Visualizing Uncertainty Metrics Across Multiple Attribute Resolutions

Kate Carlson [a]*, Barbara P. Buttenfield [a], Yi Qiang [b]

[a] Department of Geography, University of Colorado, Boulder, USA
[b] School of Geosciences, University of South Florida, USA
* Kate.Carlson@colorado.edu

Keywords: Uncertainty, Wetlands, Classification, Attribute Resolution, Confusion Matrix

Introduction

Uncertainty is inherent within geospatial data. Knowing the levels of uncertainty and its patterns of variation across a region of interest is vital for correct analysis and decision making. Quantification and communication of uncertainty is important to establish reliability and confidence in the analysis. The work outlined in this paper evaluates attribute uncertainty present at two levels of wetlands classification.

Uncertainty visualization techniques have typically displayed both the data and uncertainty using visual variables (Kinkeldey et al., 2014). Embedding uncertainty within the visualization rather displaying them separately has been found to facilitate decision making (Leitner & Buttenfield, 2000; Aerts et al., 2003). Several user studies have analyzed the effectiveness of different visual variables for uncertainty visualization (Leitner & Buttenfield, 2000; Drecki, 2002; MacEachren et al., 2012). Fuzziness and color value have been rated as logical and effective visual variables. Confusion matrices are an established method of quantifying and visualizing attribute uncertainty from misclassification. This technique is most commonly used in remote sensing to quantify misclassification rates. This work merges confusion matrices with visual variables to display attribute uncertainty spatially.

Wetlands provide a variety of functions and support ecosystem processes such as wildlife habitat preservation, water purification, flood control, and carbon storage (Millennium Ecosystem Assessment, 2005). However, they are under stress of degradation and destruction due to multiple human and natural factors (McCauley, Jenkins, Quintana-Ascencio, 2013; Tiner, 2005). Between 1780 and 1980, the United States lost 53% of its original wetlands (Dahl, 1990). Research on the uncertainty of wetland classification is important for understanding how wetlands are defined and delineated spatially, as well as reliably identifying attribute errors arising from temporal changes.

Wetlands are defined as areas where the ground is saturated with water for varying periods of the year. This definition includes marshes, swamps, bogs, and fens and influences the types of animal and vegetation communities living within the area (Cowardin et al., 2005). Hydrologic, geomorphologic, chemical, and biological factors influence how a wetland system is classified. The National Wetlands Inventory (NWI), created by the US Fish and Wildlife Service (FWS), maintains the largest database of wetlands in the United States using a Cowardin classification method (Dvorett et al., 2012). The classification includes five main wetland systems: palustrine, riverine, lacustrine, marine, and estuarine (Cowardin et al., 2005). This classification is appropriate for tracking wetland loss over time; however, it does not allow for the assessment of how well a wetland can perform ecological functions such as water purification or storage. (EPA, 2002; Smith et. al., 2013). A hydrogeomorphic (HGM) classification system is a commonly used alternative to analyze wetland functionality (Brinson, 1993; EPA, 2002). The HGM classes reflect wetlands, terrain, and soil characteristics: riverine, depression, slope, organic soil flat, mineral soil flat, estuarine fringe, and lacustrine fringe (Smith et. al., 2013). This research will compare the Cowardin (referred to for brevity as NWI) and HGM classification systems to quantify uncertainties present at two attribute resolutions and within different landscape types.

Methods

The main study area in inland Louisiana covers 2,000 square miles (1.28 million acres) (Figure 1A). This area includes ten parishes, the city of Baton Rouge, the Mississippi River, and Lake Maurepas. According to the NWI classification, three wetland types are found within this area: 566,902 acres of palustrine (44.3%), 217,458 acres of riverine (17.0%), and 68,720 acres (5.4%) of lacustrine wetlands. Covering nearly 67% of the study area, wetlands clearly dominate, underscoring the need for reliable attribution of wetlands.

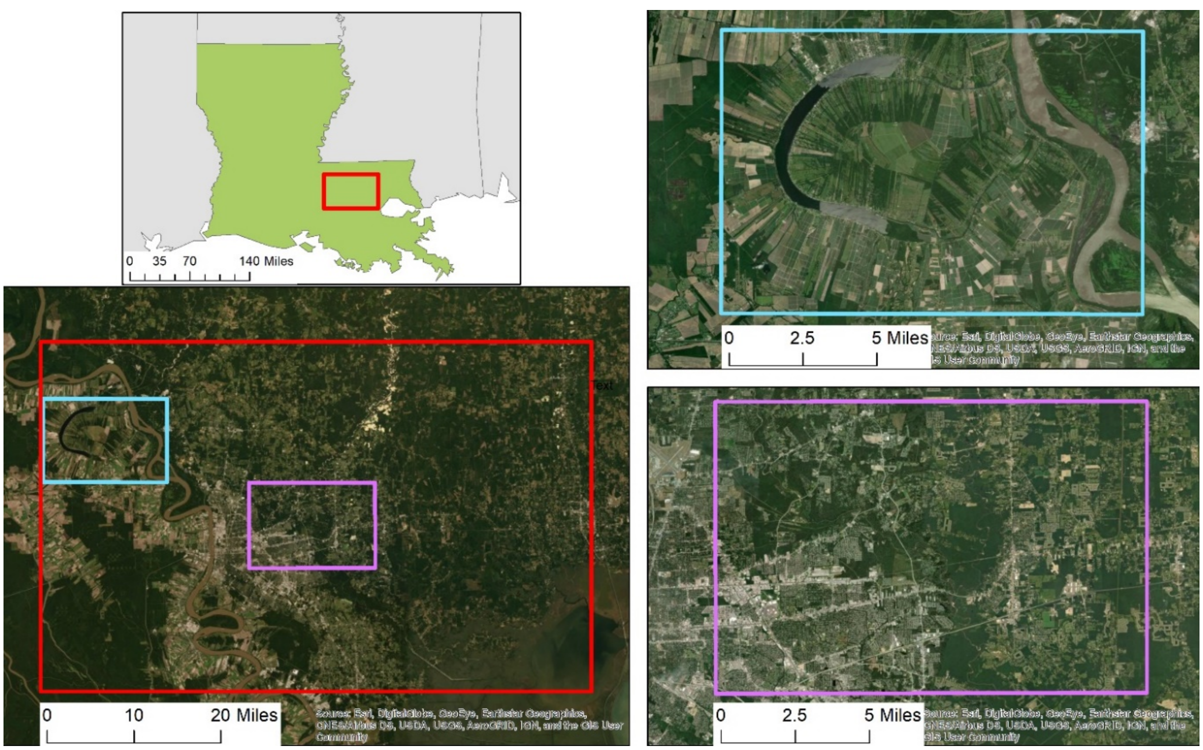

Figure 1 A: The red outline shows the 2,000 square mile study area located in Louisiana. Blue and purple outlines are subsets of the study area B: Shows the study area subset focused on False River. C: Shows the second study area subset focused on Denham Springs.

Two smaller subsets were chosen to focus on different landscape types in the study area. The blue outlined subset (Figure 1B) contains False River, an oxbow lake, in the western section as well as the Mississippi River in the eastern portion. This area has complex topography resulting from a Mississippi River meander. Land use patterns around the oxbow indicate shared access to fresh water, likely to support agriculture. The purple outlined subset (Figure 1C) is located east of Baton Rouge and is centered over the settlement of Denham Springs including areas of urban-rural interface.

Table 1- Data Layers

Data Layer	Source	Resolution	Years collected
Elevation	USGS	3 meter	2016, 2018
Hydric soils	NRCS	1:20,000	2019
Hydrography	NHD-USGS	1:24,000	2000-2019
Wetlands	NWI-FWS	1:65,000	1970s,1980s, 2010s

The data sources for this project are outlined in Table 1. The analysis compares attribute uncertainty (misclassification) for NWI and HGM wetlands classes. Through this analysis, NWI data will be treated as the validation dataset. For the initial pass, we examined a coarse attribute resolution, binary distinction between wetlands and non-wetlands. In the second pass, we analyzed attributes at a finer resolution, looking at lacustrine, riverine, and palustrine wetlands. Both the HGM and NWI systems include lacustrine and riverine classifications. They are defined by adjacency to a lake (lacustrine) or a river (riverine) with each landform providing the main input of water into the wetland (Smith et. al., 2013). HGM does not have a palustrine (vegetated wetlands) class but does have additional classes based on topography such as depressions and mineral flats. In order to compare the two classification methods, the HGM is simplified: wetlands that are not lacustrine or riverine are considered to be “palustrine”.

A buffer of 100m was applied to NHD rivers and lakes, and then overlaid with the hydric soils data. Areas adjacent to rivers (i.e., within the buffer) with hydric, predominately hydric, and partially hydric soils were classed as riverine. Areas adjacent to lakes with hydric, predominately hydric, and partially hydric soils were classed as lacustrine. The remaining areas of hydric, predominately hydric, and partially hydric soils that were not adjacent to either rivers or lakes were classed as palustrine. Areas with either non-hydric or predominately non-hydric soils were considered non-wetlands. The NWI data included open water in its wetland classification while the HGM did not. As a consequence, open water was removed from the NWI data to prevent a large number of pixels being classified as false negatives.

Uncertainty analysis was undertaken using confusion matrices, comparing HGM (test) data with NWI (validation) data to identify areas and types of misclassification at coarse and finer attribute resolutions. Pixels that are classified the same in both datasets are treated as true positives or true negatives, and all other pixels are treated as false. Figure 2 shows a confusion template to explain the methodology. The first column (red) shows false negatives (wetlands in NWI but not in HGM) for all categories while the top row (yellow) shows false positives (wetlands in HGM but not in NWI). The diagonal (orange) shows cells that agree on the presence/absence of wetlands in both data sets as well as on the type. The six gray cells indicate misclassification at the finer attribute level (wetland type) but agreement on the presence of wetlands (so no misclassification at coarser attribute resolution). Percentage values in the Results section report the proportion of pixels in each of the sixteen cases for the full area or for the subset areas, respectively.

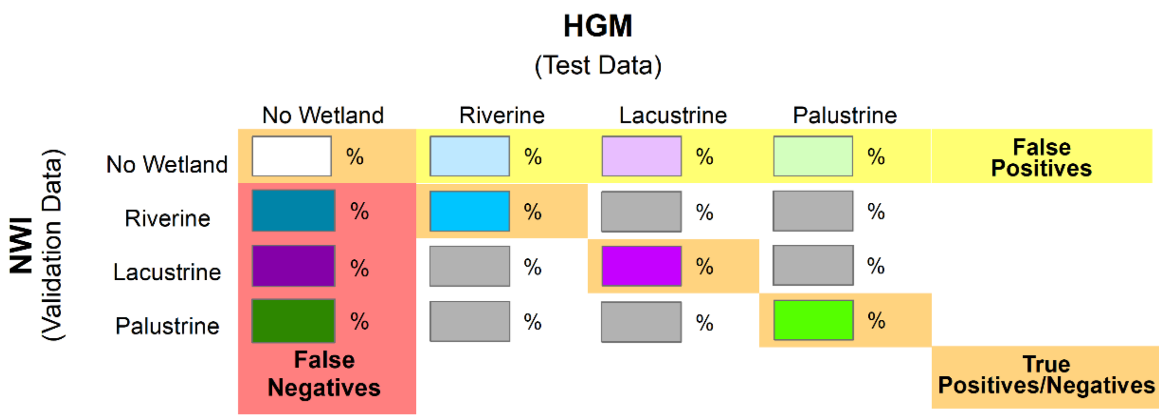

Figure 2 shows a template for visualizing two resolutions of attribute classification. The top (white) cell in the first column shows agreement at the coarser level (wetland/non-wetland), and the remaining 15 cells break out misclassification at finer attribute levels (riverine, lacustrine, and palustrine). Gray cells show misclassification occurring only at finer attribute levels.

This symbology was developed using HSV rather than RGB color palettes to consistently control hue, value, and saturation. The cells are color-coded, with hue (blue, green, purple) referring to wetland type. Differences in hue allow for easy identification of various classes and do not connote order or importance. Value and saturation are used to easily distinguish among false positives, false negatives, and true positives. Low value and bright saturation annotate false negatives while high value and pale saturation annotate false positives. The false negatives are highlighted more than false positives as the NWI classification is considered the validation dataset (i.e., ground truth). False positives are the least prominent to imply less confidence in the HGM as it is created from multiple datasets each with their own uncertainty. Medium value and bright saturation indicate true positives where NWI and HGM classifications agree. The symbology emphasizes true positives as they are areas of higher certainty. The graphic facilitates interpretation and offers visual logic in viewing a large study area with a complex multi-scale pattern of attribute uncertainty.

Results

Results for the second pass with two attribute resolutions are discussed here. Figure 3 shows the confusion matrix and a map of attribute uncertainty for the entire study area. The large number of very small wetlands creates a finely textured uncertainty surface, and palustrine wetlands appear to dominate. The HGM misclassified 14.7% of palustrine wetlands as non-wetland. The gray cells report areas misclassified at the finer attribute resolution, showing that palustrine wetlands were the most frequently misclassed with 0.6% and 0.1% being classed as riverine and lacustrine respectively.

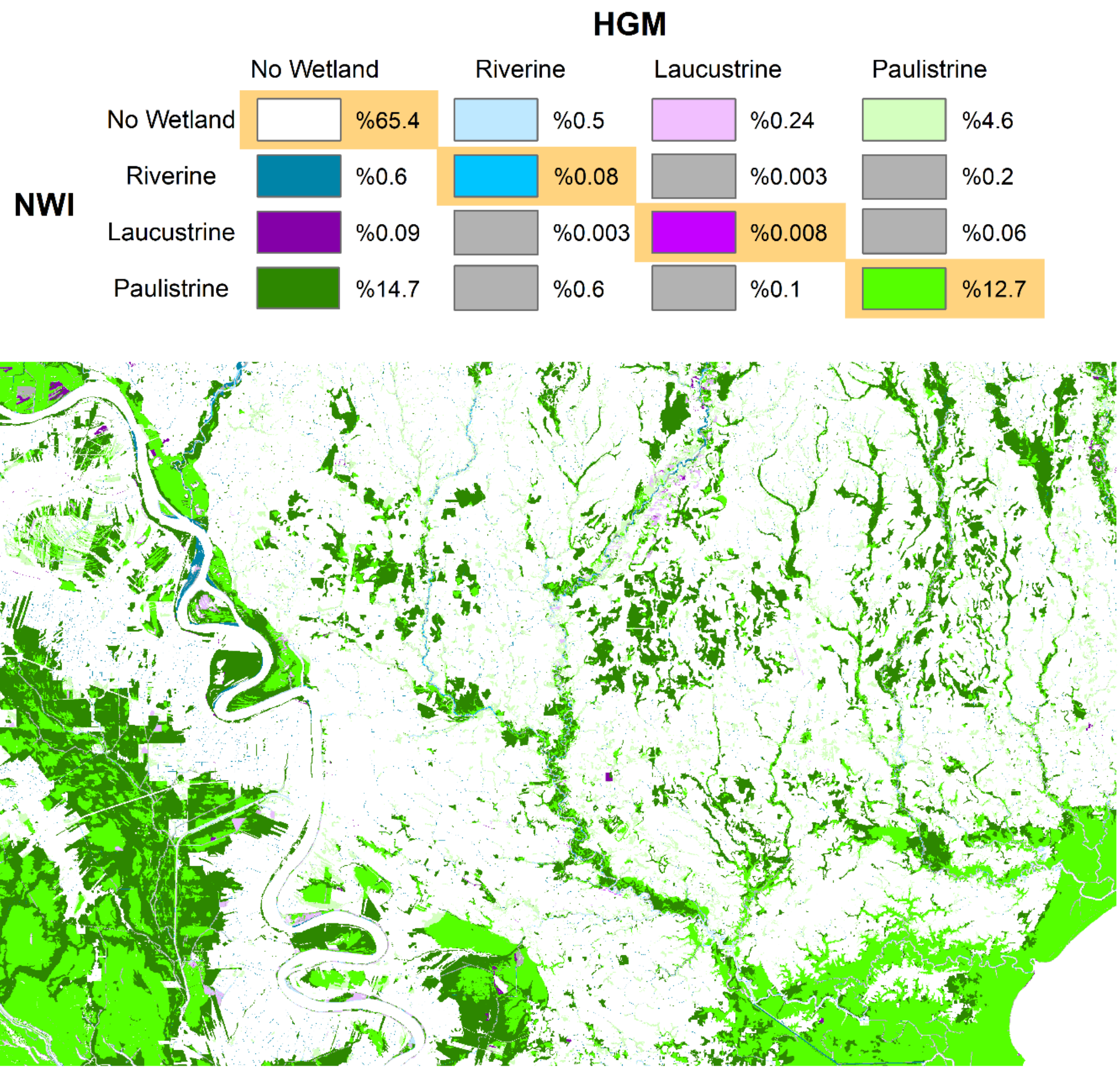

Figure 3: This map compares the NWI (Cowardin) and HGM classification method for the entire study area. Open water areas have been removed. The diagonal values labelled in orange boxes are the true negatives and positives. The first column shows false negatives while the top row shows false positives. The grey cells indicate areas of misclassification but agreement on the presence of wetlands.

Additionally, the HGM was sensitive towards palustrine wetlands, classifying 4.6% of the study area as palustrine rather than non-wetland. This is highlighted in the Denham Springs subset (Figure 4) where 7.2% of the subset area was classified as palustrine by the HGM but are considered to have no wetlands by NWI. Figure 4 also shows a large dark purple area in the southeast corner which, after viewing areal imagery, appears to be a square man-made lake. This area is classified as a lacustrine wetland by NWI but not a wetland by HGM. While NWI was treated as the validation dataset and ground truth, there are errors within the dataset that contribute to the overall uncertainty surface.

NWI \ HGM	No Wetland	Riverine	Laucustrine	Paulistrine
No Wetland	%75.6	%0.4	%0.08	%7.2
Riverine	%0.5	%0.26	%0	%0.3
Laucustrine	%0.1	%0	%0.005	%0.005
Paulistrine	%11.5	%0.36	%0.018	%3.6

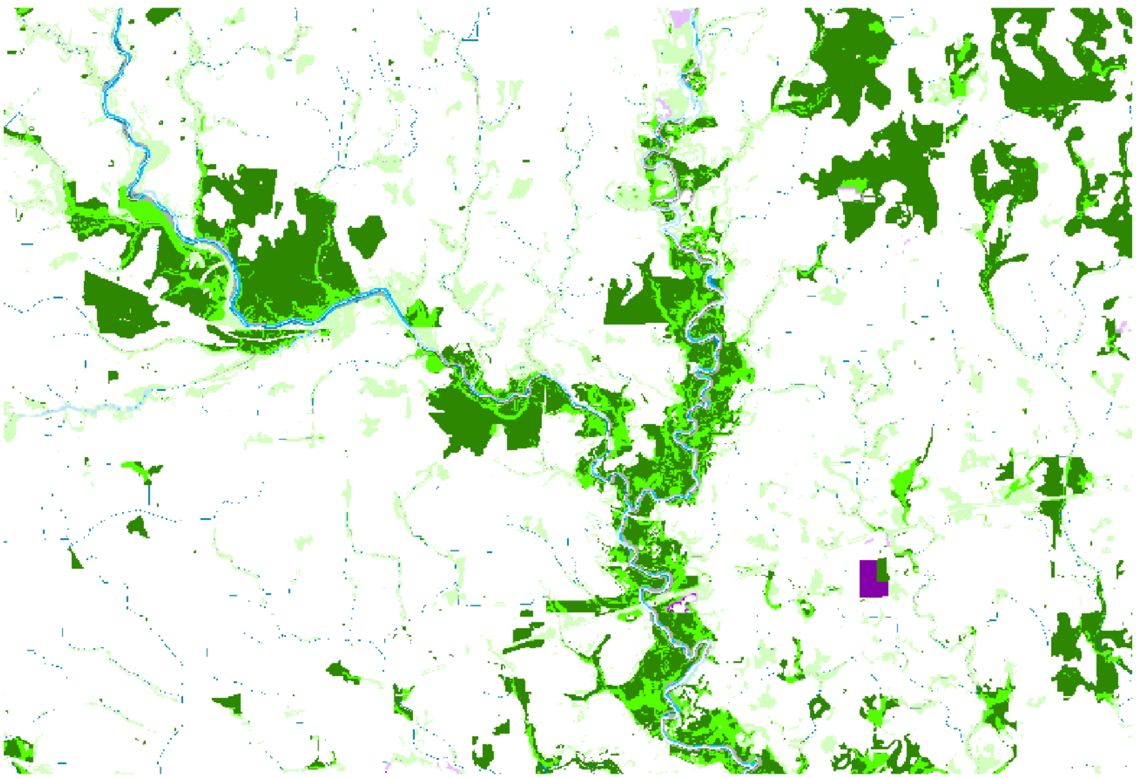

Figure 4: Subset of Figure 3 focusing on Denham Springs and comparing the NWI (Cowardin) and HGM classification method. The diagonal values highlighted in orange are the true negatives and positives. The first column shows false negatives while the top row shows false positives. The grey cells indicate areas of misclassification but agreement on the presence of wetlands.

Figure 5 shows the area around False River with interesting implications of temporal change in wetlands. The NWI data was collected in the 1970s and 1980s with a small portion collected in the 2010s, but the HGM was compiled from more recent data. The central area of Figure 5 lies between False River and the Mississippi River. The land nearest to the oxbow shows land use patterns that indicate agriculture, and those areas are classed as palustrine in the NWI dataset; but the HGM considers those areas to have no wetlands. This could be an example of wetland loss due to agricultural expansion especially since protective legislation (Swampbuster Bill, 1985) was not yet introduced when the NWI data was collected here.

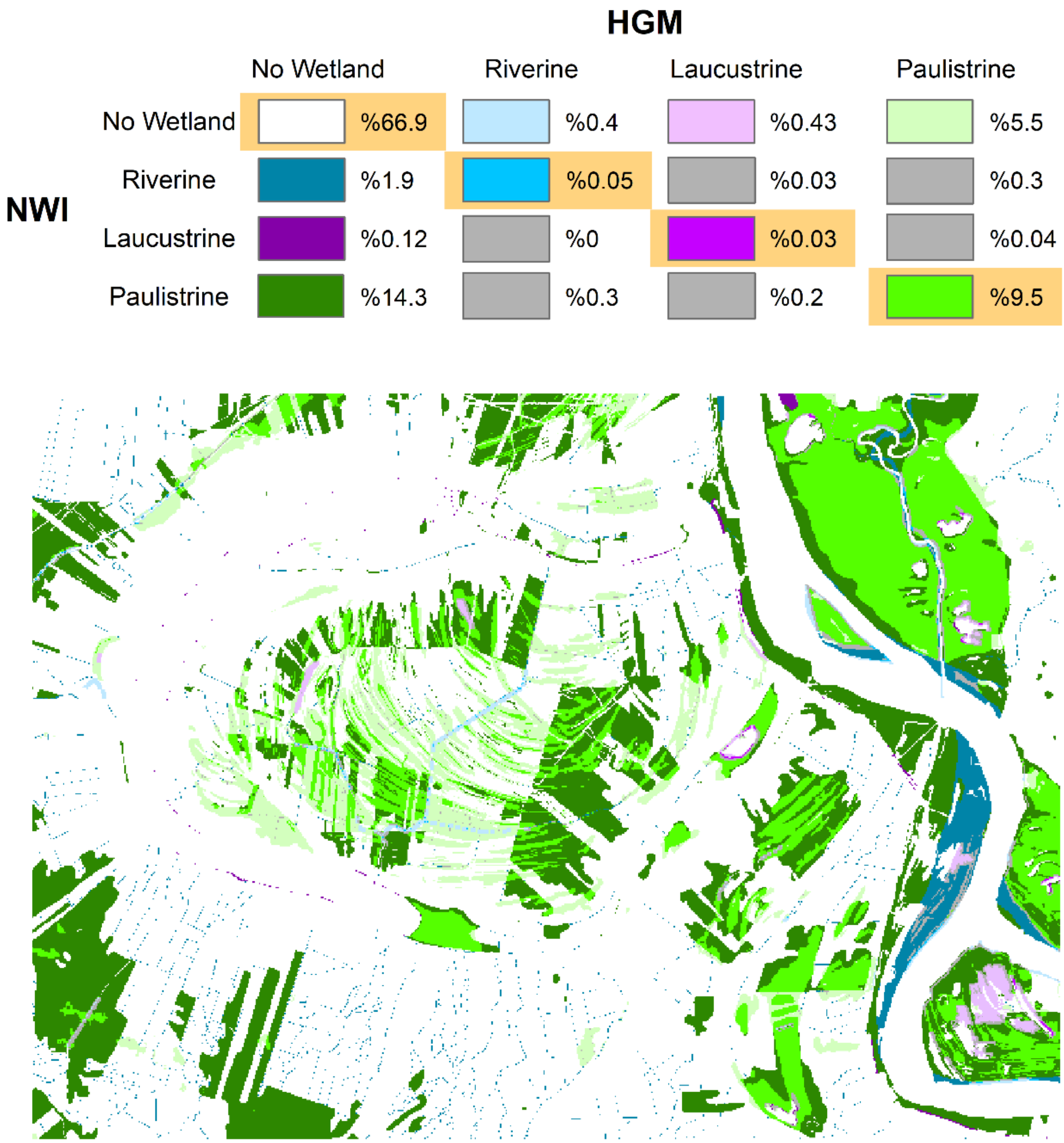

Figure 5: Subset of Figure 3 focusing on False River and compares the NWI (Cowardin) and HGM classification method. The diagonal values highlighted in orange are the true negatives and positives. The first column shows false negatives while the top row shows false positives. The grey cells indicate areas of misclassification but agreement on the presence of wetlands

Discussion and Prospects

This paper reports on an innovative method to visualize attribute uncertainty at two levels of resolution within a confusion matrix, using visual variables to distinguish among true and false, positive and negative wetlands classifications. In ongoing research, we are developing a pyramid data framework that highlights uncertainty variations across multiple spatial and temporal resolutions. The framework permits localized searching across the uncertainty surfaces described here, as for example focal spatial and temporal misclassifications. Upon completion, the tool will be made available as an open-source software tool. This tool could benefit land managers in focusing attention on areas of high uncertainty, and to quantify the degree to which further analysis is warranted. Additionally, this framework could support data collection updates in areas of higher uncertainty. Additionally, we are examining a second study area, and extending uncertainty analyses with a direct method for temporal classification.

Acknowledgements

This research is funded by NSF Award MMS-1853866 "Cross-Scale Spatiotemporal Modeling Using an Integrated Data Framework", a collaboration of the University of South Florida - Tampa, University of Colorado - Boulder, and University of Hawaii – Manoa. Support from NSF is gratefully acknowledged, as well as reviewer comments of an earlier draft that have strengthened the paper.

References

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

Brinson, M. M. (1993). *A hydrogeomorphic classification for wetlands*. (WRP-DE-4, U.S). Vicksburg, MS: Army Engineer Waterways Experiment Station.

Cowardin, L. M., Carter, V., Golet, F. C., & Laroe, E. T. (2005). Classification of Wetlands and Deepwater Habitats of the United States. In: J. H. Lehr & J. Keeley (Eds.), *Water Encyclopedia* (p. sw2162). New York: John Wiley & Sons, Inc.

Dahl, T. E. (1990). *Wetlands Losses in the United States 1780's to 1980's*. Washington, D.C.: U.S. Department of the Interior, Fish and Wildlife Service. Retrieved May 1, 2020, from http://www.fws.gov/wetlands/Documents/Wetlands-Losses-in-the-United-States-1780s-to-1980s.pdf.

Drecki, I. (2002). Visualisation of Uncertainty in Geographical Data. In W. Shi (Ed.), *Spatial Data Quality*. Taylor & Francis.

Dvorett, D., Bidwell, J., Davis, C., & DuBois, C. (2012). Developing a Hydrogeomorphic Wetland Inventory: Reclassifying National Wetlands Inventory Polygons in Geographic Information Systems. *Wetlands*, *32*(1), 83–93.

EPA. (2002). *Methods for Evaluating Wetland Condition: Wetlands Classification*: (EPA-822-R-02-017). Washington, D.C.: Office of Water, U.S. Environmental Protection Agency.

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

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

MacEachren, A. M., Roth, R. E., O'Brien, J., Li, B., Swingley, D., & Gahegan, M. (2012). Visual Semiotics Uncertainty Visualization: An Empirical Study. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2496–2505.

McCauley, L. A., Jenkins, D. G., and Quintana-Ascencio, P. F. (2013). Isolated Wetland Loss and Degradation Over Two Decades in an Increasingly Urbanized Landscape. *Wetlands* 33(1): 117–27.

Millennium Ecosystem Assessment. (2005). *Ecosystems and Human Well-Being: Wetlands and Water Synthesis*. Washington, D.C.: World Resources Institute. Retrieved May 1, 2020, from https://www.millenniumassessment.org/documents/document.358.aspx.pdf

Smith, R. D., Noble, C. V., & Berkowitz, J. F. (2013). *Hydrogeomorphic (HGM) Approach to Assessing Wetland Functions: Guidelines for Developing Guidebooks (Version 2):* (ERDC/EL TR-13-11). Washington, D.C.: U.S. Army Corps of Engineers.

Swampbuster Bill (1985). Food Security Act. Washington, D.C.: 99th Congress. Public Law 99-198. https://www.govinfo.gov/content/pkg/STATUTE-99/pdf/STATUTE-99-Pg1354.pdf

Tiner, R. W. (2005). Assessing Cumulative Loss of Wetland Functions in the Nanticoke River Watershed Using Enhanced National Wetlands Inventory Data. *Wetlands* 25(2): 405–19.