Fall 12-20-2013

Techniques for the Visualization of Positional Geospatial Uncertainty

Brent A. Barré
U.S. Naval Research Laboratory, Stennis Space Center, bbarre@uno.edu

Techniques for the Visualization of Positional Geospatial Uncertainty

A Thesis

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Master of Science
in
Computer Science

by

Brent A. Barré

B.S. University of New Orleans, 2011

December, 2013

Acknowledgement

I would like to thank the following people for their influences on this thesis, in no particular order:

- Dr. Elias Ioup of the Naval Research Laboratory (NRL at Stennis Space Center), for providing this research idea and for his generous, adept guidance throughout its exploration.

- Dr. Mahdi Abdelguerfi of the University of New Orleans (UNO), for his expert guidance toward the content and timeliness of this work.

- Dr. Md Tamjidul Hoque of UNO, for his thoughtful feedback on revising this thesis.

- Dr. Mark Livingston of NRL in Washington DC, for his expert advice from the visualization field and for his human testing advice and approval.

- Dr. John Sample of NRL Stennis, for allowing me the opportunity to complete my master's degree as part of my NRL internship.

- Norman Schoenhardt of NRL Stennis, for willingly teaching me most of what I know about WMS visualization.

- Dr. Noelle Brown of NRL Stennis, for working closely between NRL and LSU in order to conduct the excellent human testing for this thesis.

- Dr. Melissa Beck of LSU, for providing her lab space, equipment, and research assistants for human testing at LSU.

- Rebecca Goldstein of LSU, for coordinating the human testing sessions and communicating the results with Dr. Brown.

- Joseph Delaune and Laura Heisick of LSU, for running the human testing participants through the experiments.

- Todd Lovitt of NRL Stennis, for sharing his implementation of the Lohrenz clutter model (C3) and for his hands-on help with running it.

- Michael Trenchard of NRL Stennis, for his detailed explanations and help in using the C3 program and interpreting its output.

- My family, especially my parents, Sidney and Lou Anne Barré, whose work-ethic and support inspired and allowed me to get to this point in my education.

Table of Contents

Abstract

Geospatial data almost always contains some amount of uncertainty due to inaccuracies in its acquisition and transformation. While the data is commonly visualized (e.g. on digital maps), there are unanswered needs for visualizing uncertainty along with it. Most research on effectively doing this addresses uncertainty in data *values* at geospatial positions, e.g. water depth, human population, or land-cover classification. Uncertainty in the data's geospatial *positions* themselves (positional uncertainty) has not been previously focused on in this regard. In this thesis, techniques were created for visualizing positional uncertainty using World Vector Shoreline as an example dataset. The techniques consist of a shoreline buffer zone to which visual effects such as gradients, transparency, and randomized dots were applied. They are viewed interactively via Web Map Service (WMS). In clutter testing with human subjects, a transparency-gradient technique performed the best, followed by a solid-fill technique, with a dots-density-gradient technique performing worst.

Keywords: geodesy, cartography, geospatial, geographic information system (GIS), digital mapping, positional uncertainty, over-zoom, circular error, visualization, decision making, mission planning, vector data, World Vector Shoreline, Web Map Service

1. Introduction

Digital map users range in experience from, for example, the layperson using Google Maps for the first time, to the expert military strategist using an aggregation of maps to plan a mission. Across this entire spectrum, however, users are often unaware of *positional inaccuracy* in underlying map data because they cannot see it. When a map contains positional inaccuracy, the geographical coordinates of a given map feature can differ from the coordinates of that feature on the actual surface of the earth. A critical manifestation of this problem occurs in military mission planning. Consider a scenario where a map of slightly inaccurate shoreline locations is used in planning for ship navigation. A plotted course that passes near land on the map may pass on top of land in reality. As a possible consequence, a ship navigator traveling on such a course would have to notice the problem and reroute. A solution would be to give the planner a visual indication of where uncertainty lies, added to the map that he started with. This is merely one example; there is a wide variety of map data, its uncertainty, and the visualization thereof. This example also falls within the research area of geographic decision making; previous research on decision making is visited briefly in section 2.3.

Digital geographic maps are created (visualized) from different types of data. The following are data examples:

1. Vector data: points, lines, and areas, given in longitude and latitude coordinates, representing features on the earth.
2. Imagery data: images of the earth's surface obtained by satellite or aircraft.
3. Gridded data, e.g. bathymetry: underwater depth measurements at longitude and latitude coordinates separated by certain distance intervals.

Different datasets have different types and degrees of uncertainty and will often provide accuracy requirements from which uncertainty can be derived. To give an idea of these different uncertainty types, the following are examples of datasets and their accuracy requirements. They correspond to the data types listed above: World Vector Shoreline is a vector dataset, Landsat 7 is an imagery dataset, and NOS Bathymetry is a gridded dataset. Each accuracy requirement is first labeled by its uncertainty type (either "positional" or "depth" in these examples).

1. World Vector Shoreline: positional uncertainty: "Requirement for this data is that 90% of all identifiable shoreline features be located within 500 meters (2.0mm at 1:250,000) circular error of their true geographic positions with respect to the preferred datum (WGS 84)" (Soluri and Woodson, 1990).
2. Landsat 7 imagery: positional uncertainty: "Geometric accuracy of the systematically corrected product should be within 250 meters (1 sigma) for low-relief areas at sea level" (USGS, 2013).
3. NOS Bathymetry:
    a. Positional uncertainty: "…the Total Horizontal Uncertainty (THU) in position of soundings, at the 95 percent confidence level, will not exceed 5 meters + 5 percent of the depth".
    b. Depth uncertainty: given as a complex formula, in meters, at the 95 percent confidence level. See citation for more information.

    (NOAA, 2013).

Notice that bathymetry data can contain two types of uncertainty: positional and depth ("a" and "b" above). Depth uncertainty falls under a category called *attribute* uncertainty, which refers to uncertainty in the *values* at geospatial positions (depth in the above case), separate from

uncertainty in the positions themselves. The distinction between attribute uncertainty and positional uncertainty is what drives the contribution of this thesis. At the same time, uncertainty can be classified much more extensively than this. Previous research has focused on classification and is summarized in section 2.2.

Visualization is referred to here as the process of using a computer to draw a representation of information onto a display. The rendering of geospatial data onto a digital map falls under this category. Similarly, the inaccuracy, or uncertainty, in such data can also be visualized. However, not all ways of visualizing a particular uncertainty are effective ones. This becomes even more of a challenge when uncertainty is visualized on the same display as the data; the data visualization itself can already be quite cluttered. Further, it is hard to know how visualization techniques will be interpreted by users. If they do interpret an uncertainty display as intended, will this lead them to make better decisions? Will it provide too much information for them to reasonably process? This thesis defines "effective" based on questions like the above. Previously created uncertainty visualization techniques and their experimental results are summarized in section 2.4.

This thesis proposes techniques for effectively visualizing positional geospatial uncertainty. For the demonstration of methodology, an example dataset that would portray the scenario described in the first, motivation paragraph of this chapter was desired. World Vector Shoreline (WVS) was a chosen due to its mission-planning purposes. Also, this data is simply focused on the positions of shorelines, not attributes or values at those positions. This makes it a clear example for demonstrating *positional* uncertainty visualization. It should be emphasized, however, that WVS is merely an example dataset. The techniques proposed in this thesis could be adapted to vector datasets in general (i.e. points and areas in addition to lines).

For the visualization methodology, rendering via Web Map Service (WMS) allows for interactive map controls (e.g. panning and zooming) over the Internet from a WMS client of choice (Gaia in our case). The basics of WMS are as follows. When pan or zoom controls are used in a WMS client, a request is sent to the WMS server for a new map image. The WMS server then runs whatever image creation code was implemented for the request. This code creates and returns the image, which the WMS sends to the client. In my case, the WMS server calls my implementation for drawing WVS data and its uncertainty visualizations. Further, visualization was done this way because WMS is a commonly used, industry-standard map image protocol.

Referenced throughout this introduction, chapter 2 provides a summary of existing research on geospatial uncertainty visualization. Chapter 3 then contrasts this with the proposed visualization methodology of this thesis, thus introducing the contribution. The actual proposed methodologies, i.e. uncertainty modeling and visualization techniques for World Vector Shoreline data, are then put forth in chapter 4. The human-subjects testing of the methodologies, and the results thereof, are discussed in chapter 5. In conclusion, chapter 6 relates the findings of this thesis to previous work and discusses directions for the near-future.

2. Background

2.1 Overview

How should the uncertainty of data, in general, be visualized? There are several studies across different fields of research, each offering insight as to the effectiveness of certain visualization techniques with certain types of data. However, the visualization community is far from a definitive solution to this question (MacEachren et al., 2005). Within the specific field of *positional* geospatial uncertainty visualization, the amount of existing research is surprisingly small. There is, however, a more substantial body of research in the more general area of geospatial uncertainty visualization. Perhaps the most comprehensive overview of the state of the art in this area is given in the study by MacEachren et al. (2005), which draws from over 90 different studies. He provides a summary and analysis of the following subjects within geospatial uncertainty visualization: 1) typologies of uncertainty, visual variables, and visualization techniques, 2) the impact of uncertainty visualization on decision making, and 3) the results of several studies that tested the effectiveness of visualization techniques. The remainder of chapter 2 addresses MacEachren's discussion of these subjects and how they relate to this thesis.

2.2 Typology of Uncertainty

Spatial data transfer standards (SDTS), begun by the U.S. Geological Survey (USGS) in 1980, defined categories of data quality for cartographic data sets. The categories, specified to be encoded in metadata, are the following:

- **Lineage**: a description of the source material from which the data were derived and

- the methods of derivation, including all transformations involved in producing the final digital files (USGS 1997, p. 15);
- **Positional accuracy**: must include the degree of compliance to the spatial registration standard; measures can include: deductive estimate, internal evidence, comparison to source, or independent source of higher authority (USGS, 1997, p. 15);
- **Attribute accuracy**: both measurement accuracy (for features measured on a continuous scale) and class assignment accuracy (for categorical features) are included here (USGS 1997, p. 16);
- **Logical consistency**: here, the objective is to describe the fidelity of relationships encoded in the data structure of the digital spatial data (USGS 1997, p. 16);
- **Completeness**: the goal here is to describe the relationship between the objects represented and the abstract universe of all such objects. Includes issues such as selection criteria (e.g., size thresholds for spatial features, frequency counts for attributes), definitions used, and other mapping/abstraction rules (USGS 1997, p. 17).

Of particular relevance to this thesis is the positional accuracy category, which defines that the metadata "must include the degree of compliance to the spatial registration standard; measures can include: deductive estimate, internal evidence, comparison to source, or independent source of higher authority" (USGS, 1997, p. 15). The World Vector Shoreline accuracy specification (introduced in chapter 1 and used in chapter 4) is similar to this definition. It is the basis of my own uncertainty visualization measurements. In other work, Buttenfield and Weibel (1988) took the data quality categories listed above and conceptualized a framework of representations/visualizations deemed appropriate for them. The result was a matrix that matches each category to the visual variables appropriate for use in the category's visualization,

depending on data type (see Table 1). These are geospatial data types; that is, they deal with the geographical positions, or features, on the earth and their attributes, or values. The data types in this matrix (as summarized by MacEachren) are:

- **Discrete**: point and line features;
- **Categorical**:
    - area features assigned to categories through aggregation, or
    - overlay or attributes assigned to classes through partitioning and enumeration;
- **Continuous**: surfaces and volumes.

Data Quality / Data Type	Positional Accuracy	Attribute Accuracy	Logical Consistency	Completeness	Lineage
Discrete Points and Lines	Size Shape (Error ellipses) (Epsilon bands)	Value Color Saturation (Feature node checks)	Color mixing Redundancy by overprinting Slivers by solid fills (Topological cleaning)	Mapping Technique Density traces Marginalia Generalization algorithm Mapping tolerance Buffer size	Mapping Technique Minimum Bounding Rectangles
Categorical Aggregation & Overlay (Tesselation, tiling, Areal coverages)	Texture Value (Certainty of boundary location)	Color mixing (Attribute code checks) (Topographic classifier)	lack error models	Mapping Technique Missing values Logical adjacency surface Marginalia Discrete model weights	
Partitioning & Enumeration (Metric class breaks)	not meaningful	Size = height (Blanket of error)	Size = height (Maximum likelihood prism maps)	Mapping Technique Missing values Misclassification matrix Classing scheme OAL/TAI	Marginalia Source of data Scale/Resolution Date Geometry
Continuous Interpolation (Surfaces and volumes)	No clear distinction b/w the two Value Color Saturation (Continuous tone vignettes) (Continuous tone isopleths)		Size = line wt Color Shape = compactness (TIN links)	not possible by definition Mapping Technique Surface of search attenuation Marginalia Interpolation algorithm	
Graphical Syntax ← →				**Graphical/Lexical Syntax** ← →	

Table 1. The Buttenfield-Weibel (1988) matrix (adapted here) matches data type and quality to visual variables determined to be most appropriate for visualization.

The following is an example that is applied in the methodologies of this thesis: the matrix recommends visualizing positional accuracy for points and lines with the use of error ellipses and the visual variables of size and shape. This matrix serves as a general guideline. The effectiveness of a particular visualization technique when applied to particular data can only be fully determined through testing. Visual variables, visualization techniques, and testing are discussed in section 2.4.

2.3 Decision Making

Perhaps the most important goal of uncertainty visualization is to help its users make better decisions. Most studies here assume that inclusion of uncertainty information is helpful to decision makers, but as MacEachren (2005) points out, this assumption has been tested very little. Tversky and Khaneman (1974) demonstrated that humans tend to rely on heuristics rather than statistics when provided with statistical uncertainty information. This finding prompts important questions for uncertainty researchers, which MacEachren (2005) poses as the following: "(1) will experts revert to lay strategies of applying heuristics when statistical evidence is not available; (2) will providing information about data uncertainty in an explicit visual way help a lay or expert map reader make different decisions; and (3) if they do make different decisions, will provision of information about data uncertainty lead to better, more correct, decisions or simply cause analysts to discount the unreliable information, whether doing so is the best strategy or not". Moving forward, the idea of visual variables, in light of current visualization techniques will be further discussed. This serves as a review of the state of the art, as well as the foundation for the proposed techniques.

2.4 Visualization Techniques and Testing

In theory, there is an endless number of ways to visualize a given entity. As far as uncertainty goes, most, if not all visualization techniques can be created through the manipulation of the "visual variables" defined by Bertin (1983) and MacEachren (1992). They are as follows: location, size, color value, grain/texture, color hue, orientation, and shape (Bertin); color saturation, crispness, transparency, and resolution (MacEachren). There are

relatively few studies that have actually tested the effectiveness of uncertainty visualization techniques in terms of these variables. Many of them are summarized by MacEachren et al. (2005). However, perhaps because of such a wide range of visualization possibilities and data types upon which to apply, results of these studies have only been relative. What is meant by "relative" here is that, given an uncertainty visualization technique study, its researchers attempt to rank the techniques that they test only within the realm of their own experiment. However, it is posited here that perhaps a more absolute ranking of techniques can begin to be formed, based on trends that emerge from individual technique studies. Such a ranking provides a useful palette of techniques to use for the purposes of this thesis.

First, let us consider the two most basic categorizations of techniques: intrinsic and extrinsic. These categorizations were first given by Howard and MacEachren (1996) and are regularly used throughout uncertainty visualization literature. As stated by MacEachren et al. (2005), "intrinsic approaches are those that change appearance of an object, while extrinsic approaches use additional symbols to provide information about an object". Thus, one of the effects of extrinsic techniques, by using additional symbols, is the addition of "clutter" to the display. In this thesis, the techniques created should be applicable to data sets that are already quite cluttered spatially. Thus, for these purposes, an important goal is to limit any further clutter from uncertainty visualization, and thus, to exclude extrinsic techniques.

Table 2 shows several geospatial uncertainty visualization study references with their corresponding, intrinsic-only, technique rankings. Most of these studies were summarized by MacEachren et al. (2005); we created this table out of the results of the studies for easier comparison. The studies are numbered (blue) to match them to their citations at the end of the references section. The cell values are of the form: [rank] of [number of techniques tested]. For

an example of this table's usage, take the first cell value, "2nd of 2". This is taken to mean that in study [21] (which is Schweizer and Goodchild, 1992) 'Color Saturation' came in second place out of the two uncertainty visualization techniques tested in that study. Duplicate rankings for the same study indicate ties or combined techniques. Some rankings are missing because extrinsic and animation techniques are omitted here. The purpose of Table 2 is to gain an overview of technique effectiveness in order to make better-educated selections for future applications/tests. An example of what can be gathered from this table is that the "color saturation" technique frequently comes in last place -- it is proving to be one of the worst techniques. "Color value", on the other hand, frequently comes in first place -- it is proving to be one of the best techniques. As explained earlier, we will not consider extrinsic techniques due to their added clutter, and so extrinsic technique (e.g. glyphs) results are omitted from this table. Also, because our techniques will target WMS visualization, and because WMS involves displaying of one static map image per request, technique results involving animation (i.e. serial, random, or blinking/flickering animation) are also omitted from this table.

		Uncertainty Visualization Techniques (intrinsic only)					
		Color Saturation	Color value/ Lightness/ Darkness	Texture	Opacity	Legend	Grayscale
Studies	[21]	2nd of 2	1st of 2				
	[15]	3rd of 3		1st of 3			
	[5]	5th of 5			2nd of 5		
	[6]		1st of 2			2nd of 2	
	[9]	3rd of 3	1st of 3	2nd of 3			
	[10]		3rd of 5	1st of 5			
	[7]	1st of 2					
	[3]						1st of 3
	[1]		1st of 2				

Table 2. Uncertainty visualization technique effectiveness rankings from several different studies. Study citations are listed in the References section.

In general, the methodologies of these studies can be summarized as follows: 1) identify a data type, dataset, map, or mapping system whose use would benefit from uncertainty visualization; 2) devise uncertainty visualization techniques for portraying uncertainty in target maps, test maps, or map symbols from 1); and 3) perform experiments in which human subjects use maps both with and without the different uncertainty visualization techniques, comparing the results. Table 3 shows the data/map types (from methodology "step 1)") on which the uncertainty visualization techniques from Table 2 were tested -- it contains the same studies as Table 2.

		Data/Map Type
Studies	[20]	Thematic data on choropleth maps
	[15]	Health data on choropleth maps
	[5]	Land-cover classifications on choropleth maps
	[6]	Disease reports on graduated circle maps
	[9]	Land-cover classifications on choropleth maps
	[10]	Fields of scalar values
	[7]	Land-cover classification on choropleth maps
	[3]	Land-cover classification on choropleth maps
	[1]	Urban growth data on choropleth maps

Table 3. Data/map types on which the uncertainty visualization techniques from Table 2 were tested. All of the data used here exemplifies attribute uncertainty, not positional uncertainty. Study citations are listed in the References section.

While the selection of data/map types are secondary to the purpose of the studies (i.e. testing technique effectiveness), for our purposes it is actually important to note the commonality between the data. The term that stands out is *choropleth maps*. Most of the choropleth maps in these studies fit the definition of *simple choropleth maps* given in the book *Elements of Cartography* (Robinson et al., 1995). Simple choropleth maps "symbolize magnitudes of statistics as they occur within boundaries of unit areas (such as counties, states, or other kinds of enumeration districts)" (Robinson et al., 1995). Generalizing the studies' data/maps more broadly, they all fall under the category of *thematic mapping*, which "concentrates on the distribution of a single attribute or the relationship among several" (Robinson et al., 1995). The keywords in these definitions (in the context of this thesis) are "magnitudes" and "attributes", which can be described more generally as *values* at geospatial positions. Thus, by definition, the current state of tested geospatial uncertainty techniques concerns only uncertainty in *values* on thematic maps.

For visual reference, the following are some examples of previously created uncertainty visualization techniques and their uses of visual variables. Figure 2, Figure 3, and Figure 4 are from three different studies above.

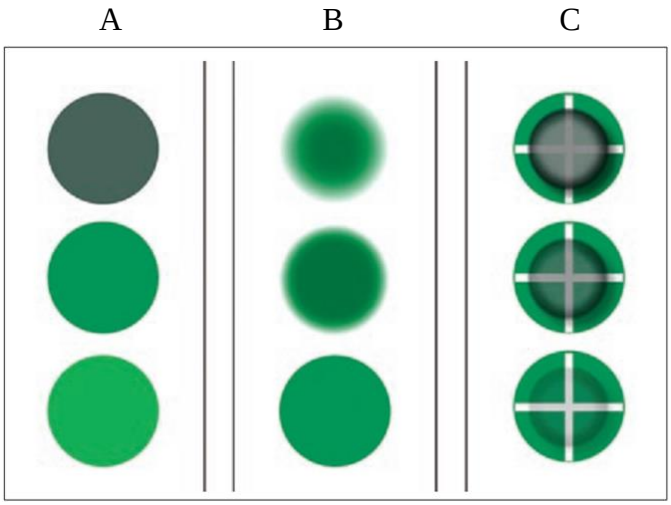

Figure 1. Point symbols depicting uncertainty decreasing from bottom to top. This is shown via decreasing from bottom to top: (A) saturation; (B) crispness of symbol edge; (C) transparency of symbol. In (C), transparency is applied to the smaller symbol in the foreground (MacEachren, 2005).

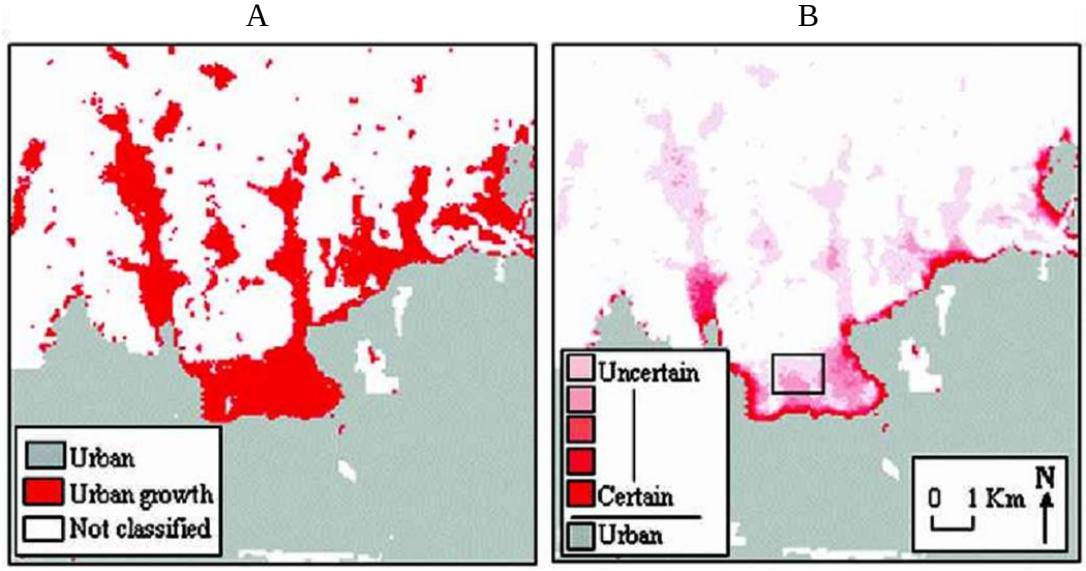

Figure 2. (A) An urban growth model depicted without uncertainty and (B) with uncertainty represented by introducing decreasing color value (B) (Aerts, 2003).

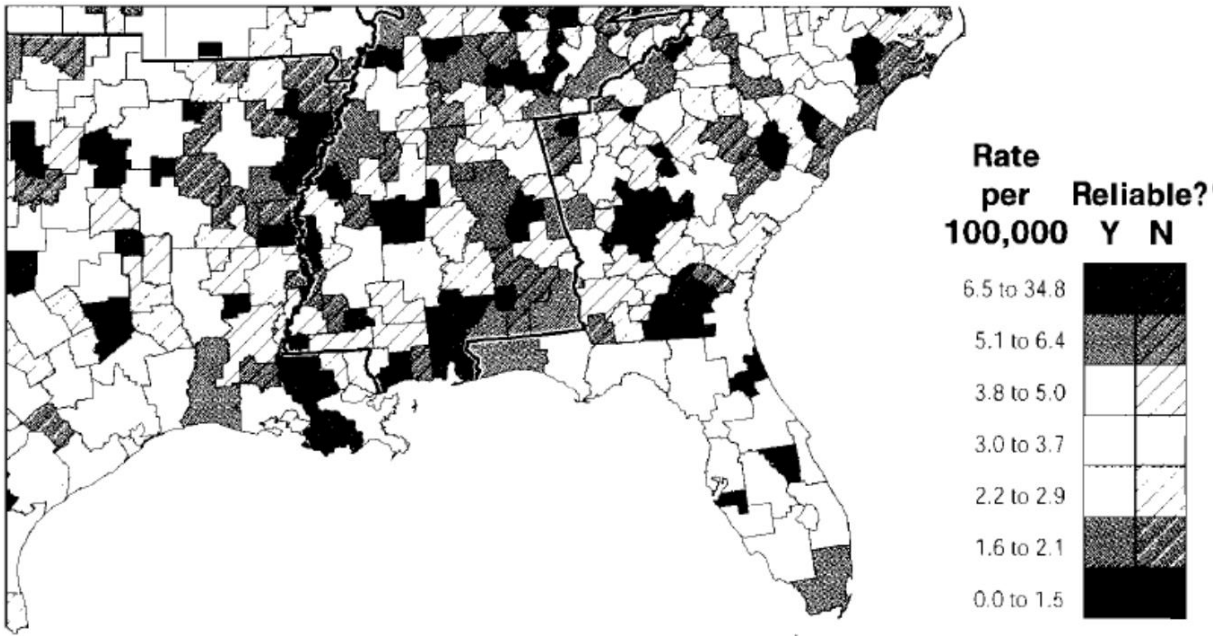

Figure 3. Map of death rates, with a striped texture introduced to depict reliability/uncertainty (MacEachren, 1998).

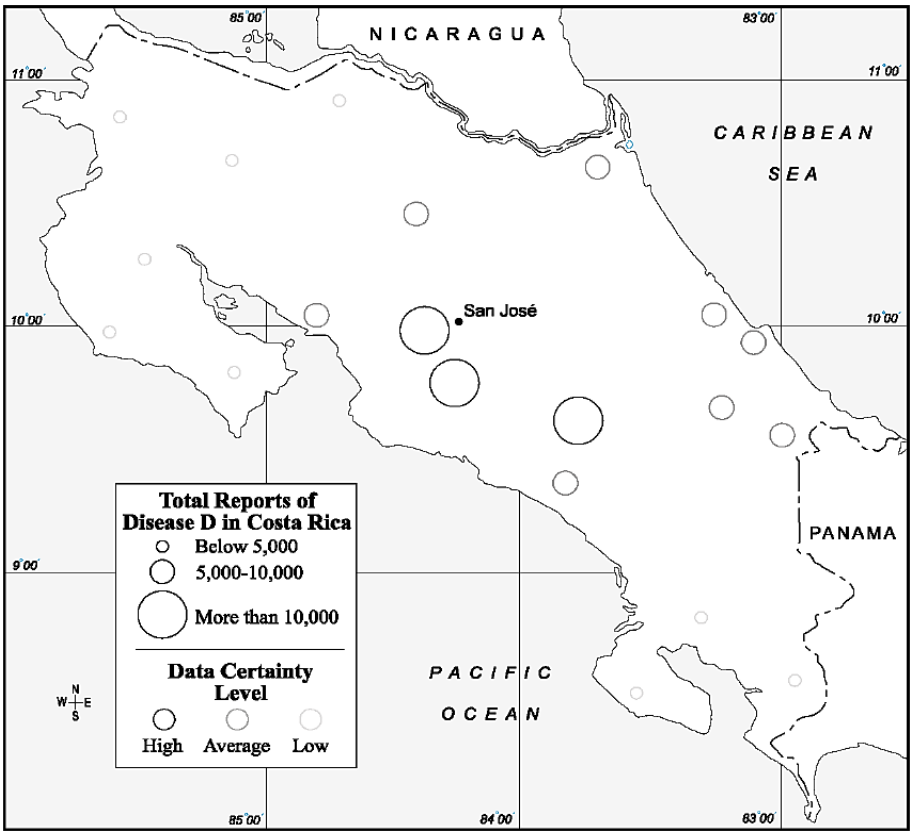

Figure 4. Circle sizes indicate number of disease reports, with circle lightness/darkness introduced to indicate certainty level (Edwards & Nelson, 2001).

3. Contribution

Consider the types of data used in the studies and examples discussed above. When looking at the studies as a whole, it becomes apparent that each of their data is visualized on thematic maps, e.g. using color or circle sizes on geographical maps to represent things like heath data, numerical data, and land cover classifications. Thus, it becomes apparent that these studies and current studies in general tend to visualize uncertainty in the *values* at geospatial positions. It seems there are no current studies that attempt to visualize uncertainty in *positions* themselves. That is the theoretical problem addressed by this thesis: how is uncertainty in geospatial positions, rather than uncertainty in data values at geospatial positions, effectively visualized?

Having the theoretical problem and an identification of effective techniques from the preceding chapter, these concepts are applied to the example addressed in this thesis: visualization of positional uncertainty in World Vector Shoreline data. However, the methodology is meant to be extended as a solution to the uncertainty visualization problem for vector data in general. It has the potential to add important visual information to multiple widely-used vector datasets. The testing of the proposed techniques will also be contributing to the body of research on effectiveness of uncertainty visualization techniques in general. The advantage gained from effective visualization of uncertainty, as mentioned in section 2.3, is the aid to the decision making of the user – the mission planner in this case. As shown later, the proposed techniques are tested in a usability experiment with human subjects.

4. Methodology

4.1 World Vector Shoreline

As mentioned previously, the data of interest of this study is World Vector Shoreline (WVS). It is composed of, among other things, lines defined by latitude and longitude coordinates that represent the coastline of the entire world (see Figure 5).

Figure 5. My visualization of WVS 1:250M-scale data.

It also provides these world coastlines at different map scales, with the highest (most detailed) being 1:250,000. The following are the data scales, along with abbreviations and the colors in which they are visualized in Figure 6:

- 1:250,000 (250K) = red.
- 1:1,000,000 (1M) = cyan.
- 1:3,000,000 (3M) = blue.
- 1:12,000,000 (12M) = magenta.
- 1:40,000,000 (40M) = green.
- 1:120,000,000 (120M) = orange.

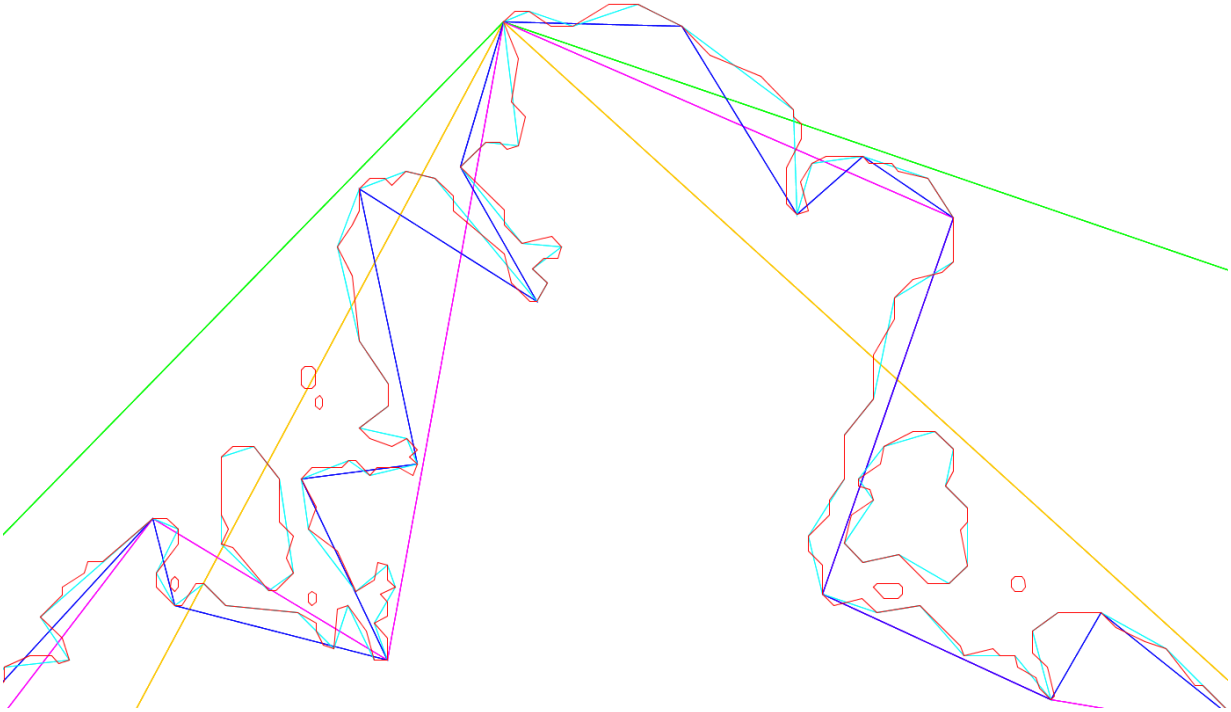

Figure 6. Higher WVS scales (e.g. red) more closely approximate the true shape of the shoreline.

Figure 6 is an example map location with all WVS scales shown simultaneously. Notice that the line approximations fit the shoreline more coarsely as scale moves from highest (red) to lowest (orange). Note that "red" is referred to as the highest scale because its value as the fraction "1/250,000" is greatest. This is an important consideration -- as scale decreases, uncertainty

increases. The formula for this relationship is discussed in section 4.4. See also http://shoreline.noaa.gov/data/datasheets/wvs.html for more WVS-specific information.

4.2 The Over-Zoom Problem

This section discusses perhaps the most widespread contributor to vector data uncertainty in GISs: the over-zoom problem. Vector data is usually compiled from a static source map or maps (e.g. paper maps) of certain scales. The scale of the vector data is derived from the scale of its source. For example, WVS data was compiled from raster data and hard copy sources "at a preferred scale of 1:250000" (Soluri & Woodson, 1990). Thus, WVS data has a scale 1:250,000 (see previous section; its smaller scales were probably approximated from this original scale). In other words, the WVS approximation of shoreline geometry is acceptably accurate for viewing up to that scale. However, in GISs, Web Map Service in our case, the user is free to specify *any* scale at which to view the data. This is often referred to as zooming in or out. When a map is "zoomed", the viewing scale changes, but the data scale does not. This allows the user to zoom in *indefinitely* past the scale of the data. As they do this, the data geometry is redrawn the same way but at higher and higher precision (i.e. finer and finer geographical coordinates on-screen) that is further and further past the amount of accuracy in the data. The following figures illustrate this concept by first contrasting it with zooming into a static (e.g. paper or image) map.

Figure 7. WVS 1:250,000 scale data for Oahu, Hawaii, shown in WMS roughly at its intended viewing scale (1:250,000)

Figure 8. When zooming into the static image from Figure 7, Pearl Harbor shorelines become thickened and blurred. This shows that precision has been increased past the accuracy in the original image.

Figure 7 and Figure 8 are examples of what is actually an advantage of static maps, such as paper maps (or images, as shown in the figures). With a paper map, the user has an immutable view of the data, shown as intended. The data is shown only at a scale where its depiction is accurate. They can magnify the map, but that will not mislead them into overestimating the accuracy of the data.

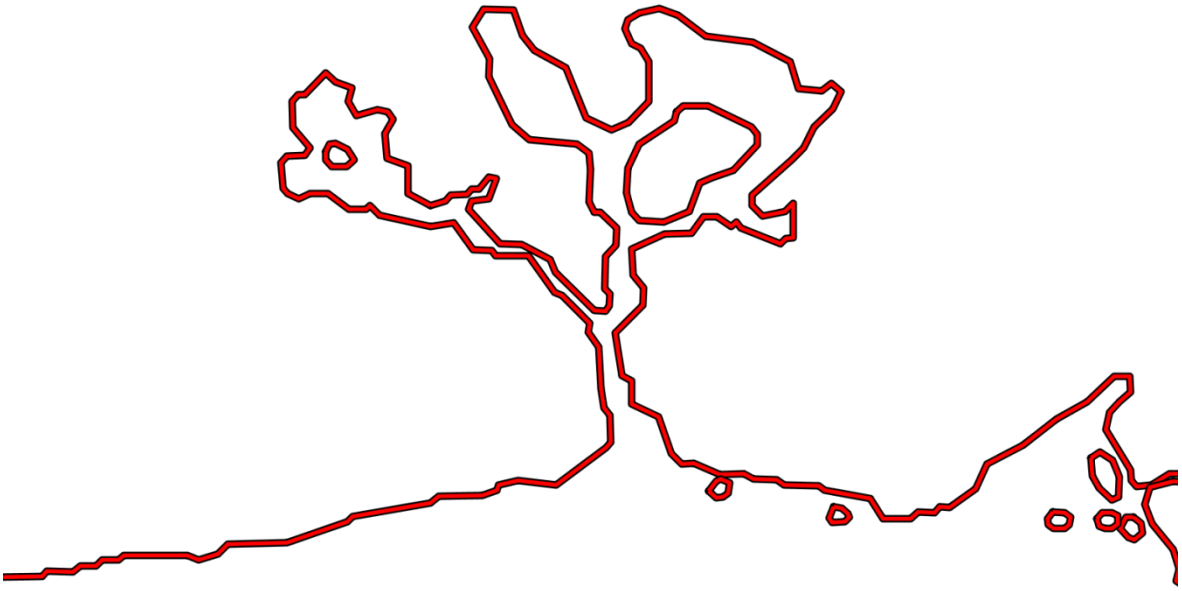

Figure 9. When dynamically zooming into WVS at Pearl Harbor in WMS, the data is precisely redrawn. This precision is misleading: Figure 8 is a better depiction of the data's accuracy.

In contrast, Figure 9 shows no loss in representation quality; if anything, the shoreline separation looks even clearer than in Figure 7. This is because the GIS client continues to redraw the same line visualization at exact coordinates, despite this being less and less truthful as it zooms past the intended scale of data viewing. This is very misleading to users. They could easily take Figure 9 as the true position of the shoreline. However, showing this over Landsat 7 satellite

imagery, which is twice as accurate (250m error) gives a different picture (see accuracy specifications in chapter 1).

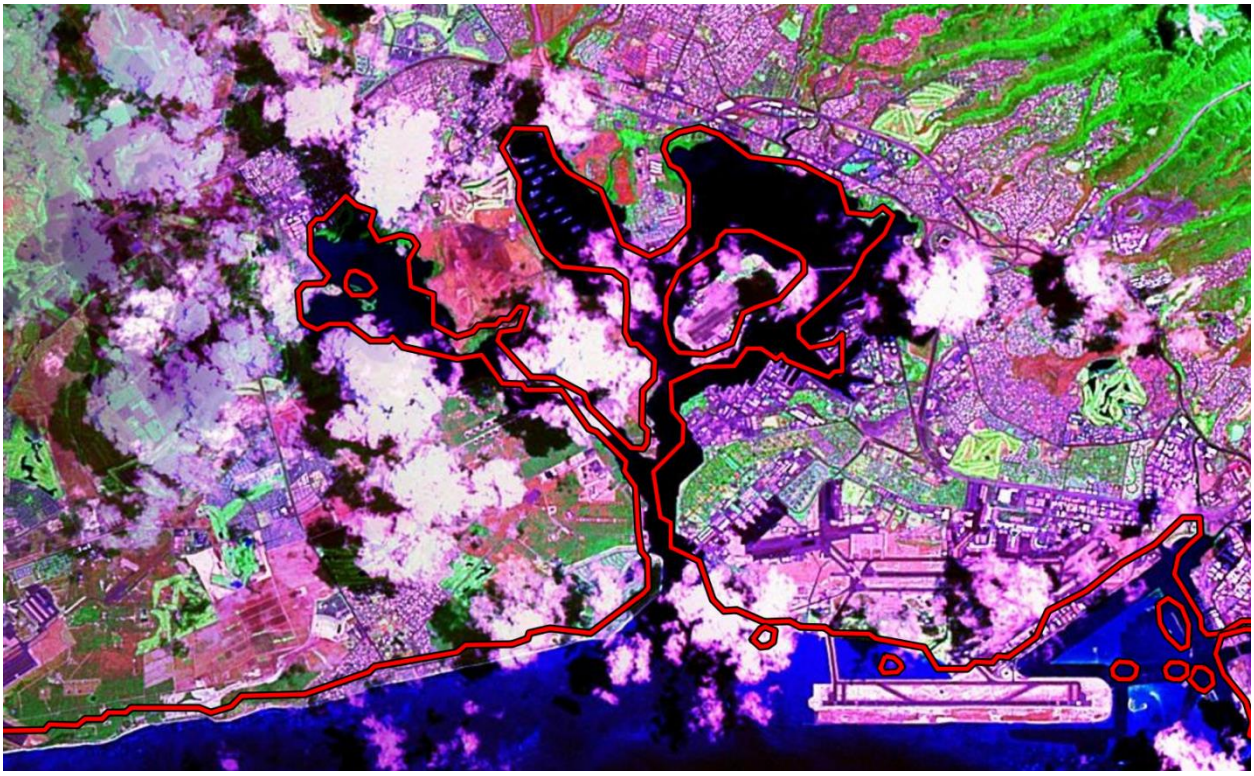

Figure 10. A visualization of WVS data shown over more accurate Landsat 7 imagery.

The data and imagery are both geo-referenced to the same coordinate system (EPSG:4326), so if they were both perfectly accurate, the shorelines would line up perfectly. However, they do not match -- some of the WVS water inlets are almost totally obstructed by land, according to the imagery. This is obviously problematic for plotting a boat's course into Pearl Harbor. If WVS is the only data on-hand, a person plotting a course could easily be unaware of this problem. This motivates creating an uncertainty model for World Vector Shoreline.

4.3 An Uncertainty Model for World Vector Shoreline

World Vector Shoreline contains no data for the actual representation of uncertainty, but an accuracy statement can be found on NOAA's website (http://www.ngdc.noaa.gov/mgg/coast/wvs.html): "requirement for this data is that 90% of all identifiable shoreline features be located within 500 meters (2.0mm at 1:250,000) circular error of their true geographic positions with respect to the preferred datum (WGS 84)" (Soluri and Woodson, 1990). Note that in terms of the over-zoom problem, even at the intended scale of 1:250000, the data is not perfect. However, visualizing the above error at all zoom levels will account for this imperfection as it is magnified by the over-zoom problem. The uncertainty model presented in this thesis is derived from the WVS statement as follows. First, the statement can be thought of as, "for any WVS data point, the real world shoreline will lie within a circle whose center is the data point and whose radius is 500m". However, the WVS shoreline is represented as a collection of line segments, not just points. Thus imagine tracing the center of a 500m-radius pen tip over the entire shoreline represented by these line segments (see Figure 11). The resulting line width of 1000m can be visualized as a buffer zone that represents a high chance (90%) that the real world shoreline lies somewhere within it.

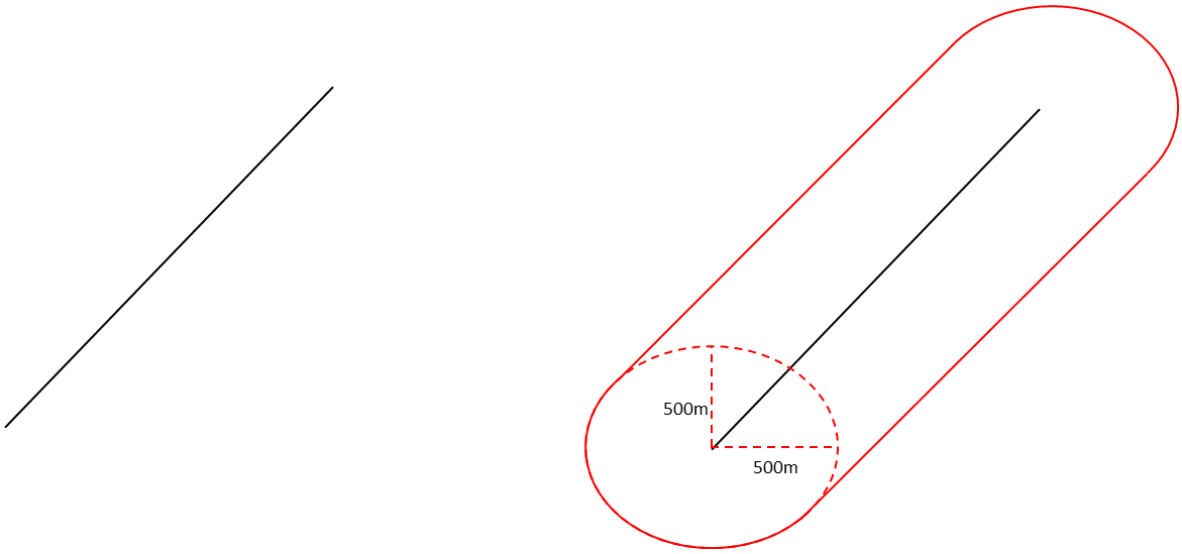

Figure 11. Left: WVS line segment. Right: WVS line segment with diagram of uncertainty model.

Because the "pen tip" is round, the join points of the thickened lines are rounded (the subtle effect of this is present e.g. in Figure 12). Also, there is an important detail in Figure 11 worth questioning here. If the radii are the same length, why does the pen tip look like an ellipse rather than a circle? This is because of the horizontal stretching due to map projection (EPSG:4326 projection in this case, when WMS is used). The stretching is least noticeable near the equator and most noticeable near the poles of the earth. Having the uncertainty model, a basic visualization of it can be added to Figure 9. The result is Figure 12.

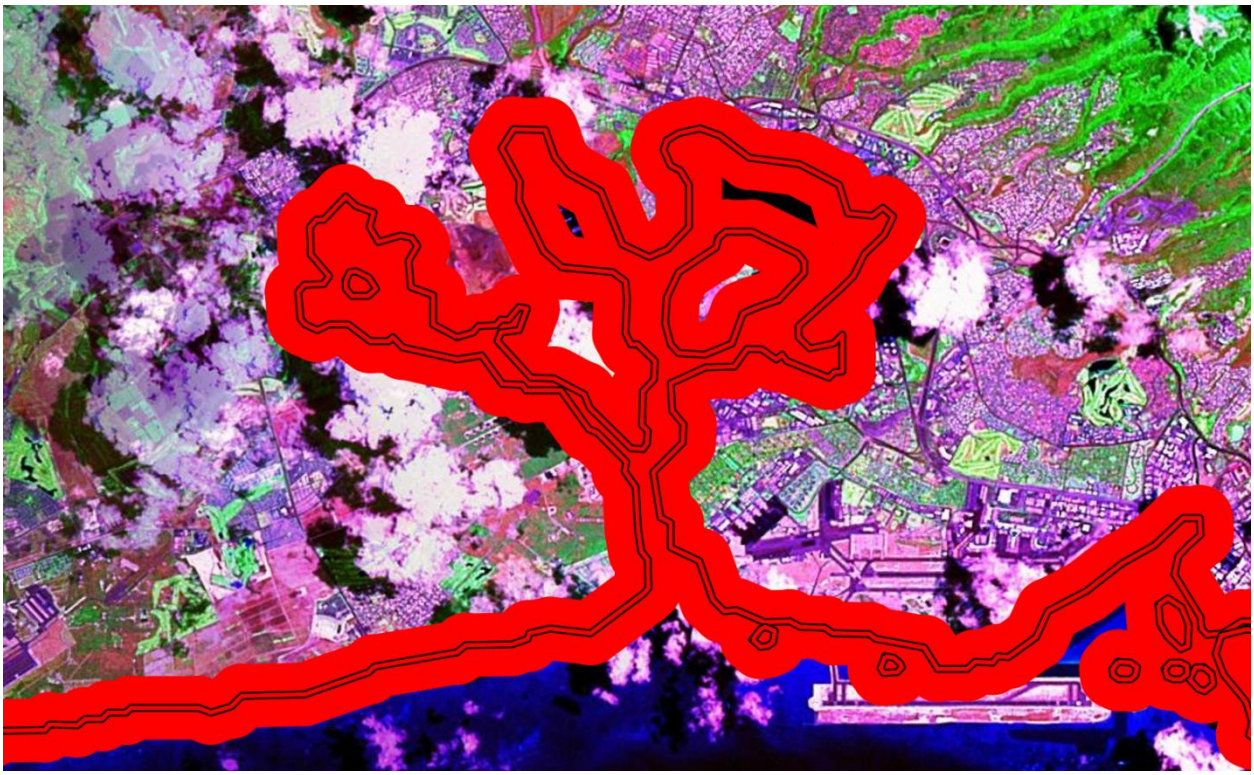

Figure 12. WVS data (center red line) shown over the uncertainty model (thick red line), shown over Landsat 7 imagery (background).

Thus, applying the outlines of the above model and filling them in, most of the shoreline discrepancies between the WVS and the more-accurate imagery are covered by the uncertainty model. In other words, the real shoreline position is very likely to lie within the positional error buffer of the uncertainty visualization.

4.4 Error Meters-to-Pixels Conversion

In order to be shown on a WMS-based display, the distances from the model above are converted from meters to degrees and then from degrees to pixels. This is computed upon every WMS image request, before the uncertainty visualization is drawn. The following describes the conversions.

1. Error (Meters) for Any Data Scale:

The WVS accuracy statement gives an error of 500m for the 1:250K scale, but it does not give the error for the other data scales. However, this can be calculated. The error distance in meters, 'edm', should increase proportionally as data scale decreases, which gives us the formula:

```
edm = (500 / 250,000) x s

where
s = data scale denominator (e.g. 250K => 1/250,000 =>
denominator: 250,000)
```

Programmatically, when the WMS server gets a request for a WVS uncertainty image, the WVS scale fraction is extracted from the request and used as the input, 's', in the above formula.

2. Meters per Degree:

The number of meters in one degree of latitude or longitude depends on the latitude at which the measurement is taken. Because a WMS-based display is used, rectangular areas anywhere on earth i.e. any latitude can be viewed. The rectangle's height defines a latitude range, not a single latitude value. To work around this, the mean latitude inside the rectangle, i.e. the latitude line that divides the rectangle into a top and a bottom half, is used. Thus this mean latitude, in radians, is the input variable, 'L' in the formulas below. The following formulas were obtained by viewing the source code from a National Geospatial-Intelligence Agency (NGA) webpage calculator, in a web-browser. They give the meters per degree, 'mpd', in the latitude and longitude directions, respectively:

$$mpd_{lat} = m_1 + m_2\cos(2L) + m_3\cos(4L) + m_4\cos(6L)$$

$$mpd_{lon} = p_1\cos(L) + p_2\cos(3L) + p_3\cos(5L)$$

where latitude calculation constants

$m_1 = 111132.92;$

$m_2 = -559.82;$

$m_3 = 1.175;$
$m_4 = -0.0023;$

and where longitude calculation constants
$p_1 = 111412.84;$
$p_2 = -93.5;$
$p_3 = 0.118;$

(NGA)

3. Degrees per Pixel

A WMS map view is basically a map image, ‘i’, bounded by an invisible geographic rectangle, ‘r’. The geographic dimensions of this rectangle are specified in (usually automatically) the WMS request. Thus degrees-per-pixel were calculated by the following:

$$dpp_x = (r_{xmax} - r_{xmin}) / i_{width}$$

$$dpp_y = (r_{ymax} - r_{ymin}) / i_{height}$$

where
dpp = degrees-per-pixel in the x- or y-direction of the image space.
r = bounding rectangle defined by an x (longitude) and y (latitude) min and max.
i = width or height pixel dimensions of the map image.

4. Meters per Pixel

Putting steps 2 and 3 together, produces the meters-per-pixel, ‘mpp’ in the x- and y-directions of the image space, respectively:

$$mpp_x = mpd_{lon} \times dpp_x$$

$$mpp_y = mpd_{lat} \times dpp_y$$

5. Error Distance in Pixels

Finally, putting steps 1 and 4 together gets the error distance in pixels, ‘edp’, in the x- and y-directions of the image space, respectively:

$$edp_x = edm / mpp_x$$

$edp_y = edm / mpp_y$

These calculations are applied to convert the error model from section 4.3 into the image space. Thus, edp_x and edp_y define the horizontal and vertical radii, respectively, of the error ellipse from that section, in pixels. This is the proposed model for WVS uncertainty in pixel geometry.

4.5 Uncertainty Visualization Techniques

Visual variables (listed in section 2.4) which to incorporate into the uncertainty visualization techniques of this thesis are now to be chosen. First, the nature of the uncertainty model forces the utilization of size, resolution, shape, and position for the following reasons. For size and resolution, since the uncertainty model is defined by a distance, it must be ensured that the distance represented by the on-screen size of the uncertainty line width is consistent. "Consistent" means that the representation must be accurate no matter what the WVS data scale level or viewing zoom level (e.g. in a WMS client). This is explained further in the zoom-progression figures at the end of this section. For shape, because the data consists of lines and the uncertainty distance is a radius, the uncertainty model will consist of connected, thickened, straight lines that are rounded at their join points (see the pen tip effect described in section 4.3). For position, the uncertainty model's line thickness will always convey a positional buffer zone (within which the shoreline's position is uncertain). Note, the geographical placement of the line is given by the data itself, so position in this sense is not an uncertainty variable. Next, aside from the visual variables inherent to the model, it is expected that the probability of real world shore line positioning decreases as the line width extends from its center. An intuitive technique to convey this information is to use outward gradients in the line width that gradually tweak the visual variables of color lightness, opacity, or both. Finally, it is desirable to further convey to

the viewer that the line width indeed represents uncertainty. To do this, the visual variable of texture was chosen to be applied to the line width. The first texture technique consists of a randomization of dots to which the above gradient idea is also applied. The second texture technique consists of a cross-hatched texture that gradually blurs to create the gradient effect.

For the purposes of modeling and displaying these techniques in practice, a WMS Java servlet was created and titled as the Uncertainty Visualization Toolkit (UVT). For each WVS data scale, it provides rendering of map layers that reflect the techniques and visual variables described above. The following figures show the proposed uncertainty techniques for WVS 1:250,000 scale data. However, the UVT has functionality for showing any scale. The center data line is omitted in these figures to focus on uncertainty. It has been shown in previous figures and does not change. Each technique/layer is shown on top of Landsat 7 imagery for reference.

Figure 13. "Solid": Renders a simple, solid shoreline width buffer. This more-primitive technique was started with to simply put the uncertainty model from section 4.3 into visualization.

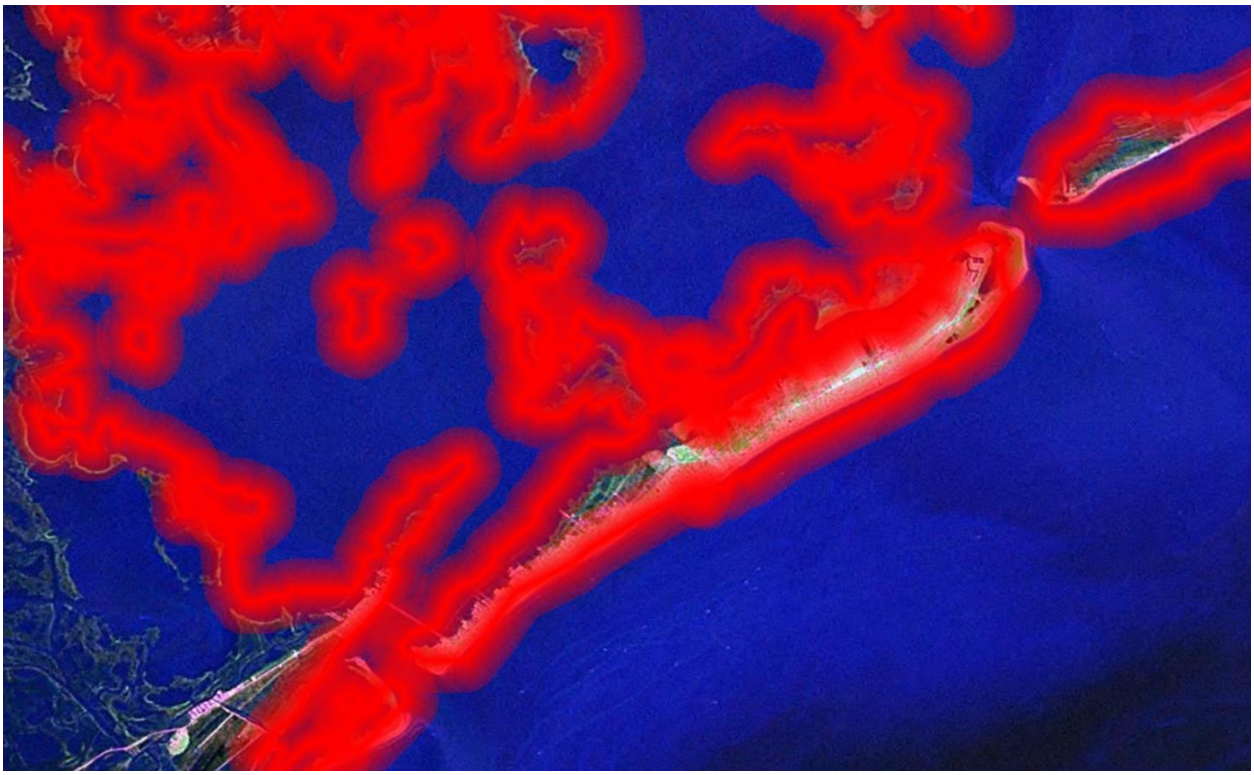

Figure 14. “Transparency Gradient”: Renders a shoreline width buffer that is a solid color at the center and gradually fades to transparent at the edges. Indicates decreasing certainty towards the edges.

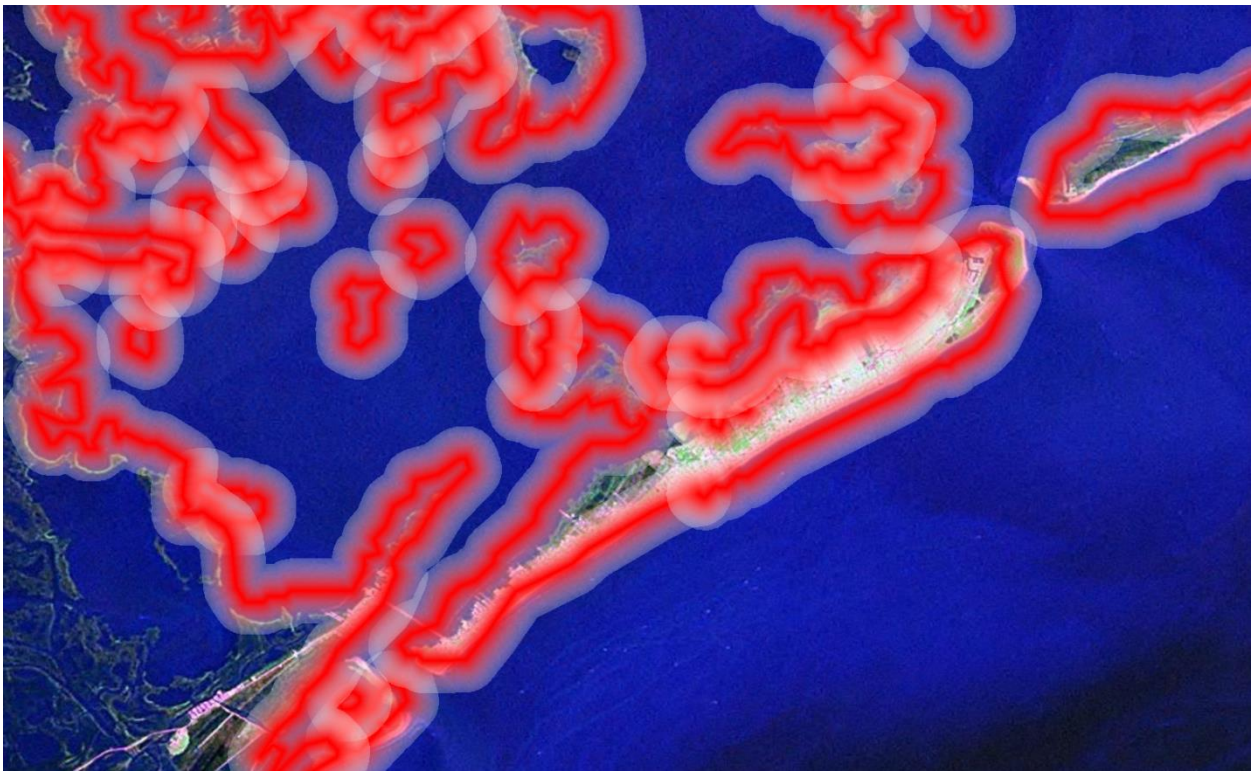

Figure 15. “White Transparency Gradient”: Same as 3, but also graduates color saturation.

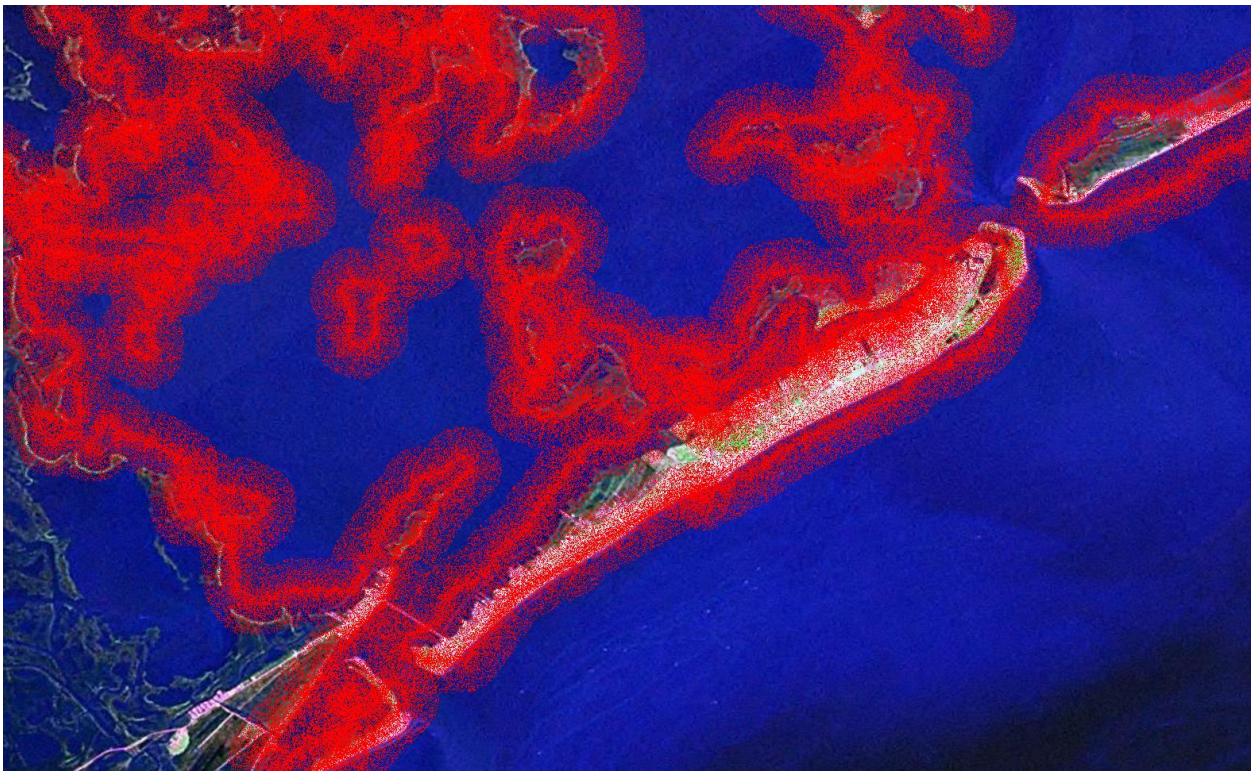

Figure 16. “Dots Texture Gradient”: Renders a shoreline width buffer that has randomized placements of dots that gradually fade from dense at the center of the shoreline width to sparse at the edges. Indicates decreasing certainty towards the edges, with an added visual noise effect to further indicate positional uncertainty.

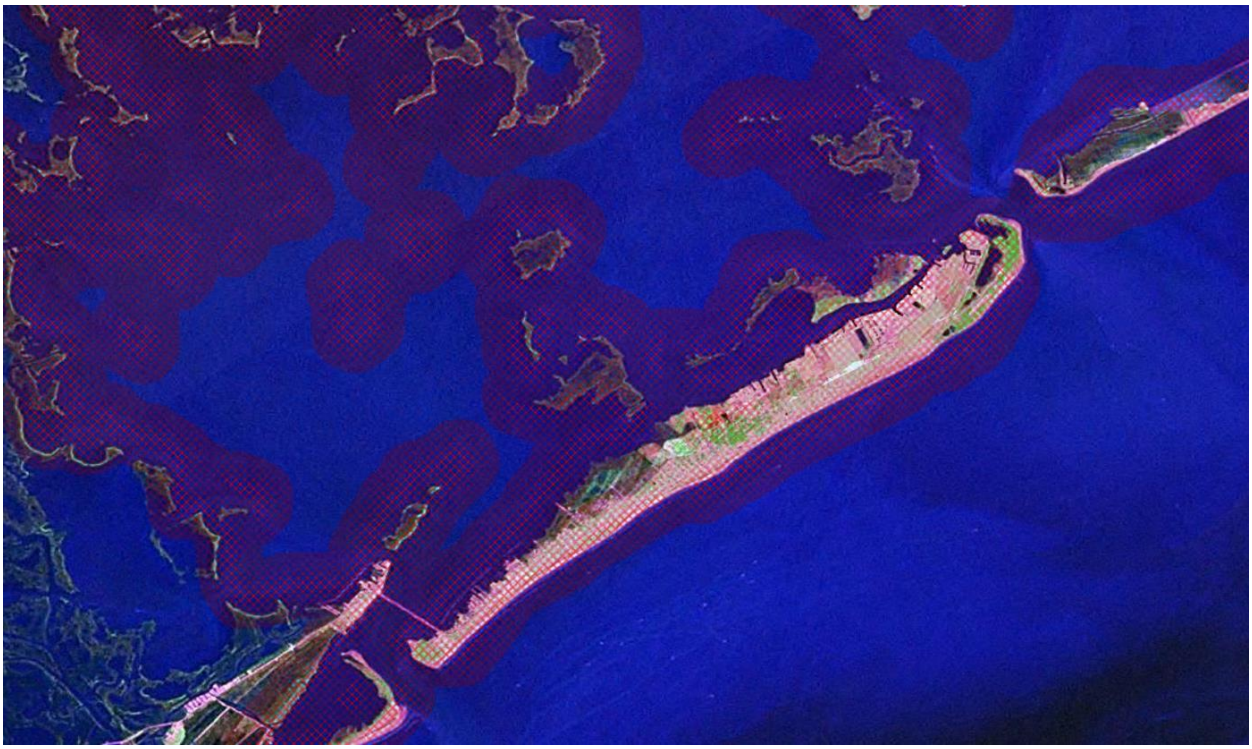

Figure 17. “Grid Texture Gradient”: Renders a shoreline width that is a cross-hatch texture that gradually fades from crisp at the center to blurred at the edges.

Note: The result here is too dark, and time constraints prevented fixing the problem. Something in the blurring process caused this problem.

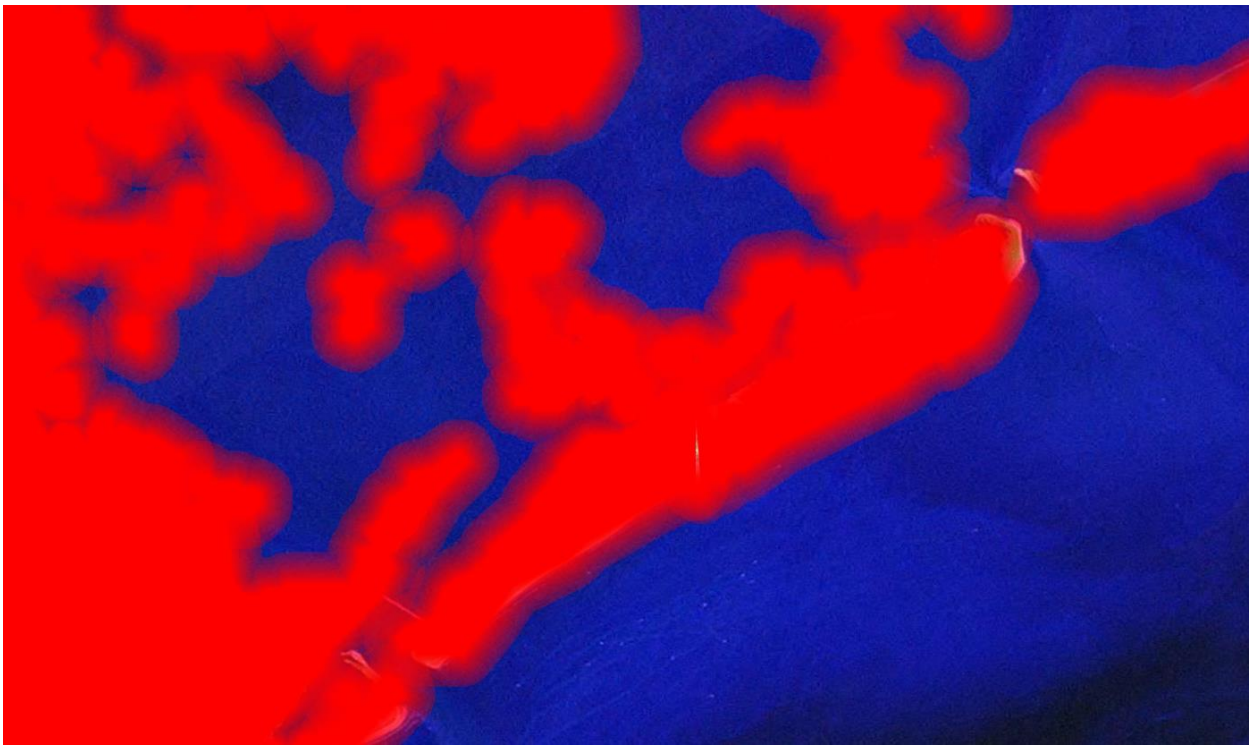

Figure 18. "Land Filling": Land areas are filled in, covering up the inland half of the buffer. The fill can be shown with any of the preceding techniques. It is shown here with "Transparency Gradient". Filling land areas may make more sense to a maritime course plotter who may be more concerned only with seaward half of the uncertainty.

Because these techniques were implemented as a WMS, they can be panned and zoomed on the fly. The following zoom progression example illustrates the size and resolution of the techniques, as mentioned in the beginning of this section. Because the same size and resolution is common to all of the techniques, this zoom progression only uses "Transparency Gradient" as an example.

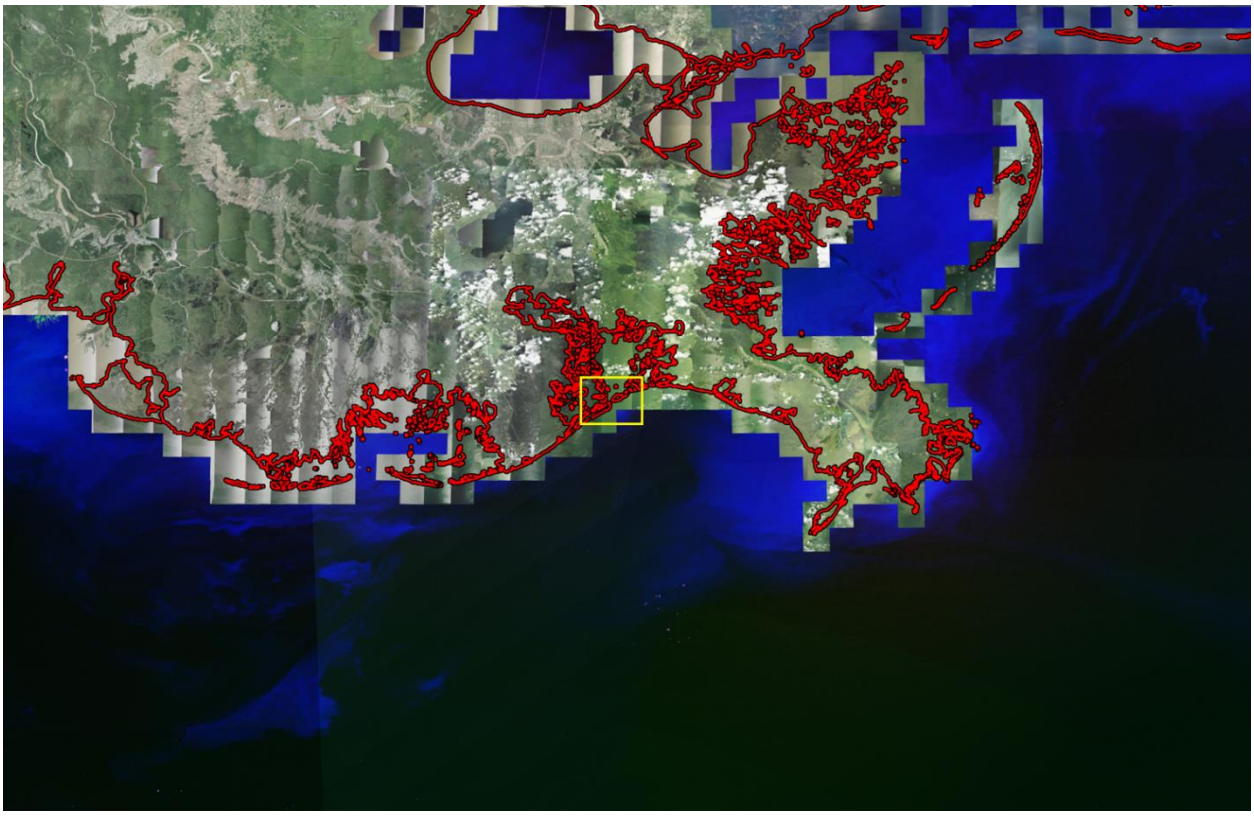

Figure 19. Because the 500-meter uncertainty is represented by such a small pixel-line-width at this zoom level, it is not visible here. Only the WVS data line (red) is visible. The yellow box is zoomed-into to get the next figure.

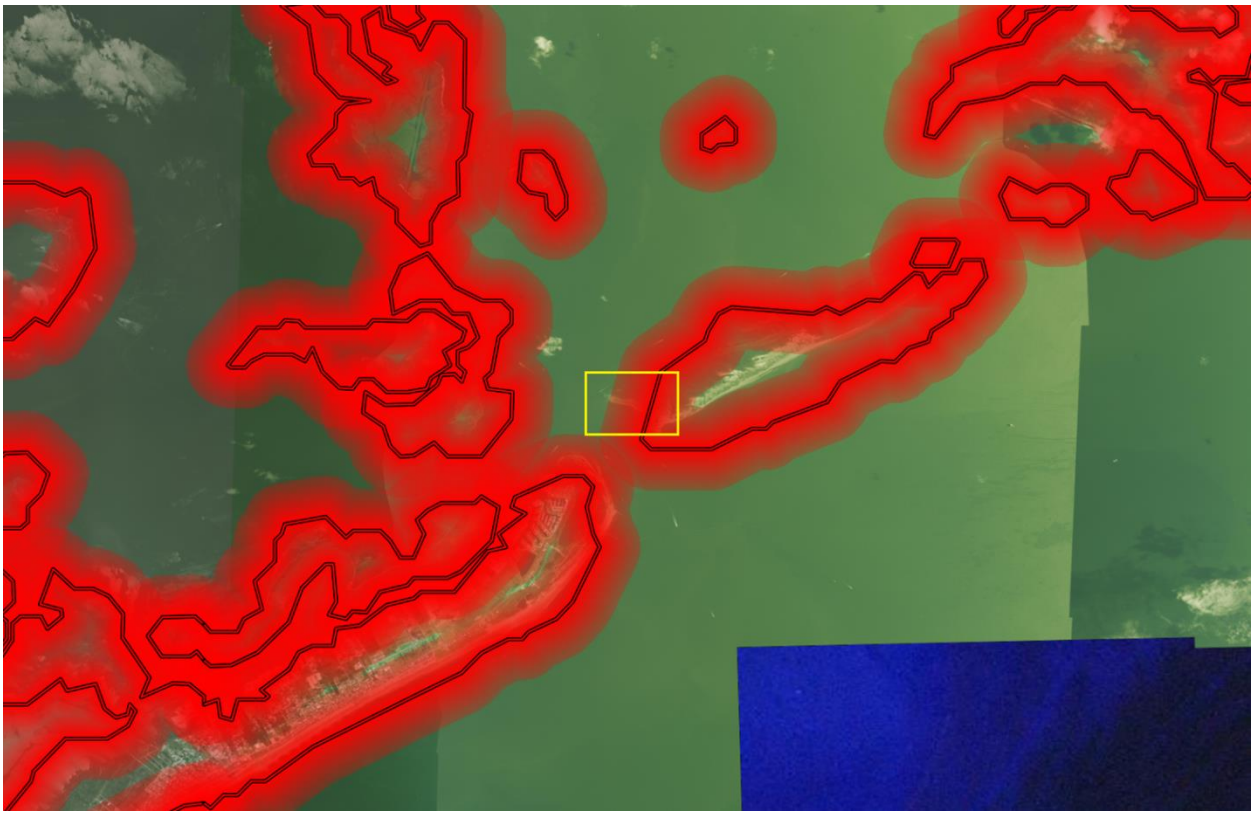

Figure 20. Uncertainty (thick red gradient) is now zoomed enough to see, but it still represents 500 meters. The data line remains the same size, however, because it represents a single shoreline position, not a range of positions like the uncertainty line. The yellow box is zoomed into to get the next figure.

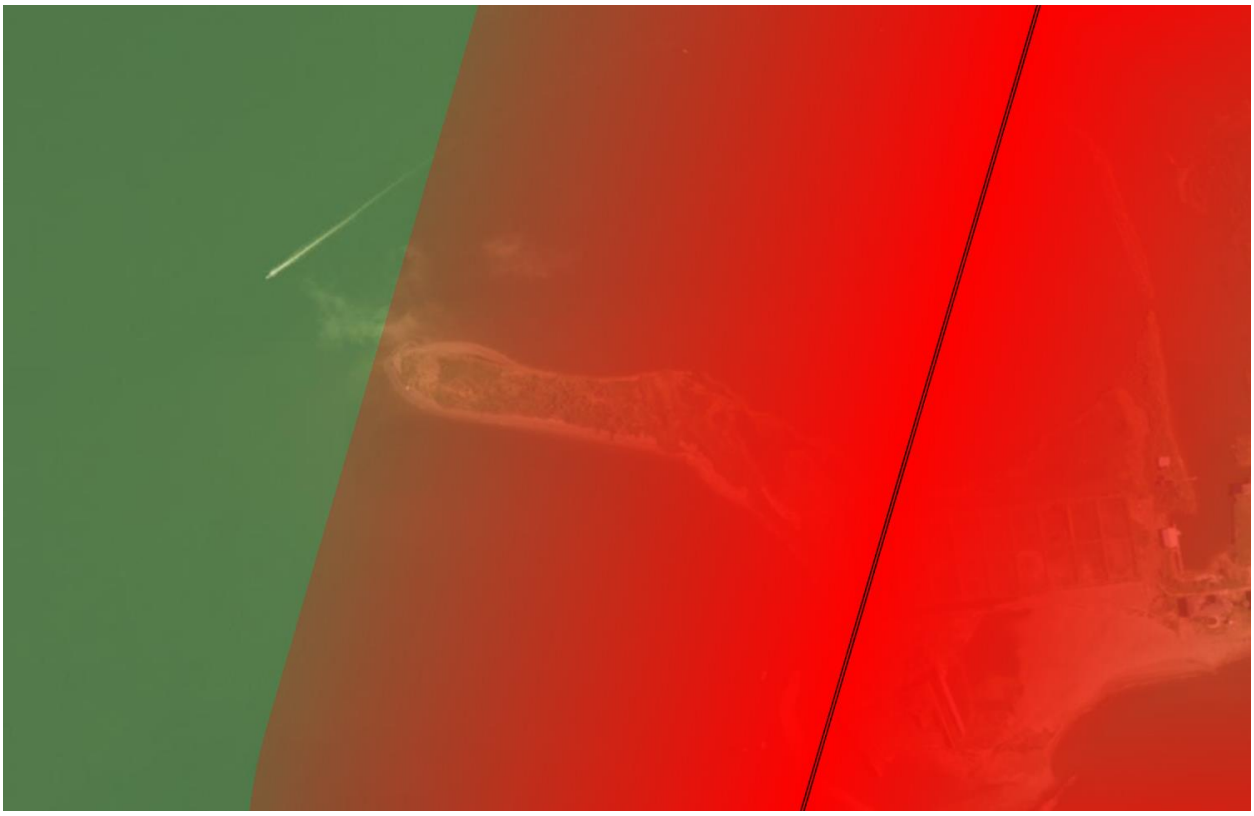

Figure 21. At this zoom level, the 500-meter uncertainty is so large that it nearly fills the image. Again, the size of the data line remains the same. Also, notice that the uncertainty distance accounts for the protrusion of land behind it, in the center of the image.

As for visual variables, zooming shows that the variable of size changes proportional to the viewing zoom-level. Also, the visual variable of resolution becomes constant in this example, at 500 meters. Resolution becomes a variable when the other WVS scales come into play because, as explained before, as scale decreases uncertainty increases.

4.6 Visualization Geometry/Implementation

Rather than provide the lengthy rendering code as to how the UVT was implemented, only key concepts are discussed here. All visualization was implemented using Java's Graphics2D and related classes. In section 4.2 a "pen tip" metaphor was used to explain the

uncertainty model. Rendering this model for the simpler, non-gradient techniques uses a similar yet different idea: the Graphics2D Java class uses a line-width input to draw along the entire line geometry. However, in section 4.4, the conversions arrived at radii defining the uncertainty model's error ellipse in pixels. Thus the error ellipse must be converted to an error line-width (in pixels). This is solved for a given shoreline segment by getting the length of the ellipse diameter that is perpendicular to the segment. The following is the formula for getting a certain ellipse diameter, as depicted by Wolfram MathWorld:

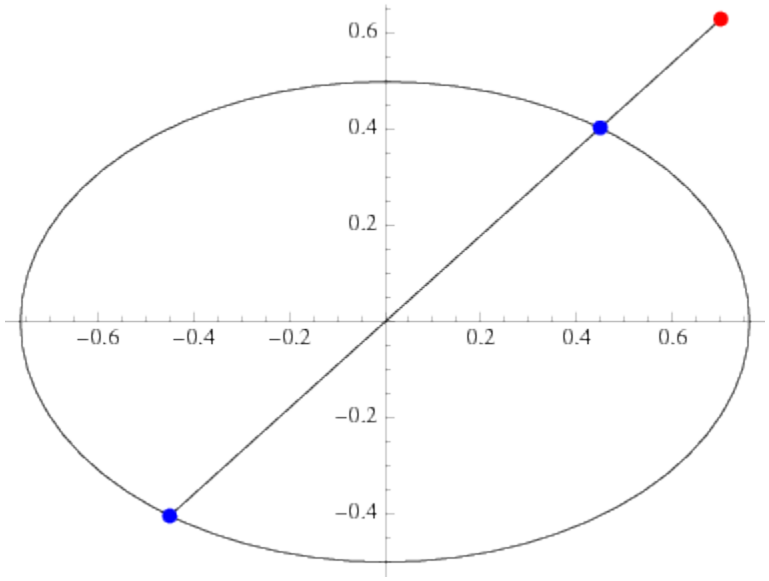

Figure 22. The intersection of an ellipse centered at the origin and semiaxes of lengths a and b oriented along the Cartesian axes with a line passing through the origin and point (x_0, y_0) is given by simultaneously solving the equations

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad y = \frac{y_0}{x_0} x$$

for x and y, yielding the points

$$x = \pm \frac{a\,b}{\sqrt{a^2\,y_0^2 + b^2\,x_0^2}} x_0 \qquad y = \pm \frac{a\,b}{\sqrt{a^2\,y_0^2 + b^2\,x_0^2}} y_0.$$

(Weisstein, 2013)

Thus the distance between these two (blue) points on the error ellipse is taken as the error buffer width in pixels.

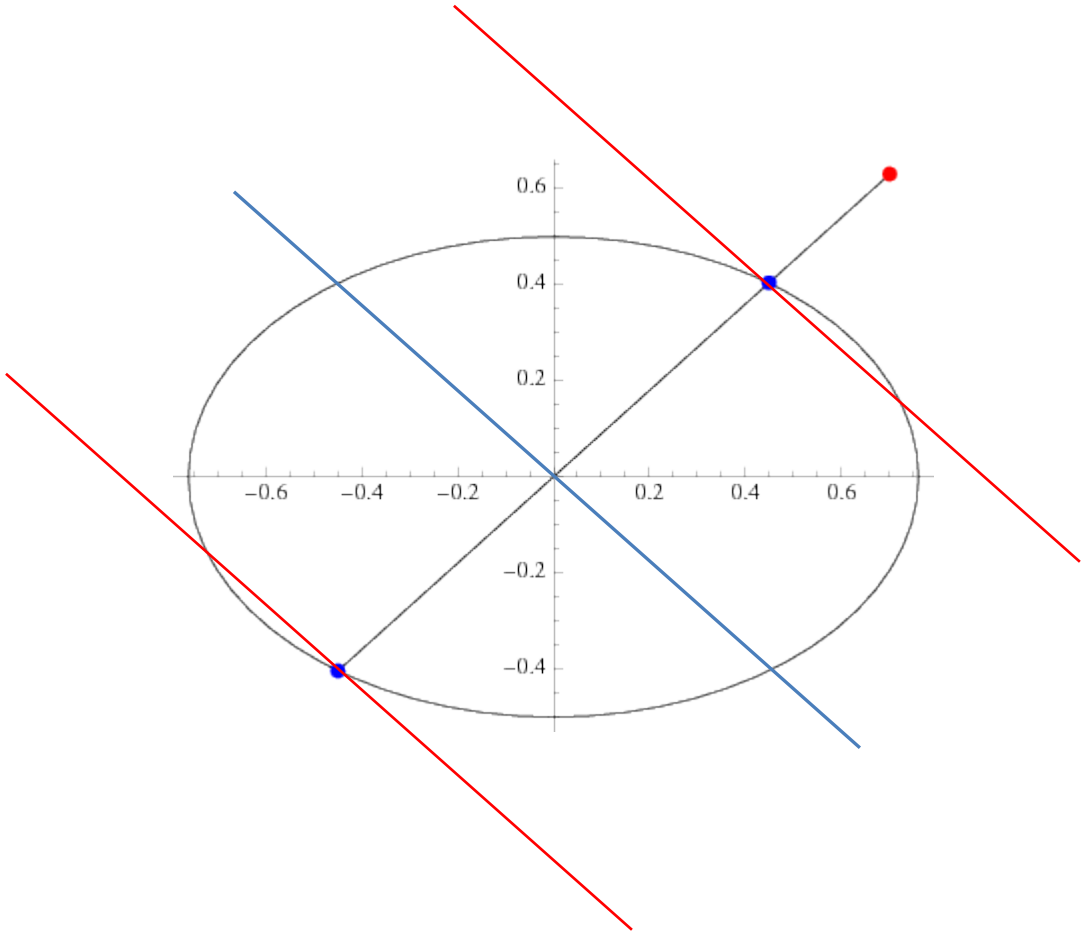

Figure 23. Diagram of shoreline uncertainty visualization shown with the ellipse diameter that defines its width. The blue line represents a shoreline segment. The red lines represent the edges of its uncertainty visualization, or error buffer.

For the gradient techniques, an additional algorithm had to be created and applied to the above geometry. There is no built-in Graphics2D method of drawing gradients along line geometry like was needed. Consider the "Transparency Gradient" technique for example. This effect appears to consist of a single fading, widened line. However, it is actually composed of

several, decreasing line widths with increasing opacities, drawn on top of each other. A pseudo-code algorithm for this example is given in Figure 24.

```
INPUT: lineCoords, lineWidth
lineCoords ← coordinates of lines to draw (in pixels)
lineWidth ← thickness of lines to draw (in pixels)
alpha = 0
alphaIncrement = 1 / floor(lineWidth / 2)
WHILE lineWidth > 0
        set to draw with an opacity of alpha
        set to draw with a line width of lineWidth
        draw a line along lineCoords using above settings
        increase alpha by alphaIncrement
        decrease lineWidth by 2
END WHILE
set to draw with an opacity of 1 (totally opaque)
set to draw with a line width of 1 pixel
draw a line along lineCoords using above settings
```

Figure 24. Algorithm for drawing a line-width transparency gradient.

First, note that the *lineCoords* input represents e.g. WVS data that has already been converted from geographic coordinates to pixel coordinates. Also, *lineWidth* represents e.g. the error distance in meters converted to pixel distance as shown in section 4.4 and the beginning of the current section. The algorithm starts at the full line-width and no opacity. It then draws along the geometry multiple times with a decreasing line width and increasing opacity. This occurs until the smallest possible width (1 pixel) and the greatest possible opacity (not transparent at all) is drawn. The *lineWidth* is decreased by 2 each iteration so that 1 pixel of each side of the last line will show, after being drawn over by the next line. This is what accomplishes the gradient effect. The extra drawing step after the loop is needed because when starting with an even line-width the loop stops after drawing a line-width of 2. The final, totally opaque, 1-pixel line would be skipped if not for the extra step. It should also be noted that the actual implementation used differs slightly from the above in that it starts with *alpha* = .25 and *alphaIncrement* = .75 /

floor(*lineWidth* / 2). This still makes the line's outer edges very transparent, but not to the point where they are invisible.

The same algorithm was used for other gradient techniques, except that different visual variables were graduated. For the "White Transparency Gradient" technique, color saturation is increased while drawing inward, in addition to opacity. For the "Dots Texture Gradient" technique, because space between dots is somewhat randomized, "max space" is the variable decreased while drawing inward. The "Grid Texture Gradient" technique employs the same idea of drawing shrinking, overlapping lines. However, gradually applying this blur is more complex than graduating a single numeric value. Convolution filtering, an image processing concept, was attempted. Because the algorithm may be flawed (the result is too dark; see Figure 17), further explanation is not provided at this time.

5. Human Testing of Techniques

5.1 Overview

To test the effectiveness of the proposed visualization techniques, a usability experiment was conducted. The experiment consisted of LSU students who were timed as they searched for targets in test maps. Due to time constraints, three of the techniques were used in the test maps: "Solid", "Transparency Gradient", and "Dots Texture Gradient". Because map clutter has previously been found to affect visual search performance (Lohrenz & Beck, 2010), clutter was added to the maps as an independent variable. Further, clutter testing is important because these techniques could be applied to existing maps that can already be quite cluttered. The Lohrenz color-clustering metric, detailed in the next chapter, was used as an objective measure of clutter. Expectations before the experiment were that as clutter increases, search response times (RTs, i.e. for targets in maps) also increase. Thus, information (e.g. simulated pressure, wind, and ENC data (NOAA)) was added on top of the visualization techniques in the test maps to simulate common types and amounts of clutter. It was also expected that the visualization techniques themselves would contribute their own clutter.

5.2 Preliminary Clutter Evaluation

Because the experiment focused on the effects of clutter, clutter scores had to first be computed for the many test map images. Although the clutter evaluation results were used in the human-testing part of the experiment (discussed in the sections hereafter), they can also be used as a standalone, analytic evaluation. Clutter in the test maps was measured using the clutter model developed by Lohrenz and Gendron (2008). For a given image, it models "local or global

[clutter] as a function of local or global color density and saliency" (83). The authors define "*local* color density as the density of clustered pixels for a specific feature and *global* color density as the weighted average of densities for all features in the display" (83). Also, they define *local saliency* as "how clearly one color or feature 'pops out' from surrounding features in an image" and *global saliency* as "the weighted average of all the local saliencies for all features in the image" (84). This is a general description of the model; for the full algorithm, see Lohrenz & Gendron, 2008. The actual software implementation of the algorithm that was used was previously written by Todd Lovitt of NRL (2013). In the clutter model, the final, weighted clutter score for an image can range 0-15, with higher scores meaning higher clutter. However, in practice clutter scores tend to range from 0 to around 10 (Michael Trenchard, Naval Research Laboratory, written communication, October 24, 2012). Because we used a large number of test maps (60; discussed in the next section), a very small sub-sample is shown here to briefly convey some important analytic results. The following figures are examples of test maps for the three uncertainty visualization techniques along with their clutter scores.

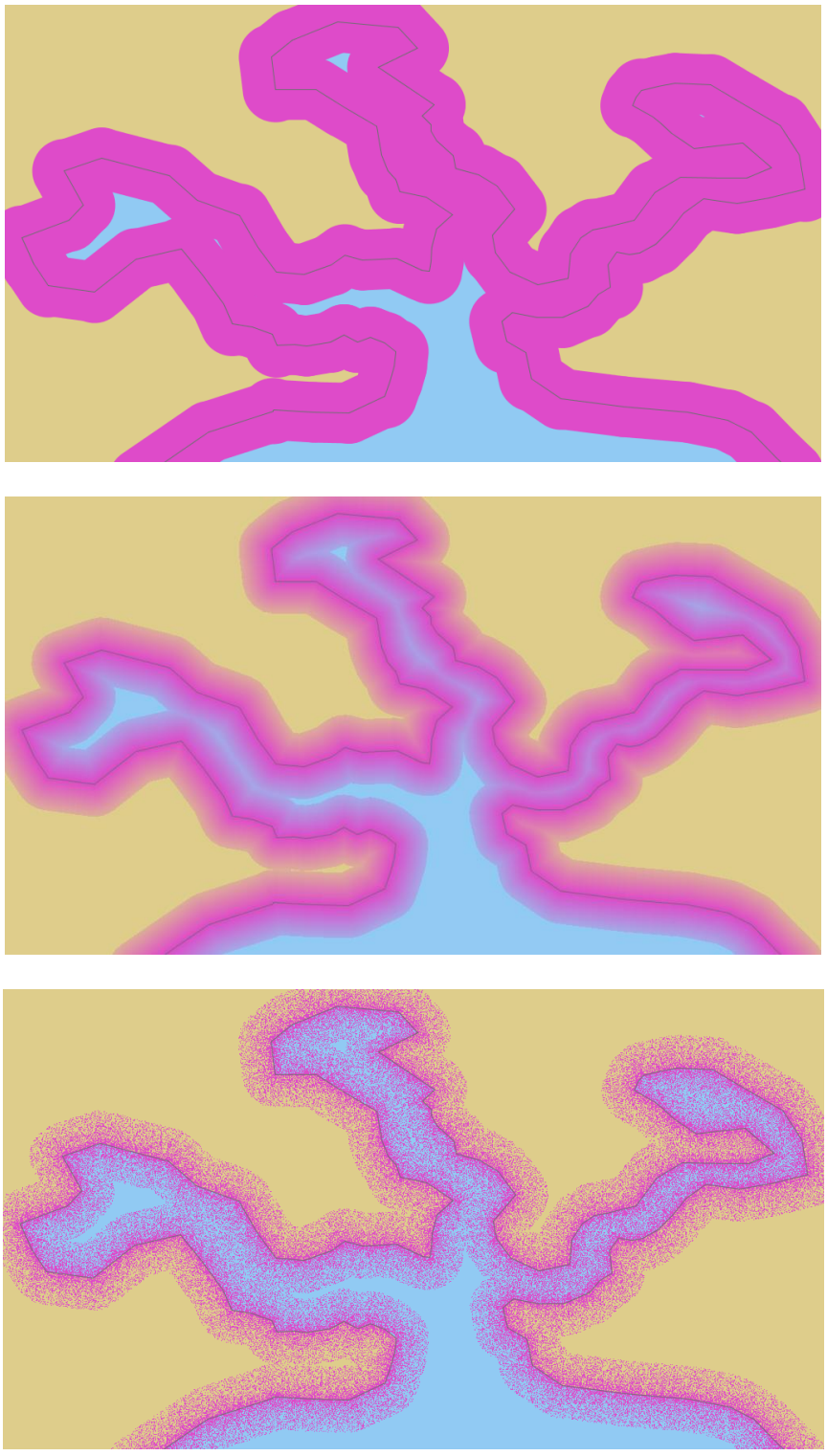

Figure 25. Example test maps and their clutter scores. Top: "Solid" technique: clutter score: 0.49. Middle: "Transparency Gradient" technique: clutter score: 0.24. Bottom: “Dots Texture Gradient” technique: clutter score: 4.35

Clutter is very low for the “Solid” and “Transparency Gradient” techniques but is significantly higher for the “Dots Texture Gradient”. The slightly higher score for “Solid” versus “Transparency Gradient” could be an effect of the clutter algorithm picking up on higher saliency for “Solid”. “Solid” should be more salient because of the abrupt change from tan to pink, whereas “Transparency Gradient” is a gradual change. “Dots Texture Gradient” could have a much higher score because the dots introduce *many* more abrupt changes between tan and pink colors.

5.3 Design

The subjects were eleven LSU students who participated in exchange for extra credit or partial course credit. The independent variables were the following: local clutter around the target (high, low), uncertainty visualization technique (dots, solid, transparent), simulated uncertainty range (250, 500, 750, 1000 meters), and simulated-data-overlay type (none, pressure, wind, ENC, all). The pressure and wind data types were created totally by hand – they were not obtained from real data. However, the ENC data was obtained from a real ENC display image for Pearl Harbor. ENC was used because it, being a vector dataset, is a potential real-world application area for the techniques tested. Using an image editor, the ENC data was reshaped to fit a simulated Pearl Harbor test map. This distinction between data types had an impact on results, described in the discussion section. To create the test maps, first, a map was created using a shoreline similar to Pearl Harbor, with three water inlets extending inland, each opening into a harbor. Each combination of uncertainty technique and range was applied to this map, creating twelve different maps. To add clutter and increase search difficulty, each data type listed above was overlaid onto these maps, creating 60 maps (images). Thus, there were 20 maps for

each technique. Figure 26 shows the least cluttered (“image 1”) and most cluttered (“image 20”) maps for the “transparent” technique.

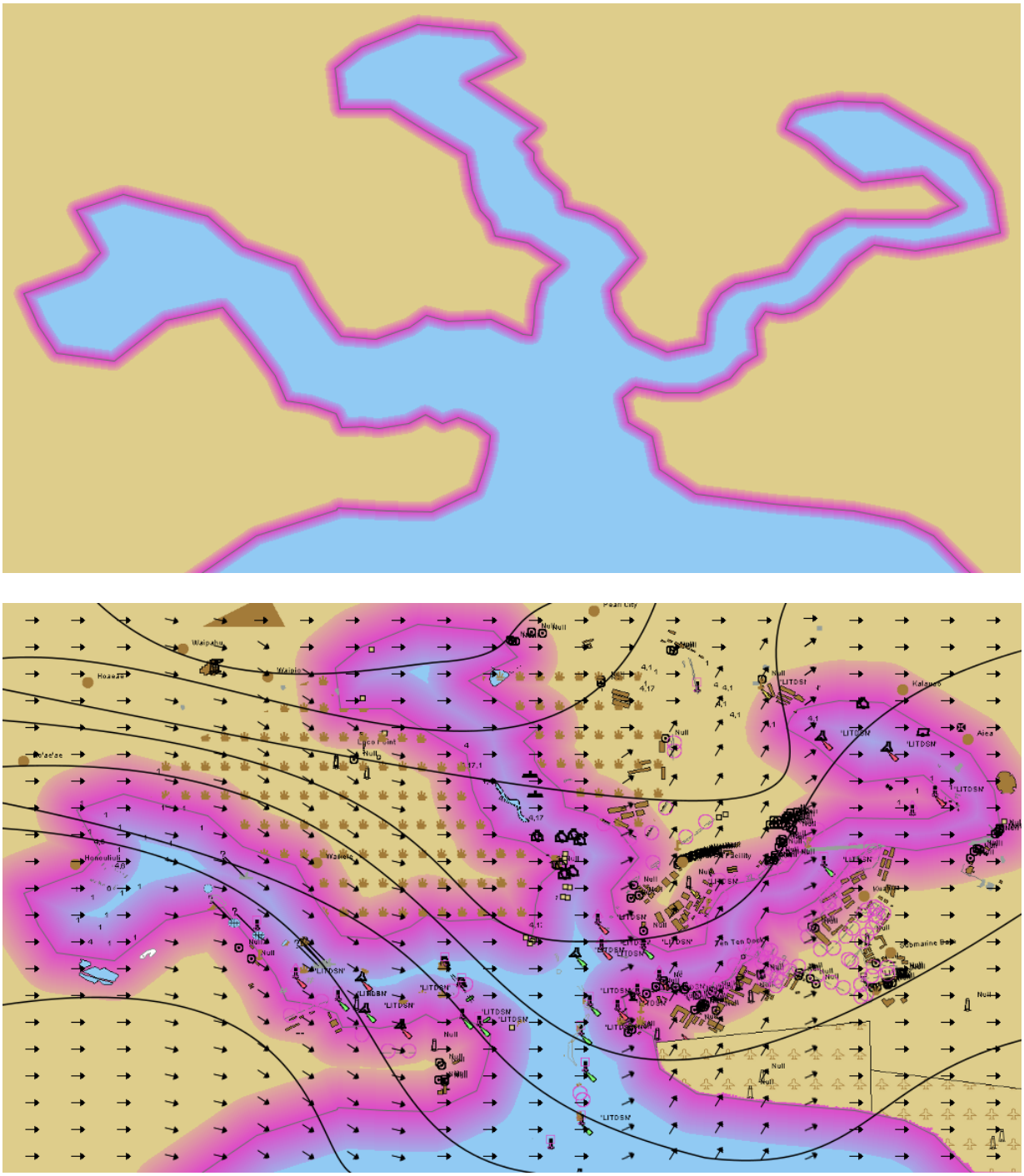

Figure 26. Top: Test map showing the "transparent" technique at uncertainty range "250", with no data added. Bottom: Test map showing the same technique at uncertainty range "1000" with all data types added (pressure, wind, and ENC).

Global clutter values were computed for each of the 60 maps to determine if any uncertainty visualization technique produced a greater level of clutter. A one-way Analysis of Variance (ANOVA) was computed on the global clutter values with technique type as the factor. The "dots" technique contained significantly higher global clutter, while there was no significant difference between "solid" and "transparent" global clutter.

To provide local clutter scores for selecting target locations, each map was divided into a 10 x 6 grid, and a clutter value was computed for each cell. For each map, one cell containing high clutter and one cell containing low clutter were selected to contain a target. This comprised the "local clutter location" variable listed above. Separate maps were created so that each target was shown by itself, for each of the 60 maps, giving 120 maps with one target each. To focus on the difference between visualization techniques, target locations were the same across all three techniques. However, target locations were unique across the other variables (giving 40 unique target locations across all 120 maps). Separate ANOVAs were run on the high and low target clutter values. For the high-clutter target locations, targets in the "dots" maps had significantly more local clutter that both "solid" and "transparent". For the low-clutter target locations, the "dots" local clutter was only significantly higher than the "transparent" local clutter. Three distractors were also placed on each map, spaced away from each other and from the map's target. The test participants were not told about the distractors.

5.4 Procedures

In the first part of the experiment, participants were shown a map with each combination of visualization technique (dots, solid, transparent) and range (100, 200, 300, 400). They were then given a questionnaire in which they indicated the map that they liked the most. This

indication was based on no prior knowledge about the maps. After this, the maps and the concept geospatial uncertainty were explained. The target search task then began, in which the participants searched for the targets described in the above section. The maps with targets were shown one at a time. On each map, when they visually located the target, they pressed a button on their controller. This recorded their *search* duration. They then clicked on the target using a mouse, after which the next map was displayed. The maps were ordered so that very similar information (e.g. the same uncertainty range) was not shown multiple times in a row. This was to avoid speeding/slowing of search times due to priming or interference effects. The participants then completed another questionnaire. Finally, the maps from the first part of the experiment were displayed again. The participants chose the map that they felt best displayed the information and would be best to use if given the opportunity.

5.5 Results

Overall accuracy was very high: on average, subjects accurately located 93% of the targets, equivalent between all visualization types. A 2 X 3 X 4 X 5 repeated measures ANOVA was run on *search* duration with target location (high local clutter, low local clutter), visualization type (dot, solid, transparent), uncertainty range (250, 500, 750, 1000), and data type (map only, pressure, wind, ENC, all) as the factors. The following summarizes the main effects of the different factors/variables on the search RTs.

Search Duration Main Effects:

1. *Visualization technique*. RTs for the "dots" technique were significantly slower than for the "solid" and "transparent" techniques. "Solid" and "transparent" were not significantly different from each other.

2. *Uncertainty range*. The only significant difference between uncertainty ranges was that the largest range (1000) had significantly slower RTs.
3. *Data type*. RTs were significantly slower for maps with ENC data and with all data types (wind, pressure, ENC). There was no significant difference between these two variables.
4. *Target location*. As expected, RTs were significantly slower when the target was located in a high cluttered region of the map, compared to low cluttered regions.

The following summarizes how the search RTs were affected by the variables' statistical interactions with each other. Variables that did not interact are not listed.

Search Duration Variable Interactions:

1. *Uncertainty range and target location.* RTs were significantly slower for maps containing both the highest uncertainty range (1000) and high-clutter target locations.
2. *Data type and target location.* RTs were significantly slower in maps with ENC data and maps with all data types, when the target was in a location of high clutter.
3. *Visualization technique and uncertainty range.* RTs were significantly slower for the "dots" technique at the "1000" uncertainty range.
4. *Visualization technique and data type.* RTs were significantly slower for the "dots" technique when ENC data and all data types were included in the maps.
5. *Uncertainty range and information type.* RTs were significantly slower when ENC data and all data types were included in maps of the "1000" uncertainty range.

The following summarizes the subjects' map preferences that were indicated before and after the target search task.

	Dots	Solid	Transparent
Pre-task	3	5	2
Post-task	1	1	7

Subjects first slightly preferred the “solid” technique. After learning about the maps and using them, they clearly preferred the “transparent” technique. No collective preference for uncertainty range was shown.

5.6 Discussion

The “dots” technique introduced the most clutter to the map and was the driving force behind slowed search times. This technique, due to its random, noisy effect, seems to be inherently cluttered. This could explain why it significantly slows search times when it is more prominent (i.e. at the “1000” range). Map clutter also increased with uncertainty range and addition of data. The main effect of data type suggests that ENC data made the target search more difficult due to a) the ENC data being more realistic than the wind and pressure data, b) the ENC data introducing more clutter than the wind and pressure data, or c) both. This finding has foreseeable implications: applying these techniques to ENC data for real-world use should be approached with caution regarding clutter. This may be critical because ENC displays can be even more cluttered than the one used in this experiment. The main effects were driven by the variable interactions. Detection was slower when there was more information surrounding the target, making it harder to distinguish from other objects. Interactions 1 through 4 further demonstrate that search times slow down when clutter is added, i.e. in the form of data overlays.

While these overlays increased latencies for each technique, the effect was most drastic for the "dots" method. Overall, the map-reading difficulty introduced by the "dots" technique makes it the worst candidate of the three for uncertainty visualization, in terms of clutter (objectively). The "solid" and "transparent" techniques were roughly tied in this regard. However, from the pre- and post-map selection, the subjects' preference changed from "solid" to "transparent" (subjectively) after they learned about maps. Thus it could be said that the "transparent" technique performed the best in this experiment, followed by "solid", with "dots performing worst.

6. Conclusion

For visualizing positional shoreline uncertainty, our usability experiment found a technique employing variable opacity to be most effective. Table 4 compares the results of the chapter 5 experiment with results of previous visual variable experiments. Table 2, explained in the background chapter, is repeated here, with the experimental results of this thesis added as a separate line at the bottom (red). This line only includes techniques in which a visual variable was added, namely, the "transparent" (opacity) and "dots" (texture) techniques. Because the inherent variables (size, resolution, shape, and position; section 4.5) are present in all three techniques, they are not included here. Thus, since the "Solid" technique contains no added visual variable, it is not included in bottom line of the table.

		Uncertainty Visualization Techniques (intrinsic only)					
		Color Saturation	Color value/ Lightness/ Darkness	Texture	Opacity	Legend	Grayscale
Studies	[21]	2nd of 2	1st of 2				
	[15]	3rd of 3		1st of 3			
	[5]	5th of 5			2nd of 5		
	[6]		1st of 2			2nd of 2	
	[9]	3rd of 3	1st of 3	2nd of 3			
	[10]		3rd of 5	1st of 5			
	[7]	1st of 2					
	[3]						1st of 3
	[1]		1st of 2				
	Barré			3rd of 3	1st of 3		

Table 4. Repeated here is Table 2, explained in section 2.4, with the experimental results of this thesis added as a separate line at the bottom (red).

The results of this thesis somewhat agree with previous work in terms of the effectiveness of opacity. Only one study (Drecki, 2002) in MacEachren's survey (2005) tested opacity. However, Drecki ranks opacity highly (2^{nd} of 5 techniques), as does this thesis (1^{st} of 3 techniques). These promising but few results could motivate further testing with opacity. On the other hand, the results of this thesis somewhat disagree with previous work in terms of the effectiveness of texture. The three previous studies here that tested texture have ranked it highly (1^{st} or 2^{nd}) while this thesis ranks it last of three techniques (by a large margin). Perhaps this is because clutter is such a large component of the "dots" technique, while the other studies did not focus on clutter.

Although this thesis can be compared to other studies in terms of visual variables, it sets itself apart in two main ways. First, the model created here visualizes uncertainty in the (x, y), or *positional,* dimension which no other studies have done. Further, this model can help mission planners to visually, automatically, and interactively realize the presence and distance of map uncertainty. Second, the experiment conducted for this thesis focuses on the effect of clutter, which few other studies in the area of geospatial uncertainty visualization seem to have done. While such an experiment is very useful, it also leaves much room for future experimentation. For example, because this thesis lends itself to applications like naval mission planning, an obvious future experiment could involve participants plotting ship courses on test maps. This could test the impact of different uncertainty visualization techniques on criteria like safety and decision making. NRL plans on starting experiments like these in the near future.

On the subject of future directions, there are also plans for extending these techniques to more data. If used for real applications, the techniques should be extended to point and area geometry, which are present in most vector datasets. As for datasets, there is current interest in applications for vector chart products like NGA's Digital Nautical Chart (DNC), for naval

navigation. Also, we have begun work on visualizing uncertainty for bathymetry data and Landsat 7 imagery. For bathymetry data, our preliminary work has focused on visualizing varying depth uncertainty. To do this we place the data on top of a black and white checkered pattern and then vary the visualization opacity based on the uncertainty (Figure 27).

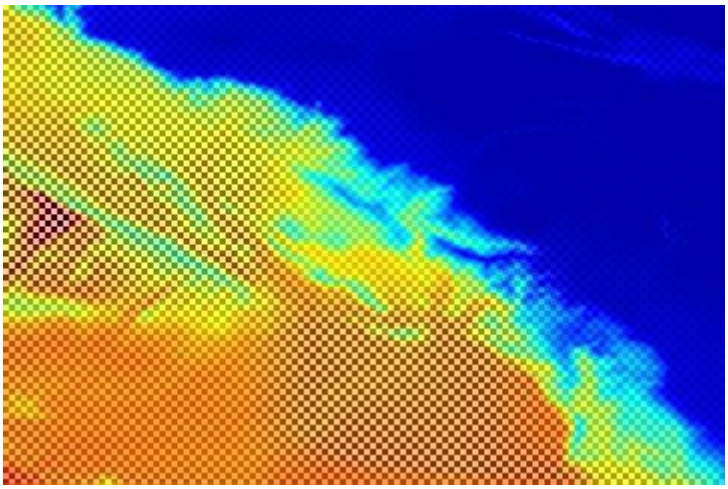

Figure 27. Visualization of varying depth uncertainty. At areas of higher uncertainty, the checkerboard pattern shows through by adding transparency to the bathymetry visualization (colors based on depth).

In preliminary work with Landsat 7 imagery, because imagery fills the map image with data (as opposed to vector data which contains features and blank space), we began by overlaying the entire image with error ellipses defined by the Landsat 7 accuracy specification (250m error; see Figure 28). The overlay is aligned the same way to every map image, but the error ellipses resize as the user pans and zooms in WMS. A graticule was also added to the overlay to provide positional reference as to the sizes of the error ellipses.

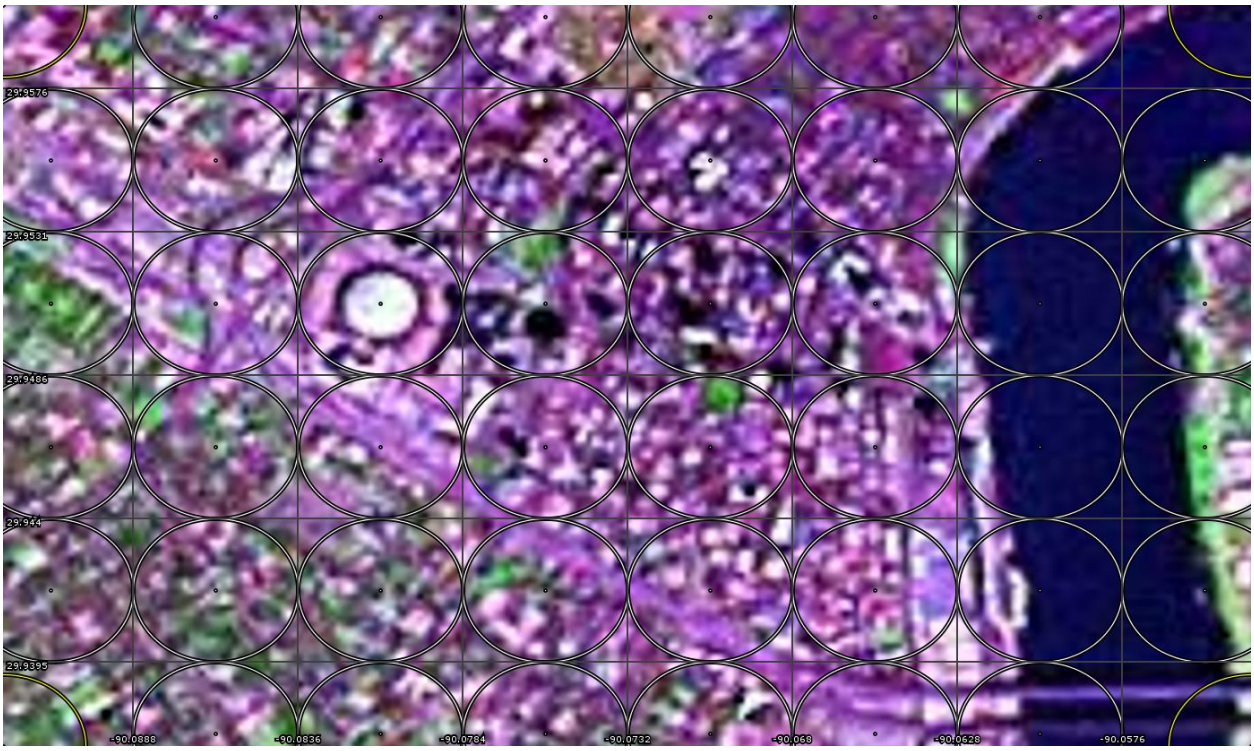

Figure 28. Landsat 7 imagery overlaid with a graticule and error ellipses whose size is defined by the Landsat 7 accuracy specification (250m error).

Thus, there is now some groundwork in place for the visualization of positional geospatial uncertainty. It is expected that work in this area expands as it continues to satisfy needs in digital mapping.

References

[1] Aerts, J. C. J. H., K. C. Clarke, and A. D. Keuper. 2003. Testing popular visualization techniques for representing model uncertainty. Cartography and Geographic Information Science 30(3): 249-61.

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

[3] Blenkinsop, S., P. Fisher, L. Bastin, and J. Wood. 2000. Evaluating the perception of uncertaintyin alternative visualization strategies. Cartographica 37(1): 1-13.

[4] Buttenfield, B. P., and R. Weibel. 1988. Visualizing the quality of cartographic data. Presented at Third International Geographic Information Systems Symposium (GIS/LIS 88), San Antonio, Texas.

[5] Drecki, I. 2002. Visualisation of uncertainty in geographical data. In: W. Shi, P. Fisher, and M. Goodchild (eds), Spatial data quality. London: Taylor & Francis. pp. 140-159, color plates: 141-143.

[6] Edwards, L. D., and E. S. Nelson. 2001. Visualizing data certainty: A case study using graduated circle maps. Cartographic Perspectives 38:19-36.

[7] Evans, B. J. 1997. Dynamic display of spatial data reliability: Does it benefit the user? Computers & Geosciences (special issue on Exploratory Cartographic Visualization) 23(4): 409-22.

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

[9] Leitner, M., and B. P. Buttenfield. 2000. Guidelines for the display of attribute certainty. Cartography and Geographic Information Science 27(1): 3-14.

[10] Livingston, M. A., Decker, J., & Ai, Z. (2011). An Evaluation of Methods for Encoding Multiple, 2D Spatial Data. *SPIE-IS&T, 7868*, 1-12.

[11] Lohrenz, M. C., & Beck, M. R. (2010). Evidence of clutter avoidance in complex scenes. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting.* San Francisco, CA.

[12] Lohrenz, M. C., & Gendron, M. L. (2008). A 3D Clustering Algorithm to Model Clutter in Electronic Geospatial Displays. *The Journal of Management and Engineering Integration, 1*(2), 83-88.

[13] Lovitt, T. (2013). Color Clustering & Statistics [computer program]. Naval Research Laboratory, Stennis Space Center, Mississippi.

[14] MacEachren, A. M. 1992. Visualizing uncertain information. *Cartographic Perspectives* (13): 10-19.

[15] MacEachren, A. M., C. A. Brewer, and L. W. Pickle. 1998. Visualizing georeferenced data: Representing reliability of health statistics. Environment and Planning: A 30: 1547-61.

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

[17] NGA. (n.d.). *Length of a Degree of Latitude and Longitude*. Retrieved September 19, 2013, from National Geospatial-Intelligence Agency: http://msi.nga.mil/MSISiteContent/StaticFiles/Calculators/degree.html

[18] NOAA. (2013). *Specifications and Deliverables.* Retrieved September 19, 2013, from Office of Coast Survey: http://www.nauticalcharts.noaa.gov/hsd/specs/specs.htm

[19] NOAA. (n.d.). *Electronic Navigational Charts: NOAA ENC®*. Retrieved October 7, 2013, from NOAA Office of Coast Survey: http://www.nauticalcharts.noaa.gov/mcd/enc/

[20] Robinson, A. H., Morrison, J. L., Muehrcke, P. C., Kimerling, A. J., & Guptill, S. C. (1995). Elements of Cartography. John Wiley & Sons, Inc.

[21] Schweizer, D. M., and M. F. Goodchild. 1992. Data quality and choropleth maps: An experimentwith the use of color. In: Proceedings, GIS/LIS '92, San Jose, California. ACSM andASPRS, Washington, D.C. pp. 686-99.

[22] Soluri, E. A., & Woodson, V. A. (1990). World Vector Shoreline. *International Hydrographic Review, LXVII*(1).

[23] Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185: 1124-31.

[24] USGS. 1997. *Spatial Data Transfer Standard (SDTS): Logical specifications*. U.S. Geological Survey, Virginia.

[25] USGS. (2013, September 3). *Landsat Processing Details*. Retrieved September 19, 2013, from USGS: http://landsat.usgs.gov/Landsat_Processing_Details.php

[26] Weisstein, E. W. (2013, September 16). *Ellipse-Line Intersection*. Retrieved September 19, 2013, from Wolfram MathWorld: http://mathworld.wolfram.com/Ellipse-LineIntersection.html

Vita

The author was born in Metairie, Louisiana, where he graduated from Archbishop Rummel High School in 2006. He obtained his bachelor's degree in computer science from the University of New Orleans in 2011. Prior, in 2010, he began a computer science internship at the Naval Research Laboratory, Stennis Space Center, in its Geospatial Computing Section. He continued this internship while continuing at the University of New Orleans in pursuit of a master's degree in computer science.