The Role of Mapping Typological Uncertainty in Decision Making

Robert Roth
Advisor: Mark Harrower

Thesis submitted in partial fulfillment
of the requirements for the degree of:

Master of Science
(Cartography and Geographic Information Systems)

at the

UNIVERSITY OF WISCONSIN-MADISON

2007

Table of Contents

Acknowledgements

I would like to acknowledge first and foremost my primary academic advisor Dr. Mark Harrower of the University of Wisconsin-Madison Geography Department. The quality of this work is a reflection upon his continued guidance, encouragement, and constructive feedback throughout the writing process. I would also like to thank my committee members Drs. Amy Burnicki, David Hart, and A-Xing Zhu for their role in the research. Aaron Erkenswick of the University of Wisconsin-Madison Cartography Laboratory deserves great applause for his “saving” of the online digital survey by offering many hours of programming and web support. Finally, I would like to thank Ann Barrett, Melissa Kraemer-Badtke, and the entire WLIA organization for their assistance in organizing the qualitative focus groups.

"Circumnavigate this body
Of wonder and uncertainty
Armed with every precious failure
And amateur cartography"

-taken from the song "Aside" by the Weakerthans

CHAPTER 1: Introduction

Despite the astounding advancements in the GISciences over the past several decades, no geospatial dataset will ever be perfect. The certainty of geospatial data is compromised by imprecise or imperfect collection techniques, improper spatial and temporal resolutions, incompletion due to time and cost constraints, and an imperfect understanding of the phenomenon being measured, among other issues. Slight uncertainties propagate as the geospatial data is manipulated, transformed, and combined with other uncertain data, providing less than ideal conditions for decision making and data exploration. It is argued that "uncertainty is an intrinsic property of knowledge and not just a flaw that needs to be excised" (Couclelis 2003). As it is impossible to eliminate all uncertainty from a dataset, it is important to understand the nature of such uncertainty, how it aggregates, and how it influences the decisions made based off of graphical representations of the dataset in map form.

A central difficulty in the understanding and representation of geospatial uncertainty is the many forms that uncertainty can take. It was recognized as early as Sinton (1978) that uncertainty has multiple components which need to be treated differently during measurement and representation. However, uncertainties derived from positional, attribute, or temporal imperfections may be easier to conceptualize, measure, and represent graphically than uncertainties produced by the theoretical model of the measured phenomenon, the credibility of the data provider, or the dependence of one data source on other data sources. A typology first offered by Thomson et al. (2005), and recently extended by MacEachren et al. (2005), included nine different categories of uncertainty in geospatial data: 1) accuracy/error, 2) precision, 3) consistency, 4) completeness, 5) lineage, 6) currency, 7) credibility, 8) subjectivity, and 9) interrelatedness[1].

Investigating typological differences is necessary due to the potential impact on decisions made based off of graphical representations of this geospatial data in map form. Research concerning the influence of uncertain representations on decision making has already been conducted in the GISciences using such domains as soil science (Fisher 1993), water quality (Howard and MacEachren 1996), climatology (Pang 2001), environmental policy (Cliburn et al. 2002), historical GIS (Plewe 2002), oceanography (Djurcilov et al. 2002), and urban planning (Aerts et al. 2003a), among many others, with the end goal of generalizability to all domains that use geospatial data for support of making decisions. Determining if different types of uncertainty influence decision making differently (and in what ways) can help inform how the data is represented graphically, providing more effective representations of the phenomenon to decision makers.

This research involved the specific domain of floodplain mapping, a field where incorrect decisions based upon uncertain data can mean a loss of millions of dollars or even human lives. Geospatial floodplain data contains all of the components of uncertainty discussed in the MacEachren et al. (2005) typology. The inclusion of uncertainty representations aims to help decision makers safely place structures in the landscape as well as aid in assessing insurance rates based on the relative flood risk of existing structures, allowing insurance companies to charge high premiums only to those

[1] These different types will be fully defined in Section 2.2.3 and in the glossary.

who rightfully carry the associated risk (helping the owner) while also reducing the frequency that companies are pinned with unexpected claims (helping the firm).

Adopting the MacEachren et al. (2005) typology, this research explored the way in which different types of uncertainty influence both the decisions that were made (i.e. the outcomes) as well as why these decisions were made (i.e. the processes). The research aimed to answer the following questions[2]:

(1) Does graphically representing different types of uncertainties influence the ***decision*** that is made as well as the ***speed*** and ***confidence*** of this decision?

(2) Which type of uncertainty elicits particular decision responses, as well as the most immediate and confident decisions? Which the least?

(3) How much of the variation in the decision outcome is explained by the ***expertise level*** of the decision maker or the ***decision difficulty***?

(4) Which type of uncertainty is the most ***influential*** on the decision making process? Which is the least influential?

(5) *Why* is uncertainty used in decision making the way that it is?

(6) Is the MacEachren et al. (2005) typology a valid categorical model of geospatial data uncertainty? Are there any categories to remove or new categories to add?

To answer these questions, the research was divided into two stages: the first employing a quantitative online survey and the second employing a qualitative focus group. The online survey was used to determine if different categories in the MacEachren et al. (2005) typology had any effect on decision making (Question #1 above), and if so, which type of uncertainty elicited particular decision responses and how these types influenced the immediacy and confidence of the decision (Question #2 above). The survey also spoke to the importance of the expertise level of the decision

[2] The italicized key terms are fully defined in Chapter 4 and in the glossary.

maker and the difficulty of the decision when making choices off of uncertainty data (Question #3 above). The second stage of research, the group interviews, attempted to push the quantitative findings of the first stage and examined which of the uncertainties was most influential in making decisions (it may not be the same as the those that yield the most accurate and confident decisions) and *why* different uncertainties, when represented graphically, influenced the decision making process in different ways (Questions #4 and #5 above). Finally, the second stage of the research investigated if the MacEachren et al. (2005) typology is a valid model for understanding and categorizing geospatial data uncertainty (Question #6 above).

CHAPTER 2 – Relevant Literature

The topic of uncertainty in geospatial data has drawn increasing attention in cartography and GIS over the past decade and a half due to its important role in decision making. Much of the early research on uncertainty is complicated by vague terminology, where terms like error, quality, reliability and validity are defined in a way that only partially matches the modern usage of 'uncertainty' (Edwards and Nelson 2001). Illustrating this point, a core GIS text of the late 1990s, *Principles of Geographical Information Systems*, by Peter A. Burrough and Rachael A. McDonnell (1998), does not once mention the term uncertainty. Discussion of uncertainty is provided briefly with every technique, often explained using the terms reliability and validity. In contrast, the current core GIS text, *Geographic Information Systems and Science* by Longley and his colleagues (2005), reserves an entire chapter early in the text to explain the nature of uncertainty and related statistical measures to quantify uncertainty levels. The presentation of this material so early on in the text (Chapter 6) allows the concepts to resonate throughout the rest of the material when specific spatial analysis techniques are introduced, serving as an overarching reminder that there are no perfect data or processing techniques. There is a similar finding in the core cartography textbooks, as the current primary text, *Thematic Cartography and Geographic Visualization* by Slocum and his colleagues (2003), is the first major cartography text to include a full chapter on uncertainty.

Longley and his colleagues (2005) begin their definition of uncertainty as the inability to perfectly reconcile representations of the landscape with the actual reality of the landscape. In this respect, uncertainty is generated by any imperfect match between the collected or modeled data and the geographic phenomenon being measured. Such a definition is easily confused with the well accepted definition of ***accuracy*** from Heuvelink (1998) as the difference between the reality and our representation of reality. One distinguishing aspect of the Longley et al. (2005) definition of ***uncertainty*** from the Heuvelink (1998) definition of accuracy is the inclusion of the "user" in measurement of uncertainty, stating that "uncertainty may thus be defined as a measure of the user's understanding of the difference between the contents of a dataset, and the real phenomena that the data are believed to represent." To avoid confusion, the terms ***quality***, ***reliability***, and ***ambiguity*** will not be acknowledged individually and are assumed to be part of (although not synonymous with) the larger definition of uncertainty, following the recommendations from Longley et al. (2005).

Such a definition fits well with the departure about a decade ago from the communication model of cartography. The communication model asserts that information is disseminated, or communicated, through the map objectively to the map user (Board 1967, Koláčný 1969). Proper communication of the information meant that every map reader would receive the same message. This model of communication was replaced by MacEachren's (1995) cartographic model of representation. In map representation, the mapmaker converts the world into a set of symbols, which are then placed on the map, and the map user reassembles these symbols to construct meaning. Rather than information being communicated through the map, it is reassembled by the individual, allowing past experiences and biases to influence the end message. Such a user-centered model of cartography fits well with the Longley et al. (2005) definition of uncertainty. The investigator takes the position that the message is inextricably linked to both the data and interpreter and that uncertainties can arrive at both levels.

In description and support of the definition of uncertainty to include the user's understanding, Longley et al. (2005) provides a schematic to conceptualize the four suggested levels in which our representation of the landscape can deviate from the real world. The four described levels are: 1) conception (how a phenomenon is defined as a variable), 2) measurement and representation (how data on this variable is collected, organized, and stored), 3) analysis (how data is used to construct information), and 4) interpretation, validation, and exploration (how information is used to construct knowledge about the original phenomenon, retroactively informing and revising the initial two levels). At each level, a filter is present that acts to remove direct correspondence between phenomenon and representation. The Longley et al. (2005) conceptual schematic of uncertainty is provided in Figure-1.

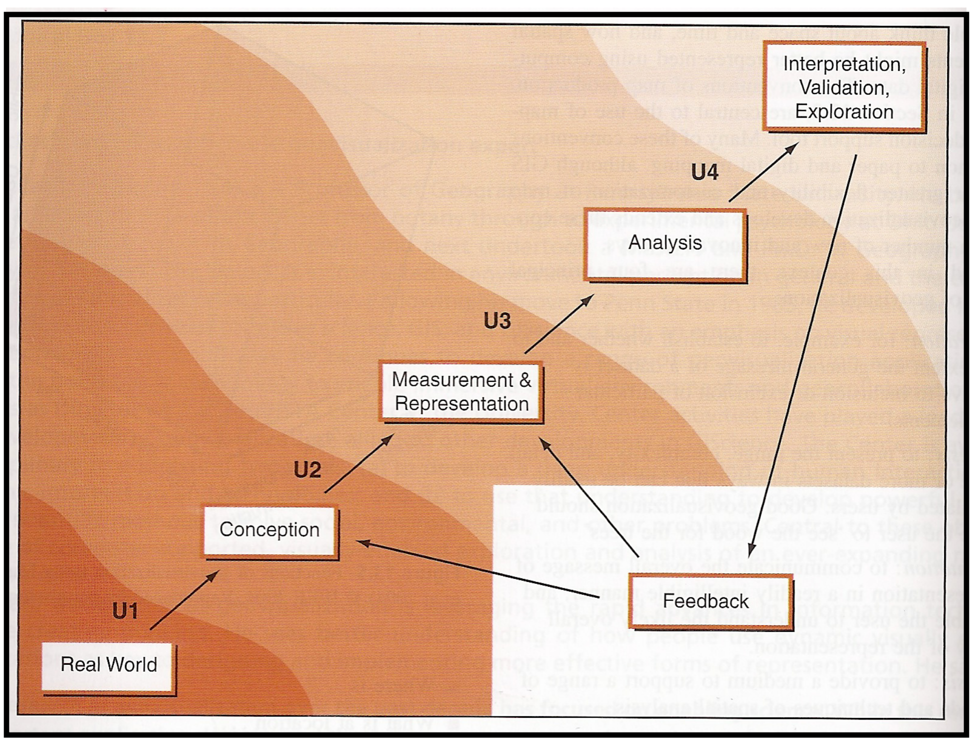

Figure-1: A conceptual schematic of uncertainty from Longley et al. (2005, p296). This schematic illustrates the four levels in which uncertainty can enter the representation.

This definition is controversial in that the uncertainty is no longer internal to the dataset. By including the user's interpretation of the data, perceptions and opinions external to the dataset itself are included as a mode in which the representation deviates from reality. While much of the study on uncertainty is focused on the second stage, the measurement and representation of the phenomenon, the investigator felt it necessary to broaden the definition due to the emphasis of the researching questions on the decision making process. This research adopts the Longley et al. (2005) definition of uncertainty, including the user (defined in the research as the ***decision make***r) as an important aspect of the uncertainty.

2.1 Graphic Representations of Uncertainty

Much of the work in cartography on the subject of uncertainty deals with finding ideal graphic representations for its depiction. The first known cartographer to study the issue was J.K. Wright during the Second World War (McGranaghan 1993). His early investigation into the certainty of map information led him to recommend the inclusion of certainty information using textual labels directly on the map called a 'legend statement' or having a split display with a smaller map of the same region to symbolize the certainty of the original data, called a 'reliability diagram'. The legend statement and reliability diagram are termed 'traditional accuracy indicators', and are characterized by the separation of data and its accompanying uncertainty into two displays (Edwards and Nelson 2001). The naming of these representation techniques predates the widespread use of the term uncertainty, but are taken to mean such by the investigator. Figure-2 shows an example of both a legend statement and reliability diagram:

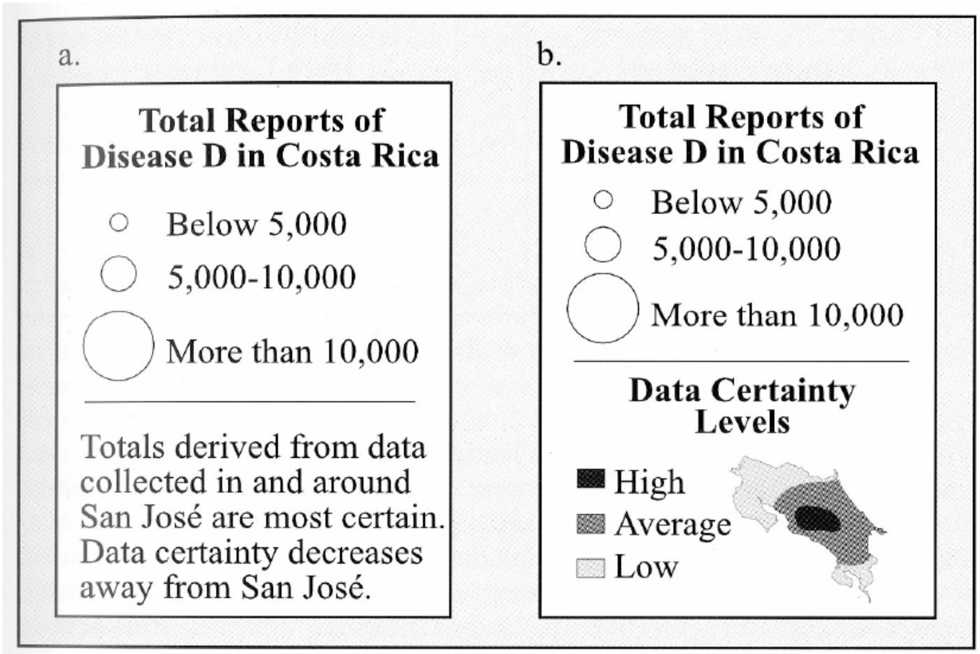

Figure-2: An example from Edwards and Nelson (2001, p21) of (a) the legend statement and (b) the reliability diagram. Both of which are considered 'traditional accuracy indicators'.

Wright was also the first to suggest a form of integrated symbolization, what he termed the 'broken contour', where the same graphic symbol shows both the thematic data and the degree of uncertainty associated with it. The location of the broken contour on the page represented the data surface itself and the value and texture of the line represented its certainty (Wright 1942). Such broken line symbolization is still used to represent the positional uncertainty of fault lines and intermittent streams on geological and topographic maps respectively (Fisher 1993).

A common usage of Bertin's (1983) set of visual variables (location, shape, size, color hue, color value, grain, and orientation) is for the symbolization of uncertainty. Manipulating these visual variables allows for an integrated representation of uncertainty on maps, as one variable can be used to represent the statistical data and a second used to represent the degree of uncertainty. Such depictions have been termed 'static verity visualization', as the use of multiple visual variables allows for the simultaneous depiction of the data and its uncertainty (Beard and Buttenfield 1991, MacEachren 1992, Pang et al. 1997, Kyriakidis 2003). It was initially argued that the visual variables of color hue, color value, and texture (a variation of Bertin's grain) were most appropriate for integrated representations (Davis and Keller 1997).

Since Morrison's (1974) suggestion, color saturation (the third dimension of color, sometimes called 'purity') has been readily accepted as another basic graphic variable. Saturation was initially triumphed as a good method for uncertainty representation, as it was theorized that brighter, more saturated colors showing highly certain data would draw the eye away from the duller, less saturated pastels showing low data certainty (MacEachren 1992, 1995). Figure-3 gives an example of using saturation to represent uncertain information.

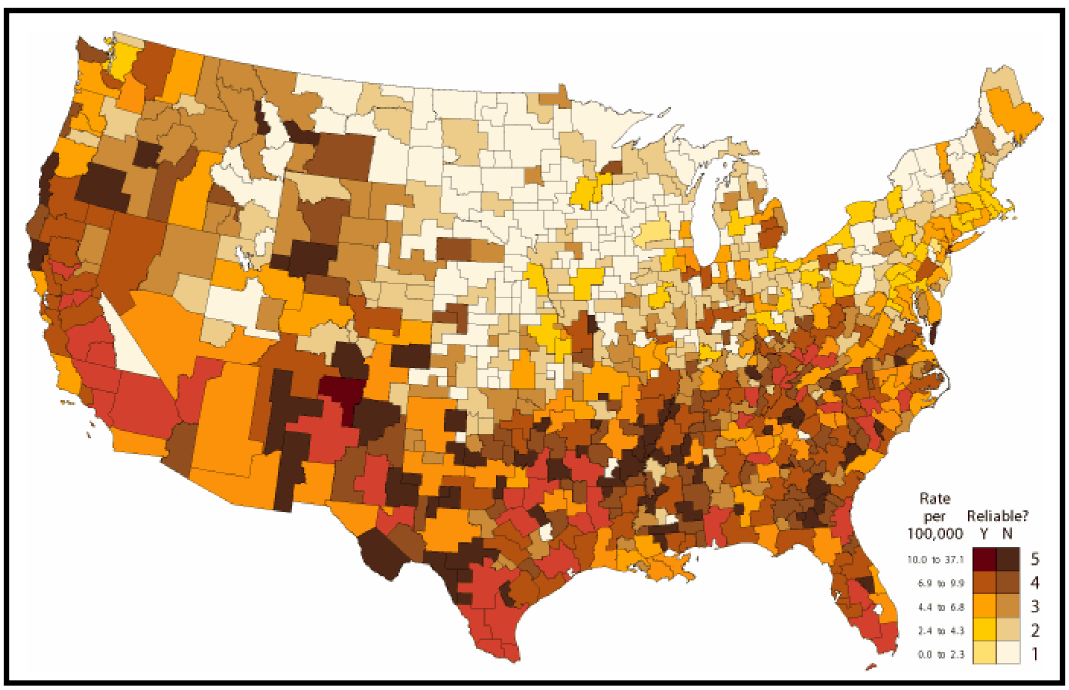

Figure-3: An example from Brewer (1994) showing the use of color saturation to represent uncertain data. Here, enumerations units with less certain data are given the same hue as their certain counterparts, but are less saturated.

Representing uncertainty with saturation allows the data itself to be integrated into the symbol using one of the other dimensions of color (color hue or color value). However, empirical studies have shown that saturation is not effective in communicating uncertainty when a different dimension of color is used to represent the data, as the viewer conflates saturation with the other represented dimension of color, creating the appearance of only a single variable upon which to make decisions (Schweizer and Goodchild 1992, MacEachren 1998 et al., Drecki 2002). Instead, color value and texture prove to be much more effective (Leitner and Buttenfield 2000). Figure-4 shows an example of using texture to represent uncertain data.

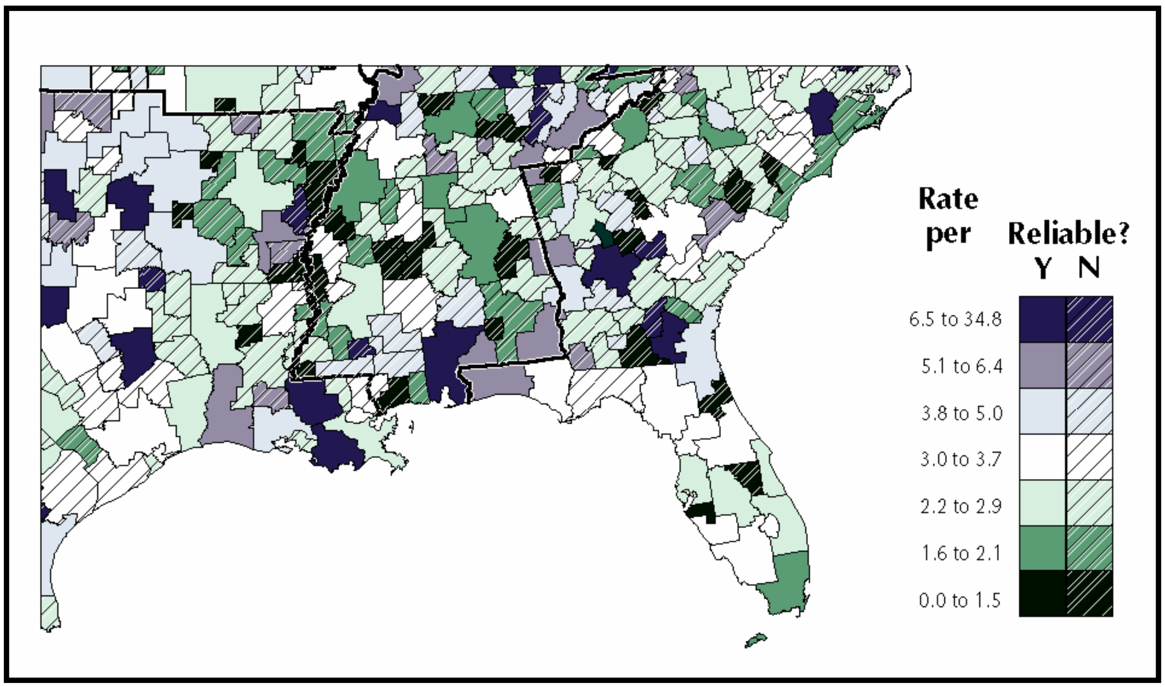

Figure-4: An example from MacEachren et al. (1998, p1552) using texture to represent uncertainty. Here, the hatching (a form of texture) is used to show enumeration units with uncertain data.

MacEachren (1995) suggested the addition of crispness, resolution, and transparency to the list of visual variables. These three variables, described together as 'clarity', deal specifically with the representation of uncertainty on integrated displays. Crispness uses a gradient fill (manipulating the visual variable color value) to represent the probable location of a linear edge when its exact position is uncertain. With crispness, the probability of the edge position is higher where the gradient fill is darker. The variable crispness was renamed in MacEachren (1995) from the van der Wel (1994) usage of the term 'focus' to prevent mistaking it with the term 'data focusing', however the two terms are still used interchangeably in the literature (Edwards and Nelson 2001). Figure-5 provides an application of MacEachren's visual variable crispness.

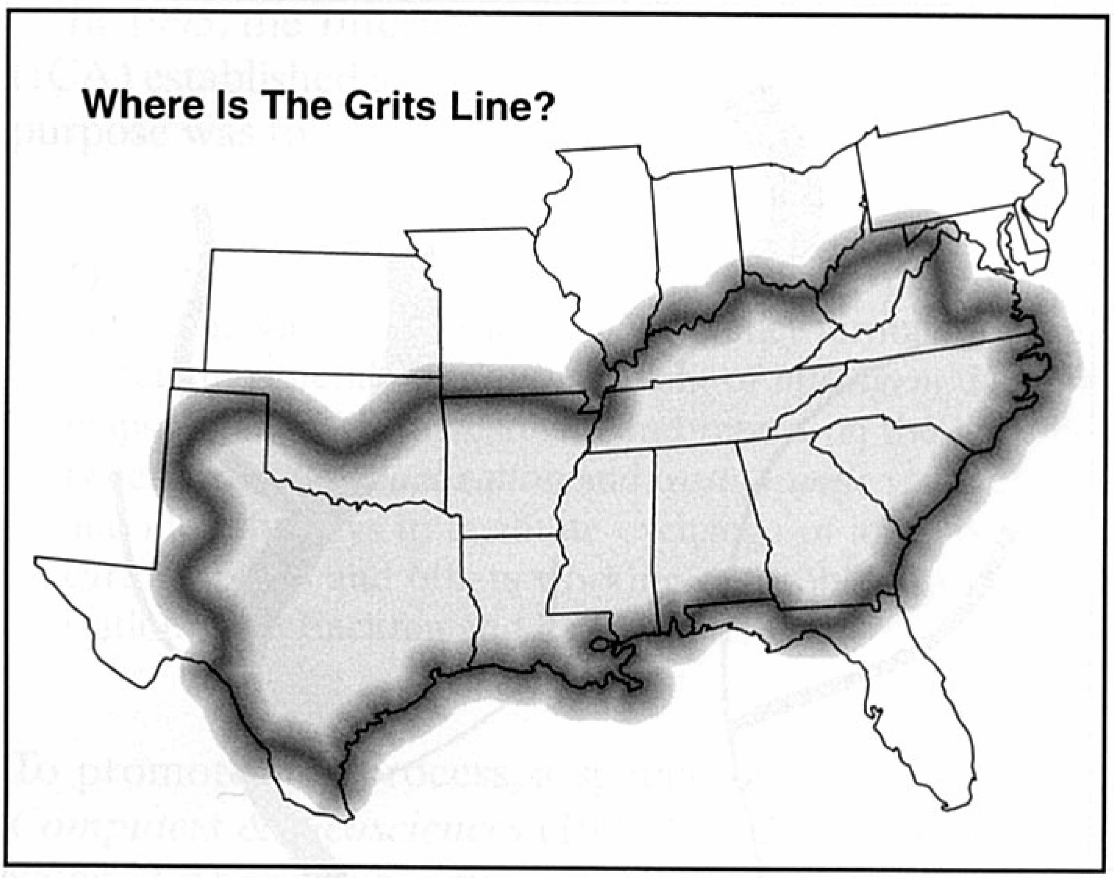

Figure-5: An example of MacEachren's crispness to represent uncertainty, reprinted in Slocum et al. (2005, p421). Here, color value is used to represent the uncertain location of the grits line.

Resolution displays the spatial precision of the data and is typically (although not solely) applied in raster format. When the uncertainly level is high, the spatial resolution is coarsened (using larger pixels), generalizing the data in a way that does not "over-sell" the certainty (MacEachren 1995). Figure-6 provides an example of the use of resolution to represent uncertainty.

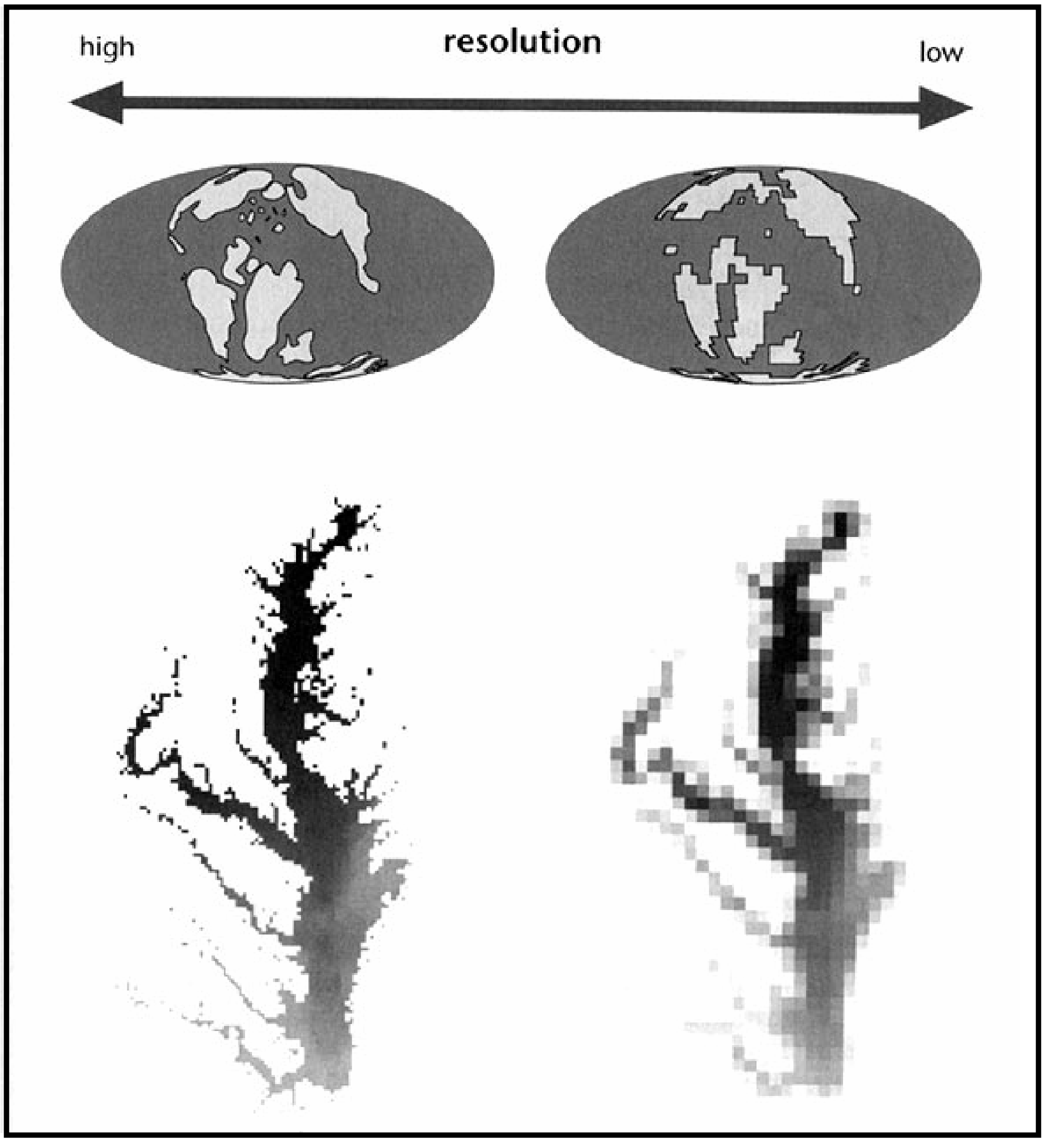

Figure-6: Two applications of the visual variable resolution, printed in MacEachren (1995, p277). Here, the pixilation of linework is varied depending on the certainty of the data.

Finally, transparency (or ‘fog’ as it is sometimes termed) is a non-opaque overlay where the spatial data is covered in hopes to “obscure the map theme in proportion to the uncertainty” (MacEachren 1995). Using this graphic variable, only highly certain areas are viewed with complete clarity. Figure-7 provides an example of transparency.

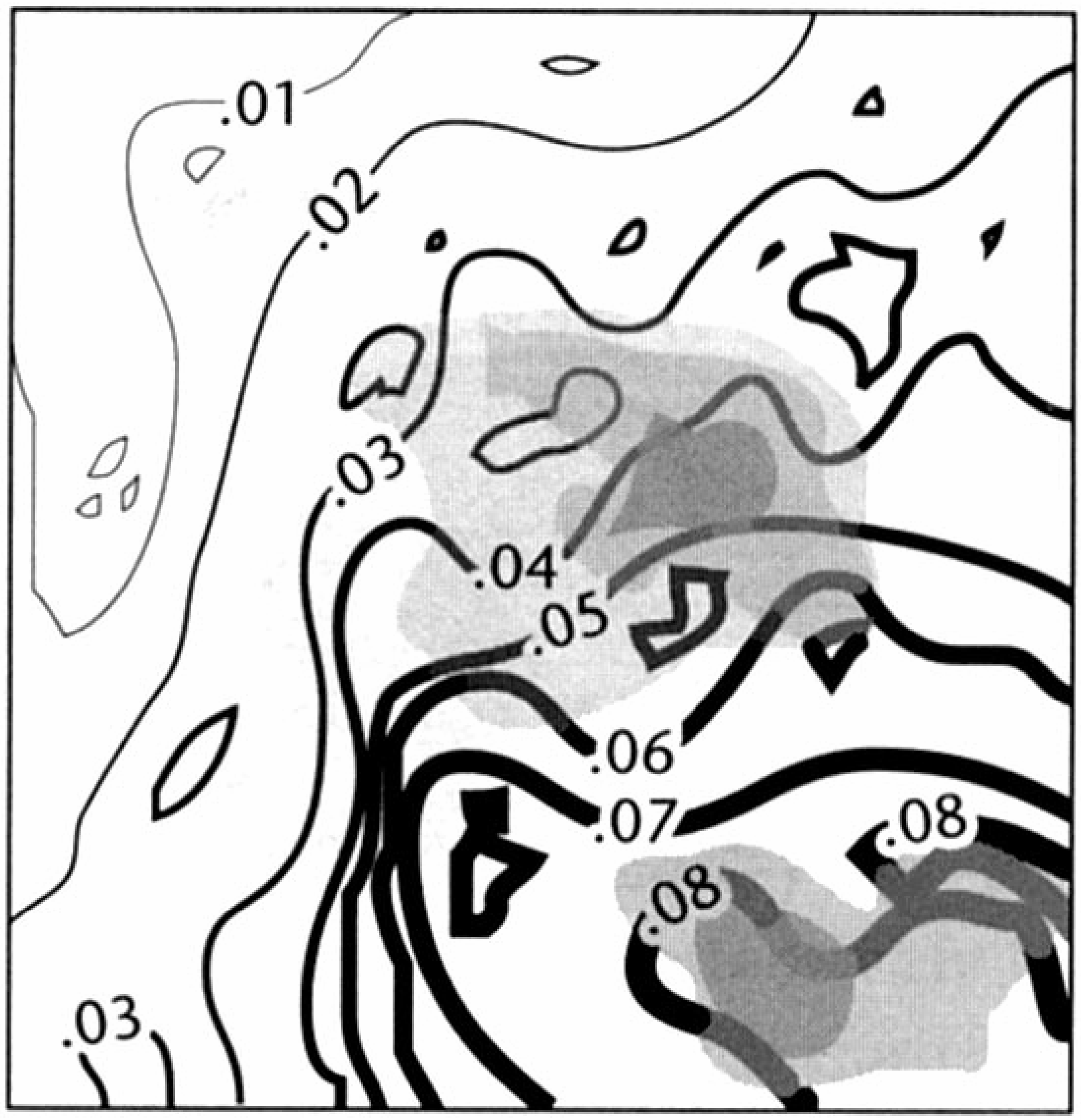

Figure-7: An example from MacEachren (1995, p278) showing the use of the visual variable transparency to encode the level of certainty in the data. Here, areas are overlaid with varying degrees of transparency to mask the linework when uncertain.

Many other techniques have been developed to represent uncertainty on static maps, including realism by McGranaghan (1993), graphic scripts (Monmonier 1993), glyphs (Pang 2001, Djurcilov et al. 2002), and opacity (Drecki 2002), among others. It is largely agreed, however, that dynamic visualization of uncertainty improves the user's understanding of data certainty more so than static displays. Animation was first used by Gershon (1992) as a way to represent positional uncertainty by framing through multiple realizations of the same linework. Animation was used in a slightly different way by Fisher (1993) for soil mapping and land cover classification of remotely sensed imagery, with the probability of each pixel belonging to a particular class represented by the dynamic visual variable duration (i.e. the higher the probably of membership, the longer it would be the classified color onscreen). The probabilities of each pixel classification were determined by the inclusion rate (the percentage of misclassified locations on the map when compared to ground truth). Similar studies of animation for multiple realizations of the same image have been completed by Davis and Keller (1997), Ehlschlaeger et al. (1997), Bastin et al. (2002), Kyriakidis (2003), and Dooley and Lavin (2007). Rather than animating each realization in a dynamic display, the 'Monte Carlo' method generates 500 to 1000 realizations, altering a single stochastic variable each time, and then quantifies the error among all images to generate a single static uncertainty display (Aerts et al. 2003b). Mowrer (1997) argued that this method has a broad range of applications because it does not specify a particular data type or end use. The Monte Carlo method of representing uncertainty has been implemented and applauded by numerous other studies (Lee et al. 1992, Dungan et al. 1993, Journel 1996, Fisher 1998, Heuvelink 1998, Kyriakidis et al. 1999).

Interactive displays for the visualization of uncertainty can be separated into two categories: intrinsic and extrinsic approaches (MacEachren et al. 2005). Intrinsic displays alter the symbolization used to depict the actual thematic data based on its centainty. The idea of an intrinsic display was first suggested by Paradis and Beard (1994) with the development of a ‘data-quality filter’. The ‘data-quality filter’ included three parameters: 1) the quality component (the type of uncertainty), 2) the quality component measurement, and 3) a threshold. On selection of these parameters, the display is updated such that only the geographic entities with certainty in the chosen category above the chosen threshold are shown. Howard and MacEachren (1996) implemented this concept, creating an interactive visualization tool called R-VIS (standing for reliability visualization) that allowed the focusing of interpolated data based on a user-defined certainty threshold. The usages of ‘date quality’ and ‘reliability’ predate the blanket term uncertainty, but are taken to mean such by the investigator. Figure-8 shows a series of screenshots from the R-VIS software.

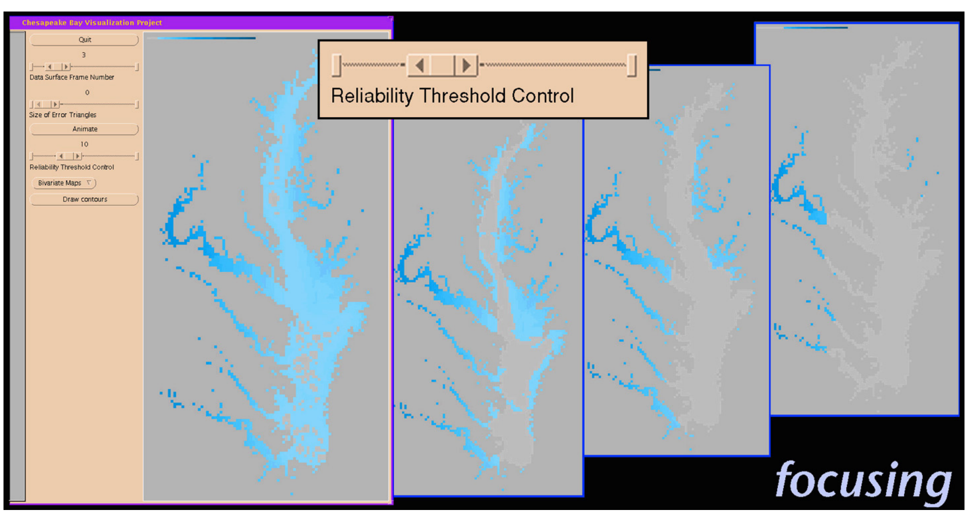

Figure-8: A series of screenshots of the R-VIS software, printed in Howard and MacEachren (1996, p72). As the reliability threshold is increased, data below the new certainty threshold is removed from the display.

Conversely, explicit displays symbolize the degree of certainty by placing a separate symbol atop the display of the thematic data. Cliburn et al. (2002) proposed the application of interactive extrinsic bar glyphs that represent the uncertainty levels across a surface. The interactivity limits the confusion associated with the large field of uncertainty indicators, as users could select individual points or subsets to further interrogate and clarify the display. It is important to note that the intrinsic/extrinsic binary of interactive displays differs from the separated/integrated binary of static displays in that both intrinsic and extrinsic interactive displays produce an integrated symbolization where data and its certainty are on the same map.

2.2 Decision Making under Uncertain Conditions

It long has been asserted that the user of a map or GIS tool approaches it with the assumption that the data displayed is fully certain (Wright 1942, Goodchild 1991, McGranaghan 1993). However, it is unclear how the representation of uncertainty in these tools changes this perception. As Harrower (2003) points out, a fundamental question that needs to be addressed is if "incorporating uncertainty information acts to *clarify the map*, as reported by Leitner and Buttenfield (2000) and Edwards and Nelson (2001), or *clutter the map*, as suggested by McGranaghan (1993)?" It was at first thought that uncertainty information was much like any other type of spatial information in that its inclusion only made the map more complex, cluttered, and difficult to use. Such a position led McGranaghan (1993) and Beard and Mackaness (1993) to caution representing uncertainty information, warning that it may be necessary to limit the symbolization of uncertainty so that the actual thematic data does not becomes clouded and useless. Leitner and Buttenfield (2000) took the first step to test this hypothesis, demonstrating that the integration of uncertainty representations actually decreases the time it takes to make decisions by clarifying the underlying thematic data and by increasing the confidence the decision maker has in his or her decision.

To construct knowledge of how graphic representations of uncertainty are used in decision making, the GISciences have turned to the discipline of risk management and insurance. Geospatial uncertainty can be thought of as a form of risk when the decisions made off the uncertain information carry negative consequences. Agumya and Hunter (2002) incorporated theory from risk management by implementing a fitness for use approach to using uncertain data. To determine if the geospatial data can be used for decision making, the certainty of the data is first assessed to see if it is fit for use by weighing the potential consequences of possible decision outcomes against the likelihood of the outcome's occurrence. Because of this, there are three components that are necessary for the estimation of risk: (1) the data itself (which carries a given degree of uncertainty), (2) the probability of a risk occurring, and (3) the consequences of that risk if it occurs. This estimated risk is then compared against a 'threshold' level of acceptable risk to determine the appropriate response.

Agumya and Hunter (2002) identified four possible options to deal with uncertain data: (1) risk avoidance (not using the geospatial data in the decision), (2) risk reduction (reducing the severity of the consequences, reducing the probability of the consequences occurring, or reducing the degree in which the decision is based off the data), (3) risk retention (accepting all data as long as it surpasses a threshold of certainty), and (4) risk transfer (reducing liability of the decisions made off of the uncertain geospatial information – e.g. via insurance). It is important to explicitly represent the degree of uncertainty on maps so that the user can make informed decisions, even if the added ambiguity makes it more difficult to make a confident decision. Figure-9 illustrates the Agumya and Hunter (2002) decision making model.

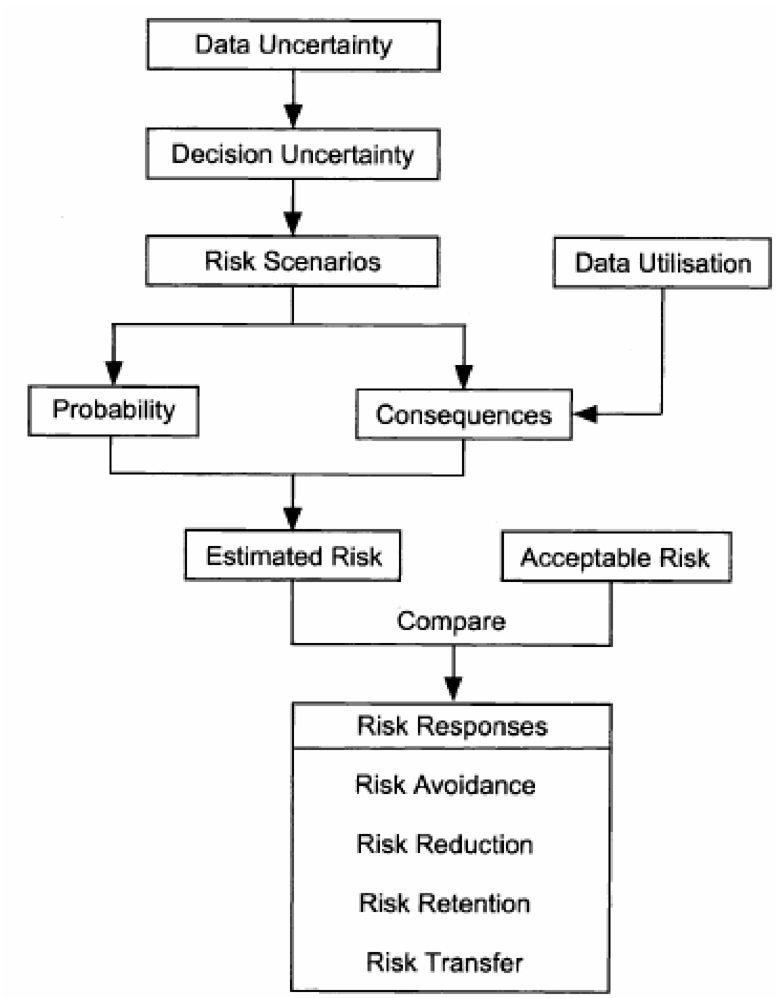

Figure-9: The decision making process under uncertain conditions, printed in Agumya and Hunter (2002, p407). This framework shows how decision makers come to conclusions by weighing the risk of a possible incorrect decision that was based upon uncertain data.

Depending on the domain, there are several factors that influence which of the above risk response options are chosen. Building upon the decision making theory in risk management, the GIScience literature has pointed to three possible factors that influence decision making under uncertain conditions:

(1) the level of expertise,

(2) the difficulty of the decision, and

(3) the type of uncertainty that is being represented.

2.2.1 Factor #1: Level of Expertise

A first possible factor in applying risk theory to making decisions under uncertain conditions is the level of expertise the decision maker has on the topic represented. In the age of the democratization of the GISciences, tools for the display and exploration of geospatial data are now available to non-experts (Rod et al. 2001, Wood 2003). Couclelis (2003) asserts that such tools are mechanisms that allow users to construct knowledge from geospatial information. It is hypothesized by Couclelis (2003) that different types of data uncertainties will affect the knowledge construction of novices and experts differently. Following this thread of thought, it is important to understand how types of uncertainties influence decision makers of different levels of expertise so that novices and experts alike can come to similar conclusions about the represented geospatial information (even if they came to these similar conclusions in a dissimilar manner).

Three studies concerning the representation of uncertainty have specifically looked at the difference between experts and novices. The first study, conducted by Evans (1997), did not focus on decision making explicitly but rather compared several depictions of uncertainty. It was determined, however, that most users could at least understand and utilize uncertainty regardless of level of expertise with the subject. A study by Kobus et al. (2001) specifically examined decision making, recording the impact of representations of uncertainty on the speed and accuracy of employing military tactics. It was determined that more experienced officers made decisions under uncertain conditions more quickly than their inexperienced counterparts. Interestingly, under certain conditions, there was no significant difference of decision making speed between experienced and inexperienced officers. A final study by Aerts et al. (2003) examined visualizing uncertainties in an urban growth model called SLEUTH. Unlike other experiments, they only tested expert urban planners using an easily distributed online survey, but still discriminated subjects based on their level of visualization experience. Several differences between experts and novices were uncovered, including the ability to discern spatial patterns and the preference of bi-color schemes for portraying uncertainty.

2.2.2 Factor #2: Decision Difficulty

A second identified factor is the difficulty of the decision (sometimes called 'task difficulty' in the literature). Beard and Mackaness (1993) were early to recognize the existence of different types of uncertainty assessment tasks, termed 'data quality assessment tasks' by the authors but taken to refer to the broader term of uncertainty by the investigator. Three different levels of assessment were identified: (1) notification, (2) identification, and (3) quantification. Notification indicates the presence of uncertainty in the dataset, while identification indicates the type of uncertainty and where it is located. Quantification describes the magnitude of certainty in the dataset, linking numbers to the degree of certainty. While Beard and Mackaness (1993) ranked difficulty based on the initial assessment of a dataset (i.e. it is easier to determine if uncertainty is present than it is to quantify it), it can be logically transferred to the difficulty in decisions made from representations of that dataset. The Evans (1997) study cited in Section 2.2.1 also commented on decision difficulty, reporting that most users could understand and utilize uncertainty regardless of decision difficulty, a finding that suggests at least notification of uncertainty is not determined by the complexity of the task.

The most involved study on the effect of difficulty on decision making under uncertain conditions comes from Leitner and Buttenfield (2000). While their broader project goal was to determine if representing uncertainty graphically aided or hindered decision accuracy, speed, and confidence, they also examined how the decision difficulty affected these three aspects of decision making. Leitner and Buttenfield defined two tasks: (1) selecting the optimal location for a park based on predetermined planning criteria and (2) selecting the worst location for an airport using the opposite of the given planning criteria for the park. Because the subjects needed to reverse the criteria for the airport siting decision, it was deemed more difficult in comparison to the park siting decision. One of the most significant findings is that decision time decreased when uncertainty information was included for easy tasks, but not for difficult tasks. This is a clear indication that the difficulty of the task impacts the importance of uncertainty on the decision.

2.2.3 Factor #3: Type of Uncertainty

A final characteristic stressed in the literature that affects the way decisions are made based off of uncertainty representations of geospatial data is the type of uncertainty itself. In almost all studies conducted on the representation of uncertainty, only a single graphic depiction is used to encode all of the many types of uncertainties the geospatial data may contain. However, MacEachren et al. (2005) has indicated that a challenge before the field of cartography and geographic visualization is "Developing representation methods for depicting multiple kinds of uncertainty". While Leitner and Buttenfield (2000) and Edwards and Nelson (2001) have empirically demonstrated that representing uncertainty acts to clarify the underlying data, it is not clear if representing multiple types of uncertainty will have the same clarifying effect or instead confuse the map reader and hinder decision making. To answer the call of MacEachren et al. (2005), it is necessary to first examine if the representation of multiple types of uncertainty acts to clarify the decision and then, if so, examine the nature of each type.

Although not a true typology of geospatial data uncertainty, MacEachren (1992) acknowledged early on that there are three different aspects to geospatial data: (1) positional/locational, (2) attribute, and (3) temporal. Uncertainty can be introduced into the data at each one of these aspects. Figure-10 provides examples of uncertainty for the above three aspects of geospatial data.

	Locational	Attribute	Time
Precision	state birth rates (vs. county)	soil order	mean monthly rainfall (vs. daily)
Accuracy	*position* of hotel on road map	*total #* lung cancer cases per county	*date* of the last tornado

Figure-10: MacEachren's (1992, p12) discussion of three aspects of geospatial data in relation to uncertainty. Here, an example of accuracy and precision is provided for the positional/ locational, attribute, and temporal components of geospatial data.

The first typology of uncertainty offered in the literature adopted the Spatial Data Transfer Standard (SDTS) categorization of the U.S. Federal Information Processing Standard (FIPS) (Buttenfield 1993). The SDTS lists five types of possible uncertainties: (1) positional accuracy, (2) attribute accuracy, (3) logical consistency, (4) completeness, and (5) lineage. Paradis and Beard (1994) discussed this typology in their 'data-quality filter', theorizing that each type of uncertainty could be used individually to filter the dataset. Zhu (2005) offers an interesting categorical analysis of what he terms "aspects of data accuracy." Elements of the accuracy of a dataset included in the Zhu (2005) typology are: (1) accuracy, (2) precision, (3) resolution, (4) consistency, and (5) completeness. It is important to note that this typology does not match the adopted definition of uncertainty from earlier in the chapter, as each listed element is internal to a single dataset. MacEachren et al. (2005) provides an updated typology based off of the components suggested by Thomson et al. (2005). The following topology, taken verbatim from MacEachren et al. (2005), is the most contemporary and extensive list of uncertainties, and has been adopted for this research:

(1) Accuracy/error: difference between observation and reality, usually estimated based on knowledge of the measurement/estimation device and of phenomena in the work.

(2) Precision: exactness of measurement/estimate, derived from parameters of the measurement, estimation device, and/or procedure.

(3) Completeness: extent to which information is comprehensive.

(4) Consistency: extent to which information components agree. This is a more general definition than that found in formal standards for spatial data.

(5) Lineage: conduit through which information has passed. This is a complex category that has at least the following subcomponents: number of individuals, organizations, processes through which information moves; specification of which individuals, organizations, or processes.

(6) Currency: time span from occurrence through information collection/processing to use. The certainty that information is "current" will be a function of both time span and context, e.g., year-old data about vehicles parked in a factory loading bay is less certain to be current than year-old data about location of the factory.

(7) Credibility: combination of factors such as reliability of information source. Certainty may be based on past experience, e.g., the analyst is correct 85 percent of the time, or on categorization of the source, e.g., U.S. analyst versus a non-U.S. informant; motivation, experience, or other factors.

(8) Subjectivity: the extent to which human interpretation or judgment is involved in information construction. This component of uncertainty is, of course, difficult to assess—and that assessment will have some level of subjectivity.

(9) Interrelatedness: source independence from other information. This is a common standard used in the news media to assess certainty that a story is authentic.

CHAPTER 3 – The Case Study

Floodplain maps are a formalized cartographic tool used for the evaluation of flood liabilities. Urban planners use these maps to assess potential hazards of construction in a watershed before areas are developed. Similarly, federal emergency agencies and private insurance firms use these maps to evaluate the risk scenarios of structures that are already in the landscape, whose construction cannot now be avoided. Such a mapping application provides a powerful situation in which to investigate the impact of uncertainty representations on the decision making process and its relation to risk management and insurance. Decisions made off of these maps have real world importance, with millions of dollars and lives hanging in the balance. Understanding how uncertainty can be successfully integrated into these maps facilitates the making of informed decisions about placing and insuring structures in the landscape, helping both the individual owner and the insurance firm.

The Great Midwest Flood of 1993 illustrated the importance of obtaining geospatial information with a high degree of certainty. The flood, occurring along the Mississippi and Missouri Rivers between May and October of 1993, is an example where an incomplete understanding of a phenomenon, fueled by uncertain data, caused decision makers to grossly underestimate the 'worst-case' flooding scenario (FEMA 2003). Due to these miscalculations, over one-thousand levees failed across the Midwest, allowing the flood waters to aggregate downstream. As a result, an estimated 12 to 16 billion US dollars in damaged was caused, dislocating fifty-four thousand people from their homes, and taking the lives of fifty individuals (Strahler and Strahler 2002). This catastrophe demonstrated the real-world risk of making an incorrect decision (or in this case, a set of incorrect decisions) based upon uncertain data.

3.1 Domain Concepts and Terminology

The term ‘stream’ is defined as “a long, narrow body of flowing water occupying a trenchlike depression, or channel, and moving to lower levels under the force of gravity”, while the term ‘river’ tends to be used only when describing a stream of significant flowage (Strahler and Strahler 2002). The stream drains the unabsorbed, excess water from the two adjacent hillslopes, transporting the water downstream into a larger system. As the discharged water moves to continually lower levels, it becomes organized into a hierarchical drainage system of stream tributaries which flow into larger rivers, eventually discharging into the river’s mouth (Knighton 1998). When this drainage system is constrained by ridges (called drainage divides), the system is termed a ‘drainage basin’ or ‘watershed’.

The water level in the stream’s channel, which is a function of the raw volume of water that is being discharged and the size of the channel itself, fluctuates during periods of heavy precipitation or extreme drought. The term ‘flood’ can be defined as any occurrence when the size of the stream’s channel is no longer large enough to accommodate the volume of water that is being discharged (Strahler and Strahler 2002). When the waters exceed the channel, they spill onto flat, surrounding lowland called the ‘active floodplain’. The active floodplain is defined by the adjacent lands that are expected to flood on an average of once a year. In contrast, the ‘genetic floodplain’ is defined as all adjacent lowland areas that become periodically inundated by water over the course of roughly 100 years (Alexander and Marriott 1999).

Despite the ever-present risk of flood damage, floodplains are one of the most heavily settled areas in America. Although only five percent of the United States is located in a floodplain (either fluvial or coastal), these lands hold approximately 25% of the nation's population (Krimm 1998). Settlement on a floodplain has two major economic advantages: (1) fertile lands for agricultural activity and (2) a natural mode of transportation for trade (Alexander and Marriott 1999). Periodic flooding acts to constantly replenish nutrients in the floodplain's soil, making the areas extremely productive for the cultivation of crops. Also, the overland flow of water off of the surrounding bluffs concentrates water on the lowlands, providing natural irrigation when there is not a flood event. Finally, the proximity of the floodplain to a stream or river provides a natural means for the exchange of goods cultivated or produced in the floodplain. Many of the cities incurring the most damage from the Great Midwest Flood of 1993, such as Cape Girardeau, MO, Grafton, IL, and St. Louis, MO, have economies reliant upon the river for transportation and trade, illustrating how the river is both a source of prosperity and destruction for the surrounding floodplain (FEMA 2003).

3.2 Application of the Agumya and Hunter (2002) Model

Assessing the flood risk of a site in the floodplain is a decision task that fits well with the Agumya and Hunter (2002) decision making model based on risk management (see Figure-9). In order to estimate the risk of a flood, three components are necessary:

(1) geospatial data,

(2) the probability of that the identified event will occur, and

(3) the consequences of the event.

In the domain of floodplain mapping, the geospatial data include a delineation of the floodplain extent and a point coordinate of the site in question. Regarding the delineation of the extent, most floodplains exhibit a common topographic anatomy, as shown in Figure-11.

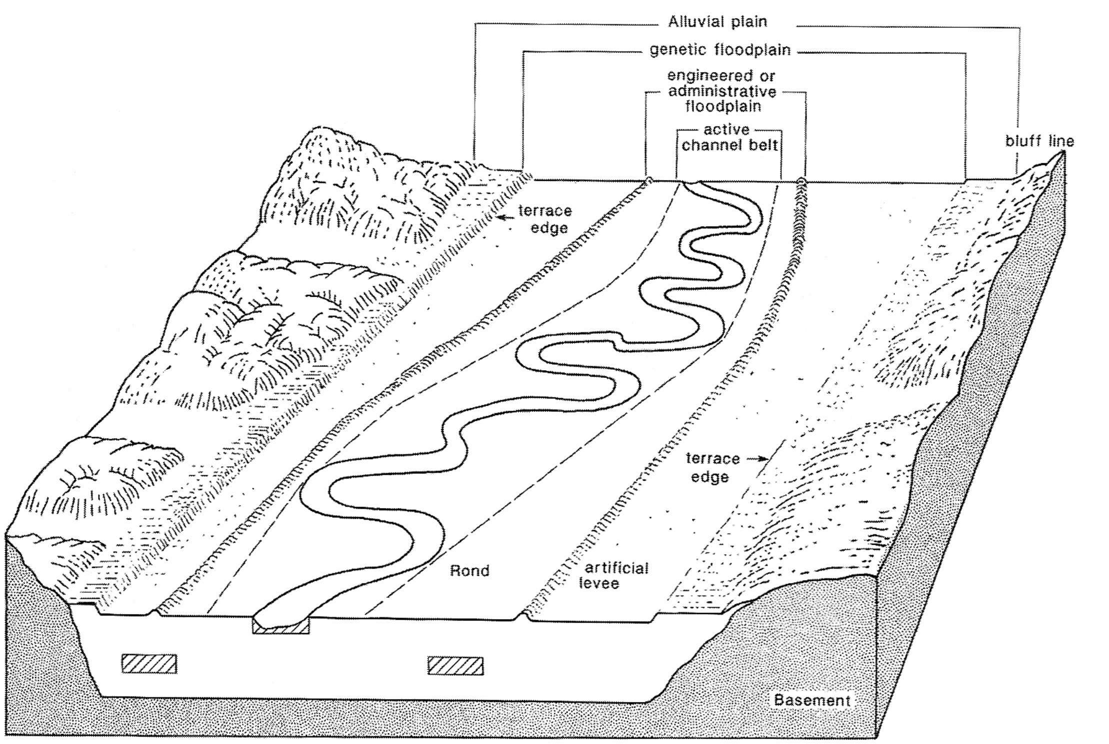

Figure-11: The topographic anatomy of a floodplain, printed in Alexander and Marriott (1999, p4). The boundary of the genetic floodplain is typically marked by a terrace edge.

The above diagram illustrates how a stream or river meanders over the landscape to elongate the course of the discharging water, slowing the flow to a more stable energy level. Important to this diagram is the demarcation of the genetic floodplain based on the lowest terrace edge. A 'terrace' is a flat section of land that is marked by a steep descent at its edge and is caused by a coupling of a river's incision into the landscape with periods of massive sedimentation during flooding (Knighton 1999). It is not uncommon to have multiple terraces before finding the hard bluff line, with the terrace of highest elevation being the most ancient. Because the terrace of lowest elevation confines the genetic floodplain, the simplest method of floodplain delineation is to accept the contour of the inner-most terrace as the floodplain line (Alexander and Marriott 1999).

A second method to delineate the extent of a flood is to use historical data (Thorndycraft et al. 2002). The most common usage of historical flood data is to plot the magnitude of a flood against the frequency that the magnitude occurs in the historical record (Strahler and Strahler 2002). Because floods of large magnitudes occur much less frequently than small flood events, there is typically an inverse relationship between magnitude and frequency. The frequency is then used to describe the probability of a particular magnitude occurring in a given year. For instance, the 100-year flood represents a magnitude that has a 1% probability of being equaled or exceeded each year. A common misconception is that the 100-year flood only occurs every one hundred years; a 100-year flood can occur in continuous years as well as multiple times in the same year. Because the genetic floodplain represents a flood that occurs approximately every one hundred years, the floodplain can also be defined as the extent of the 100-year flood.

A final method for delineating the floodplain extent is to examine the stream discharge itself. Stream discharge is defined as the volume of water flowing through a particular areal cross-section of the stream over a given unit of time, and is measured in cubic meters or cubic feet per second (Strahler and Strahler 2002). A hypothetical level of discharge can be simulated against the average or expected discharge to determine if the water will spill over the channel banks. The discharge rate associated with the level of exceedance that inundates the adjacent lowlands can then be used to determine the extent of the floodplain. The use of a hypothetical exceedance of discharge reduces the reliability on an accurate historical record.

The second two aspects necessary for the Agumya and Hunter (2002) decision making model, the probability of occurrence and the consequences of occurrence, are directly related to the probability of the occurrence of a given flood magnitude and the degree of damage a flood of the same magnitude would generate. The probability value used in the decision is a function of the structure's position in the landscape relative to the delineated floodplain boundary in the dataset. The consequence, when thinking purely in monetary terms, is a function of the magnitude of the given flood and the value of the damaged structures. The consequence value requires the utilization of an outside data source (see Figure-9), in this case some sort of property value assessment. In this research, the probability of a flood and the consequence are held constant so that only the difference in submitted data uncertainty is reflected in the estimated risk.

The National Flood Insurance Program provides a real-world application for the Agumya and Hunter (2002) model. The National Flood Insurance Program offers proactive protection to individuals living in the floodplain through the sale of federally backed flood insurance (Krimm 1998). The National Flood Insurance Program has the responsibility of assessing the flood risk of structures that are currently in the landscape as well as promoting sound construction practices to reduce future property damages.

Figure-12 illustrates the decision making process of the National Flood Insurance Program according to the Agumya and Hunter (2002) model:

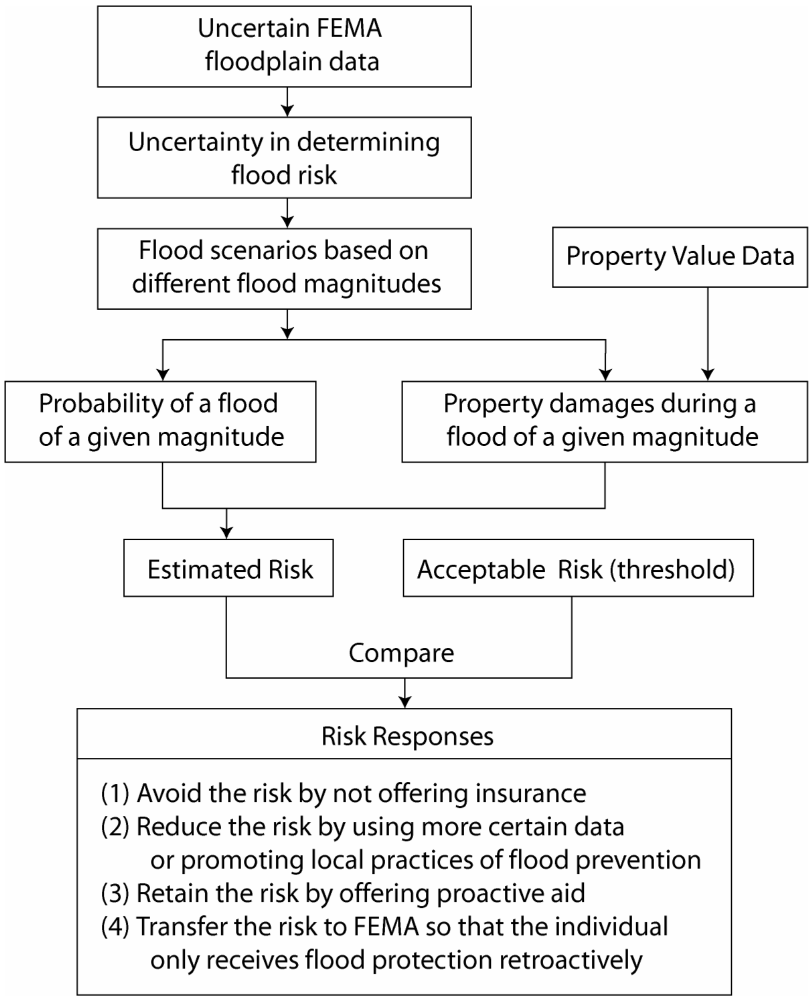

Figure-12: Applying the Agumya and Hunter (2002) decision making model to floodplain mapping.

The decision begins with floodplain data provided by FEMA (Krimm 1998). Because of an array of uncertainties associated with the data (examples provided in Section 3.3), the decision response using the FEMA data will be uncertain. Using the data, a series of flood scenarios are constructed for a specific site depending on the magnitude of the flood event. Once each flood scenario is identified, the probability of that flood event occurring is weighed against the damage that the flood event with cause.

The evaluation of the probability of the flood against the consequences of the flood generates the estimated risk of a structure in the landscape. This estimated risk is then compared with the acceptable level of risk determined by the National Flood Insurance Program. The acceptable risk level is a 'threshold' determined by planners that ranges between a zero-risk event (a non-flood or flood that causes no damage) to the 'worst-case scenario' (representing the gravest possible situation where property damage and lethality are peaked) (Kondo 1998). After this comparison, the National Flood Insurance Program has four options: (1) avoid the risk by not offering insurance, (2) reduce the risk by either improving the certainty of the data on which the decision is made or by reducing the consequences of the risk by promoting local practices of flood prevention, (3) retain the risk by offering flood insurance to the individual, or (4) transfer the risk to FEMA so that the individual only receives retroactive protection in the form of disaster relief aid.

3.3 Application of the MacEachren et al. (2005) Typology

Floodplain data contains all of the components of geospatial uncertainty outlined by the MacEachren et al (2005) typology. According to category theory, the groupings in a theoretical typology must be both mutually exclusive (a single instance belongs in only one category) and collectively exhaustive (no instance can be assigned to a category) (McGrew and Monroe 2000). However, it is important to note that when this typology is applied to floodplain mapping, the categories are not mutually exclusive or collectively exhaustive. In several situations, an uncertainty category from the MacEachren et al. (2005) typology is completely reliant upon the degree of uncertainty in another defined category. For this research, the term ***first-order component*** is defined as an uncertainty type about the data itself (either the positional, attribute, or temporal value), while the term ***second-order component*** is defined as a derived uncertainty type that is contingent upon, at least in part, the degree of uncertainty in a first-order component. A discussion on the methodological impact of first-order versus second-order components on the research is provided in Section 4.1.2.

For floodplain data, this research identifies first-order components as (1) accuracy/error, (2) completeness, (3) credibility, (4) currency, (5) precision, and (6) subjectivity, while second-order components include (1) completeness (found at both levels), (2) consistency, (3) interrelatedness, (4) lineage, and (5) subjectivity (found at both levels). This division is only theorized for the domain of floodplain mapping, and it is unsure how it would transition to other case studies. Figure-13 summarizes the divide between first-order and second-order uncertainties.

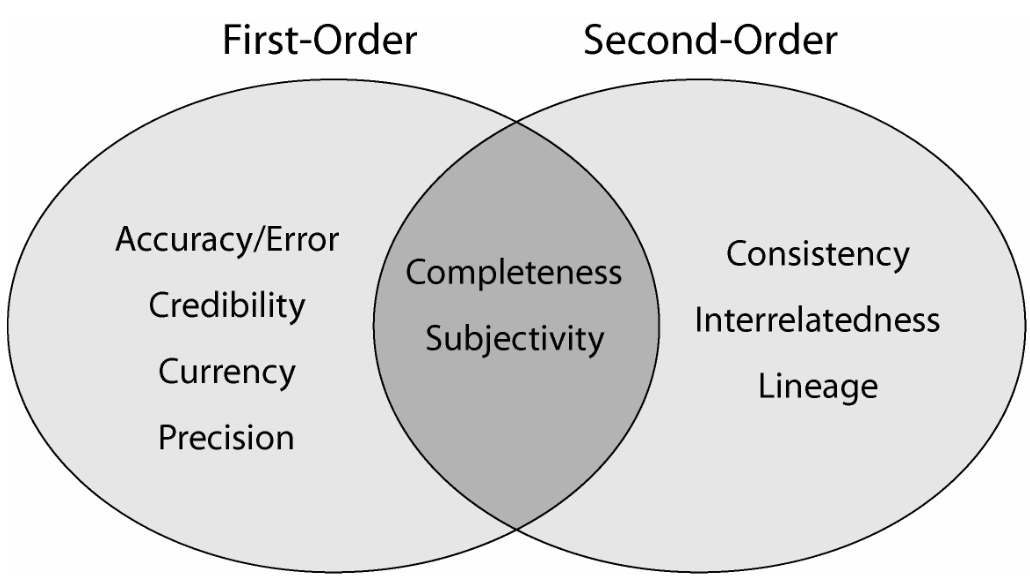

Figure-13: A Venn Diagram dividing the MacEachren et al. (2005) typology into first-order and second-order components

3.3.1 First-Order Components in Floodplain Mapping

(1) Accuracy/error: MacEachren et al. (2005) defines the first type of uncertainty as the difference between the observation (the floodplain delineation) and the reality (the actual floodplain). It is important to note that accuracy and error are not interchangeable terms. While ***accuracy*** is used to describe the difference between observation and reality (see Chapter 2), ***error*** is viewed as the "inverse of accuracy" and is defined as "the discrepancy between the attribute value in the database and the actual attribute value" (Zhu 2005). Error, in a scientific sense, has also been defined as the "differences between observers or between measuring instruments" (Longley et al. 2005). In this regard, error is the degree of repeatability in the measurement, either among observers or among measuring devices, and not related to the ability to correctly capture the underlying reality. While the Zhu (2005) definition justifies the conflation of the terms accuracy and error, the Longley et al. (2005) definition suggests this conflation is perhaps inappropriate. For simplicity, the Zhu (2005) definition is adopted in this research.

In its application to floodplain mapping, positional accuracy/error is the ability to draw the observed floodplain boundary in the correct location. When defining the floodplain based on the historical record, there is also a very important temporal accuracy/error component to the certainty of the data. The ability to correctly predict the probability of a flood occurrence of a given magnitude relies on the matching of a flood to a date (Strahler and Strahler 2002). This is especially true when attempting to determine changes and trends in the flood tendency over time. Accuracy/error in floodplain mapping, like uncertainty derived from inaccuracy or error in all other domains, is very difficult to determine because the reality is never fully known (if reality was known, then there would be no need for a measurement of the accuracy or error). Because of this, accuracy/error uncertainty is best reported at a ratio level as the probability or likelihood of the position of a line in a given location and is typically represented using the visual variable crispness (Longley et al. 2005, MacEachren et al 1995). The calculation of these accuracy/error probabilities is modeled for the full spatial extent based upon sampled control points. Such quantification of uncertainty using probability theory has several drawbacks, as described by Zhu (2005). Because the type of uncertainty is important to the study, and not the mode in which the uncertainty was statistically generated, the probability of a flood event, derived by the frequency of the event in the historical record, was used a proxy for the category accuracy/error. Discussion of the limitations of using this proxy as the accuracy/error measure is provided in the concluding chapter.

(2) Completeness: When used as a first-order component of uncertainty, completeness describes the degree to which the dataset is comprehensive in location, attribute, and time. The linework that FEMA provides is not continuously exhaustive for all of the United States, leaving locations unmapped. Because FEMA provides data in 'panels', it is not uncommon to have an abrupt and arbitrary edge to floodplain data, leaving a portion of the floodplain unmapped (Krimm 1998). Completeness is also an issue for non-FEMA sources, as floodplains often do not correspond to the arbitrarily placed enumeration units that restrict the jurisdiction of state or local data commissions.

(3) Credibility: Credibility relates to the reliability of the data provider or source. The most credible source for floodplain data was expected to be FEMA.

Because the collection of detailed geospatial floodplain information is important for the allocation of disaster prediction and relief, FEMA has spent over 1 billion US dollars on floodplain data and currently maintains over 90,000 map panels (Krimm 1998). However, FEMA only generates new linework or reviews existing linework if a community petition is granted or if significant changes in the area have occurred (e.g. rapidly increased development, a recent flood of unexpected magnitude, severe change in flood boundaries, etc.). Because of this, the best available data may be from a local or state source.

This example of credibility of the data source addresses the definition of uncertainty as "a measure of the user's understanding of the difference between the contents of a dataset, and the real phenomena that the data are believed to represent." Unlike accuracy/error and completeness, which are measured at a ratio level internally, the inclusion of credibility information is categorical internally (as the source is a metadata recording with no attached credibility judgment), but becomes ordinal in nature during the final filter of user interpretation described in Figure-1 (as the user attaches judgments of credibility based on experience with the source, certification of source, etc.). The inclusion of credibility as an uncertainty type decisively separates the MacEachren et al. (2005) typology from that described by Zhu (2005) due to the inclusion of external evaluations of uncertainty on a dataset.

(4) Currency: Currency deals with the time span between collection of the data and presentation of the data for decision making. Unlike temporal accuracy, discussed above, currency focuses on the degradation in the data's relevancy over time, rather than accurately pinpointing the dataset's creation date. Drainage networks are dynamic in nature, with the location of the floodplain characterized by massive movements over time (Knighton 1999). Because the floodplain can shift greatly in a short amount of the time, the more current the geospatial information, the more certain a decision maker can be when assessing a location's risk of flood. While data generated from FEMA would likely be taken as the most credible, it is often characterized by a large degree of currency uncertainty because the linework is not updated at a regular interval. Further, even after existing linework is revised by FEMA, the processing time for surveying of the floodplain and publication of the updated data is on average 58 months (Krimm 1998).

(5) Precision: The precision component involves the exactness of the measurement. Typically, the precision is described as the number of digits used to report a measurement (Longley et al. 2005). However, the description of precision as the "degree of detail that can be recorded" calls into question a parallel issue of resolution (Zhu 2005). There is a considerable amount of research that separates the definition of the term ***resolution***, as the level of spatial detail, from ***precision***, as the exactness of measurement. Resolution deals with a fundamental cause of uncertainty: it is impossible to continuously collect data on a phenomenon at every single point in space. Because of this, it is necessary to collect data over subdivisions of space (pixels, enumeration units, etc.). The size of these spatial units is typically a function of the scale at which the data is collected. The aggregation of measurement into larger spatial units filters detail from the representation of reality, introducing uncertainty. In accordance with category theory, this issue of resolution is incorporated into the uncertainty category precision in order to remain collectively exhaustive. Discussion of limitations in collapsing precision and resolution into a single category are provided in the concluding chapter.

The most applicable element of precision/resolution in floodplain mapping deals with positional delineation of the floodplain. Positional precision/resolution involves the level of detail in the demarcation of the floodplain line and is a function of the scale at which the data is produced. As Knighton (1999) points out “Channeled flow occurs over a large range of spatial scales, from small headwater streams to major rivers.” The scale that is chosen for the generation of the linework dictates the level of detail that is captured. Floodplain boundaries that are delineated at too coarse of a scale run the risk of over-generalizing the landscape, smoothing over important nuances that may place a structure inside or outside of the floodplain.

An example of uncertainty derived from positional imprecision is documented in a review of the ‘Floodplain Redelineation Project’ in Winnebago County, WI (Lulloff 1994, LICGF 1998). In the early 1990s, Winnebago County desired to use FEMA Flood Insurance Rate Maps (FIRMs) with its new large-scale, digital planimetric, topographic, and cadastral mapping. Due to the fact that the FEMA data was compiled at a coarser scale, the discrepancies due to precision (as well as currency) were so severe that the river channel often jumped over the floodplain linework. This ultimately led to a re-delineation of the floodplain linework to match the local base mapping.

(6) Subjectivity: When used as a first-order component, subjectivity relates to the amount of human decision involved in the generation of the data. Although the subjective interpretations or judgments should be informed by experts in the domain, they can be at times arbitrary. An important uncertainty in defining the floodplain is the locational/positional subjectivity of the method chosen to draw the line. As Alexander and Marriott (1999) point out, a floodplain can be defined in many different ways, and the only accepted definition by scholars is deliberately vague. As discussed in Section 3.2, the delineation of the floodplain can be completed by using the terrace contour, by examining the probability of flooding using the historical record, and by simulating a hypothetic level of discharge exceeding the average flow rate. Each of these will produce a different floodplain extent, introducing a degree of subjective uncertainty into the linework. Similar to the first-order component credibility, subjectivity is categorical in nature when recorded internally to the geospatial data, but becomes ordinal when interpreted by the decision maker (as previous experiences and technique biases come into play).

3.3.2 Second-Order Components in Floodplain Mapping

(1) Completeness: Completeness as a second-order component of uncertainty looks at the comprehensive coverage of the first-order uncertainty components in space and time. Completeness in this respect translates to the availability of metadata explaining the first-order uncertainties and represents the existence or non-existence of a particular first-order component at a given location. Completeness as a second-order component of geospatial data uncertainty enters into floodplain mapping through five of the six first-order components: the completeness of accuracy/error information, the completeness of credibility information, the completeness of currency information, the completeness of precision/resolution information, and the completeness of subjectivity information. It is theorized that examination of the completeness of the sixth first-order type, completeness, is a redundant, self-referencing activity and therefore not a valid type of geospatial data uncertainty. Each of the five cases of completeness as a second-order uncertainty only determines if the first-order uncertainty is provided throughout the full areal or temporal extent (if at all). This existential quality can extend to the other second-order components as well, making the category completeness either an overarching uncertainty type of the other eight or an internal type that exists within each of the eight, and therefore not part of the MacEachren et al. (2005) typology.

(2) Consistency: The consistency component of geospatial data uncertainty looks at how a particular first-order component varies over space and time. While completeness as a second-order component looks at the existence of a first-order component, consistency examines the variation in quality of a first-order component over space and time. In floodplain mapping, consistency addresses each of the first-order components, excepting completeness (since consistency is an evaluation on data that has already been deemed to exist through completeness). A particularly prevalent uncertainty of floodplain data is derived from the consistency of credibility. It is not uncommon for FEMA to incorporate the data generated by a local or state commission directly into their own linework to avoid redundancy of data collection (Grimm 1998). Although this occurs only when the local or state data is agreed to be of much higher quality, it still introduces an inconsistency of source.

(3) Interrelatedness: The interrelated component involves the extent to which source data layers are dependent on each other. An ideal application of floodplain delineation would use source layers that contained equal degrees of each of the first-order uncertainties (e.g. same degree of accuracy/error, same level of precision, same source, same time period, etc.). However, because this is rarely the case, the weight of a particular dataset needs to compensate for the permutation of uncertainty through different steps of the data creation (Longley et al. 2005). This is a difficult component to assess because it is a function of each of the first-order components of uncertainty.

(4) Lineage: Lineage refers to the channel through which the data passed from collection to use. It appears that the second-order component lineage is a combination of the first-order components credibility and currency. Lineage realistically assumes that the data creation does not occur in a single instance in time, and because of such, can be spread over time among multiple data collectors, manipulators, and presenters. In floodplain mapping, lineage would represent the tracking and articulating of each step in the data's lifecycle, addressing both the time a particular step was completed (the currency of the step) and the firm that completed the step (credibility). Such a full recollection of the data creation process provides a more detailed account for decision making than the first-order components of currency and credibility alone.

(5) Subjectivity: When used as a first-order component of uncertainty, subjectivity involves any judgment that the client made until the completion of the map product. Subjectivity as a second-order component of uncertainty gets at the core of the inclusion of the "user's understanding" in the evaluation of uncertainty (Longley et al. 2005). Here, the decision maker uses personal experience to weight the importance or restriction of each of the first-order components. While examples of the first-order components accuracy/error and precision/resolution will likely lead to a uniform subjectivity response because they are internal to the dataset, other first-order components may lead to a more ambiguous weighting. For example, the data creator may have made the first-order subjective choice to map the floodplain based on the terrace contour, but the map reader may be critical of this decision, and in a second-order subjective choice, choose to place more weight on the floodplain linework generated from the historical record or a hypothetical discharge exceedance. While this component of uncertainty is clearly not internal to the dataset itself, it is included because the Agumya and Hunter (2002) model includes all aspects of the decision making process.

CHAPTER 4 – Methodology

To answer the research questions posed in the opening chapter, two separate stages of research were conducted. The first stage employed an online, interactive survey aimed at acquiring initial data. Because little empirical research has been published on typological differences of uncertainty, the first purpose of the survey was to determine if different categories of uncertainty in the MacEachren et al. (2005) typology have a differing effect on decision making. If there was a significant difference in the recordings for one or all of the categories, the survey then served the secondary purpose of matching uncertainty types with particular decision responses, decision speeds, and confidence levels. The survey was designed to obtain repeatable results with recorded data at the ratio level so that significant differences can be quantified.

The second stage of the research employed several group interviews designed much like a focus group. The qualitative results of the focus groups were used to supplement the quantitative survey data and help explain patterns in the survey results. Questioning in the focus groups was posed to help find which category of uncertainty was most influential in the decision making process. The results of the survey were then compared to the offerings in the interview session to see if the most influential categories of uncertainty on decision making were the same that yielded the most accurate and confident decisions. This analysis aimed to help mapmakers prioritize the different types of uncertainty for graphic representation when space on the map is limited. The focus group also looked deeper into the decision making process, asking *why* particular categories of uncertainty were more influential than others. Insight gained from this questioning aimed to go beyond the *outcomes* of decision-making and instead examined the *process* of decision making itself, an agenda Harrower (2004) called "…a promising and important new direction for research in GIScience." Finally, the focus groups investigated the validity of the MacEachren et al. (2005) typology.

4.1 Quantitative Online Survey

The survey was built using the Flash Macromedia software and proctored over the Internet. The major drawback to conducting research online is that the participants are no longer in a controlled environment, meaning that the participant can be interrupted midway, have his or her attention split on another website or activity, and be assisted by another person during the survey. However, like the Aerts et al. (2003a) study, the use of an online survey was justified because of the desire to question a large amount of domain experts, located across the United States, in a very short period of time. It was determined that evaluating domain experts in an uncontrolled environment would produce a more realistic description of expert decision making under uncertain conditions than evaluating easy-to-access university students in a controlled environment. Each subject of the online survey was asked individually to participate via email so that the survey response rate could be recorded. In this contact email, participants were asked to simulate a controlled environment while taking the survey and were reminded of this again during the introductory section of the survey itself. The survey results were stored anonymously in an external database, making all results confidential and untraceable back to the participant.

The design of the survey consisted of three components:

(1) the maps used in the survey,

(2) the legends describing symbology on the maps, and

(3) the survey questions associated with each map-legend pair.

4.1.1 The Map Component of the Survey

The maps used in the survey were based on the floodplain of the Willamette River in Albany, Oregon, located 70 miles south of Portland. The floodplain itself was modeled using floodplain data produced by Rich Catlin (2002) of the City of Albany Planning Division. The data was produced in evaluation of a 1996 flood event and delineated the floodplain in great detail. Although the attribute data was available for this area, categorical uncertainty information was incomplete. Because of this, the final uncertainty linework on the survey was hypothetical. Supplemental data also from the City of Albany was utilized to conform to the local terrain when making the hypothetical linework. This map is representative of the common technique for representation of the floodplain boundary. The floodplain line is depicted using a light blue fill overtop structural information of the area, with the river itself depicted in a darker blue. The edge of the floodplain boundary is shown as a hard line, with no graphic uncertainty representations included. Interestingly, this map provides a unique example of a floodplain map that includes uncertainty information as a legend statement in fine print along the bottom of the map page. Figure-14 shows this floodplain data in a 2002 map produced by Rich Catlin for display purposes.

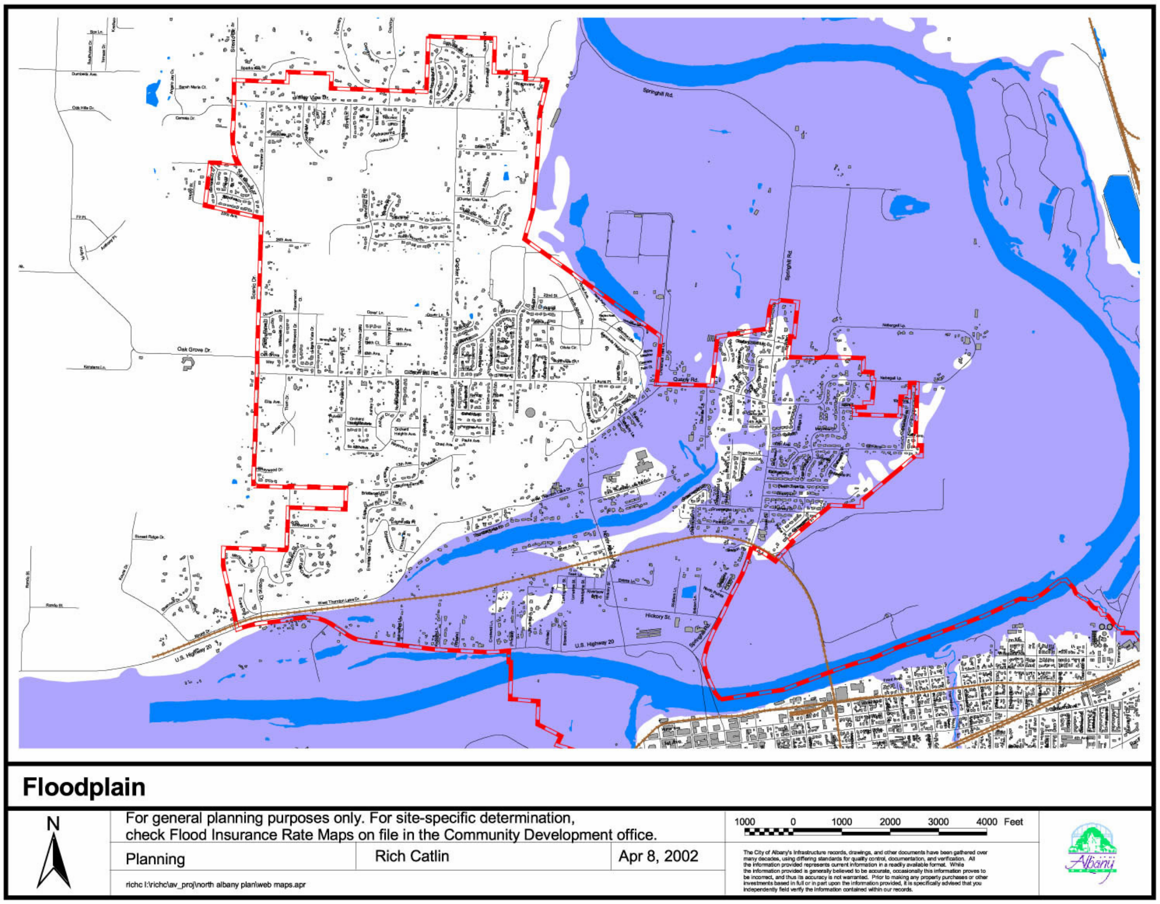

Figure-14: A map from Catlin (2002) showing the Willemette River floodplain at the city of Albany, Oregon.

On each map, three depictions of a single uncertainty type were represented in three different hues, chosen from a safe color scheme on ColorBrewer.org to ensure legibility of the lines for the color deficient. This representation of uncertainty is different that the multiple realization technique of the 'Monte Carlo' method in that the three variations signified three different datasets (each with its own description of uncertainty) rather than signifying three versions of the same dataset. No map contained information on multiple types of uncertainty. The three depictions were not animated and remained visible on the map throughout its entire usage. Each depicted variation of uncertainty was explained by an accompanying legend, following the Leitner and

Buttenfield (2000) study. The variations themselves are defined in Section 4.1.2. Three sites, represented by a black circle, were placed on the map relative to these depictions. For each type of uncertainty used in the survey, there were four accompanying versions of the map: three representing only one of the above sites individually and a fourth representing all three at one time. For the remainder of the document, the first three versions are referred to as ***one-site maps*** and the fourth version is referred to as a ***three-site map***. The sites were labeled using a 14pt black font encased in white so that they could be referenced by the survey questions.

The four map versions were designed by the investigator to exhibit a varying degree of difficulty. The three-site map was expected to be the most difficult for participants, due to the added complexity of a ranking task described by Leitner and Buttenfield (2002). The individual one-site maps were also designed to have a varying level of difficulty. The first two one-site versions were designed so that the represented site was within only a single depiction of the floodplain. The containing floodplain depiction was different for these two versions, with one version showing the site contained only by the most certain depiction (labeled site-C on the three-site maps) and the second version showing the site contained only by the least certain depiction (labeled site-A on the three-site maps). The third one-site version placed the site within two depictions of the floodplain boundary (labeled site-B on the three-site maps). It was expected by the investigator that the map versions would vary in difficulty in the following order, from easiest to most difficult: the site-B one-site map, the site-C one-site map, the site-A one-site map, and the three-site map. Also, because the same sites represented on the one-site maps were placed together on the three-site maps, the expected risk ranking on the three-site maps was B-C-A.

The background of the map was represented in varying shades of gray corresponding to the amount of overlap or agreement among the three depictions. Areas within the floodplain on all three lines were colored the darkest shade of gray and areas outside of the floodplain on all three lines were colored the lighted shade of gray. Areas inside the floodplain on one or two of the depictions were given intermediate shades of gray according to their place in the ordinal color scheme. Because the uncertainty dataset was hypothetical, this sequential grayscale color scheme was designed to suggest only a nominal degree of certainty; no ratio level data on certainty was represented. The river and surrounding oxbow lakes were represented in black. The reasoning for using black rather than the more traditional blue was to suggest that the river was the highest category in the grayscale representation of floodplain certainty. The river linework was drawn directly from the Catlin (2002) data source. Figure-15 shows an example map from the survey:

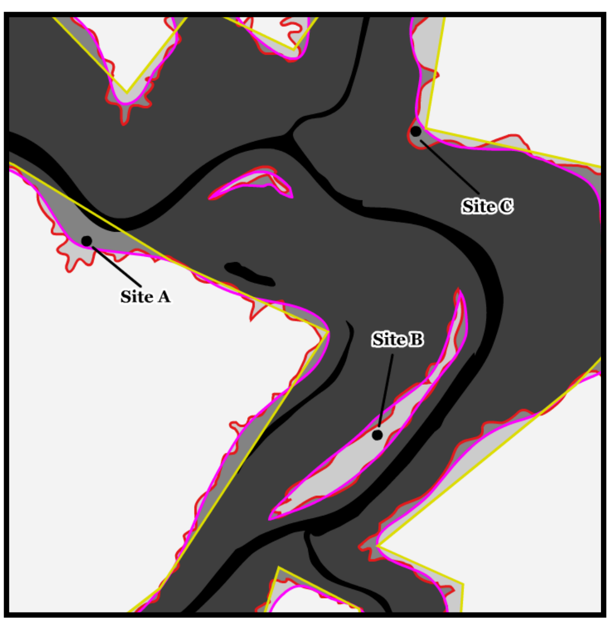

Figure-15: An example three-site map, taken from the online survey, showing three datasets with a varying level of detail (precision/resolution)

Several controls were implemented in the map component to avoid the use of previous knowledge of the area and to discourage learning throughout the course of the survey. To disguise the geographic location of the dataset, all signifiers were removed from the map, including information on roads, houses, geographic markers, and labels. The lesser known Willamette River floodplain at Albany, Oregon was chosen over the Mississippi River floodplain, described in the opening of Chapter 3, to decrease the likelihood that subjects would have prior knowledge of the area before participation. The use of hypothetical uncertainty information further reduced the influence of prior participant knowledge, as it was ensured that the participants have never seen the uncertainty dataset upon which they were making decisions. To control learning during the survey, each newly loaded map was randomly rotated around its center. The order of maps was also randomized for each participant, making it unlikely that two participants saw the same order of maps. The randomization of rotation and order were learning controls also implemented in the Leitner and Buttenfield (2000) study.

4.1.2 The Legend Component of the Survey

The second component of the survey, the legend, described to the participant what is being represented and was therefore the focal way in which typological uncertainty was communicated to the participants. The upper portion of the legend described the type of uncertainty represented. Each description is associated with a variation of the uncertainty type that is represented in the map.

Selection of the appropriate types to be included in the study and their associated variations was complicated by two issues:

(1) the division between first-order and second-order types of uncertainty

(2) the level of measurement of each depicted uncertainty type (nominal, interval, and ratio)

The explicit division between first-order and second-order uncertainty types, explained in Section 3.3, was necessary due to the lack of mutually exclusive and collectively exhaustive categorization in the MacEachren et al. (2005) typology. It was important to only test uncertainty types that were not contingent upon another type of uncertainty (first-order components) for two reasons. The first was to maintain a constant level of map complexity. Depictions of second-order components required the synthesis of several first-order components, creating a much more complex map, and therefore, a more complex legend. Such variations in complexity would create a varying level of decision difficulty, introducing a confounding variable. The second reason for displaying only first-order components was to maintain independent random sampling during statistical analysis. The testing of second-order components introduced conditional dependency, complicating the comparison of groupings with common statistical techniques like analysis of variance (ANOVA). Because of these two reasons, only first-order uncertainty types from the MacEachren et al. (2005) typology were tested (accuracy/error, completeness, credibility, currency, precision/resolution, and subjectivity).

The second issue, level of measurement, addressed unit scale associated with each variation. As described in Section 3.3, several of the first-order uncertainty types are commonly measured at the ratio level. Because other first-order uncertainty types, like credibility and subjectivity, are intrinsically categorical to a dataset and extrinsically ordinal to the decision maker, all of the types of data needed to be degraded to the nominal level to maintain a constant level of map complexity. The change in level of measurement can be expressed in terms of the Beard and Mackaness (1993) 'data-quality assement task' typology as a move from quanitification to identification of uncertainty. Such degradation was especially difficult with the category accuracy/error, as it is typically reported probabilistically at the ratio level. Because of this, the proxy of the flood frequency was used; the investigator acknowledges that this solution was not a representation of uncertainty per se, but would yield more interesting results than the ordinal descriptors of 'highly accurate', 'moderately accurate', and 'not accurate'. Is it important to note that any statistical findings associated with the category accuracy/error would need to remain speculative due to the use of the proxy. Table-1 describes the six legends for the six different first-order uncertainty types and the corresponding variations represented in the map:

Uncertainty Type	***Variation #1***	***Variation #2***	***Variation #3***
Accuracy/Error Proxy	0.2% annual chance of flood	1% annual chance of flood	5% annual chance of flood
Completeness	Data/No Data Line	Floodplain/Non-Floodplain Line	<unavailable>
Credibility (Source)	Local Source	Statewide Source	Federal Source
Currency	Data collected in 2005	Data collected in 1995	Data collected in 1985
Precision/Resolution	High Detail	Intermediate Detail	Low Detail
Subjectivty	Floodplain defined by the inner-most terrace	Floodplain defined by the historical flood record	Floodplain defined by a simulated discharge level

Table-1: The uncertainty descriptions used in the legend component of the digital survey. Each variation was associated with one of the depictions shown on the map by labeling a particular hue in the legend. It is important to reiterate that these variations are based off of a definition of uncertainty that includes the user's interpretation of the dataset.

Because understanding the legend is vital to understanding the map, and therefore correctly assessing the risk of a structure, several precautions were taken to ensure that the legend was read with every map-legend pair. Firstly, before the participants were given the survey questions, they were required to complete a brief training session. In the training session, the importance of the legend was communicated and the participants were warned to first analyze the newly loaded legend before examining the map. Secondly, as a result of a pilot survey, the legend was placed at the top, left corner of the screen. Eye-movements studies have shown that Western readers tend to process a page from the top, left corner to the bottom, right corner, similar to the way text is read in a book (Slocum et al. 2003). Placing the legend at the top, left corner put the legend in the highest priority location on the screen and ensured that the participant did not need to scroll the browser to view the legend. Thirdly, a loading screen was programmed to darken the screen after the completion of questioning for each map to signal that a new map was being loaded. During this time, the previous map and legend were removed, providing a visual cue that a new legend was being loaded onscreen. Finally, the questions panel on the survey was disabled until the legend was fully loaded and visible. This prevented responding to a particular map before the new legend was loaded and visible. Also described in the legend component of the survey were the varying shades of gray for the depiction overlap, the river, and the circular symbol for the sites. Figure-16 shows the matching legend for the precision/resolution map shown in Figure-15.

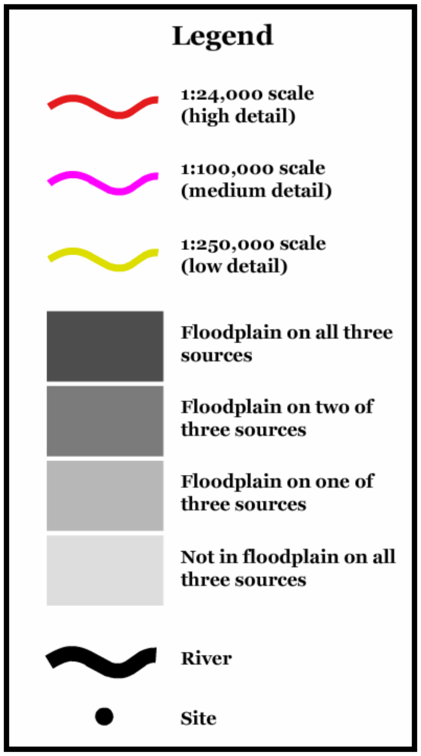

Figure-16: The matching legend component to the precision/resolution map shown in Figure-15. Each legend describes the three depictions of the floodplain, the grayscale basemap shading, the river, and the site symbol.

4.1.3 The Question Component of the Survey

The final component of the survey was the questioning. Like the Aerts et al. (2003a) study, the survey began with a short 'background' section to determine the level of expertise of the participant. The research focused on two specific 'expert' groups: (1) GIScientists who prepare and reads maps daily for decision makers and (2) domain experts who specifically use floodplain maps for decision making. The background survey presented two triplets of questions aimed at determining the expertise level in each of the two expert groups. The first question in each triplet surveyed the degree of education or training in the topic ('Have you taken a course of map design or map use?' or 'Have you taken a course in flood events, hydrology, or risk management?'), the second question in each triplet surveyed the degree of work experience in the topic (Does your current or previous job require you to design or use maps in anyway?' or 'Does your current or previous job require you know about flood events, hydrology, or risk management?'), and the final question in each triplet surveyed the self-reported expertise level in the topic at the ordinal level (novice, intermediate, and expert). The term ***expertise level*** was defined as the response to these initial questions and these tabulations were used to determine if there were any significant differences in performance across expertise level in education, work experience, and self-reporting.

After completion of the background survey and the training session, the series of related maps and legends were loaded in a randomized order. For each map-legend pair, the participants were required to answer three questions. The first question varied based on the number of sites included in the map. When one-site maps were displayed, the first question required the participant to assess the flood risk of the site on a scale of 1-5 ('1' being safely located and '5' being insecurely located). When three-site maps were displayed, the first question required the participant to rank the three locations based on their relative flood risk. The term ***decision*** or ***decision task*** was defined as the response to this first survey question. Because of the nature of uncertainty, there is no 'correct' response for this first question, and instead all results were compared to the average response.

The second and third questions for each map-pair were follow-up questions for the initial risk assessment. The second question asked the participant to rank the difficulty of the risk assessment on a scale of 1-5 ('1' being an easy siting decision and '5' being a difficult siting decision). The term ***decision difficulty*** or ***perceived decision difficulty*** was defined as the response to this second question. Although the four versions of each uncertainty map were intended to create a varying degree of decision difficulty throughout the survey much like the Leitner and Buttenfield (2000) study, this second question allowed for the recording of a perceived decision difficulty, independent of the investigator-defined difficulty grouping. The third question asked the participant to rank his or her confidence in the risk assessment on a scale of 1-5 ('1' being least confident that the decision was correct and '5' being most confident that the decision was correct). The term ***confidence*** was defined as the response to this third question.

Tallying each question triplet for each map-legend pair, along with the six initial background questions, leads to a total of seventy recorded user-input questions. However, the time that the participant took to answer each question in the main survey was also recorded. The term ***decision speed*** was defined as the time (in seconds) that the user took to answer an individual question. For the first question in the triplet, the time was defined as the number of seconds from when the map completed loading to the selection of an answer. For the second and third question in the triplet, the time was defined as the number of seconds between the response input of the previous question and the response input of the current question. It is acknowledged that the recording of time in an uncontrolled environment is highly problematic, and it will be suggested that any significant findings in time be verified under a controlled laboratory experiment. The final recorded variable from the online survey is the order in which the map-legend pairs were shown to the participant, pushing the total recorded variables to 135 (six background survey, sixty-four user-input in the main survey, sixty-four time-based in the main survey, and the order of the map-legend pairs).

4.2 Qualitative Focus Groups

The second section of the research, the focus groups, was used to complement the digital survey, attempting to move beyond the *outcomes* of the decision making process and attempt to learn something about the *process* itself. The focus groups were organized to be composed of 2-4 subjects and to take one full hour. Unlike the digital survey, participation in the focus groups was restricted to domain experts in floodplain mapping. Reasoning for the exclusive use of domain experts was three-fold. Firstly, the years of experience in floodplain mapping helped the domain expert comment on how geospatial data uncertainty enters into floodplain mapping, and further, which type of geospatial uncertainty was most influential in the decision making process and why. Secondly, interviewing domain experts allowed for the teasing out of any disconnects between suggestions in the academic literature and actual conduct by practitioners. Finally, each subject was compensated $20 for their participation. Because the funding backing this section of the research was limited, it was deemed more cost-effective to limit the participation to experts.

4.2.1 Interview Organization

The interview began with the completion of a consent form providing a brief explanation of the study and outlining the rights of the research participant. In the consent form, the participants were informed that the research was authorized by both the University of Wisconsin-Madison, that they could refrain from answering any question, and that the interviews would be voice recorded with digital devices. Participants were also asked in the consent form to confirm that they have completed the online digital survey as a precursor to the interviews. Finally, the same biographical survey described in Section 4.1.3 was circulated to ensure that those tested were experts.

Questioning during the interview was kept informal and open-ended, following Eyles' (1988, as cited in Valentine 1997, p111) idea that an interview should be a "conversation with a purpose," rather than a rapid-fire series of predetermined questions. Further, all questions were posed with the goal of neutrality so to not lead the participants, with any questions concerning the research angle or the researchers' opinions deferred to the final ten minutes of the interview session. The following are the three primary themes that were covered during the interviews. Again, the ordering of these questions was fluid and, because of the open structure, many of these prewritten questions were not discussed in all of the sessions.

(1) Overview of Uncertainty in Floodplain Mapping: Each session opened by first discussing uncertainty generically to get an overview on the subjects' understanding of the topic. Each participant was asked to define uncertainty in geospatial data and provide real-world examples. It was hoped that having everyone provide a unique definition would emphasize the idea that uncertainty can take many forms, although no reference to a typology was made by the investigator. Following the overview, the discussion was turned to how uncertainty enters into the specific domain of floodplain mapping. Example questions posed here include: 'How has uncertainty entered into floodplain mapping in your experience?', 'Has uncertainty changed the way you collect floodplain data or make floodplain maps?', and 'How do you represent uncertainty in your floodplain maps?' The purpose of these questions was to provide initial insight into how domain experts deal with uncertainty in floodplain maps as well as to have the subjects volunteer different types of uncertainty before the MacEachren et al. (2005) typology was addressed. During this opening section of questioning, the investigator did not prompt the participants with a definition of the term uncertainty or suggestions of different types of uncertainty.

(2) Discussion of the MacEachren et al. (2005) typology: At approximately twenty minutes into the interview, a sheet explaining the MacEachren et al. (2005) typology (as shown in the conclusion of Chapter 2) was circulated and the participants were given several minutes to examine it. The only instruction provided while circulating the typology sheet was to try to think of examples of each of the types in the domain of floodplain mapping. Once the reading of the list was completed, the participants were asked to provide real-world examples of each type, both hypothetical and observed. Participants were not discouraged from offering multiple examples of a single type or from offering examples from a type that was already discussed. After this brainstorming, the subjects were asked if the typology reflects the real-world. Example questions posed here include: 'Do any of these types not apply to the uncertainty that is found in floodplain mapping?', 'Are there any types of uncertainty in floodplain mapping that is not included in this list?', and 'Does knowing about this list change the way you would work with uncertainty in floodplain mapping?' Throughout the second section of questioning, the investigator only answered questions that asked for clarification of a particular uncertainty type definition; no examples were provided.

(3) Determining the Influence of Each Type of Uncertainty: The final section of the interview aimed at understanding which uncertainty types were relied upon during decision making and why these types were used with such importance. Several variations of the same question were asked during this session, including: ‘When making a decision off of a floodplain map, which uncertainty type(s) would you weight most heavily?’, ‘When making a decision off of a floodplain map, which uncertainty type(s) do you think is (are) the most important to represent?’, and ‘If you were to be informed that uncertainty exists in a particular type(s), which informed type(s) would cause you to approach a decision task with the most caution?’ While posing the same question in different ways was used to stimulate discussion, responses to any of these questions were used to define the term ***influence***. It is important to note that for each question posed, participants were also asked to list the least influential type(s). Finally, the participants were asked if the most influential type would change based on a particular situation or context, as well as if there would be a change in influence when the consequences of an incorrect decision increased or decreased. Subjects were then asked to explain this shift with real-world examples. Although the idea of decision difficulty was never discussed, this line of question was used to see if the term ‘influence’ varies according to the difficulty of the decision.

The final five to ten minutes of the interview were reserved to allow the participants to ask questions about the research theory and methodology. During this time, data CDs containing important literature on geospatial data uncertainty and documents describing the research in more depth were distributed as intellectual compensation in addition to the monetary compensation provided. It is not until this time that the primary research questions posed in the opening chapter were revealed, as well as any initial findings from the online survey or preliminary hypotheses generated by the investigator. The interview sessions were concluded precisely sixty minutes after their initiation.

CHAPTER 5 – Results and Discussion

5.1 Results of the Quantitative Online Survey

A link to the online quantitative survey was emailed to a total of 135 potential participants and was completed by 56 participants, producing a surprisingly high survey response rate of 41.5%. Among those asked were university faculty and graduate students studying GIScience and fluvial processes, as well as private, state, and federal level professionals working in the GISciences and floodplain mapping. Due to the anonymous nature of the survey, it is uncertain how well each of these groups was represented. The survey was left online for a total of four weeks from mid-February 2007 thru mid-March 2007.

As explained in Section 4.1.1, there were four different map versions for each type of uncertainty tested. The first three versions represented only a single site on the map and required the participant to assess flood risk of the portrayed site. Conversely, the fourth version portrayed all three of the sites from the first three versions together on a single display and required the participant to rank the sites based on their relative risk of flooding. Because of the increased task complexity of the fourth version from the first three, summary statistics for the terms decision task, decision speed, decision difficulty, and confidence were calculated separately for one-site and three-site maps in Sections 5.1.1 and 5.1.2. It is not until the decision difficulty was taken into account in Section 5.1.3 that the results from the two groups were pooled.

As explained in Section 4.1.3, assessment questions concerning the three-site maps required the participant to rank the sites (consistently labeled ‘Site A’, ‘Site B’, and

'Site C') on an ordinal scale from low risk to high risk. To represent the decision task responses for three-site versions in a tabular fashion, the response was divided into three variables, termed ***site-A ranking**, **site-B ranking**,* and ***site-C ranking*** relating to the label given to the site on the map. These rankings were converted to numbers, similar to the assessment questions for the one-site maps, so that the site with the lowest risk was given a '1', the site with an intermediate degree of risk was given a '2', and the site with the highest risk was given a '3'. The answer deemed correct by the investigator prior to the proctoring of the survey was B-C-A (from lowest risk to highest risk) on all six maps, although it was not assumed that this ranking would be reflected in the collected data. Because of this, the expected site-A ranking was a value of '3', the expected site-B ranking was a value of '1', and expected site-C ranking was a value of '2'. Finally, it is important to note that the decision speed associated with the three-site versions still aggregated the time taken to rank each of the three sites into a single value, as the time spent for an individual ranking was dependent on the order in which the sites were ranked.

After each participant selected their answer for decision task, decision difficulty, and confidence, an accompanying speed was also recorded, representing the amount of time, in seconds, that the user needed to complete the survey question. As described in section 4.1.3, recording speed in an uncontrolled environment is problematic, and was the most significant drawback to disseminating the survey online. As expected, two participants exhibited a single, dramatic increase (>10x the average response time) in the recorded speed sometime during the survey. This spike in the speed recordings likely represented a period when the participant shifted attention away from survey (e.g. phone call, browsing another website, etc.). Both of these speed spikes were associated with questions concerning the decision task variable while the one-site map versions were displayed; speed recordings for the decision difficulty and confidence responses did not exhibit any unexpected spikes nor did the decision task responses associated with the three-site map versions. As of such, summary statistics and hypothesis tests were conducted for an adjusted decision task speed variable that removed the two spikes, decreasing the sample size from 56 to 54 for all adjusted speed statistics. The speed recordings for the responses to the decision difficulty and confidence questions were not included in the Section 5.1 analysis because they were highly varying and statistically insignificant. Because of this, the term ***decision speed*** was redefined to describe only the amount of time a participant needed to answer the initial decision task question. Finally, all statistics provided in the following sections on decision speed are in the units of seconds.

5.1.1 Examining the Results across Uncertainty Type

Table-2 provides the descriptive statistics for the decision task, decision speed, decision difficulty, and confidence variables for each of the six first-order types of uncertainty displayed in the survey. Provided in Table-2 are individual calculations of mean and standard deviation for each uncertainty type and an overall calculation independent of uncertainty type. The summary statistics were calculated separately for responses to one-site and three-site maps.

	Variable	Acc/Err Proxy		Completeness		Cred./Source		Currency		Prec/Res		Subjectivity		Overall	
		avg	std	avg	std	avg	std	avg	std	avg	std	avg	std	avg	std
One-Site	Decision Task	3.065	1.027	3.232	1.257	3.256	1.027	3.321	1.128	3.173	1.163	3.339	0.110	3.241	1.121
	Decision Speed	26.218	20.693	68.235	628.761	25.688	23.161	42.607	216.027	24.115	18.518	27.285	18.756	35.858	271.733
	Adjusted Decision Speed	26.272	20.971	21.006	13.437	25.453	22.809	25.731	21.101	23.904	18.691	26.976	18.568	24.890	19.543
	Decision Difficulty	2.089	0.867	2.571	1.329	2.280	0.978	2.220	0.975	2.232	0.954	2.387	0.941	2.297	1.027
	Confidence	3.738	0.937	3.214	1.350	3.595	0.98	3.619	1.002	3.732	0.994	3.613	0.935	3.585	1.055
Three-Site	Site-A Ranking	2.679	0.636	2.339	0.769	2.625	0.620	2.661	0.581	2.536	0.660	2.554	0.570	2.565	0.648
	Site-B Ranking	1.214	0.530	1.375	0.676	1.161	0.458	1.304	0.658	1.124	0.494	1.089	0.394	1.226	0.549
	Site-C Ranking	2.107	0.493	2.143	0.672	2.214	0.530	2.036	0.571	2.214	0.624	2.357	0.520	2.179	0.576
	Decision Speed	44.790	30.141	32.605	24.306	44.979	33.693	42.359	35.022	40.793	30.335	54.862	38.329	43.398	32.706
	Decision Difficulty	2.357	1.086	2.821	1.377	2.839	1.125	2.732	1.07	2.589	1.075	2.911	1.133	2.708	1.156
	Confidence	3.482	1.191	2.893	1.289	3.250	1.179	3.393	1.123	3.500	1.095	3.357	1.086	3.313	1.172

Table-2: The summary statistics for the decision task, decision speed, decision difficulty, and confidence variables, separated across uncertainty type. Shown in the table are the mean ('avg') and standard deviation ('std'). The values for decision task, decision difficulty, and confidence are reported on a scale of '1' to '5' and the values for decision speed and adjusted decision speed are in seconds.

When running individual statistical Z-tests, no single map displayed significance in decision task, decision difficulty or confidence. This result is logical, as a Z-test relies heavily on the variance of the population. Because the entire population contained responses to questions of varying difficulty, the variation of the responses when grouped solely by uncertainty type was expected to be quite large. For example, the spread of the data for the one-site decision task variable produced a variance so large (1.121) that finding statistical significance was impossible (with a 95% confidence interval of 0.999-5.483 when the range of answers is only from 1-5). Rather than comparing each grouping against the entire population, each variable instead was tested using an analysis of variance (ANOVA) hypothesis test. The ANOVA statistical test required the acceptance of two primary assumptions (Aczel and Sounderpandian 2006):

(1) the populations were normally distributed and randomly sampled

(2) although the means of each population may have been different, the variances of each population were equal

Figures in Table-2 are provided mostly for summary purposes, but also to validate the assumption of homogeneity of variance. It was also assumed that the data conforms to a normal distribution due to the large sample size.

A one-way ANOVA hypothesis test, using uncertainty type as the grouping factor, was conducted for each of the eleven variables summarized in Table-2. For each test, the null hypothesis was that there was no difference among the six uncertainty groupings[3]. Table-3 summarizes the results of these ANOVA hypothesis tests. Findings that are significant at alpha = 0.05 are bolded.

	Variable	***Overall***	
		F-stat	p-value
One-Site	Decision Task	1.801	0.110
	Decision Speed	0.712	0.614
	Adjusted Decision Speed	1.990	0.078
	Decision Difficulty	**4.436**	**0.001**
	Confidence	**5.703**	**0.000**
Three-Site	Site-A Ranking	2.105	0.064
	Site-B Ranking	1.942	0.087
	Site-C Ranking	2.098	0.065
	Decision Speed	**2.798**	**0.017**
	Decision Difficulty	1.773	0.118
	Confidence	2.091	0.066

Table-3: A set of one-way analysis of variance (ANOVA) hypothesis tests using uncertainty type as the grouping factor. Findings that are significant at alpha = 0.05 are bolded.

[3] H_0: $\mu_{accuracy} = \mu_{completeness} = \mu_{credibility} = \mu_{currency} = \mu_{precision} = \mu_{subjectivity}$

Of the eleven ANOVA tests conducted, three were found to be significant at alpha = 0.05. It is important to note that an added five tests were significant at alpha = 0.10, and two more tests were almost significant at this alpha level. Only the unadjusted decision speed on one-site maps, included in this initial test to contrast the adjusted decision speed, was not near rejection of the null hypothesis. These hypothesis tests provided initial evidence that for simple decisions, the type of uncertainty represented significantly influenced the map reader's perceptual difficulty in making a decision (*p-value of 0.001*) and their personal confidence in the decision that was made (*p-value of 0.000*). Further, the analysis suggested that the type of uncertainty that was represented perhaps influenced the decision outcome (*p-value of 0.110*) and the time it took to arrive at this outcome (*p-value of 0.078*), although more testing is required to see if this connection is statistically significant. For more difficult tasks, the analysis suggested that uncertainty type significantly impacted the amount of time it took to make a decision (*p-value of 0.017*) and was perhaps influential on the decision outcome (*p-value of 0.064, 0.087, and 0.065 for the three site rankings*), the perceptual difficulty of making the decision (*p-value of 0.118*), and the confidence in the decision (*p-value of 0.066*). These initial findings were very suggestive that the type of uncertainty that was represented played an important role in the decision making process.

Following the ANOVA hypothesis testing, Tukey Pairwise-Comparison tests were conducted to see which of the groups from the ANOVA testing caused the rejection of the null. The null hypothesis is rejected as long as a single tested group was significantly different from any other group in the ANOVA test. The Tukey Pairwise-Comparison test examines each possible pairing of groups tested during ANOVA using a single, “family” level of significance (Aczel and Sounderpandian 2006). For each Tukey comparison, the null hypothesis was that there was no difference between the two groups at hand[4]. The Tukey test was a more appropriate way to pinpoint the significant difference than the use of multiple Z- or Student-T tests. Table-4 provides a listing of the significantly different pairings from the eleven ANOVA hypothesis tests:

	Variable	***Overall***	
		significant at alpha = 0.05	significant at alpha = 0.01
One-Site	Decision Task	none	none
	Decision Speed	none	none
	Adjusted Decision Speed	none	none
	Decision Difficulty	**completeness vs. currency** **completeness vs. precision/resolution**	**acc./err. proxy vs. completeness**
	Confidence	**completeness vs. credibility/source**	**acc./err. proxy vs. completeness** **completeness vs. credibility/source** **completeness vs. precision/resolution** **completeness vs. subjectivity**
Three-Site	Site-A Ranking	none	none
	Site-B Ranking	none	none
	Site-C Ranking	**currency vs. subjectivity**	none
	Decision Speed	**currency vs. subjectivity**	**completeness vs. subjectivity**
	Decision Difficulty	none	none
	Confidence	none	none

Table-4: A listing of the significantly different group pairings underlying the ANOVA results in Table-3. The data was produced using the Tukey Pairwise-Comparison test. Any pairing not showing on the table was found to not be significant at alpha = 0.05.

Results from the Tukey test uncovered two interesting occurrences in the survey data. For the one-site responses, the completeness group differed strongly with almost all of the other groups for both decision difficulty and confidence. When revisiting the summary statistics, completeness had a higher average decision difficulty and a smaller average confidence value than the other five uncertainty types. This evidence suggested

[4] H_0: $\mu_A = \mu_B$

that for simple decisions, the representation of completeness increased the difficulty of making a decision and decreased the confidence in this decision compared to the representation of other uncertainty types. When evaluating the summary statistics for the three-site responses, the completeness group also had one of the highest average decision difficulties and had the lowest average reported confidence, although these levels were not found to be statistically significant at alpha = 0.05. Because the null was accepted for decision task and decision speed on one-site maps, the Tukey analysis did not report any significantly different grouping in these variables at alpha = 0.05.

For the three-site responses, there was conflict between currency and subjectivity, showing significance in both the site-C ranking and the decision speed variables. When examining the summary statistics, the participants were able to rank 'Site C' most effectively when currency was represented (as the average is closest to a '2') and least effectively using subjectivity (where the average is the furthest from a '2'). Similarly, subjectivity showed the highest average value for decision speed with currency providing the second lowest. This evidence suggested that subjectivity decreased the ability to make the correct decision outcome for complex tasks, while the representation of currency increased this ability. Paralleling this, subjectivity increased the amount of time needed to arrive at a decision, while currency decreased the amount of time necessary to make a decision. These findings, however, do not transfer well to the one-site responses.

Perhaps the most interesting result from the Tukey test was the high degree of division in decision speed between the completeness and subjectivity groupings. Although the completeness group appeared to influence decision difficulty and confidence negatively, the participants were able to respond closer to the expected result in a quicker fashion when completeness was represented. Although the highly significant finding in speed may be a product of the recording of speed in an uncontrolled environment, the investigator theorizes that this result may be caused by the ability of the participants to degrade completeness down to a binary choice (either present or absence) when making a decision, while the other uncertainty types remain, at least to a degree, ratio level. This hypothesis of creating binaries or thresholding is discussed in more detail in Section 5.2.3.

5.1.2 Examining the Results across Expertise level

Following the initial look at the survey results, it was analyzed across the expertise level information collected in the background survey. As discussed in Section 4.1.3, the survey allowed for the definition of expertise level in three ways:

(1) amount of education/training

(2) amount of work experience

(3) self-reported

The data was evaluated across all three definitions of expertise level for both GIScience experts and domain experts, producing six categories of expertise. Table-5 summarizes the frequency tallies of expert versus novice in each of the six categories. A major issue discovered before analysis of the data was the under-sampling of GIScience novices with a more uniform distribution along domain expertise. The even split of domain expertise suggested that many of the domain experts also listed themselves as GIScience experts. The investigator expected the two types of expertise to be, to a degree, mutually exclusive. Due to the lack of novice sampling against which to compare the experts, all findings about the relationship between GIScience expertise level and decision making must remain speculative.

	Expertise Category	Expert	Intermediate	Novice
GIScience	Education/Training	47	n/a	9
	Work Experience	42	n/a	14
	Self-Reporting	40	14	2
Domain	Education/Training	26	n/a	30
	Work Experience	21	n/a	35
	Self-Reporting	10	34	12

Table-5: Survey participation between expert and novice. This study adopted three definitions of *expertise level* (education/training, work experience, and self-reporting), and identified two different types of relevant expertise (GIScience and domain).

Similar to Section 5.1.1, the results were split between one-site and three-site maps and ten variables were examined (all variables from Section 5.1.1 excepting the unadjusted decision speed). Descriptive statistics for one-site maps were not included due to length.

Several amendments to the ANOVA approach described in the previous section were required when examining the survey responses across the factor of expertise level. Both the interaction of expertise level on the entire population and the influence of expertise level on individual uncertainty types were of interest. Ideally, a two-way ANOVA hypothesis test would have been conducted to efficiently answer both questions. The two-way ANOVA approach required that each cell in the factorA-by-factorB matrix contained an equal sample size, but the sampling distributions among different groupings of expertise were not uniform (Aczel and Sounderpandian 2006). A multiple regression hypothesis test is typically conducted when cell sample size varies, but could not be implemented in this situation due to the qualitative nature of the expertise level variable. As of such, expertise level and uncertainty type could not be examined jointly in a single test, requiring the completion of a separate hypothesis tests based on the expertise level division for each type of uncertainty, as well as a seventh hypothesis test on the expertise level division for the entire dataset. Employing such a large number of hypothesis tests increased the probability of rejecting the null when it should be accepted by increasing the number of Bernoulli trials, and therefore increasing the likelihood of making false claims of significance. This was a second reason for making any findings speculative. Due to these issues, discussion was limited to when the data was evaluated wholly in the seventh version of testing.

Hypothesis testing of the expertise level utilized two different types of tests: difference of means tests for expertise level definitions that contained only two categories (education/training and work experience) and ANOVA tests for expertise definitions that contained three categories (self-reported). Table-6 and Table-7 describe the difference of means and ANOVA hypothesis tests respectively. For the difference of means tests, the null hypothesis was that there was no difference between expert and novice[5], and for the ANOVA tests, the null hypothesis was that there was no difference among the three self-reported expertise categories[6]. It is important to note that the expert group was treated as the lead in each hypothesis test.

---

[5] H_0: $\mu_{expert} = \mu_{novice}$

[6] H_0: $\mu_{expert} = \mu_{intermediate} = \mu_{novice}$

			Variable	Overall	
				t-stat	p-value
One-Site	GIScience Expertise	Education	**Decision Task**	**1.995**	**0.046**
			Adjusted Decision Speed	0.797	0.426
			Decision Difficulty	**-4.032**	**0.000**
			Confidence	**6.453**	**0.000**
		Work	Decision Task	1.216	0.224
			Adjusted Decision Speed	0.923	0.356
			Decision Difficulty	-0.146	0.884
			Confidence	**5.707**	**0.000**
	Domain Expertise	Education	**Decision Task**	**3.064**	**0.002**
			Adjusted Decision Speed	0.946	0.344
			Decision Difficulty	1.303	0.193
			Confidence	**2.888**	**0.004**
		Work	**Decision Task**	**3.316**	**0.001**
			Adjusted Decision Speed	**2.200**	**0.028**
			Decision Difficulty	**-2.102**	**0.036**
			Confidence	**5.942**	**0.000**
Three-Site	GIScience Expertise	Education	Site-A Ranking	1.502	0.134
			Site-B Ranking	-1.025	0.306
			Site-C Ranking	-0.607	0.544
			Decision Speed	-0.820	0.413
			Decision Difficulty	**-3.631**	**0.000**
			Confidence	**4.272**	**0.000**
		Work	Site-A Ranking	0.486	0.627
			Site-B Ranking	0.459	0.647
			Site-C Ranking	-1.093	0.275
			Decision Speed	-0.003	0.997
			Decision Difficulty	**-3.687**	**0.000**
			Confidence	**3.525**	**0.000**
	Domain Expertise	Education	Site-A Ranking	0.470	0.639
			Site-B Ranking	1.340	0.181
			Site-C Ranking	**-2.071**	**0.039**
			Decision Speed	0.978	0.329
			Decision Difficulty	0.425	0.671
			Confidence	0.863	0.389
		Work	Site-A Ranking	1.000	0.318
			Site-B Ranking	**2.169**	**0.031**
			Site-C Ranking	-1.272	0.204
			Decision Speed	1.482	0.139
			Decision Difficulty	**-2.381**	**0.018**
			Confidence	**3.280**	**0.001**

Table-6: Difference of means hypothesis testing based on the expertise definition of education/training or work experience. Participants could only select ‘expert’ or ‘novice’.

		Variable	Overall	
			F-stat	p-value
One-Site	GIScience	Decision Task	0.863	0.422
		Adjusted Decision Speed	1.435	0.239
		Decision Difficulty	**11.470**	**0.000**
		Confidence	**20.858**	**0.000**
	Domain	**Decision Task**	**5.159**	**0.006**
		Adjusted Decision Speed	**5.976**	**0.003**
		Decision Difficulty	**9.292**	**0.000**
		Confidence	**20.088**	**0.000**
Three-Site	GIScience	**Site-A Ranking**	**3.543**	**0.030**
		Site-B Ranking	1.531	0.218
		Site-C Ranking	2.557	0.079
		Decision Speed	2.830	0.060
		Decision Difficulty	**5.835**	**0.003**
		Confidence	**4.687**	**0.010**
	Domain	**Site-A Ranking**	**4.195**	**0.016**
		Site-B Ranking	2.070	0.128
		Site-C Ranking	1.198	0.303
		Decision Speed	**8.005**	**0.000**
		Decision Difficulty	2.625	0.074
		Confidence	1.196	0.304

Table-7: ANOVA hypothesis testing results using the expertise definition of self-reporting as the grouping factor. For this background question, the participant was able to rank themselves as 'expert', 'intermediate', or 'novice'. Findings that are significant at alpha = 0.05 have been bolded.

Results of the ANOVA and difference of means hypothesis tests showed a significant difference along expertise level groupings for both decision difficulty and confidence. Of the twelve hypothesis tests conducted for the variable decision difficulty in Table-6 and Table-7 (one for each of the six definitions of expertise level, performed for both the one-site and three-site maps), eight of the tests returned significance at alpha = 0.05. Of particular interest were the associated p-values with the decision difficulty tests that displayed significance. Six of the hypothesis tests returned a p-value approaching '0', suggesting an infinitesimal chance of rejecting the null hypothesis when it should be rejected. The type of expertise level definition used or the complexity of the task had an inconclusive influence on decision difficulty, as no pattern was uncovered concerning the four hypothesis tests that did accepted the null hypotheses.

The difference of means and ANOVA hypothesis tests provided an even larger amount of evidence connecting expertise level to the variable confidence. Of the twelve hypothesis tests conducted on confidence (one for each of the six definitions of expertise level, performed for both the one-site and three-site maps), ten returned a significance at alpha = 0.05, expanding upon the eight significant tests found for the variable decision difficulty. All ten of these tests reported statistical significance at alpha = 0.01, with most of the p-values approaching '0'. The two tests that did not reject the null hypothesis occurred when using domain expertise on the three-site maps, a result that is also found in the decision difficulty hypothesis tests.

When re-examining the summary statistics, the average reported decision difficulty for nearly every single definition of experience was lower than its novice or intermediate counterparts and the average reported confidence for all definitions of expertise level was higher than its novice or intermediate counterparts. In summary, the analysis suggested that an increase in expertise level significantly decreased the perceived difficulty of making a decision and significantly increased the confidence in the decision that was made. These findings are logical, as it can be inferred that a person with experience assessing flood risk would have an easier time arriving at a decision and would be more confident that the decision he or she made was correct.

Compared to decision difficulty and confidence, the analysis on decision task and decision speed exhibited a larger degree of variation and was much less conclusive. For the decision task variable, the only noticeable link to expertise level appeared for domain experts when answering questions for one-site maps. This link only slightly remains when examining decision speed for domain experts on the three-site maps. For GIScience experts on both map versions and domain experts on the three-site maps, there appeared to be no relationship between expertise level and decision task or decision speed. The contrast between decision task and decision speed versus decision difficulty and confidence exemplified a potentially dangerous issue when representing uncertainty for expert use in decision making. Although the expert may feel that it is becoming increasingly easier to make decisions under uncertain conditions as he or she gains expertise, and thus increasing the confidence in these decisions, the actual decision outcomes are not improving in a parallel fashion as experience is added. Although there is early evidence that domain experts may be improving their ability to complete simple decision tasks under uncertain conditions, this above concern remains due to the inability to transition the improvement to all definitions of expertise or to more difficult tasks.

5.1.3 Examining the Results across Decision Difficulty

Similar to the analysis on expertise level in Section 5.1.2, the online digital survey allowed for multiple definitions of the term decision difficulty. The digital survey results were analyzed using two forms of decision difficulty:

(1) self-reported

(2) investigator-defined

The first definition of decision difficulty used the participant responses for the second question in each map-legend pair question triplet, defined as the variable decision difficulty in previous sections. As discussed in Section 4.1.3, the subjects could rate the perceived difficulty on a scale of '1' through '5'. It is important to note that the responses were now pooled across the one-site and three-site maps, with exception to the decision task variable.

The number of subjects responding with a perceived difficulty in each of the five categories was highly varying, with the majority of subjects responding with a value of '2' or '3'. The frequency tally showed a comparable distribution of the decision difficulty rankings for all of the types of uncertainty excepting completeness. Compared to the other types of uncertainty, the distribution for completeness was skewed greatly to the left. This shift in frequency distribution suggests that the subjects had a greater amount of difficulty in answering questions associated with the completeness maps. This increase in decision difficulty was also reflected in Table-2.

Similar to Section 5.1.2, the non-uniform cell frequency in the factorA-by-factorB matrix required the running of a separate ANOVA test for each of the six types of uncertainty across the five decision difficulty response groups, as well as a seventh hypothesis test on decision difficulty for the entire dataset. A series of one-way ANOVA hypothesis tests were conducted, all using the null hypothesis that there was no difference in the variable being tested across the five decision difficulty groups[7]. Due to the problems with running a multitude of hypothesis tests described in Section 5.1.2, the discussion was limited to the tests that analyzed the entire dataset. Table-8 summarizes the six hypothesis tests using the factor decision difficulty conducted on the variables decision task (for one-site maps), site-A ranking, site-B ranking, site-C ranking, adjusted

---

[7] H_0: $\mu_{difficulty5} = \mu_{difficulty4} = \mu_{difficulty3} = \mu_{difficulty2} = \mu_{difficulty1}$

decision speed, and confidence. No Tukey tests were conducted due to the inherently ordinal nature of the five decision difficulty groups.

Variable	*Overall*	
	F-stat	p-value
One-Site Decision Task	0.992	0.411
Site-A Ranking	**6.445**	**0.000**
Site-B Ranking	**3.709**	**0.006**
Site-C Ranking	1.260	0.286
Adjusted Decision Speed	**5.118**	**0.000**
Confidence	**229.723**	**0.000**

Table-8: A set of one-way analysis of variance (ANOVA) hypothesis tests using decision difficulty as the grouping factor. Findings that are significant at alpha = 0.05 have been bolded.

The ANOVA hypothesis tests showed that the five categories of self-reported decision difficulty exhibit a significant difference for the variables decision speed and confidence at alpha = 0.01. This was initial evidence that an increase in the perceived difficulty of the task significantly slowed the speed in arriving at a decision (*p-value of 0.000*) and significantly decreased the confidence that the decision maker had in the made decision (*p-value of 0.000*). However, the results for decision task were much less conclusive. While the site-A and site-B ranking for three-site maps were statistically significant at alpha = 0.01, the one-site decision task variable and the site-C ranking were not significant. The investigator theorizes that the significance found for the site-A and site-B rankings perhaps were due to the small sampling size in several of the self-reported decision groups (with the size getting as small as n=2 in some cells). Testing of a larger sample is required to validate these preliminary findings concerning the connection between decision difficulty and decision task.

The second version of decision difficulty used for the research was investigator-defined. As discussed in Section 4.1.1, the investigator designed the four map versions with an assumed ordinal level of difficulty. It was expected that the four versions would increase in difficulty in the following order: the site-B one-site map, the site-C one-site map, the site-A one-site map, and the three-site map depicting sites A, B, and C together. When evaluating the summary statistics, it was interesting to find that the overall averages did not conform to the assumed ordinal ranking of map versions. As expected, responses to the site-B one-site maps had the lowest decision speed and decision difficulty, with the highest associated confidence level, and responses to the three-site maps had the highest decision speed and decision difficulty, with the lowest associated confidence level. However, the site-A and site-C one-site maps were reversed from the assumed order in the variables decision speed, decision difficulty, and confidence. Conversely, responses to the variable decision task did follow the ordinal ranking assumed by the investigator. Such a mixture may suggest that the difficulty of the site-A and site-C versions was more similar than expected by the investigator.

Because each cell in the factorA-by-factorB matrix had an equal sample size when grouping the results based on the investigator-defined decision difficulty, a series of two-way ANOVA hypothesis tests were conducted. Two-way analysis of variance testing answers the following three questions about a dataset (Aczel and Sounderpandian 2006):

(1) is there is a significant difference in the grouping of factor A?

(2) is there is a significant difference in the grouping of factor B?

(3) is there is a significant interaction between factor A and B?

The factors used in the two-way ANOVA were the four groupings of investigator-defined decision difficulty and the six groupings of uncertainty type. Because of the inclusion of uncertainty, the two-way ANOVA results helped to expand and verify the discussion from Section 5.1.1. The null hypothesis for both the decision difficulty[8] and uncertainty type[9] factors was that there was no difference among groups, addressing the first two bulleted points from above. The null hypothesis for the third bulleted question was that there was no interaction between factors. Table-9 summarizes the results for the two-way ANOVA hypothesis tests.

Variable	***Question***	***Overall***	
		F-stat	p-value
Decision Task	**Decision Difficulty**	**392.384**	**0.000**
	Uncertainty Type	1.933	0.103
	Interaction	**4.539**	**0.000**
Adjusted Decision Speed	**Decision Difficulty**	**44.448**	**0.000**
	Uncertainty Type	**6.767**	**0.000**
	Interaction	**1.926**	**0.028**
Decision Difficulty	**Decision Difficulty**	**15.111**	**0.000**
	Uncertainty Type	**3.343**	**0.010**
	Interaction	**4.329**	**0.000**
Confidence	**Decision Difficulty**	**6.500**	**0.000**
	Uncertainty Type	**8.368**	**0.000**
	Interaction	**4.144**	**0.000**

Table-8: A set of two-way ANOVA hypothesis tests using the factors investigator-defined decision difficulty and uncertainty type. Findings that are significant at alpha = 0.05 have been bolded.

The results of the ANOVA hypothesis test confirmed the findings on uncertainty type from Section 5.1.1 as well as the one-way ANOVA hypothesis testing on decision difficulty using the definition of self-reported difficulty. Significant difference at alpha = 0.01 was found for the decision speed (*p-value = 0.000*), decision difficulty (*p-value =*

---

[8] H_0: $\mu_{siteA} = \mu_{siteB} = \mu_{siteC} = \mu_{threeSite}$

[9] H_0: $\mu_{accuracy} = \mu_{completeness} = \mu_{credibility} = \mu_{currency} = \mu_{precision} = \mu_{subjectivity}$

0.000), and confidence (*p-value = 0.000*) variables when grouping the data by the factor investigator-defined decision difficulty. Similar significance at alpha = 0.01 was found for decision speed (*p-value = 0.000*), decision difficulty (*p-value = 0.010*), and confidence (*p-value = 0.000*) when grouping the data by the factor uncertainty type. Due to the parallel findings of Table-3, Table-7, and Table-8, the investigator concluded that both the type of uncertainty represented and the difficulty of the decision were significant causal variables in the perceived difficulty of making a decision and the level of confidence that the decision maker had in their decision choice. Although the findings on speed need confirmation in a controlled setting, the investigator believed that this relationship also extended to the speed in which the decision was made.

The two-way ANOVA results for the decision task also matched findings from previous analysis in this section and analysis from Section 5.1.1. The findings in Table-8 show that the difficulty of the decision had a significant impact on the decision task at an alpha of 0.01 (*p-value of 0.000)*, while the uncertainty type impacted the decision task to a lesser degree (*p-value of 0.103*), just missing significance at alpha = 0.10. When coupling the results from both definitions of decision difficulty, the investigator determined that the difficulty of the task does impact the decision outcome, although the inconsistent findings from the self-reported definition of decision difficulty required this claim to be curbed slightly. Further, when coupling the uncertainty type results from Table-8 with those in Table-3, the investigator determined that there was no connection between uncertainty type and the decision task. Although more testing is required to support this claim, the investigator has the initial hypothesis that the uncertainty type that was represented on the map has little effect on the decision outcome.

Interestingly, the interaction between decision difficulty and uncertainty type showed significance at alpha = 0.05 across all four tested variables and approached a p-value of ‘0’ in three out of four hypothesis tests. This high level of significance suggested that decision difficulty and uncertainty type need to be considered in tandem when assessing their impact on decision task, decision speed, decision difficulty, and confidence. This means that a single level of decision difficulty or a particular type of uncertainty does not consistently produce a given response in decision task, decision speed, perceived decision difficulty, or confidence. Instead, each decision making situation needs to be evaluated in a case-by-case manner, considering the decision difficulty and represented uncertainty type jointly.

5.2 Results of the Qualitative Focus Groups

Participants for the qualitative focus groups were recruited using the ‘gatekeeper’ technique, a method that utilizes a single or several individuals (the gatekeepers) within larger organizations to help recommend and gain access to many other individuals within the organization who fit the ideal participant description (Valentine 1997). An important characteristic of a gatekeeper is that he or she must have the power to provide contact information, such as email addresses and office phone numbers. For this research, the gatekeeper was Ann Barrett, the Executive Services Manager for the Wisconsin Land Information Association (WLIA). The WLIA is an organization composed of professionals working for both government and private firms dealing with land information and geospatial technologies. Ann Barrett provided a contact list of all WLIA members as well as recommendations for members working specifically in floodplain mapping. Further, Ann Barrett authorized the focus groups to be conducted during the 2007 WLIA annual conference in Appleton, WI on the 8th and 9th of March.

A call for participation in the focus groups was circulated several weeks in advance of the 2007 WLIA annual conference. Although the participation call was sent to all non-student registered members of the WLIA, it explicitly stated that only those with experience in floodplain mapping were eligible for participation in the research. The call for participation also made clear that those interested were required to complete the online digital survey before attending a focus group session. Ten qualified conference attendees replied with interest and two, one-hour focus group sessions were organized, with five participants per session. Of the ten subjects intending to participate, only six attended the sessions; two cancelled via email within twenty-four hours of their scheduled session and two were absent without excuse. Although the cancellations occurred too close to the sessions to recruit more participants, the sessions were balanced with three participants apiece, meeting the size specifications set in Section 4.2.

5.2.1 Unprompted Response to Uncertainty in Floodplain Mapping

The initial section of the interview process had the participants talk aloud about the ways in which uncertainty enters into floodplain mapping, with the hope of the participants identifying and describing different types of uncertainty without being prompted by the investigator. It is the investigator's opinion that the participants had an exceptional understanding of the uncertainty involved in making decisions off of floodplain maps. All participants readily agreed that all floodplain maps contained a degree of uncertainty and that there was "no guarantee" of the location of the floodplain depiction for decision making purposes.

In both sessions, uncertainty was initially described as the "reliability" of a floodplain line that is placed on a map. This working definition matches well with the concept of positional/locational accuracy/error and was expressed to be the most important aspect of uncertainty in floodplain mapping. However, in both sessions the concepts of precision/resolution and currency were unanimously identified without prompting as important to the overall certainty of a dataset. Precision/resolution was framed in terms of the scale at which the dataset was created and the level of detail in the dataset. Currency was defined as the date when the data was collected. Participants in both sessions emphasized that the more current the dataset the better due to the dynamic nature of the phenomenon being represented on floodplain maps. A participant in one session went as far as saying that the accuracy/error is inextricably connected to the currency of the dataset, stating that to him "currency means accuracy." Although the investigator would argue the uncertainty type accuracy/error depends on many more variables than simply currency, it was interesting to see how the other participants in the aforementioned session agreed with the assumption that more current always means more accurate. This remark revisits the drawback of the MacEachren et al. (2005) typology's lack of mutually exclusive categories and perhaps requires revision to the first-order/second-order component division. Further discussion of this concern is found in the concluding chapter. Finally, in the first session, credibility and lineage were mentioned only once by separate participants when discussing the importance of knowing who produced the dataset and, in the second session, subjectivity was briefly mentioned by a single participant in reference to the method used to produce the dataset.

This initial discussion of uncertainty prior to the introduction of the MacEachren et al. (2005) typology matched well with the Tukey results from Section 5.1.1. The participants were most comfortable identifying and discussing accuracy/error, credibility, currency, and precision/resolution and less comfortable discussing completeness and subjectivity. The Tukey hypothesis test results in Table-4 supported this division, as there were significant differences found for completeness on the one-site maps and subjectivity on the three-site maps. The inability to describe completeness and subjectivity in-depth without being prompted suggested that the participants did not regularly work with these types of uncertainties or consider them when making decisions using geospatial data, perhaps explaining why the decision difficulty responses were significantly higher and the confidence responses were significantly lower for completeness and subjectivity.

5.2.2 Confirmation of the MacEachren et al. (2005) typology

Concluding the open-ended discussion on uncertainty in floodplain mapping, a sheet detailing the MacEachren et al. (2005) typology was circulated and the participants were given several minutes to critically examine the list. It was expected by the investigator that the introduction of the new uncertainty terms would generate discussion concerning the validity of each identified type as an actual occurrence in floodplain mapping. However, each type of uncertainty listed was immediately accepted by all participants, even the categories of completeness, consistency, and interrelatedness that were not mentioned in either session before prompting. The positive reception in both interview sessions provided strong evidence that the MacEachren et al. (2005) typology was a valid categorical description of uncertainty in floodplain mapping. No evidence was provided during either session suggesting that the MacEachren et al. (2005) typology was limited to only the floodplain mapping domain.

During discussion of the MacEachren et al. (2005) typology in the second focus group session, an interesting additional category was recommended by a single participant and then agreed upon by the other participants in the session. The participant argued that any floodplain dataset needed to conform to certain levels of certainty in order to be adopted by a government agency or to be used by the decision maker. It was agreed upon by those in the session that the actual uncertainty 'threshold' varied depending upon the end user. To this end, it was argued that the party commissioning the dataset was another important category of uncertainty in floodplain mapping that was different from the MacEachren et al. (2005) definition of credibility. The provided definition of credibility from Section 2.2.3 acknowledged only the data collector or provider, with no reference to the data requester. The participants argued that the same firm could release two datasets mapping the same phenomenon, but that the datasets could have a varying degree of certainty based on the quality requirements of the client. In this example, the credibility of the dataset was equally dependent upon the data producer and the data requester. Although the participants argued that this was a separate type of uncertainty, the investigator believed that the definition of credibility should instead be amended to include both the information source and client. Other than this discussion concerning the data requester, there were no further additions to the MacEachren et al. (2005) typology.

5.2.3 Assessing Influence and Getting at the 'Why?'

The final portion of the focus groups required the participants to weight the relative influence of each type of uncertainty when making a decision off of floodplain maps. Focal questions in this portion of the interviews included "Which type of geospatial uncertainty is most important to represent on a floodplain map" and "Which type of geospatial uncertainty would influence you the most when assessing the flood risk of a site." The participant reaction to this series of questioning was the most unexpected finding of the study. Rather than immediately discussing the relative rankings of each uncertainty, participants in both sections volunteered that they would never include representations of uncertainty on a floodplain map or use any metadata recordings of uncertainty to help make the decision. This discourse provided substantial insight into how uncertain information is used in the decision-making process.

The participants argued the primary motivation for collecting uncertainty information was to allow for a final quality check of the data before being incorporated into decision making, not for supplementary explanation of the geospatial data during the decision. The use of the term ***quality*** in this regard follows the Beard and Mackaness (1993) definition of quality as "fitness for use" (i.e. the fitness that a dataset can be used in decision making). This use of uncertainty information is parallel to installing a speedometer in a car so that the manufacturer can check that the vehicle meets speed requirements after assembly, rather than providing the speedometer for the customer to more appropriately use the vehicle after purchase from the manufacturer. This application of uncertainty in decision making fits well with the concept of thresholding, as discussed in the Agumya and Hunter (2002) model of decision making. With thresholding, the decision maker marks a hard line of acceptable risk (the threshold) so that the decision task transforms from an infinite number of decision outcomes to the binary decision of acceptance or rejection. This concept of thresholding matches with the credibility discussion in Section 5.2.2, where the participants pointed out that quality requirements outlined by the client have a large influence on the end certainty of the dataset. The participants reiterated multiple times that once the linework meets the certainty requirements of the client, it is “as if it was completely accurate.” This usage of uncertainty blurs the concepts of ‘best available’ and ‘best possible’.

Such a backlash against any representation of uncertainty on floodplain maps strongly contrasted the acknowledgment of the existence and importance of uncertainty described in Section 5.2.1. However, participants in the focus groups hinted upon two possible explanations for degrading the uncertainty data down to the binary decision of ‘good’ versus ‘bad’ using thresholding :

(1) the use of thresholding speeds the decision making process by virtue of having less possible decision outcomes from which to choose, and

(2) the use of thresholding prevents any explicit representations of uncertainty on the map itself, reinforcing the probity of the dataset

The first bulleted explanation helped to interpret the results in Section 5.1.1 concerning the completeness uncertainty type. On the one-site maps, subjects reported a significantly lower response speed for decision tasks on maps representing completeness.

The investigator speculated that this reduction in speed is due to the binary characteristic of completeness. Because the uncertainty representation already existed in terms of yes-no, the decision maker did not have to mentally degrade the representation using a threshold. It is important to note, however, that subjects did not perform any better on the task itself, providing early evidence that while the thresholding technique improved decision speed, it may not lead to better decision outcomes. Finally, the investigator hypothesizes that it is this pre-existing binary quality of completeness that made it difficult to organize in the first-order versus second-order categorization in Section 3.3.2.

The implications of the second bulleted explanation for the thresholding decision making method are much more severe than for the first. A common argument in the literature is that the mapmaker does not want to represent or discuss the uncertainty of the data for fear of undermining it. Mowrer (1999) remarks that "Perhaps the worst nightmare of a natural resources manager is to appear 'uncertain' to the public, or to admit that there is 'error' in the decision process being presented." This concern was affirmed in the focus groups, with one participant arguing that the uncertainty should not be represented because it "would bring to the forefront the questionability [of the data]." The example provided was a bank using the floodplain data for insurance assessment. It was predicted that the bank would likely not use the data if it was communicated to be uncertain, and that banks typically prefer hard lines on maps to alleviate their own liability. Perhaps motivating this fear is the general opinion of the participants that most decision makers would not know how to use the uncertainty information provided. It was only after the investigator offered the idea of conducting scientific research on finding the best representation techniques that participants warmed to the idea of representing uncertainty on floodplain maps.

Following the lengthy discussion on the necessity (or lack thereof) of representing uncertainty on floodplain maps, the participants were asked the focused question "Which type of uncertainty is most influential in your assessment of the potential flood risk of a site." All participants in both sessions agreed that accuracy/error is the most influential type of uncertainty when making decisions and that it would be the type most needed when making an educated decision off of geospatial data. Precision/resolution and currency were listed as having a secondary degree of influence and the remaining six types from the MacEachren et al. (2005) typology were groups as the least influential.

CHAPTER 6 – Conclusion and Future Directions

6.1 Summary of Findings

The following is a bulleted summary of the findings from Chapter 5, organized around the questions posed from the introductory chapter. All reported results were found to be significant.

(1) Does graphically representing different types of uncertainties influence the *decision* that is made as well as the *speed* and *confidence* of this decision?

*The type of uncertainty represented does not affect the decision that is made.

*The type of uncertainty impacted the speed in which the decision was made, the perceived difficulty of making the decision, and the confidence the decision maker has in the decision.

(2) Which type of uncertainty elicits particular decision responses, as well as the most immediate and confident decisions? Which the least?

*Representation of completeness elicited the fastest decision responses, while representation of subjectivity elicited the slowest decision responses.

*Representation of completeness elicited the highest degree of perceived decision difficulty as well as the lowest degree of confidence in the decision.

(3) How much of the variation in the decision outcome is explained by the *expertise level* of the decision maker or the *decision difficulty*?

*The decision maker's level of experience impacted the perceived difficulty of making the decision and the confidence in the decision.

* The decision maker's level of experience did not impact the decision outcome or the speed in making the decision.

*The difficulty of the decision impacted the decision outcome, the speed of the decision, the perceived difficulty of making the decision, and the confidence in the decision.

(4) Which type of uncertainty is the most *influential* in the decision making process? Which is the least influential?

*Accuracy/error, above all, is the most influential type of uncertainty.

*Currency and precision are also heavily influential in the decision making process.

*Completeness, credibility, and subjectivity are the least influential types of uncertainty in the decision making process.

(5) *Why* is uncertainty used in decision making the way that it is?

*Decision making under uncertain conditions uses a thresholding model of decision making.

*Thresholding is used in decision making under uncertain conditions to speed the decision making process and to make the uncertainty of a dataset transparent

(6) Is the MacEachren et al. (2005) typology a valid categorical model of geospatial data uncertainty? Are there any categories to remove or new categories to add?

*The MacEachren et al. (2005) typology is a valid listing of uncertainty in geospatial data. No subtractions or additions are recommended.

*The MacEachren et al. (2005) typology lacks mutually exclusive categories.

In summary, the quantitative online survey found a significant disconnect between the participant perception of the decision (decision speed, decision difficulty, and confidence) and the decision outcome (decision task). The participants reported a varying level of decision speed, perceived decision difficulty, and confidence when the uncertainty type was varied. However, there was not a parallel disparity in the actually decision task variable. Similarly, the participants responded with a varying level of perceived decision speed, decision difficulty, and confidence when the level of expertise was varied, but again showed no difference in performance on decision task. A significant difference in all four variables was found only when the results were analyzed across decision difficulty. This disconnection between the decision maker's perception of the decision and the outcome of the decision is a serious issue that needs to be addressed. The primary reason for representing uncertainty is so that the decision maker can make more informed decisions based upon the geospatial data. However, this preliminary evidence suggests that such representations may increase the speed of completing the task, the perceived difficulty in completing the task, and the confidence that the decision maker has in their decision, but does little to improve the actual decision outcome. While positively influencing decision maker perception is important, the ability to achieve better, more responsible decision outcomes is the core goal.

The qualitative focus groups provided great insight into the decision making process, but also revealed a considerable hurdle that still needs to be leaped before uncertainty representation can be properly incorporated into practice. Throughout the interviews, it was established that decision makers in the floodplain mapping domain use uncertainty information as a way to establish if the data is fit for use. To see if a dataset is fit for use, a threshold of acceptable uncertainty is established, converting the uncertainty decision to a binary of 'good' or 'bad'. All participants acknowledged both the existence and the importance of understanding and recording uncertainty. Despite this, there was consistent resistance to the idea of representing the uncertainty on the floodplain map itself. This was the most unexpected finding of the study, and one that deserves a considerable amount of attention in future research.

6.2. Limitations of the Study

In critical evaluation of the research, the investigator has identified four possible limitations that need to be considered when judging the validity and generalizability of the research findings:

(1) ambiguous term definitions of uncertainty in the literature,

(2) the lack of mutually exclusive and collectively exhaustive categories in the MacEachren et al. (2005) typology,

(3) the depiction of uncertainty among three datasets rather than within a single dataset, and

(4) the proctoring of the quantitative section of the research in an uncontrolled environment.

The solution to each limitation in the research was deemed the most appropriate by the investigator given the available literature, previous scientific experiments, and the specific domain. However, argument can be built against each solution and the investigator encourages repeating the research using different assumptions. The following is a discussion on how the identified four limitations may impact the conclusions.

(1) Uncertainty Term Definition: For this research, uncertainty was defined as “a measure of the user’s understanding of the difference between the contents of a dataset, and the real phenomena that the data are believed to represent” (Longley et al. 2005). As described in Chapter 2, the inclusion of the user as another level in which uncertainty is present is controversial, with many others defining uncertainty as a quantifiable characteristic intrinsic to the dataset. Because the research focuses on the decision making process under uncertain conditions, inclusion of the interpretation/

validation/exploration filter shown in Figure-1 was deemed necessary by the investigator. A constrained perspective may also debunk the MacEachren et al. (2005) typology, as uncertainty types such as credibility, subjectivity, and interrelatedness are external user evaluations of uncertainty, at least when utilized to ordinally rank multiple datasets.

(2) Mutually Exclusive and Collectively Exhaustive Categorization: According to category theory, a typology must include categories that are both mutually exclusive and collectively exhaustive (McGrew and Monroe 2000). Early on in the research, the investigator identified the violation of mutual exclusion in the MacEachren et al. (2005) typology and separated the categories into first-order components (those not reliant upon or connected to another type of uncertainty) and second-order components (those dependent upon a first-order or other second-order types of uncertainty). The quantitative online digital survey was restricted to the former so that independent random sampling among groups could be established. However, discussion from the qualitative focus groups provided early evidence that many of the first-order components identified in Section 3.3.1 may be partially dependent upon another first-order component. Examples included the reliance of accuracy/error on the currency of the dataset and credibility on the accuracy/error and precision/resolution of a dataset. Although the investigator still hypothesizes that the first-order versus second-order division is correct, further examination is required.

Finally, the MacEachren et al. (2005) typology did not incorporate uncertainty derived from the resolution of a dataset. To remain collectively exhaustive, resolution was grouped with precision because it is the most similar category. However, Zhu (2005) provides arguments suggesting that the precision and resolution are different types of uncertainty, and accordingly, the investigator suggests the revision of the MacEachren et al. (2005) typology to include the tenth category of resolution.

(3) Uncertainty Among Datasets versus Within a Single Dataset: Depiction of uncertainty on the floodplain maps in the quantitative online survey was modeled after the Monte Carlo method of multiple realizations. However, the Monte Carlo method is different in that it generates 500-1000 realizations of a single dataset (Aerts et al. 2003b). Because the representation of uncertainty in the survey used only a single realization of each dataset, and instead compared three datasets to convey the message, it can be argued that it was a much weaker representation technique. This method of representation was chosen for two reasons. First, the maps were designed to be as simplistic as possible so that comparison of expertise level would reflect differences in the decision making process, rather than the level of familiarity with complex maps. Second, each uncertainty type needed to be symbolized in the same fashion. Because credibility and subjectivity were categorical in nature intrinsically, all six representations needed to be converted to this level of measurement, allowing the decision maker to then ordinally rank them during the interpretation/validation/exploration level shown in Figure-1. The investigator encourages the repetition of the experiment using a different method of representation to validate the findings.

(4) Research in an Uncontrolled Environment: Finally, the quantitative survey utilized the Internet for dissemination, following the Aerts et al. (2003a) study. As discussed in Section 4.1, the major drawback to conducting the research online was that the participants were no longer in a controlled environment. Because of this, the participant could be interrupted during the survey, use reference materials to aid in completing the tasks, and finish the survey with the help of others. Despite this, it was determined that evaluating domain experts in an uncontrolled environment would produce a more realistic description of expert decision making under uncertain conditions than evaluating easy-to-access university students in a controlled environment. The investigator encourages the repetition of the experiment in a controlled setting, although a shift in results is expected due to the likely under-sampling of domain experts.

6.3 Future Directions and Concluding Remarks

This research provides an initial argument that the type of uncertainty represented greatly matters in the decision making process. However, there are many more questions at this point than answers. The follow is a list of future research questions that will help to understand how typological differences in uncertainty impact the decision making process:

(1) How well do these findings transfer to representations of the identified second-order components of uncertainty and what are the implications of a division between first-order and second-order components of uncertainty?

(2) How well do these results transfer to other domains and other types of decision tasks?

(3) How would the results change if the experiment is conducted in a controlled environment?

(4) How does the MacEachren et al. (2005) typology influence the cartographic symbolization of uncertainty? How does the inclusion of multiple types of uncertainty on a single display influence the symbolization?

(5) How do we bridge the gap between an acknowledgment by experts of the importance of uncertainty depictions and the fear of experts in undermining the data by representing this uncertainty?

Uncertainty is an inherent aspect of all geospatial data. The results from this research, as well as the results from the many other projects cited in the second chapter, demonstrate that uncertainty has a statistically significant impact on the decision making process. Because uncertainty can never be fully removed from a dataset, the research agenda should be shifted, at least in part, away from the discovery of new techniques that generate more 'certain' data and instead address the role that uncertainty plays in the decision making process to the end of providing the decision maker with visuals that will allow him or her to make more informed and appropriate decisions.

REFERENCES

Aczel AD, and J Sounderpandian. 2006. *Complete Business Statistics, 6th Edition*. New York, NY: McGraw-Hill Companies, Inc.

Aerts JCJH, KC Clarke, and AD Keuper. 2003. Testing popular visualization techniques for representing model uncertainty. *Cartography and Geographic Information Science,* 30 (3): 249-261.

Aerts JCJH, MF Goodchild, and GBM Heuvelink. 2003. Accounting for Spatial Uncertainty in Optimization with Spatial Decision Support Systems. *Transactions in GIS*, 7 (2): 211-230.

Agumya A, and G Hunter. 2002. Responding to the consequences of uncertainty in geographical data. *International Journal of Geographical Information Science,* 16 (5): 405-417.

Alexander J, and SB Marriott. 1999. Introduction. In: SB Marriott and J Alexander (eds.) *Floodplains: Interdisciplinary Approaches.* London, England: The Geological Society.

Bastin L, PF Fisher, and J Wood. 2002. Visualizing uncertainty in multi-spectral remotely sensed imagery. *Computers & Geosciences*, 28: 337-350.

Beard MK, and BP Buttenfield. 1991. NCGIA Research Initiative 7: Visualization of spatial data quality. *NCGIA Technical Paper 91-26*. University of California Santa Barabara, Santa Barbara, CA: National Center for Geographical Information and Analysis.

Beard K, and W Mackaness. 1993. Visual Access to Data Quality in Geographic Information Systems. *Cartographica,* 30 (2-3): 37-45.

Bertin J. 1983. *Semiology of graphics: Diagrams, networks, maps*. (Translated from a French 1967 edition by W. Berg, editor). Madison, WI: University of Wisconsin Press.

Brewer CA. 1994. Color Use Guidelines for Mapping and Visualization. In: A MacEachren and DRF Tayler (eds.) *Visualization in Modern Cartography*. New York, NY: Elsevier Science Inc.

Board C. 1967. Maps as models. In: J Chorley and P Haggett (eds.) *Models in Geography* London, England: Methuen.

Burroughs PA, and RA McDonnell. 1998. *Principles of Geographic Information Systems.* New York, NY: Oxford University Press.

Buttenfield B. 1993. Representing Data Quality. *Cartographica,* 30 (2-3): 1-7.

Catlin R. 2002. "Floodplain" (cited map). Albany, OR: The City of Albany Planning Division.

Cliburn DC, JJ Feddema, JR Miller, and TA Slocum. 2002. Design and evaluation of a decision support system in a water balance application. *Computers and Graphics*, 26 (6): 931-949.

Couclelis H. 2003. The Certainty of Uncertainty: GIS and the Limits of Geographic Knowledge. *Transactions in GIS*, 7 (2): 165-175.

Davis TJ, and CP Keller. 1997. Modelling and visualizing multiple spatial uncertainties. *Computers Geosciences*, 23 (4): 397-408.

Djurcilov S, K Kim, P Lermusiaux, and A Pang. 2002. Visualizing scalar volumetric data with uncertainty. *Computers and Graphics*, 26: 239-48.

Dooley MA, and SJ Lavin. 2007. Visualizing Method-Produced Uncertainty in Isometric Mapping. *Cartographic Perspectives*, 56: 17-36.

Drecki I. 2002. Visualisation of uncertainty in geographical data. In: W Shi, P Fisher, and M Goodchild (eds.) *Spatial data quality.* London, England: Taylor & Francis.

Dungan JL, DL Peterson, and PJ Curran. 1993. Alternative approaches for mapping vegetation quantities using ground and image data. In: WK Michener, JW Brunt, and SG Stafford (eds.) *Environmental Information Management and Analysis: Ecosystem to Global Scales*. London, England: Taylor and Francis.

Edwards LD and ES Nelson. 2001. Visualizing data certainty: A case study using graduated circle maps. *Cartographic Perspectives*, 38 (Winter): 19-26.

Ehlschlaeger CR, AM Shortridge, and MF Goodchild 1997. Visualizing spatial data uncertainty using animation. *Computers and Geosciences,* 23 (4): 387-395.

Evans BJ. 1997. Dynamic display of spatial data-reliability: does it benefit the map user? *Computers and Geosciences,* 23 (4): 1-14.

Eyles J. 1988. *Research in Human Geography*. Oxford, England: Blackwell Publishing.

Federal Emergency Management Agency (FEMA). 2003. The 1993 Great Midwest Flood: Voices 10 Years Later. Washington, DC: The US Department of Homeland Security.

Fisher PF. 1993. Visualizing uncertainty in soil maps by animation. *Cartographica*, 30 (2-3): 20-27.

Fisher PF. 1998. Improved modeling of elevation error with geostatistics. *GeoInformatica*, 2: 215-233.

Gershon ND. 1992. Visualization of fuzzy data using generalized animation. *Proceedings, Visualization '92*. Boston, Massachusetts: IEEE Computer Society Technical Committee on Computer Graphics.

Goodchild MF. 1991. Issues of Quality and Uncertainty. In: JC Muller (ed.) *Advances in Cartography.* London, England: Elsevier Applied Science.

Harrower M. 2003. Representing Uncertainty: Does it Help People Make Better Decisions? *UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop.* Landsdowne, VA: National Conference Center.

Heuvelink GBM. 1998. *Error Propagation in Environmental Modelling with GIS.* London, England: Taylor and Francis.

Howard D, and AM MacEachren. 1996. Interface design for geographic visualization: Tools for representing reliability. *Cartography and Geographic Information Systems*, 23 (2): 59-77.

Journel AG. 1996. Modeling uncertainty and spatial dependence: Stochastic imaging. *International Journal of Geographic Information Systems*, 10: 517-522.

Kobus D, S Proctor, and S Holster. 2001. Effects of experience and uncertainty during dynamic decision making. *International Journal of Industrial Ergonomics*, 28: 275-290.

Kondo T. 1998. Requirements for crisis management in flood control plans and the design of flood control facilities. In: S Fukuoka (ed.) *Floodplain Risk Management*. Rotterdam, Netherlands: AA Balkema Publishers.

Knighton D. 1998. *Fluvial Forms and Processes: A New Perspective.* New York, NY: The Oxford University Press.

Koláčný A. 1969. Cartographic information – a fundamental concept and term in modern cartography. *Cartographic Journal*, 6: 47-49.

Krimm RW. 1998. Reducing flood losses in the United States. In: S Fukuoka (ed.) *Floodplain Risk Management*. Rotterdam, Netherlands: AA Balkema Publishers.

Kyriakidis PC, AM Shortridge, and MF Goodchild. 1999. Geostatistics for conflation and accuracy assessment of Digital Elevation Models. *International Journal of Geographical Information Science*, 13: 677-707.

Kyriakidis PC. 2003. Towards a System Approach to the Visualization of Spatial Uncertainty. *UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop.* Landsdowne, VA: National Conference Center.

Land Information & Computer Graphics Facility (LICGF). 1998. Making Waves in Floodplain Mapping: Winnebago County Modernizes Land Records. *Information Bulletin, No 2*. Madison, WI: LICGF University of Wisconsin-Madison.

Lee J, PK Snyder, and PF Fisher. 1992. Modelling the effect of data errors on feature extraction from Digital Elevation Models. *Photogrammetric Engineering and Remote Sensing*, 58: 1461-1467.

Leitner M and BP Buttenfield. 2000. Guidelines for the display of attribute certainty. *Cartography and Geographic Information Science*, 27 (1): 3-14.

Longley PA, MF Goodchild, DJ Maguire, and DW Rhind. 2005. *Geographic Information Systems and Science, 2nd Edition.* Hoboken, NJ, USA: John Wiley & Sons.

Lulloff AR. 1994. Winnebago County – Floodplain Redelineation Project. *URISA 1994 Annual Conference Proceedings*. Washington DC: Urban and Regional Information Systems Association.

MacEachren AM. 1992. Visualizing uncertain information. *Cartographic Perspectives*, 13: 10-19.

MacEachren AM. 1995. *How Maps Work*. New York, NY: The Guileford Press.

MacEachren AM, CA Brewer, and LW Pickle. 1998. Visualizing georeferenced data: representing reliability of health statistics. *Environment and Planning: A*, 30: 1547-1561.

MacEachren AM, 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.

McGranaghan M. 1993. A cartographic view of spatial data quality. *Cartographica*, 30 (2-3): 8-19.

McGrew JC, and CB Monroe. 2000. *An Introduction to Statistical Problem Solving in Geography.* Boston, MA: McGraw Hill Companies, Inc.

Monmonier M. 1993. Exploring the Quality of Enumeration-Area Data with Graphic Scripts. *Cartographica*, 30 (2-3): 28-36.

Morrison JL. 1974. A theoretical framework for cartographic generalization with the emphasis on the process of symbolization. *International Yearbook of Cartography*, 14: 115-127.

Mowrer HT. 1997. Propogating uncertainty through spatial estimation processes for old-growth subalpine forest using sequential Gaussian simulation in GIS. *Ecological Modelling*, 98: 73-86.

Pang AT, CM Wittenbrink, and SK Lodha. 1997. Approaches to uncertainty visualization. *Visual Computer*, 13 (8): 370-390.

Pang A. 2001. Visualizing uncertainty in geo-spatial data. *Proceedings of the Workshop on the Intersections between Geospatial Information and Information Technology*. Washington DC: National Academies Committee of the Computer Science and Telecommunications Board.

Paradis J, and K Beard. 1994. Visualization of Spatial Data Quality for the Decision-Maker: A Data-Quality Filter. *URISA Journal*, 6 (2): 25-34.

Plewe B. 2002. The Nature of Uncertainty in Historical Geographic Information. *Transactions in GIS*, 6 (4): 431-456.

Rod JK, F Ormeling, C van Elzakker. 2001. An agenda for democratizing cartographic visualization. *Norsk Geografisk Tidsskrift - Norweigian Journal of Geography,* 55: 28-41.

Schweizer DM, and MF Goodchild. 1992. Data Quality and Choropleth Maps: An Experiment with the Use of Color. *GIS/LIS '92,* 2: 686-699.

Sinton D. 1978. The inherent structure of information as a constraint to analysis. In: D. Dutton (ed.), *Harvard papers on geographic information systems.* Reading, UK: Addison Wesley.

Slocum TA, RB McMaster, FC Kessler, and HH Howard. 2005. *Thematic Cartography and Geographic Visualization*. Upper Saddle River, NJ: Pearson Prentice Hall.

Strahler AH, and AR Strahler. 2002. *Physical Geography: Science and Systems of the Human Environment, 2nd Edition.* New York, NY: John Wiley & Sons.

Thomson J, B Hetzler, A MacEachren, M Gahegan, and M Pavel. 2005. Typology for visualizing uncertainty. *Conference on Visualization and Data Analysis 2005 Proceedings.* San Jose, CA: IS&T/SPIE Symposium on Electronic Imaging.

Thorndycraft VR, G Benito, C Llasat, and M Barriendos. 2002. Palaeofloods, Historical Data & Climatic Variability: Applications in Flood Risk Assessment. *Proceedings of*

The Palaeoflood and Historical Evidence in Flood Risk Assessment International Workshop. Madrid, Spain: CSIC.

Valentine G. 1997. Tell me about...: using interviews as a research methodology. In: R Flowerdew and M Martin (eds) *Methods in Human Geography: a guide for students doing research projects*. Essex, England: Addison Wesley Longman Limited.

van der Wel FJM, RM Hootsmans, and F Ormeling. 1994. Visualization of Data Quality. In: AM MacEachren, DRF Taylor (eds) *Visualization in Modern Cartography*. New York, NY: Pergamon/Elsevier Science.

Wood M. 2003. Some Personal Reflections on Change…The Past and Future of Cartography. *The Cartographic Journal*, 40 (2): 111-115.

Wright JK. 1942. Map makers are human: comments on the subjective in maps. *The Geographical Review*, 32 (4): 527-544.

Zhu AX. 2005. "Research issues on uncertainty in geographic data and GIS-based analysis." In: R McMaster, L Ursery (eds) *A Research Agenda For Geographic Information Science*. New York, NY: CRC Press.

Glossary of Terms

uncertainty (p6) - a measure of the user's understanding of the difference between the contents of a dataset and the real phenomena that the data are believed to represent (after Longley et al. 2005)

---

accuracy (p6, p24, p35) - the difference between the reality and our representation of reality (after Heuvelink 1998)

completeness (p24, p36, p40) - extent to which information is comprehensive (after MacEachren et al. 2005)

consistency (p24, p41) - extent to which information components agree (after MacEachren et al. 2005)

confidence (p56) – the confidence that the decision that was made was correct (applicable to the quantitative online survey)

credibility (p25, p37) - reliability of information source (after MacEachren et al. 2005)

currency (p25, p37) - time span from occurrence through information collection/processing to use (after MacEachren et al. 2005)

decision/decision task (p55) – assessment of the flood risk of a site (applicable to the quantitative online survey)

decision difficulty (p55) – the perceived decision difficulty in completing a decision task (applicable to the quantitative online survey)

decision maker (p8) – the user that is interpreting, validation, and exploring the dataset (after Longley et al. 2005)

decision speed (p56, p64) – the time, in seconds, needed to respond to a decision task (applicable to the quantitative survey)

error (p24, p35) - the discrepancy between the attribute value in the database and the actual attribute value (after Zhu 2005)

expertise level (p55) – the degree of familiarity in completing the decision task

first-order component (p34) - an uncertainty type about the data itself (either the positional, attribute, or temporal value)

influence (p60) – the degree to which a particular type of uncertainty was relied upon during the decision

interrelatedness (p25, p41) - source independence from other information (after MacEachren et al. 2005)

lineage (p25, p42) - conduit through which information has passed (after MacEachren et al. 2005)

one-site map (p48) – a map version from the study that represents only a single site relative to the three floodplain depictions

precision (p24, p38) - the exactness of measurement (after Zhu 2005)

quality (p87) – fitness of a dataset for use (after Beard and Mackaness 1993)

resolution (p38) - the level of spatial detail (after Zhu 2005)

second-order component (p34) - a derived uncertainty type that is contingent upon, at least in part, the degree of uncertainty in a first-order component

subjectivity (p25, p39, p42) - the extent to which human interpretation or judgment is involved in information construction (after MacEachren et al. 2005)

three-site map (p48) - a map version from the study that represents three sites relative to the three floodplain depictions