Geovisualization of Florida Land Use and Land Cover using Bivariate Mapping, Statistical Legends, and Visual Analytics

Georgianna Strode, Victor Mesev, Benjamin Thornton, Marjorie Jerez, Thomas Tricarico, and Tyler McAlear

Florida State University

Contact: Gstrode@admin.fsu.edu

Abstract

The concepts of land use and land cover convey information about a landscape. Although land use and land cover are different concepts, the terms have become intermingled and often used interchangeably in GIScience, resulting in a simplified one-to-one relationship that can be visualized using univariate mapping techniques. This research posits that the relationships between land use and land cover are complex, many-to-many relationships, and can be represented bivariately.

Land cover is determined by the direct observation of the earth's surface, resulting in categories such as forest, wetland, and development. Land use by contrast is a socioeconomic interpretation of the activities that take place on that surface, with categories such as residential, commercial, and agricultural. Complications arise when these categories overlap. For example, a forested area can be used for recreation, grazing, or timber production. A residential area can be located within a forest, a grazing plain, or a developed area. A bivariate representation can be useful in conveying these overlapping categories. The research described in this article demonstrates complex interactions between land use and land cover in Florida.

Land use data were obtained from the 2014 Florida Department of Revenue database and land cover data from the 2010 National Land Cover database. The resulting geovisualization consists of three parts: 1. A bivariate map using color theory and symbology based on Gestalt principles to represent and differentiate the two data sources; 2. A two-dimensional statistical legend to inform the data combinations and data distributions; 3. Visual analytics using Sankey (flow) diagrams to show the relationships and frequencies of categorical data interactions. This combination of visuals advances cartography and GIScience and gives new perspectives to the many-to-many relationships between land use and land cover.

Keywords

Bivariate mapping, statistical legends, visual analytics, land use, land cover, Gestalt principles

Introduction

Land cover is determined by the direct observation of the earth's surface, resulting in categories such as forest, wetland, and development. Land use by contrast is a socioeconomic interpretation of the activities that take place on that surface, with categories such as residential, commercial, and agricultural. Complications arise when these categories overlap. Although land use and land cover are different concepts, the terms have become intermingled and often used interchangeably in GIScience, resulting in a simplified one-to-one relationship that can be visualized using univariate mapping techniques. Univariate thematic maps show one topic at a time, such as a map of land use or a separate map of land cover. This research posits that the many-to-many relationships between land use and land cover are complex and can be represented bivariately. Our project seeks to develop visualizations that show both datasets simultaneously while allowing them to be visualized separately.

This project uses bivariate mapping and visual analytics to convey the relationships between land use and land cover. To develop the bivariate map, we employ color, symbology, and basic Gestalt principles. Traditional colors are used where possible and color theory is used to inform new color schemes. Symbology helps differentiate the two data sets. Gestalt psychology is based on the principle that humans strive to perceive experiences as regular, orderly, symmetrical, and simple. Basic Gestalt visualization principles use laws of symmetry and proximity to convey similarity and to help the user recognize patterns and groupings. A gridded map at 1-kilometer scale is used for both data sets to normalize the data and to enhance visualization. Statistical legends and visual analytics are added to show different perspectives of the data combinations.

Background

History of Land Use and Land Cover Classifications

Land use and land cover are related but distinct concepts. Land cover is determined by direct observation of the earth’s surface, whereas land use is the socioeconomic activity that occurs on the land. Land cover can include grass, asphalt, water, or bare ground. Land use describes how people utilize the land and includes categories such as residential, commercial, or agricultural. Complex many-to-many relationships result. For example, grassland cover can occur in many land uses: residential, pasture, or sports areas. Similarly, residential land use can include a variety of land covers such as trees, grass, and asphalt (Fisher and others 2005).

Remote sensing technology of the 1970s provided the means to identify land cover through collection of timely and accurate data to complement ground survey and observation. The United States Geological Survey (USGS) developed a new land use and land cover classification system in response to the new data accessibility and the need for federal and state agencies to have a uniform national classification system (Anderson and others 1972). This classification was consistent nationally, yet flexible enough for regional adaptation. The cartographic result of the new classification system represented land use and land cover univariately.

Although the national classification system developed by USGS has been widely adapted and expanded, concerns have been raised about combining the two datasets. Critics cite lack of clarity between land use and land cover (Fisher and others 2005). Nor does the classification system always meet the needs of users. For example, planners are interested in land use, whereas environmentalists are more concerned with land cover (Fisher and others 2005).

In 1985, the Florida Department of Transportation (FDOT) revised its original land use and land cover classification system in accord with the USGS system (Anderson and others 1976). The result was the Florida Land Use, Cover and Forms Classification System (FLUCCS). The system was not designed to explore the relationships between land use and land cover across the state but instead to meet specific user requests from other state agencies, local governments, and private enterprise. This ambitious system was revised in 1999 to include additional wetland classes as well as more sophisticated forms of aerial photography, Landsat Multi-Spectral Scanner (MSS) data, and Landsat Thematic Mapper (TM). By 1999, most image data from remotely sensed images and all non-image data were stored in a Geographic Information System (GIS). In 2019, the statewide land use/land cover system is no longer maintained by FDOT. It is compiled by the Florida Department of Environmental Protection (FDEP) from maps created by each of the state's five water management districts. The Florida Fish and Wildlife Commission and the Florida Natural Areas Inventory also produce habitat, natural community, and land cover maps specific to their needs.

The research presented here is a visualization exercise for the purpose of quickly assessing general land use and land cover at a uniform scale. The bivariate mapping and visual analytics described here can be used to gain an understanding of other co-occurring variables. In no way is this work intended to replace any other mapping methods developed by state agencies. Further, this visualization can be used as a teaching tool to sharpen visual skills and cartographic ability to interpret landscapes. It is hoped that this research can contribute to cartographic knowledge and serve as an example for other bivariate subject matters.

Bivariate Mapping

Bivariate maps show two thematic phenomena on the same map. Thematic data differ from point and line data in that the geographic area is a polygon such as a county or census block. Whereas maps using point and line data can easily show multiple types of data, makers of thematic maps face challenges of conveying two types of data within the same polygon. If bivariate mapping is successful, it has potential to reveal relationships and patterns between two types of data more effectively than side-by-side univariate maps (Carstensen 1986). Bivariate visualizations are useful in revealing relationships between data pairs such as the elderly and minorities, youth in rural areas, educational attainment and income, and other investigations where it is desirable to explore associations between different variables in a single map. The maps should be constructed so that viewers can see/observe and estimate the degree of interaction between the phenomena. Ideally, any relationships between the variables should be highlighted while the data gradients of the individual datasets are preserved (Trumbo 1981, Ware 2009). In addition, bivariate mapping is a logical choice for time series maps (Schroeder 2010).

Bivariate maps can use similar basic visualization techniques as univariate maps. Sequential color schemes use a color gradation to represent increases in data values (Brewer 1994). Symbology techniques can be used to further represent data values. Symbology choices can be based on data sets and spatial arrangement. These can be visualized using variations in symbol size, color, or texture (Slocum and others 2005).

Basic Gestalt Principles

Gestalt is a psychology term signifying unified whole. Developed by German psychologists in the 1920s, Gestalt visualization principles are guidelines for viewers to organize visual elements into groups or unified wholes. Images using Gestalt principles are coherent, unified, and more easily understood. There are six principles: similarity, continuation, closure, proximity, figure/ground, and symmetry/order (Bradley 2014a). This research considers four of these principles to be important in the development and understanding of the land use/land cover bivariate map.

The four Gestalt principles used for this research are: 1. Figure/ground–Viewers are able to separate the foreground figure from the background. Additionally, viewers can theoretically toggle focus between the figure and background seeing completely different images. Evidence for this is the famous illusion showing both a young girl and an old woman at the same time. Viewers can see both images yet switch between the two images at will (Weisstein nd). Figure and background can both attract attention and one can overtake the other, creating a "reversible" effect (Bradley 2014b). 2. Similarity–When objects share similar visual characteristics (e.g. shape, size, color), viewers often see a pattern or group. The more commonality that elements share, the greater the perception of coherence. 3. Proximity or "grouping"–Close arrangement of elements can have the same effect as similarity. Objects close together can be seen as a group even if they appear different. 4. Symmetry/Order–Orderly arrangements of elements give a sense of stability, consistency, and structure (Arnheims 1976; Bradley 2014a).

Basic Color Theory

Choice of color progression can affect the understanding of the map data. A sequential color scheme uses a gradient from light to dark where lighter colors represent "less" and darker colors indicate "more." A diverging color scheme shows data extremes while minimizing the "normal" or "expected." A qualitative scheme shows categorical data through color differences. Figure 1 shows an example of each color scheme (Brewer 1984).

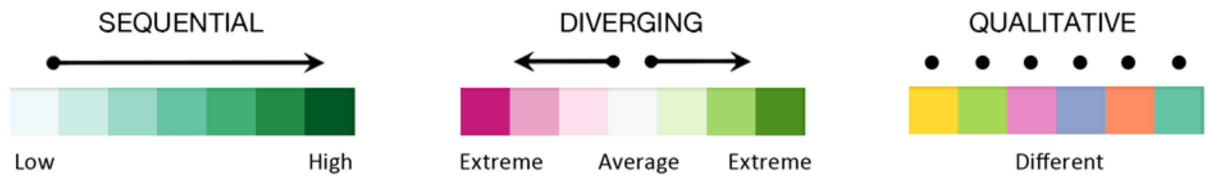

Figure 1. Examples of the three types of color distributions.

Statistical Legends and Visual Analytics

With the increased availability of large amounts of spatial data, statistical techniques are needed to enhance data understanding. Maps using data from multiple sources can overwhelm and confuse viewers. Legends are an important communication tool for conveying map meaning to the viewer. Traditional static legends are one-size-fits-all generalizations of data interactions that can be interchanged between maps. Statistical legends, on the other hand, are customized for each map. Each statistical legend conveys a quantitative summary of the data interactions.

Visual analytics are an additional technique for data understanding and organization. Traditionally pie and bar charts have been used to supplement spatial maps. Today open source tools provide the ability to develop complex analytics as shown in Figure 2 (Bostock 2015).

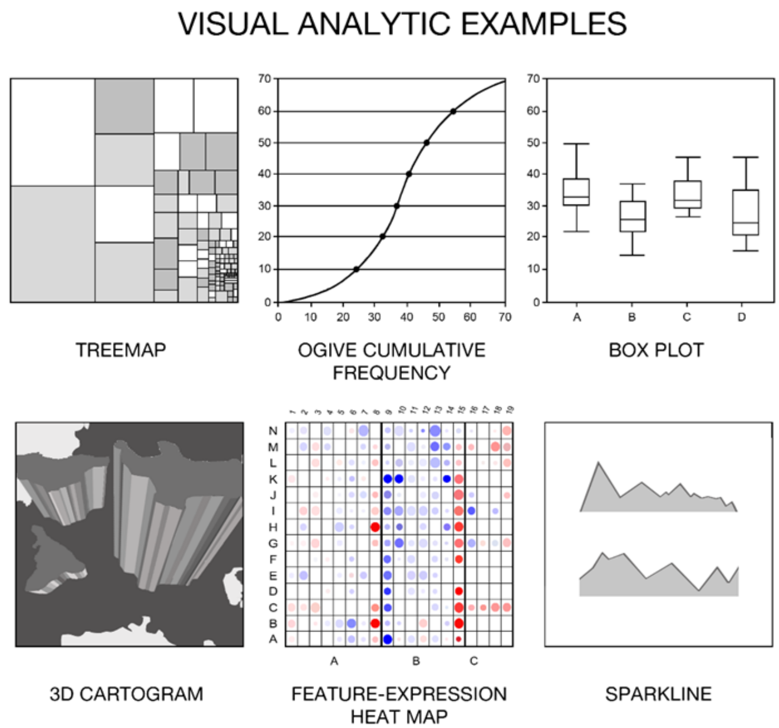

Figure 2. Examples of visual analytics.

Method

Our geovisualization method combines traditional colors, color theory, and Gestalt principles to develop a bivariate map. Traditional colors are important as they are already established and easily recognizable by users familiar with land use and land cover. Color theory is useful for guiding decisions for new color patterns to bridge the gap between the two traditional palettes. Gestalt principles serve as a platform for map development. We decided to use a gridded map to provide uniformity for two reasons: first, grids preserve data density and normalization; and second, grids enhance visualization through the Gestalt principles of similarity, symmetry, proximity.

Data Input

Land cover data used for this research are from the 2010 National Land Cover database through the Multi-Resolution Land Characteristics (MLRC) Consortium. We combined the categories of shrub, planted/cultivated, and herbaceous into one category named vegetation to simplify the visualization. The final categories for this project include: open water, wetlands, forest, vegetation, barren, and development in four intensities of open, low, medium, and high. This research also includes land use data come from the 2014 Florida Department of Revenue cadastral database. The land use categories for this project are the same as used by the property appraiser, which includes residential, commercial, industrial, agricultural, institutional, government, and miscellaneous. We selected data from the property appraiser instead of from traditional zoning maps because the cadastral (parcel-based) data have a higher resolution than traditional zoning maps, and already have a land use code assigned by the local property appraiser. Cadastral data includes sufficient information to support land information across jurisdictional boundaries (von Meyer and others 2002). It is worth noting that in some cases cadastral data can show land use as ownership and not the actual use of the land. For example, the government land use category is more accurately a statement of land ownership that does not indicate the usage. Government land could be a forest, prison, administrative, or other.

Both land use and land cover data sets are normalized to a 1-km grid using the predominant value. From a data perspective, gridded maps preserve data density and are less susceptible to the Modifiable Area Unit Problem (Openshaw 1983; Mennis 2003). From a visualization perspective, the symmetry and similarities offered by a gridded map follow Gestalt visualization principles.

Bivariate Map

Nested Symbology

Symbology is a way to represent both datasets in the same spatial unit. We overlaid a circle over a square as shown in Figure 3. The Gestalt figure/background principle tells us that the map reader can distinguish between figure and background and switch between the two. Because land cover is an observable feature and thus readily understandable by viewers, we chose to place land cover as the figure, represented by a circle. Alternatively, land use is a more abstract concept and we placed it in the background with a square.

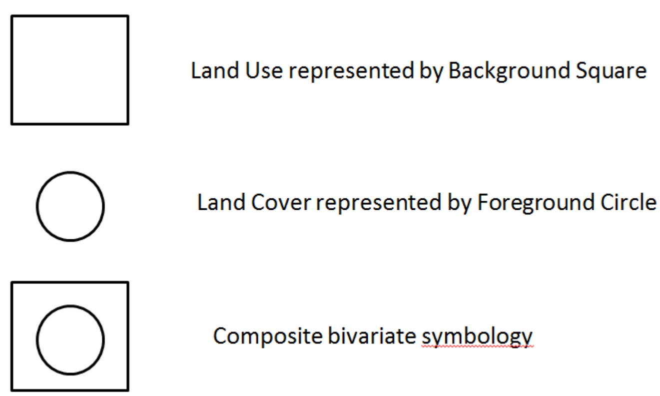

Figure 3. Symbology selections to represent two data sets.

Color

Figure 4 shows traditional land use colors (Jeer and Bain 1997) and land cover colors (USGS 2006). These colors are established, and we chose to continue use of these colors where possible. Figure 5 shows the methodology behind blending these two together while considering the two symbology choices. Impact to the land is shown through a gradation from blue water to greens and pinks to represent higher impact to the land through development. Through an internal iterative trial-and-error design process we were able to determine the most easily distinguished colors. Figure 6 shows the final static legend intended to combine the two traditional color sets in a manner that shows impact on the land.

Land Cover		Land Use
Water	Open Water	Residential, Single Family
Developed	Developed, Open Space	Residential, Multi-Family
	Developed, Low Intensity	Retail and Commercial
	Developed, Medium Intensity	Industrial
	Developed, High Intensity	Institutional and Public Facilities
Barren	Barren Land	Recreational Uses
Forest	Deciduous Forest	Industrial Utilities
	Evergreen Forest	
	Mixed Forest	
Shrubland	Dwarf Scrub	
	Shrub/Scrub	
Herbaceous	Grassland/Herbaceous	
	Sedge/Herbaceous	
	Lichens	
	Moss	
Planted/Cultivated	Pasture/Hay	
	Cultivated Crops	
Wetlands	Woody Wetlands	
	Emergent Herbaceous Wetlands	

Figure 4. Traditional land cover and land use colors.

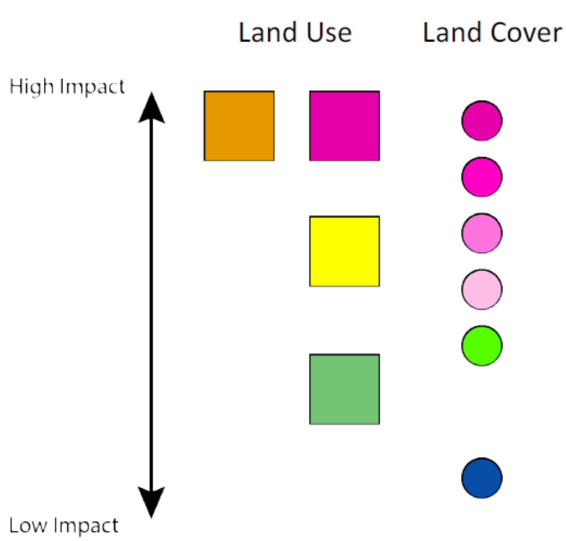

Figure 5. The logic behind color decisions is a blend of traditional land use, land cover, and degree of impact on the land.

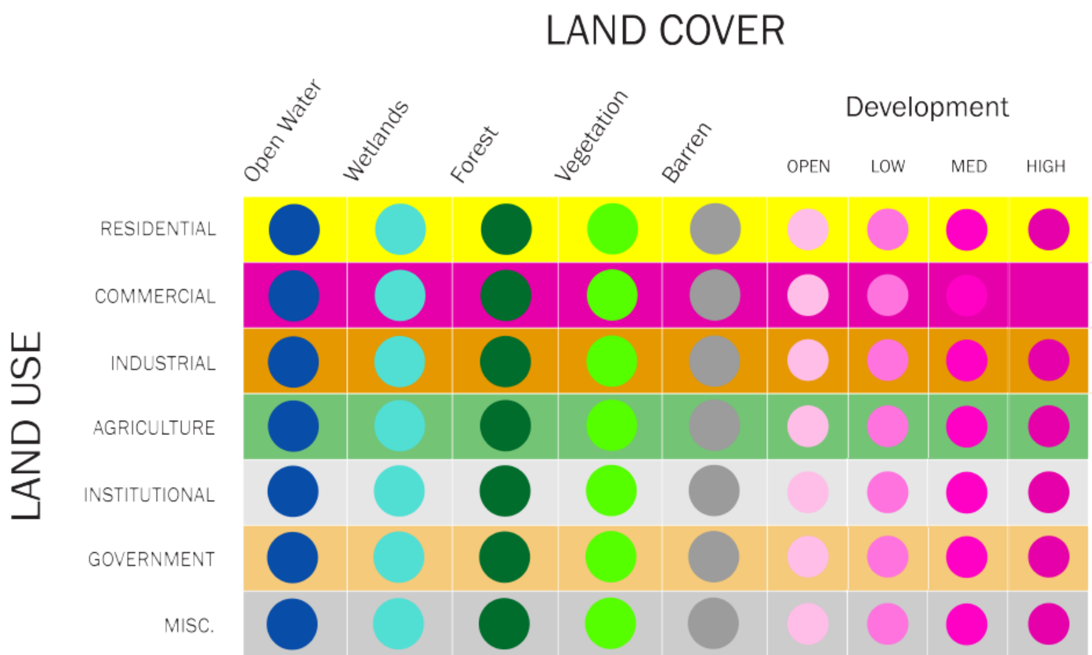

Figure 6. Static legend showing 63 combinations of land use and land cover.

Figure 7 shows an example of the bivariate map. Residential areas are identifiable by a yellow background and can have many types of land cover. Yellow backgrounds with green circles are forested residential. Areas with fuchsia background indicate high development. Medium pink circles represent medium development and light pink circles show residential areas with more open space. A green background indicates the land is used for agriculture. Within this category, there are several types of land cover: forest, vegetation, and wetland as show by the circles of dark green, medium green, and pale blue. Governmental lands are shown with a tan background with land cover of open water (blue), vegetation (medium green), and wetland (pale blue).

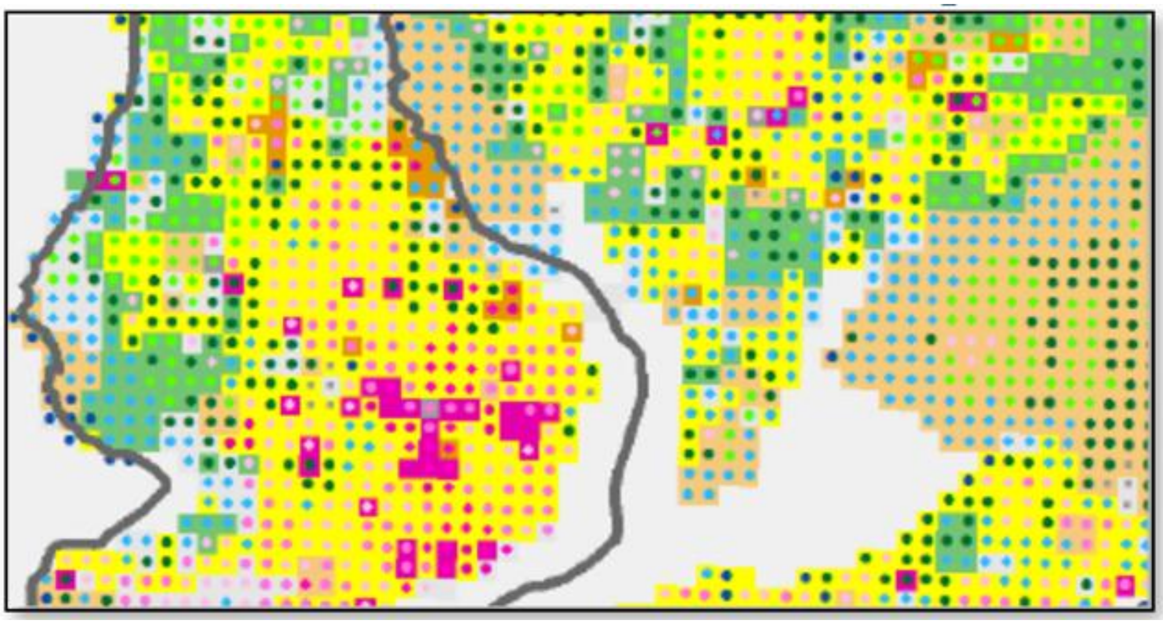

Figure 7. Bivariate map using color and symbology. Pinellas and Hillborough counties.

Statistical Legends

Statistical legends show the quantity of the combinations of data interactions. Unlike the static legend, statistical legends will vary from map to map. Figure 8 shows the statistical legend for land use and land cover for Florida. This legend shows that most land use is agricultural, and that vegetation is the most common land cover.

Visual Analytics

Sankey diagrams show relationships and frequency. We used D3, an open source software, to create these images. We used the same colors as the bivariate map. We edited the original code to add a gradation technique to further illustrate the data combinations. Wider lines represent more data combinations and thinner lines show fewer interactions. Figures 9 and 10 illustrate the complex many-to-many relationships between land use and land cover in Florida. A specific land cover can be found in many land uses. Similarly, a land use can pair with multiple land covers. For example, agriculture is the most prevalent land use with predominantly wetland and planted/cultivated land cover. A second Sankey diagram (Figure 10) shows secondary categories to further illustrate the finer divisions for each data type. The land use categories of government, industrial, and miscellaneous were grouped into a new category of “Other” for simplification. The secondary diagram shows further divisions, such as agricultural land use being comprised of timberland, grazing land, cropland, and commodity. A land cover of planted can be subdivided into pasture and cultivated crops.

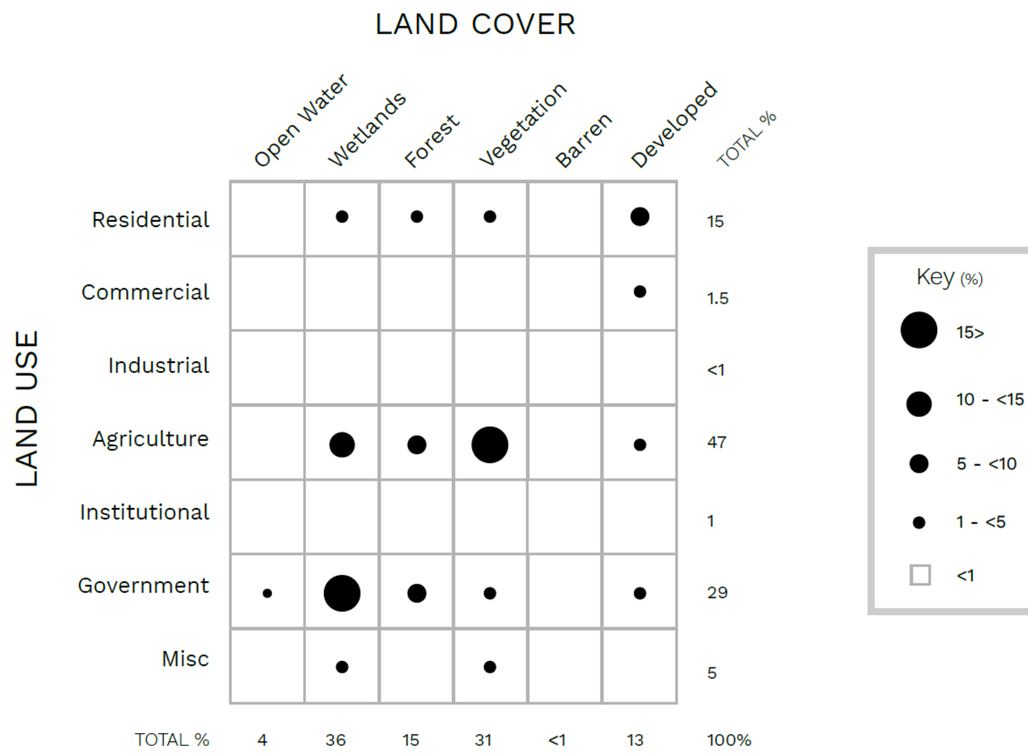

Figure 8. Statistical legend quantifies the Florida data distributions.

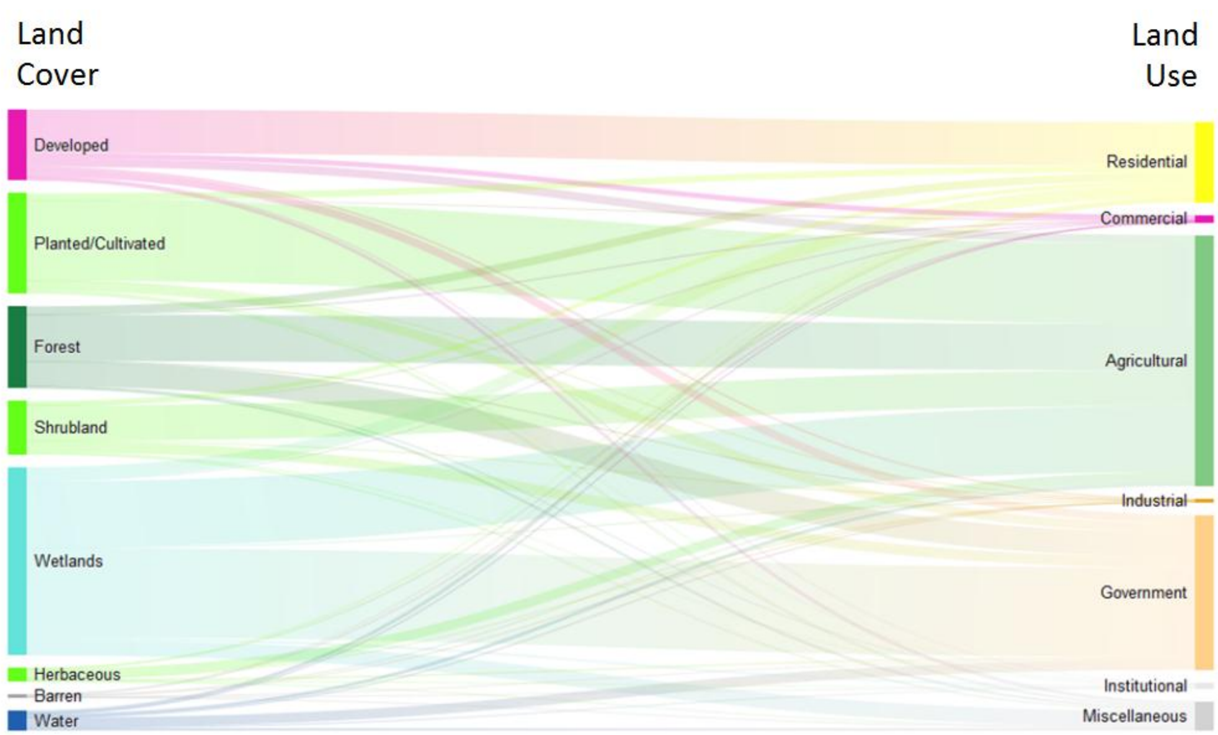

Figure 9. Primary Sankey diagram for Florida Land Cover / Land Use.

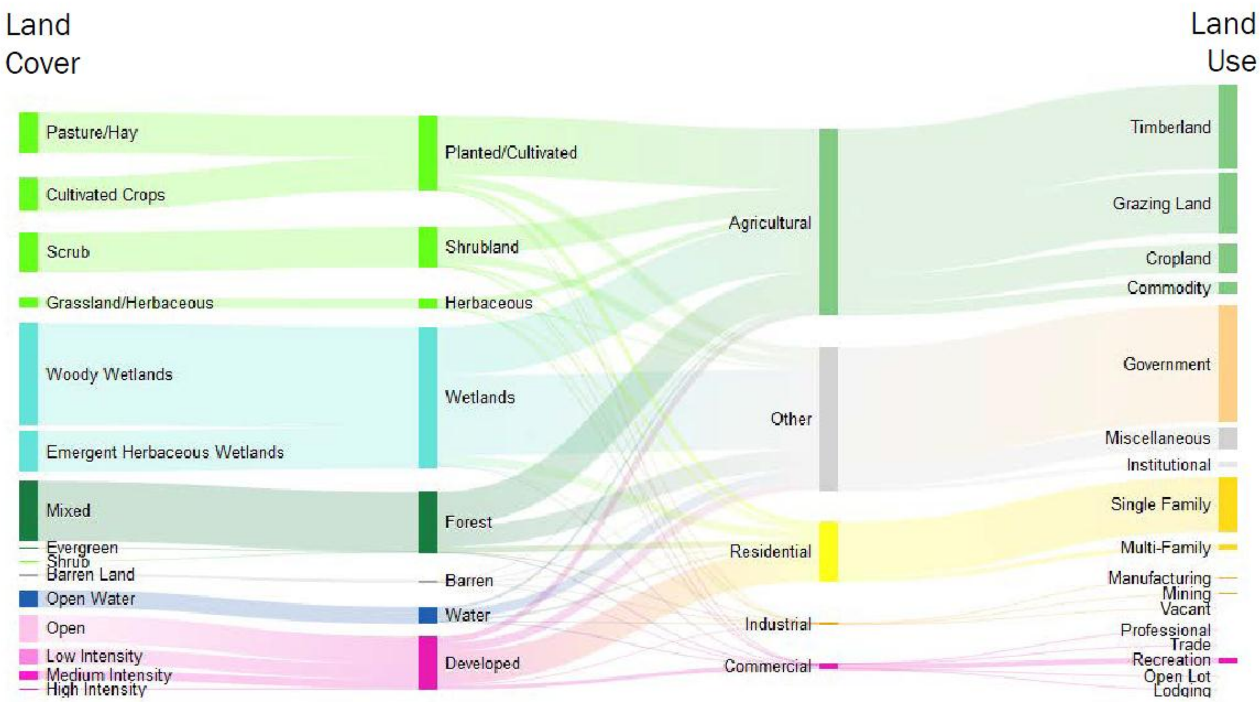

Figure 10. Secondary Sankey diagram for Florida Land Cover/Land Use.

Results

The final product shows a bivariate map, statistical legend, and visual analytics. Each item serves its own purpose. Collectively they contribute to an innovative perspective on the relationships between land use and land cover. The bivariate map shows the spatial distribution of the data interactions. The statistical legend and Sankey diagram illustrate the statistical distribution of the data combinations in different ways. The statistical legend may be more useful for those seeking a quick summary of data whereas the primary Sankey diagram may appeal to more visual readers.

Gestalt principles seem to work effectively in the bivariate map. Map readers should be able to switch between figure and background, thus allowing them to distinguish land use and land cover. The principle of similarity assures that objects of similar visual characteristics will be seen as belonging together to allow readers to recognize patterns emerging from similar colors and shapes. The proximity, or "grouping," principle allows map readers to group objects close to one another.

Discussion

The methods of land use/land cover visualization presented here are in no way intended to replace methodologies developed by other agencies for a specific purpose. This project has potential to help users of land use and land cover by calling attention to the differences between the data types. Clearly separating these two topics through distinct data displays could stimulate thoughtful conversation concerning usage, definitions, and interpretation. This visualization could be useful in identifying landscape characteristics easily and quickly. A second benefit is that this visualization could be a valuable teaching tool to help sharpen visual skills and cartographic ability to distinguish representations between maps, other visuals, and the actual geography of the landscape. A third benefit of this research is a contribution to the cartographic toolkit. This project provides an example that can be replicated for use with other topics.

Resources

Figure 11 shows a preview of the completed geovisualization in a poster-style format (PDF). The statewide GIS Land Use and Land Cover data used for this research are pre-styled in ArcMap format (.mdx) for independent use. Both poster and GIS data are available for download at: https://usng-gis.org/landuse.html. This geovisualization can be replicated for other states through the 1-km shapefiles provided by the U.S. National Grid Information Center at: https://usngcenter.org/portfolio-item/usng-gis-data/.

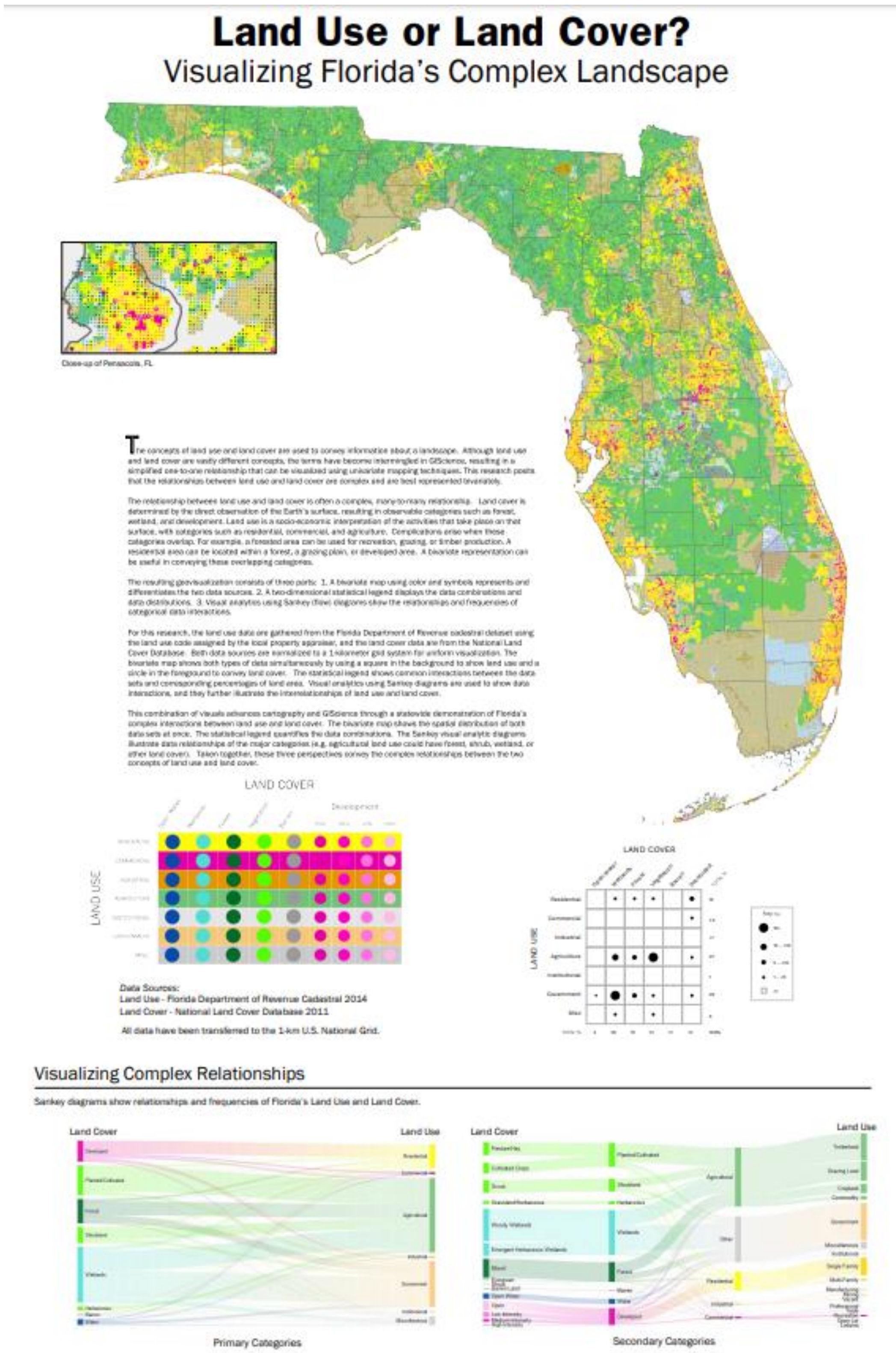

Figure 11. Final poster for Land Use and Land Cover.

Acknowledgements

The authors would like to thank two anonymous reviewers for their constructive comments.

References

Anderson, J.R., Hardy, E. E., Roach, J. T. (1972). *A Land Use and Land Cover Classification system for use with Remote Sensor Data* (Geological Survey Circular 671). USA, United States Department of the Interior, Geological Survey.

Arnheims, R. (1954, 1976). *Art and Visual Perception: A Psychology of the Creative Eye*. University of California Press.

Bostock, M. "*D3.js - Data-Driven Documents*." D3.js - Web. 03 Nov. 2015.

Bradley, S. (2014). *Design principles: visual perception and the principles of Gestalt*. Retrieved February 24, 2017, from https://www.smashingmagazine.com/2014/03/design-principles-visual-perception-and-the-principles-of-gestalt/

Bradley, S. (2014). *Design principles: space and the figure-ground relationship*. Retrieved February 10, 2017, from https://www.smashingmagazine.com/2014/05/design-principles-space-figure-ground-relationship/

Brewer, C.A. (1994). *Color Use Guidelines for Mapping and Visualization*, in Visualization in modern cartography (ed. MacEachren, Al, Taylor, D.R. Fraser. Pergamon.

Carstensen, L. W. (1986). "Bivariate choropleth Mapping: The Effects of Axis Scaling." *The American Cartographer* 13(1): 27-42.

Fienberg, S.E. (1979). *Graphical Models in Statistics*. The American Statistician. 33.4: 165-178. Accessed July 7, 2015.

Fisher, P.F., Comber, A.J., Wadsworth, R. (2005*). Land use and Land Cover: Contradiction or Complement* in Fisher, P. (Ed.), Re-Presenting GIS. Wiley, Hoboken, New Jersey.

Jeer, S. and Bain, B. (1997). *Traditional color coding for land uses,* Report – American Planning Association.

Mennis, J. (2003). *Generating Surface Models of Population using Dasymetric Mapping*. The Professional Geographer, 55(1), 31-42.

Openshaw, S. (1983). *The modifiable areal unit problem*. Concepts and Techniques in Modern Geography, vol. 38. Norwich: Geobooks.

Schroeder, J. P. (2010). *Bicomponent trend maps: a multivariate approach to visualizing geographic time series.* Cartography and Geographic Information Science 37.3: 169+.

Slocum, T.,McMaster, R., Kessler, F., Howard, H. (2005). Thematic Cartography and Geographic Visualization (2nd ed.). Upper Saddle River, New Jersey: Pearson.

Trumbo, B.E. (1981*). A Theory for Coloring Bivariate Statistical Maps*. The American Statistician, 35(4),220-226.

United States Geological Survey (2006). *Product legend*. Multi-Resolution Land Characteristics Consortium.

von Meyer, N., Ader, B., Nagy, Z., Stage, D., Ferguson, B., Benson, K., Johnson, B., Kirkpatrick, S., Stevens, R., Mates, D. *Parcel Identifiers for Cadastral Core Data: Concepts and Issues*. Federal Geospatial Data Committee Cadastral Committee. July 2002.

Weisstein, E. W. "Young Girl-Old Woman Illusion." From *MathWorld*—Wolfram Resource. http://mathworld.wolfram.com/YoungGirl-OldWomanIllusion.html

Ware, C. (2009). *Quantitative Texton Sequences for Legible Bivariate Maps*, *ICCC Transactions on Visualization and Computer Graphics*. 15.6: 1523-1529.