., and M. F. Goodchild. 1992. "Data Quality and Choropleth Maps: An Experiment with the Use of Color." In *GIS/ LIS '92*, San Jose, CA, November 10–12, 686–699.

Senaratne, H., L. Gerharz, E. Pebesma, and A. Schwering. 2012. "Usability of Spatio-Temporal Uncertainty Visualisation Methods." In *Bridging the Geographic Information Sciences:*

15th AGILE International Conference, edited by J. Gensel, D. Josselin, and D. Vandenbroucke, Avignon, France. 24–27 April 2012, Berlin: Springer, 3–23.

Severtson, D. J., and J. D. Myers. 2013. "The Influence of Uncertain Map Features on Risk Beliefs and Perceived Ambiguity for Maps of Modeled Cancer Risk from Air Pollution." *Risk Analysis* 33 (5): 818–837.

Shattuck, L. G., N. L. Miller, and K. E. Kemmerer. 2009. "Tactical Decision Making under Conditions of Uncertainty: An Empirical Study." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 53 (4): 242–246.

Shekhar, S. and H. Xiong. 2008. *Encyclopedia of GIS*. New York: Springer.

Shi, W. 2010. *Principles of Modeling Uncertainties in Spatial Data and Spatial Analyses*. Boca Raton: CRC Press/Taylor & Francis.

Shneiderman, B., and C. Plaisant. 2006. "Strategies for Evaluating Information Visualization Tools: Multi-Dimensional In-Depth Long-Term Case Studies." In *BELIV 2006 Proceedings of the 2006 AVI Workshop on beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, Venice, Italy: ACM, 1–7.

Simpson, A. L., B. Ma, E. C. S. Chen, R. E. Ellis, and A. J. Stewart. 2006. "Using Registration Uncertainty Visualization in a User Study of a Simple Surgical Task." In *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI 2006)*, 1–6 October 2006, Copenhagen, Denmark, 397–404.

Slocum, T. A., D. C. Cliburn, J. J. Feddema, and J. R. Miller. 2003. "Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance." *Cartography and Geographic Information Science* 30 (4): 299–317.

Smith, J., D. Retchless, C. Kinkeldey, and A. Klippel. 2013. "Beyond the surface: current issues and future directions in uncertainty visualization research." In *Proceedings of the 26th International Cartographic Conference (ICC 2013)*, Dresden, August 25–30.

Snyder, L. V. 2006. "Facility Location under Uncertainty: A Review." *IIE Transactions* 38 (7): 547–564.

St. John, M., J. Callan, and S. Proctor. 2000. "Tactical Decision-Making under Uncertainty: Experiments I and II." San Diego, CA: SPAWAR Space and Naval Warfare Systems Center, Technical Report 1821.

Štěrba, Z., C. Šašinka, Z. Stachoň, P. Kubíček, and S. Tamm. 2014. "Mixed Research Design in Cartography: A Combination of Qualitative and Quantitative Approaches." *Kartographische Nachrichten* 64(5), 262–269

Tabachnick, B. G., and L. S. Fidell. 2013. *Using Multivariate Statistics*. 6th ed. Boston, MA: Pearson Education.

Taubenböck, H., T. Esch, A. Felbier, M. Wiesner, A. Roth, and S. Dech. 2012. "Monitoring urbanization in mega cities from space." Remote Sensing of Environment 117, 162–176.

Thomas, J. J. and K. A. Cook. Eds. 2005. *Illuminating the path. The research and development agenda for visual analytics*. 1st ed. Los Alamitos, CA: IEEE.

Thomson, J., E. Hetzler, A. MacEachren, M. Gahegan, and M. Pavel. 2005. “A typology for visualizing uncertainty.” In *Proceedings of SPIE 5669, Visualization and Data Analysis 2005,* edited by R. F. Erbacher, J. C. Roberts, M. T. Gröhn, and K. Börner, San Jose, CA: IS&T/SPIE, 146–157.

Tomaszewski, B.M., A.C. Robinson, C. Weaver, M. Stryker, and A.M. MacEachren. 2007. “Geovisual analytics and crisis management.” In *Proceedings of the 4th International Information Systems for Crisis Response and Management (ISCRAM) Conference*, Delft, The Netherlands, May 13–16.

Tomaszewski, B.M., J. Szarzynski, and D. Schwartz. 2014. “Serious Games for Disaster Risk Reduction Spatial Thinking.” In *Extended Abstract Proceedings of the Eighth International Conference on Geographic Information Science (GIScience 2014)*, edited by K. Stewart, E. Pebesma, G. Navratil, P. Fogliaroni, and M. Duckham, September 23–26, Vienna, Austria.

van der Wel, F. J. M., L. C. van der Gaag, and G. H. Gorte Ben. 1998. Visual exploration of uncertainty in remote-sensing classification. *Computers and Geosciences*, 24 (4), 335–343.

van der Wel, F. J. M., 2000. *Assessment and visualisation of uncertainty in remote sensing land cover classifications*. Ph.D. Thesis. Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht, The Netherlands.

van Wijk, J.J., et al., 2010. Evaluation. In: D.A. Keim, et al., eds. *Mastering the information age. Solving problems with visual analytics*. Goslar: Eurographics Association, 131–144.

von Landesberger, T., et al., 2012. Visual analytics methods for categoric spatio-temporal data. In: *Proceedings of IEEE conference on Visual Analytics Science and Technology (VAST)*, edited by G. Santucci and M. Ward, 14–19 October, Seattle, WA, 183–192.

Viard, T., G. Caumon, and B. Levy. 2011. “Adjacent versus Coincident Representations of Geospatial Uncertainty: Which Promote Better Decisions?” *Computers & Geosciences* 37 (4): 511–520.

Vullings, L. A. E., C. A. Blok, C. G. A. M. Wessels, and J. D. Bulens. 2013. “Dealing with the uncertainty of having incomplete sources of geo-information in spatial planning.” *Applied Spatial Analysis and Policy* 6 (1), 25–45.

Wang, T.D., K. Wongsuphasawat, C. Plaisant, and B. Shneiderman. 2011. “Extracting insights from electronic health records: Case studies, a visual analytics process model, and design recommendations.” *Journal of Medical Systems* 35 (5), 1135–1152.

Winer, B. J. 1962. “Latin Squares and Related Designs.” In *Statistical Principles in Experimental Design*, edited by B. J. Winer, New York: McGraw-Hill Book, 514–577.

Wittenbrink, C. M., A. T. Pang, and S. K. Lodha. 1996. “Glyphs for visualizing uncertainty in vector fields.” *IEEE Transactions on Visualization and Computer Graphics* 2 (3): 266–279.

Wray, K. (2007). ‘Using the creative design process to develop illustrative rendering techniques to represent information quality’, Journal of Young Investigators, 17.

Zhang, J., and M. F. Goodchild. 2002. *Uncertainty in Geographical Information*. London: Taylor & Francis.

Zhang, Q., C. Blok, and X. Tang. 2008. "Animated representation of uncertainty and fuzziness in Dutch spatial planning maps." In *XXIst ISPRS Congress Technical Commission II*, Beijing, China, July 3–11, 1043–1048.

Zuk, T. and S. Carpendale. 2006. "Theoretical analysis of uncertainty visualizations." *Proceedings of SPIE 6060, Visualization and Data Analysis 2006*, edited by R. F. Erbacher, J. C. Roberts, M. T. Gröhn, and K. Börner, San Jose, CA: IS&T/SPIE, 606007.

Zurita-Milla, R., C. Blok, and V. Retsios. 2012. "Geovisual analytics of Satellite Image Time Series." In *Proceedings of the 2012 International Congress on Environmental Modelling and Software*, Leipzig, Germany, July 1–5, edited by Seppelt, R., Voinov, A.A., Lange, S., Bankamp, D., International Environmental Modelling and Software Society (iEMSs): Leipzig, Germany, 1431–1438.