[DC-01-042] Changes in Geospatial Data Capture Over Time: Part 2, Implications and Case Studies

Abstract

Advances in technological approaches and tools to capture geospatial data have contributed to a vast collection of applications and enabled capacity for new programs, functions, products, workflows, and whole national-level spatial data infrastructure. In this entry, such outcomes and implications are described, focusing on developmental changes in specific application areas such as land use & land cover inventory, land parcel administration, and business, as well as examples from federal agencies, including the US Geological Survey, the Census Bureau, US Fish and Wildlife Service, and the US Department of Agriculture. These examples illustrate the diverse ways that the dramatic changes in geospatial data capture methods and approaches have affected workflows within agencies and have spatially empowered millions of users and the general public. For additional information on specific technical changes, see Part 1:

Keywords: aerial imagery and photogrammetry, land use change, NSDI, wetlands

Author & citation

Cowen, D. J. and Anderson, K. E. (2021). Changes in Geospatial Data Capture Over Time: Part 2, Implications and Case Studies. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2021 Edition), John P. Wilson (ed.).
DOI: 10.22224/gistbok/2021.1.5

Explanation

1. The Current Geospatial Environment
2. Case Study: Land Use, Land Cover, and Inventory
3. Case Study: Authoritative Land Parcel Systems
4. Case Study: Big Business
5. Case Study: Architecture, Engineering, and Construction
6. Case Study: the US Census Bureau
7. Case Study: The Evolving Role of the US Geological Survey
8. Case Study: US Fish and Wildlife Service – National Wetlands Inventory
9. Case Study: the US Department of Agriculture
10. Conclusion

1. The Current Geospatial Environment

The dramatic technological changes in the way that geospatial data can be captured and automatically processed and made interoperable, described in Part 1 of this entry on on geospatial data capture, have contributed to the larger geospatial revolution that has impacted the way individuals live, work, and play in a geospatially enabled society. Specific outcomes and consequences stemming from the pervasive collection of geospatial data include the following list of observations around outcomes and their implications.

- Geospatial data are essential, accessible, and valuable
- Citizens depend on accurate and current geospatial data in their daily lives
- Citizens are armed with surveying devices
- Citizens are sensors who are capturing features and reporting on conditions
- Smart phones can create a trace and unwittingly be tracked
- Personal information is georeferenced
- Major IT companies and automobile manufacturers continuously collect huge amounts of geospatial data for navigation systems and location-based services
- Utility companies are expected to have smart geospatially enabled infrastructure
- Local governments provide on-line access to authoritative property information and spatially enable a full range of services
- Surveillance systems require and collect geospatial data
- Data are updated in real time from passive and active sensors
- UAVs and small satellites quickly acquire new information
- Countless maps and photos have been scanned and georeferenced
- Geospatial data are easily discovered and acquired through web sites
- Several global base maps and air photo mosaics are available through various applications
- Maps are produced from geospatial data rather than being the source for data
- Collaborative applications are conducted with Big Data through cyberGIS
- The disciplines of land information systems (LIS) and GIS have merged
- CAD and GIS are integrated in Building Information Models (BIM) that combine interior and exterior spatial models
- The Internet of things has a geospatial dimension

Some changes and events have been specific to the federal government.

- High resolution air photo coverages are often created through public private partnerships rather than Federal programs
- Local government geospatial data is used to update federal data, i.e Census
- The Federal government has made a long-term commitment to produce high resolution elevation data
- The Geospatial Data Act of 2018 (GDA) Recognizes responsibilities beyond the Federal government.

In the United States, the ability to capture and use geospatial data is influenced by the overarching National Spatial Data Infrastructure (NSDI) that governs data access, distribution, and use. This major milestone in the effort to promote data sharing came from a 1994 Executive Order, number 12906. The NSDI is comprised of 17 National Geospatial Data Asset (NGDA) themes, eight of which are “Framework” data themes that serve as the foundation for many GIS applications (FGDC 2002). These eight Frame data themes are addresses, cadastral, elevation, geodetic control, governmental unit, hydrography, orthoimagery, and transportation. The data should be available via a national GeoPlatform, intended to provide free and easy access to comprehensive data sets.

In reality, the Federal Government continues to grapple with producing and maintaining the NSDI, for numerous reasons. Coordinating new acquisition, maintenance of data, and easy distribution of data is a massive undertaking, both across agencies all at the federal level and between the federal level and other levels, such as states and regional entities. Numerous groups have researched the needs and shortcomings of the NSDI, including the

National Research Council (1993, 1995, and 2003), the Federal Geographic Data Committee (FGDC 2000), the Government Accounting Office (2004), The Congressional Research Office (date), and The National Geospatial Advisory committee (NGAC 2009). The Coalition of Geospatial Organizations (COGO, https://cogo.pro/) has published two report card assessments of the NSDI overall, and at both times, sub-par grades were awarded. There have also been several congressional hearings including the 2003 House Government Reform Subcommittee: “Geospatial information: Are we headed in the right direction or are we lost”.

Twenty-four years after the initial executive order, Congress passed The Geospatial Data Act of 2018 (GDA) that “codifies the committees, processes, and tools used to develop, drive, and manage the National Spatial Data Infrastructure (NSDI) and recognizes responsibilities beyond the Federal government for its development.” This formally acknowledges the complex mixture of needs and resources from various levels of government as well as the private sector.

2. Case Study: Land Use and Land Cover Inventory and Analysis

In the mid-1960’s Roger Tomlinson created the Canada Geographic Information System (CGIS) to manage the Canada Land Inventory (Tomlinson 1967). Using state of the art scanning and digitizing procedures the system assembled more than 3,500 maps. The CGIS is considered the first major GIS and demonstrated that it was feasible to perform several polygon-based analytical tasks on a large scale. It provided an impetus to conduct land cover inventories on a large scale, even if hardware and software were challenged. In the 1970s several federal agencies and forestry companies decided to inventory natural resources, and land use. Notable among these early GIS adopters were the US Geological Survey, the Bureau of Land Management, Department of Agriculture and the Fish and Wildlife Service.

Some early forms of raster data simply involved assigning values to cells on a transparent grid overlaid on a map or photo. These cells were only referenced to a Cartesian coordinate system. While tabulation of cell values could provide an inventory, producing pleasant output was a challenge (Figures 1 and 2).

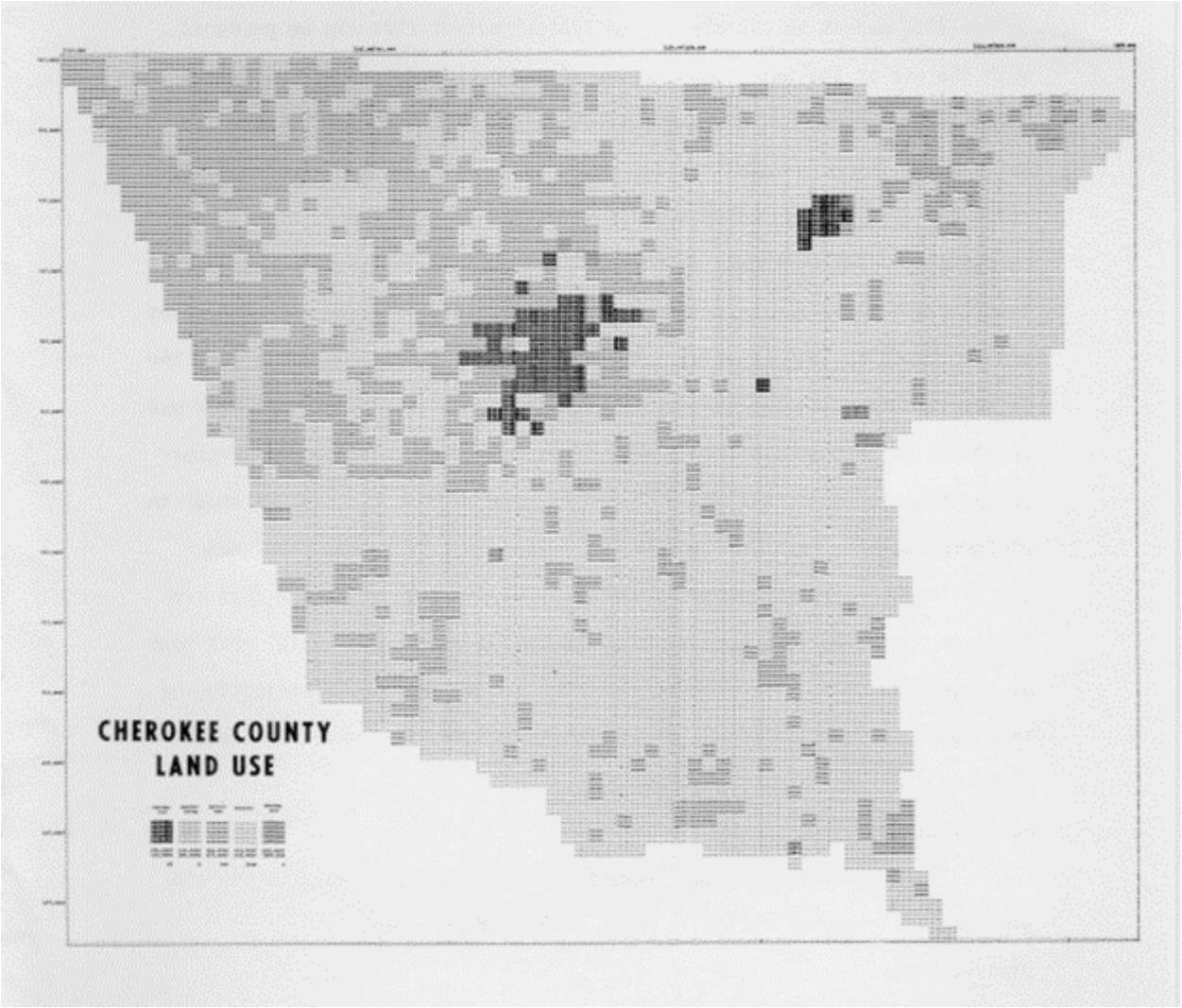


Figure 1. Line printer output for land use in Cherokee County, South Carolina, circa 1976. Note that each 200-foot cell is half an inch wide, the smallest square on a line printer. Source: authors.

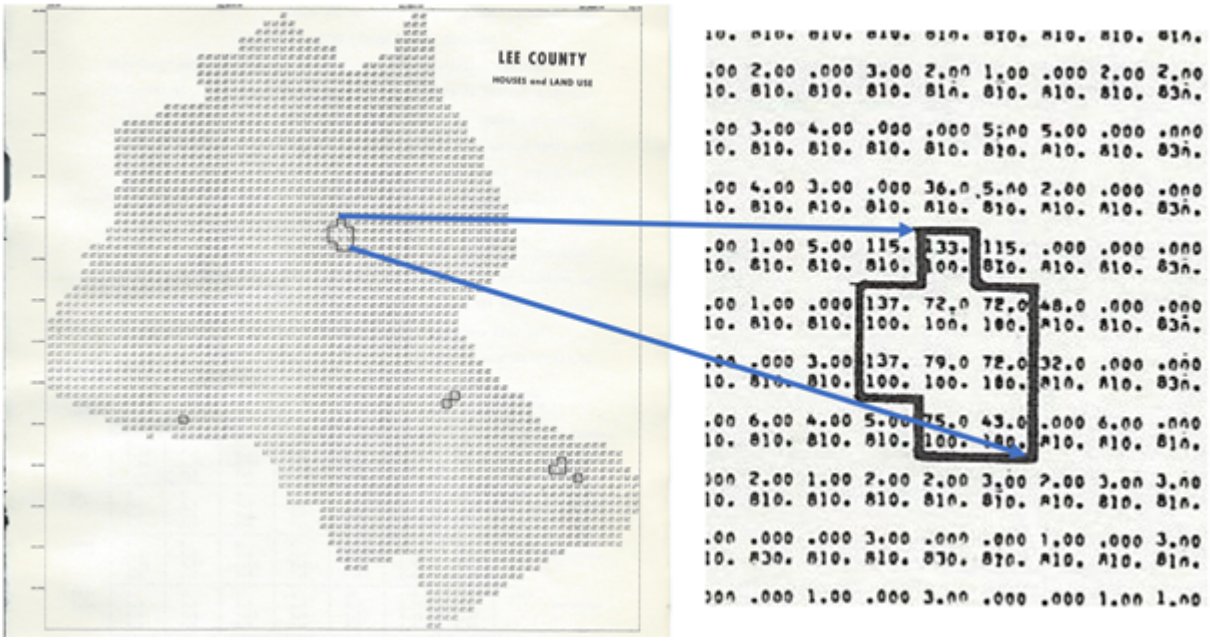


Figure 2. Counts of the State House and land use codes for cells in Lee County, South Carolina, circa 1976 .Source: authors.

Maps of square cells reproduced on a line printer were huge or distorted and pen plotters were slow and costly. Furthermore, generating color output in the mid-1970s required access to expensive state-of-the-art raster display devices (Figure 3).

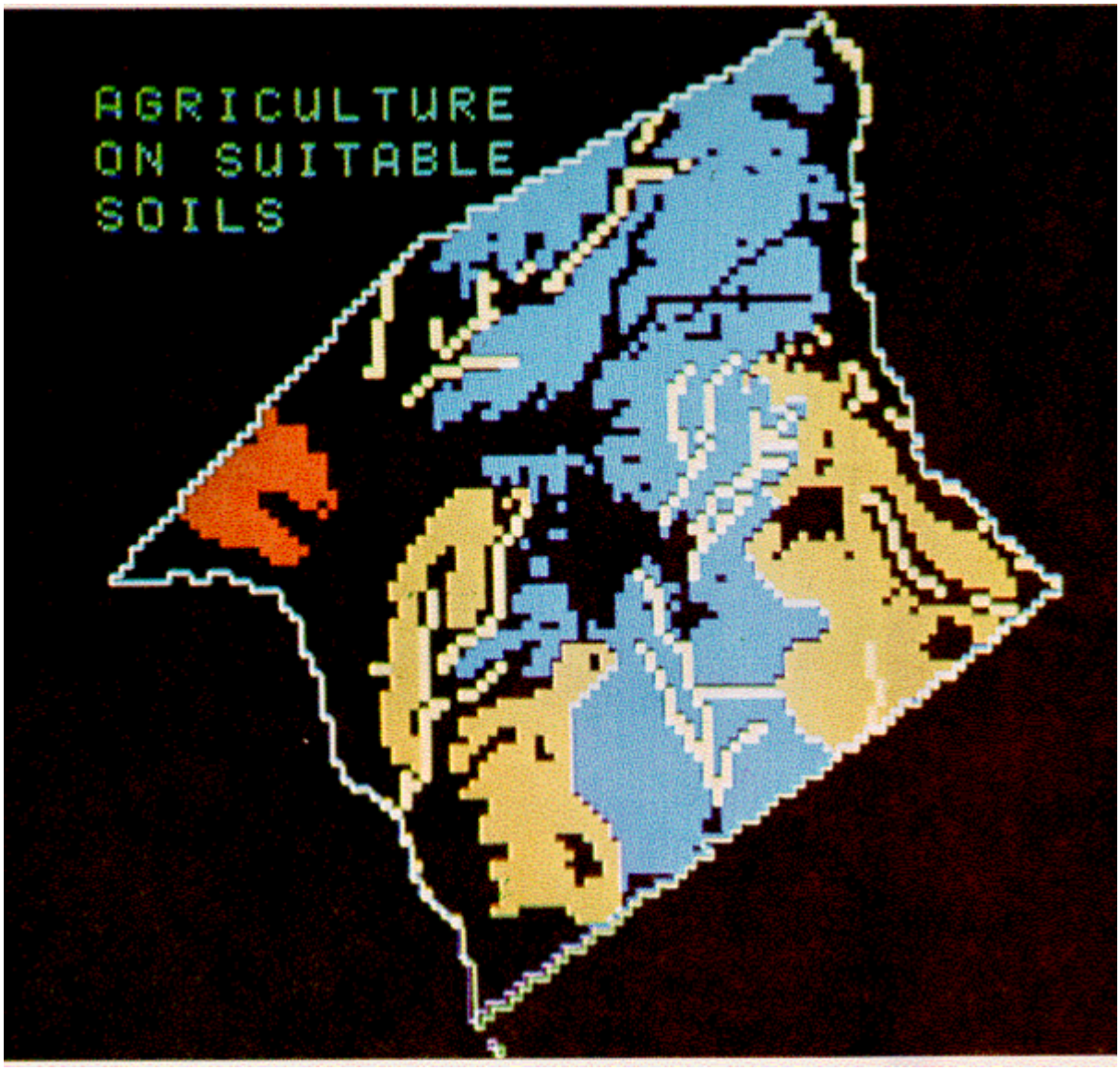


Figure 3. Color raster image display of agricultural land use on suitable soils for Anderson County, South Carolina, circa 1976. Data generated by General Electric company. Image source: authors.

One of the first areas of concern for early GIS developers was to address the needs from landscape architects and planners regarding suitability for development. A common approach to aid the decision-making was to produce maps of positive and negative factors on transparent media that could be overlaid to highlight suitable areas and justify decisions. According to Steinitz et al. (1976) the concept of map overlay can be traced back to at least the year 1912 when Manning (1913) prepared maps of alternative transportation patterns in Billerica, Massachusetts. The concept was popularized by Steinitz at Harvard and Ian McHarg at the University of Pennsylvania. McHarg's 1969 book "Design with Nature" became the bible for this type of suitability analysis (McHarg 1969). The process of "light table gymnastics" was greatly affected by the selection of factors and even the shade of gray to use. The methods for producing and using hand drawn overlays was an important research topic for landscape architects and planners who tried to develop rigorous procedures for the production of the overlays (Hopkins 1977). Nevertheless, the

activity was highly subjective and dependent on the overlay of a series of rather crude transparent overlays. Crude hand drawn transparent suitability maps have been replaced by sophisticated raster based Multi-criteria decision-making tools. There are no limits to the number of inputs or how they are weighted. Stakeholders can generate their own suitability models and be involved in the planning process. Analytical operations for working with a series of raster themes was well defined by Tomlin and others in the late 1970s. Tomlin's Map Analysis Package (MAP) (Tomlin 1990) provided a simple map algebra syntax to perform suitability analysis. His grid cell tools were widely adopted and form the core of many commercial and open source software systems used today.

The alternative spatial data model used by the CGIS maintained the fidelity of the individual point, line, and polygon features. It was also critical for dealing with topologically structured features such as transportation networks and boundaries. While the raster structure forced data layers into a fixed framework, vector-based integration of several layers presented many technical challenges. Overlay of themes with differences in scale, positional accuracy, and detail inevitably resulted gaps and overshoots that required extensive editing to create "clean" coverages. In the development of software tools to resolving these discrepancies was a core requirement for GIS data integration (Jensen et al. 2004). However, mature vector-based software tools were not available until the mid-1980s.

The concept of a land unit was a common approach for handling vector-based overlay analysis (Zonneveld 1989). Some landscape architects and planners viewed the landscape as a series of areas (polygons) that share similar factors such as slope, soils, vegetation, and land use. In practice, land units initially derived by interpretation from aerial photography or other form of remote sensing. A series of experts (soil scientists, geomorphologists, biologists etc.) worked together to identify the boundaries of land units through field observation. Thus, land unit polygons were composites of several factors developed through a consensus process that also resolved boundary issues as part of the procedure. It should be noted that the US Department of Agriculture still utilizes the concept of the Common Land Unit to refer to approximately 35 million individual agricultural fields that can be identified on aerial photographs (USDA 2017)

Several remote sensing systems have effectively automated the common land unit by combining imagery with vector themes such as soil and terrain.

3. Case Study: Authoritative Land Parcel Systems

A second major wave of GIS adoption focused on the administrative and financial needs of local government. As former Maryland Governor, Martin O'Malley has stated a GIS is successful only if it can "show me my house" (O'Malley 2009). Clearly, the ability to find, visualize, and retrieve information about specific structures on any computing device has made spatial data an indispensable part of everyday life. Even in the mid-1970s many local governments placed a high priority on the creation and management of their parcel records. Even forty years ago enlightened local government officials understood the importance of a well-designed data structure built on geodetic control (Figure 4). Several counties invested in expensive workstations to manage land records. In 1980, the National Research Council even called for a National Multipurpose Cadastre (National Research Council 1980). The public now expects, its local government to provide online maps and

information relating to the use, value, and ownership of property.

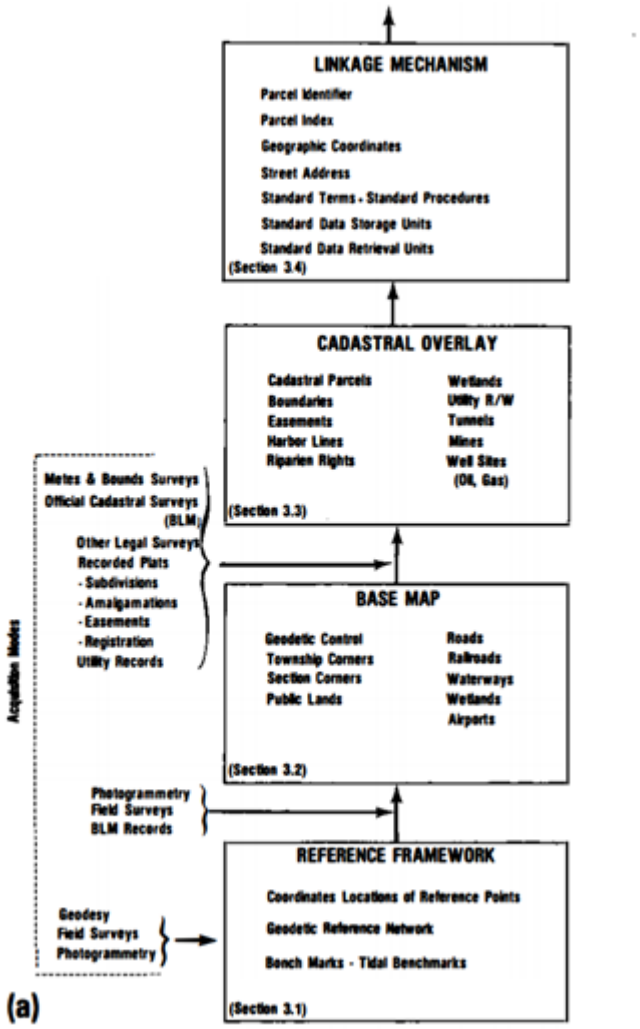


Figure 4. Foundation for a multipurpose cadastre, 1980. This demonstrates the importance of an accurate geodetic reference network to establish authoritative land records systems that include street addresses. Source: National Research Council (1980). Used with permission.

Multipurpose land record requires the capture of detailed features and links to authoritative information. Conceptually, land ownership is represented as a wall-to-wall coverage of mutually exclusive and non-overlapping polygons. Throughout the decade of the 1980s local governments struggled with ways to build parcel systems. Most of the initial programs focused on converting existing tax maps into an integrated parcel coverage. Unfortunately, tax maps are merely sketches of property lines and do not have legal status (Commonwealth of Massachusetts 1999). Nevertheless, digital versions of tax maps were often used to provide visual reference for a host of financial, administrative, and planning functions that did not require a high degree of positional accuracy (Figure 5). For legal purposes, the boundaries of a parcel are described by surveyors' metes and bounds on ownership deeds. It has been estimated that parcel data is the foundation for more than 30 administrative functions. Initially, tax maps were manually digitized or scanned and

vectorized. The need for authoritative parcels led to the development of special Coordinate Geometry (COGO) software-built polygons from the survey notes on a deed. By linking the property corners to surveyed markers each parcel can be fixed to real world coordinates. Local governments now manage these systems as a parcel fabric that often incorporates high resolution (6 inch) aerial photography and LiDAR. Maintaining these systems requires a trained staff, sophisticated computing environment and investments in high resolution data.

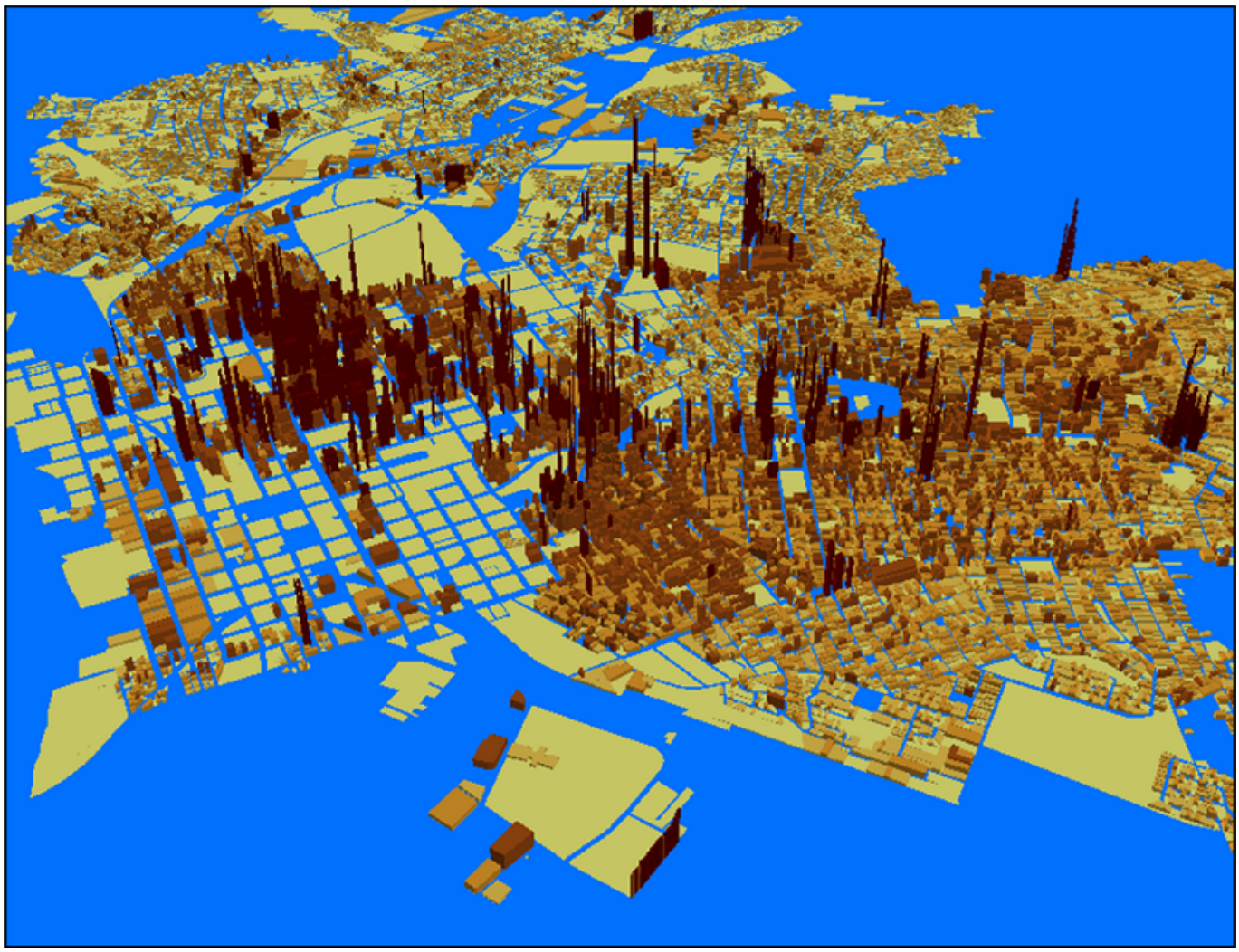


Figure 5. Parcel-based map of assessed value in Columbia, South Carolina. The map demonstrates the large number of tax-exempt properties associated with state offices, the University of South Carolina, and religious institutions. Source: authors.

Furthermore, parcels provide an extremely useful address database for geocoding (Figure 6). Using the site address of the parcel or its centroid it is straightforward to find matches in a list of addressees. A simple relational join can attach coordinates to an address. This procedure is a major tool for converting text to geospatial data. It is more accurate than interpolation from address ranges and forms the basis for a national address data based that will support the next generation of E911 service.

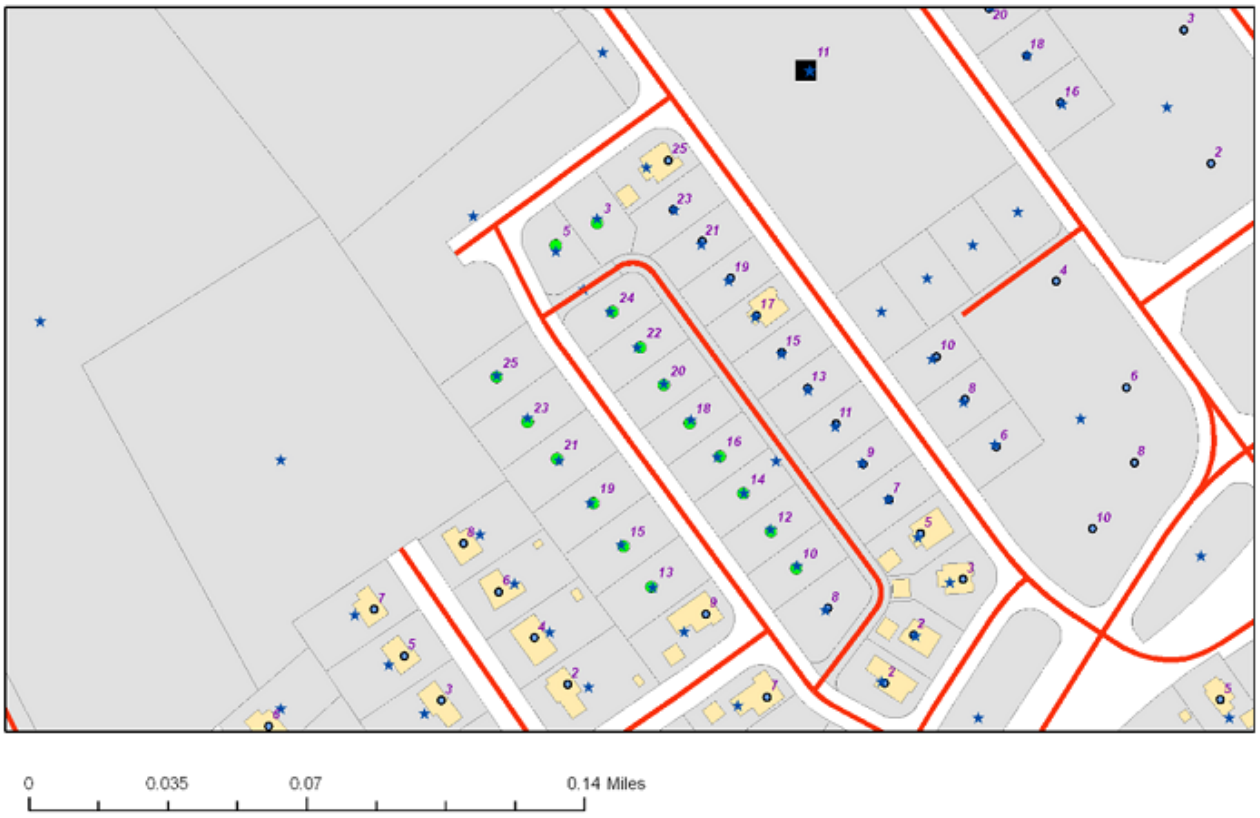


Figure 6. Parcel centroids used to create address point file for both improved and unimproved parcels. Source: authors.

It also supports reverse geocoding that associates attributed to features based on a spatial search from a user selected point. The algorithm utilizes a spatial search to locate and select address points or parcels near the point. A common application involves the selection and notification of residents near an incident. The search can be controlled by user defined parameters. Reverse geocoding is one example of creating new geospatial data through spatial search and overlay.

4. Case Study: Big Business

Automobile manufacturers and specialized GPS navigation manufacturers have required accurate streets, addresses and points of interest. The importance of this type of data was highlighted by Nokia's purchase of NavTeq for $8.1 billion in 2007. Since then NavTeq has been acquired by HERE, a consortium of European automobile manufacturers. The HERE data includes detailed lane information and huge collections of signage and pavement information. The appearance of Google Earth and Google Maps in 2005 was another game changer. They incorporated high resolution imagery as well as a detailed map base to provide the public with a free spatial search and navigation system. Google had the resources to build and maintain their own geospatial data. They deploy fleets of specialized vehicles to capture images and features.They also form partnerships with state and local governments to create high resolution imagery. They even provide 3D aerial oblique images. In 2012 Apple launched a similar service. As result, in today's world billions of

users are dependent on accurate geospatial data to find locations and navigate. Both the dedicated vehicle navigation systems and the mapping apps are pushing to improve the capture of real time information about road conditions, traffic, and weather conditions. In this manner they are using the vehicle drivers as a set of real time direct sensors. The current research and development involves geospatial data to support autonomous vehicles. Each vehicle is a data server that must process real time data feeds from active sensors to navigate through an ever-changing environment. The data must be acquired in real time and precisely. The emergence of successful prototypes of autonomous vehicles demonstrates the capabilities of the current data capture technology.

5. Case Study: Architecture, Engineering, and Construction (AEC)

Another recent series of adoption of GIS has been for architecture, engineering, and construction (AEC) companies. Agriculture, forestry and mining companies were early adopters that relied on GIS to inventory and manage their resources. Their needs were largely met with general coverage from air photos and traditional mapping. Today, they can utilize LiDAR to pinpoint specific trees, and calculate precise volumes of extracted materials. With thermal sensors on UAVs farmers can apply precise treatments to sections of their fields. These developments represent a modernization of the existing practices. Adoption of GIS by AEC companies constitutes a paradigm shift with new workflows.

In the 1970s, GIS and Computer Aided Design (CAD) had totally different software tools and even ways of representing features. Architects used CAD to automate traditional drafting operations. The drawing of a building was recorded in Cartesian coordinates. Special tools created perfect curves and joins, and the objects were typically linear or point features. Interior and exterior features were handled on different drawings. The features were not linked to a set of attributes that would facilitate filled polygons. At the same time the GIS developers began to explore ways to incorporate greater detail about the exterior of structures. These were important additions to support building inspections and utility infrastructure. This led to new models for the exterior of structures.

Most important was the Open Geospatial Consortium (OGC) standard City Geographic Markup language (CityGML). This data model can incorporate different levels of detail (LOD) for the representation of 3D urban objects as they move from the planning to the completion stage of development. As GIS tools and graphic processors improved in the way they handled 3D structures it became apparent that there was a need for the CAD drawings to be placed into geographic framework. It was also apparent that there were better tools for calculation of locations within a structure. This effort resulted in the Building Information Model (BIM) that incorporates features on the interior of buildings.

According to the OGC, a BIM links to and makes use of geospatial information such as: property boundaries, zoning, soils data, elevation, jurisdictions, aerial images, land cover, land use, etc.

“In this context ‘building' refers to the building process and BIM is a cumulative digital representation of physical and functional characteristics of a facility in the built environment. … Different stakeholders at different phases of the life cycle of a facility insert, extract, update or modify information in the BIM to support and reflect the roles of that stakeholder…. BIM is much more than the assembled 2D or 3D Computer-Aided Design

(CAD) and Facilities Management (FM) drawings created for the facility. The facility and its detailed information base need to be linked to the land on which it is sited and made available as an effective tool for AEC, owners and operators. Hence, geospatial information becomes a key component." Open Geospatial Consortium 2020.

In many ways this is taking a holistic view of space. It recognizes that everything is somewhere. The Internet of Things (IoT) has a geospatial component. Household appliances are now equipped with wi-fi that can used to fix their location. A dramatic example of this integration is the Hartsfield Jackson Atlanta International Airport that built complete 3D models of the interior and exterior of the world's busiest airport.

6. Case Study: the US Census Bureau

Five decades ago, the Census Bureau invented a geocoding system that has been used countless times to transform a set of addresses into a geospatial data. This effort was another game changer in the evolution of geospatial data capture. The US Census Bureau has always been a leader in the adoption of technology. During the 1960s, it began to explore ways to automatically locate street addresses. They first developed address coding guides that listed address ranges of streets within a census tract. For the 1970 decennial census the Bureau refined these procedures to create the Dual Independent Map Encoding system stored within a new Geographic Base Files format (GBF/DIME). These files imposed strict topological rules on street segments to form blocks. Using the address range on the left and right side of a DIME street segment it was possible to interpolate the coordinates of a street address and place it into the correct block. For the 1990 decennial census the Census Bureau and the USGS partnered to create the nationwide Topologically Integrated Geographic Encoding and Referencing (TIGER) database from the USGS 1:100,000-scale DLG data (Bureau of the Census 2015). To complete this important dataset the USGS accelerated the completion of the 1;100,000 topographic base map.

The release of TIGER line files was the catalyst for a new wave of applications that ran on inexpensive personal computers. In addition to demographic and housing applications based on census geography, TIGER provided a consistent nationwide tool to transform a set of street addresses into geospatial data. Following Office of Management and Budget guidelines, TIGER was placed in the public domain. This fueled the development of a whole new location-based industry. It was the basis for all web-based mapping systems such as MapQuest and even the street centerline for vehicle navigation. It can be argued that this marked the origins of all web-based mapping and navigation systems in use today including Google map, Open Street Map and those maintained by automobile manufacturers. Coupled with inexpensive, user friendly software on personal computers, mapping and spatial analysis became common place. The Bureau maintains a popular free online geocoding service based on interpolation along a potential address range.

TIGER had a huge impact on the popularity of GIS, however, it was created from a medium scale federal data base. At 1:100,000 TIGER line files did not align well with larger scale features created by local governments (Figure 7). Over the last three decades the Bureau has worked hard to develop its own address point files and to partner with local and state governments to improve the accuracy of its file. It has created web-based tools that facilitate a "bottom up" approach to share new features to maintain TIGER on a regular

basis.

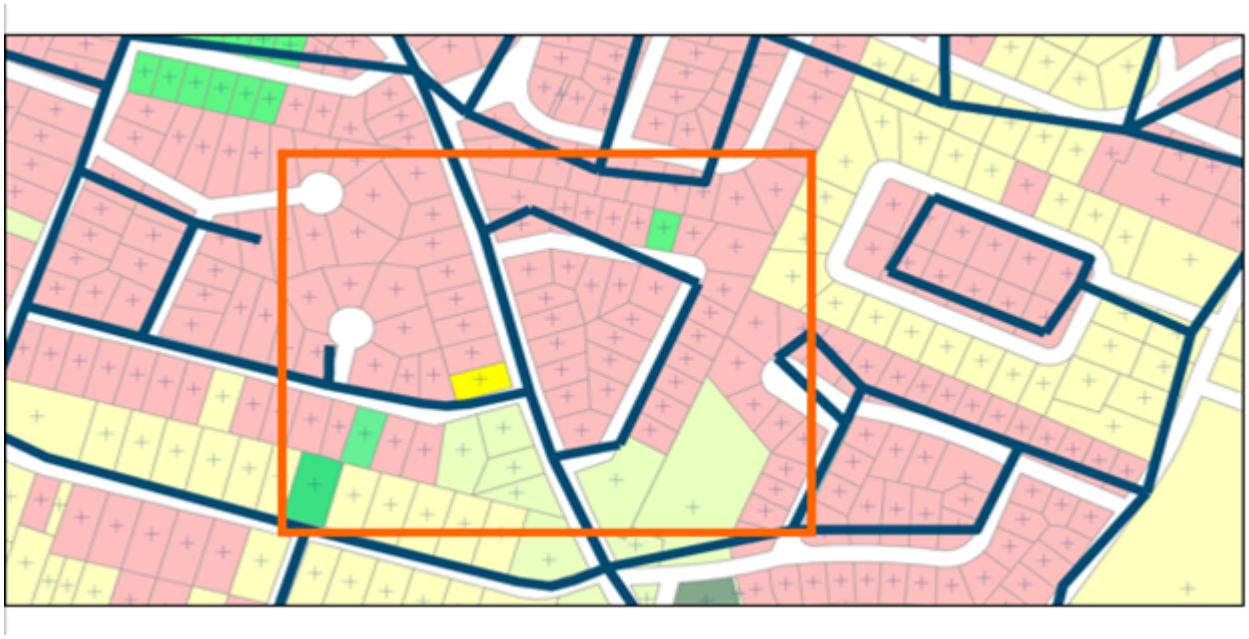


Figure 7. A comparison of TIGER streets and parcel data. Source: authors.

From a data capture perspective, it is important to compare geocoding based on the TIGER line files with parcel-based address points. A TIGER segment includes the potential range of addresses on each side of street. For many blocks there is a potential address range of a hundred addresses while the maximum actual address may only be a third of that range. For example, the street in Figure 8 has a range of 100 addresses, the address 116 is interpolated by TIGER to be located 16% of the distance from the start of the street. This point is about three houses away from the actual house. Clearly this is unacceptable for E911 and legal applications. In response to this limitation there is an effort to create a national address data base.



Figure 8. Comparison of TIGER geocoding and parcel-based addresses. The potential ending address is 100 however, the actual highest address is 29. Source: Generated by author from the Richland County, South Carolina GIS.

7. Case Study: The Role of the US Geological Survey (USGS)

In many ways the USGS has fueled the growth of GIS applications. The 125-year history of the production of its 7.5-minute topographic quadrangles provides a useful case study of the history of how maps were created and then converted into geospatial data. The production of a quadrangle required horizontal and vertical control, identification of cultural features, collection names, the capture of transportation and hydrological features as well as governmental boundaries. The production process required experts from cartography, geodetic surveying, benchmark surveying, photogrammetry, and remote sensing. As map making progressed into a digital production environment the differences between these professions blurred. Sophisticated mechanical devices were replaced by specialized software tools that are now accessible under the umbrella of GIS. The culmination of this transition is the online National Map, a suite of products and services that provide access to base geospatial information to describe the landscape of the United States and its territories. The National Map embodies 11 primary products and services and numerous applications and ancillary services" (USGS n.d.)

7.1 USGS and Map Making

When the USGS began producing topographic maps cartographers relied on extensive in situ field measurements acquired by surveyors using tape, compass traverses and aneroid barometers. Cartographers transformed field sketches into maps with contour lines. The development of plane tables and alidades that could measure angles greatly increased the ability to capture accurate contour lines. It was not until World war I that photos taken from aircraft became a viable source for map production. Furthermore, it took a urgent request from the TVA in the 1930’s to push photogrammetric approaches into large scale

production. As Usery et al. (2009) described the photogrammetric process:

The ability to view a three-dimensional terrain surface by doubly reflecting the overlap area, or stereomodel, of a pair of stereophotos in a multiplex stereoplotter effectively replaced the requirements of field sketching. An operator could fix a vertical floating mark at a preset elevation in the stereomodel and trace contours to represent the terrain. Similarly, tracing a road or other planimetric feature in the stereomodel, but allowing the mark to change elevation along the feature, provided recording of all required planimetric features for the topographic map." (Usery et al., 2009).

After World War II, the USGS refined photogrammetry-based map production and several of its employees developed specialized optical tools to improve the process. Regional mapping centers undertook the necessary field work, integrated the materials and compiled standardized paper and mylar hardcopy maps. This included photomechanical devices for stereoplotting, aerotriangulation, point measurement, and other photogrammetric operations. A major advancement was the creation of orthophotos by mechanical and photographic devices. The orthophoto process warps the source image so that distance and area correspond with real-world measurements. Developed by photogrammetrists in the 1960s the process takes overlapping stereo images and a digital elevation model to adjust for variations in terrain and tilt of the aircraft. (Figure 9).

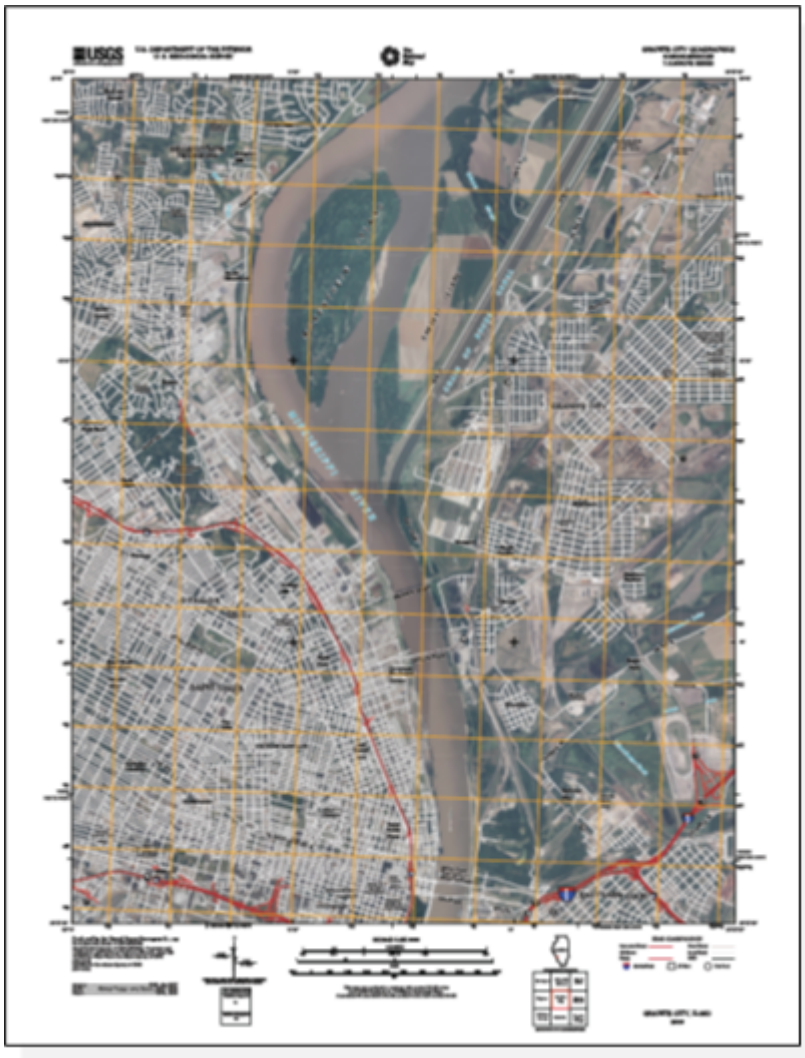
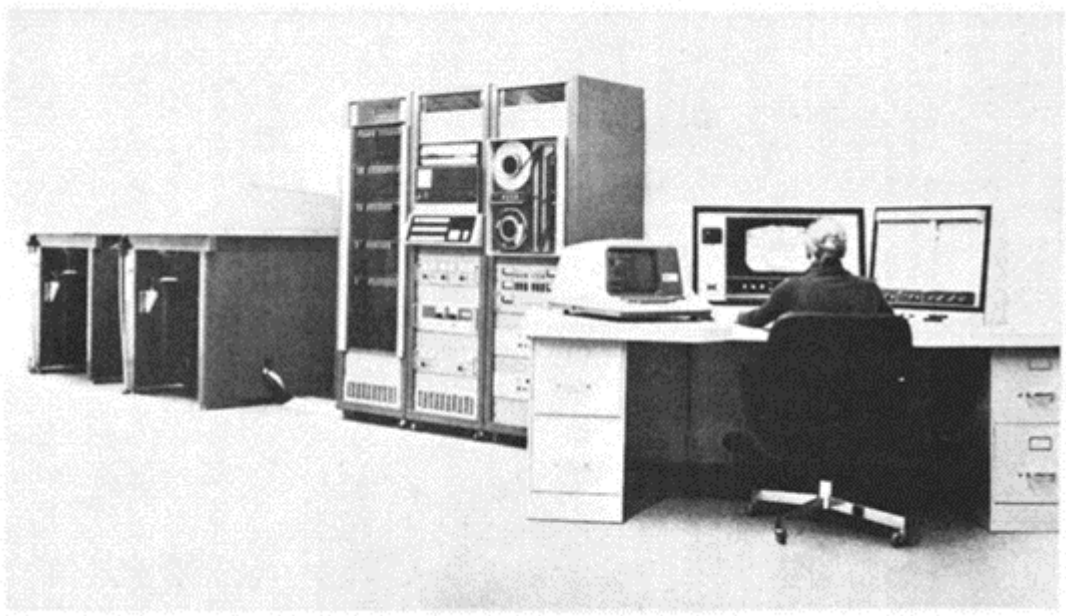


Figure 9. USGS orthophoto quarter Quadrangle. Source: USGS public domain.

The USGS started making orthophotos in 1965. When in the mid-1970s they acquired the Gestalt Photo Mapper, it greatly enhanced their production (Figures 10 and 11).



FRONTISPIECE. GPM-II. The scanners are at the left and the correlator and system computer are in the three cabinets at the center: the operator is seated at the console.

Figure 10. Gestalt Photo Mapper. Image source: ASPRS from Kelly et al. 1977, used with permission.

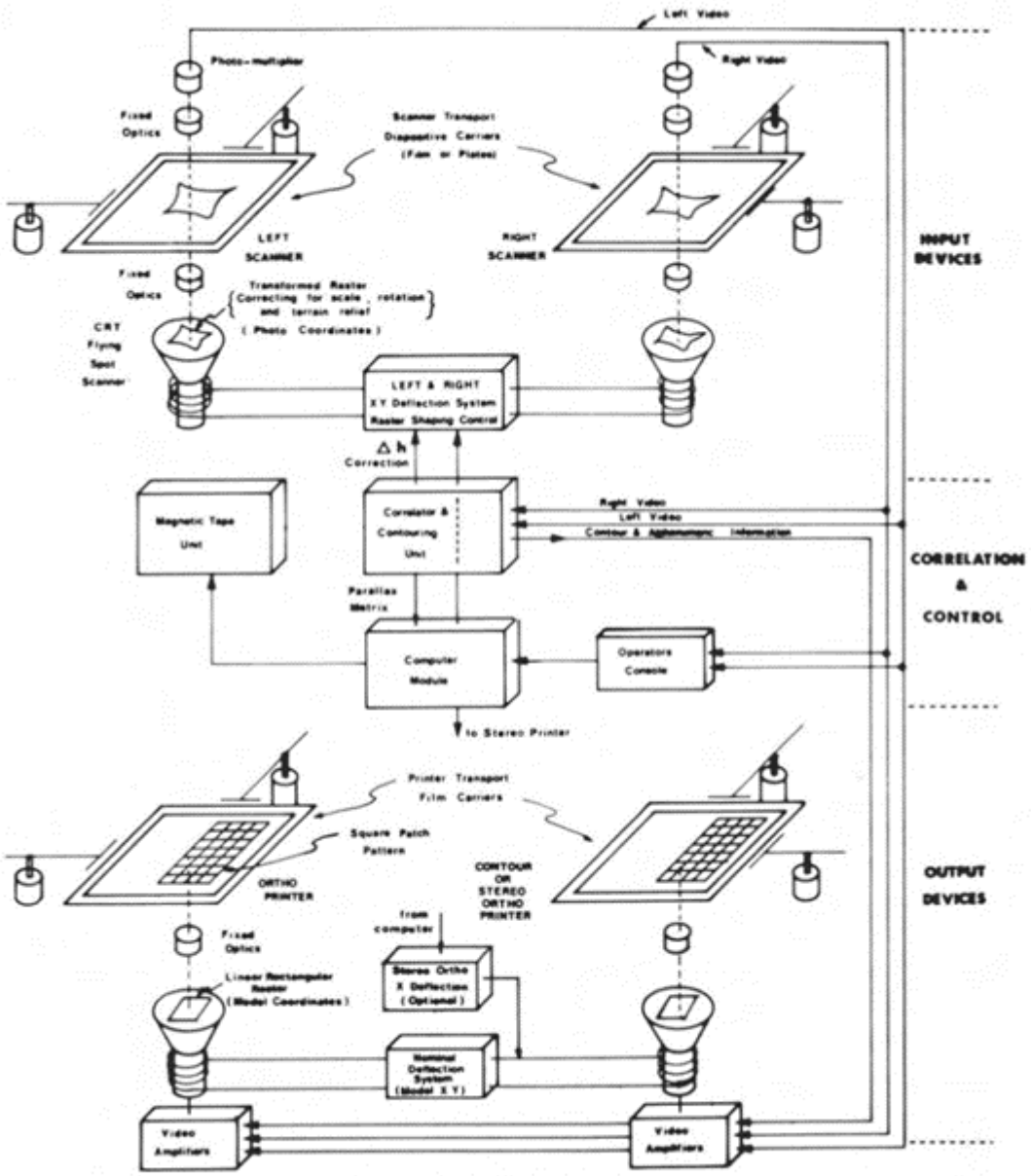


FIG. 7. GPM II block diagram.

Figure 11. Block diagram of the Gestalt Photo Mapper. Demonstrating the complex hardware and software generate orthophotos from stereo images. Image source: ASPRS from Kelly et al. 1977, used with permission.

The migration to a digital rather than photographic processes had a dramatic impact on map production. A quadrangle includes 200 features that were originally combined onto five different color plates. That meant that features such as roads, buildings, and lettering were combined on a black plate. The plates were designed for a five-color lithographic printing process. Clearly having several unrelated features on the same plate was not ideal for the transition to a digital production process. Ultimately there had to be a fundamental switch in the data model. Maps were to be produced from geospatial data, not be the source of Geospatial data. The USGS demonstrated that transition with the production of the 1:100,000 series DLGs. Instead of color separates, the maps were made designed with

digitizing in mind by combining 30 to 35 feature separates. This made scanning and vectorization more efficient. This change facilitated one of the milestones in the history of GIS. The map separates were scanned at a resolution of 1,200 dots per inch on large drum scanners. The raster data was thinned to a single pixel and converted to a smooth line. In partnership with the Bureau of the Census the vector lines were edited, attributed, and topologically structured to create the TIGER line files. This would not have been possible with the old printing-oriented process.

7.2 Scanning Existing Maps

While the USGS began the transition to a digital production environment with the 1:100,000 series there was a huge demand among the GIS community to convert the existing 55,000 topographic quadrangles into geospatial data. These maps were the de facto base maps of the Nation. They displayed the most complete set of features such as structures, transportation, hypsography, hydrology, and administrative boundaries. Digital versions of these maps provided critical base maps, data layers and reference information for adding additional layers. Since these maps were tied to a reference grid they became an important source of original geospatial data as well as being a framework for adding other themes. The first step was to use large format scanners to create digital raster graphic files (DRG) versions of the maps. (USGS 2015). Ultimately, the USGS scanned 178,000 historical topographic maps at several scales. Most of these are available for download from the USGS National Map web site in GeoTIFF and GeoPDF formats. These formats enable on-GIS users to turn map layers on and off, obtain real world coordinates, and take measurements. These raster maps are used to generate six different seamless base maps for on-line GIS environment. They also can be directly accessed by GIS platforms several web enabled formats including REST, Web Map Service (WMS), and Web Map Tile Service (WMTS).

To meet the demand for vector GIS layers the scanned raster version had to be converted to digital line graphs (DLG) (Figure 12). However, to add clarity, cartographers often adjusted the position of a feature. This degraded the positional accuracy of the theme and limited its utility.

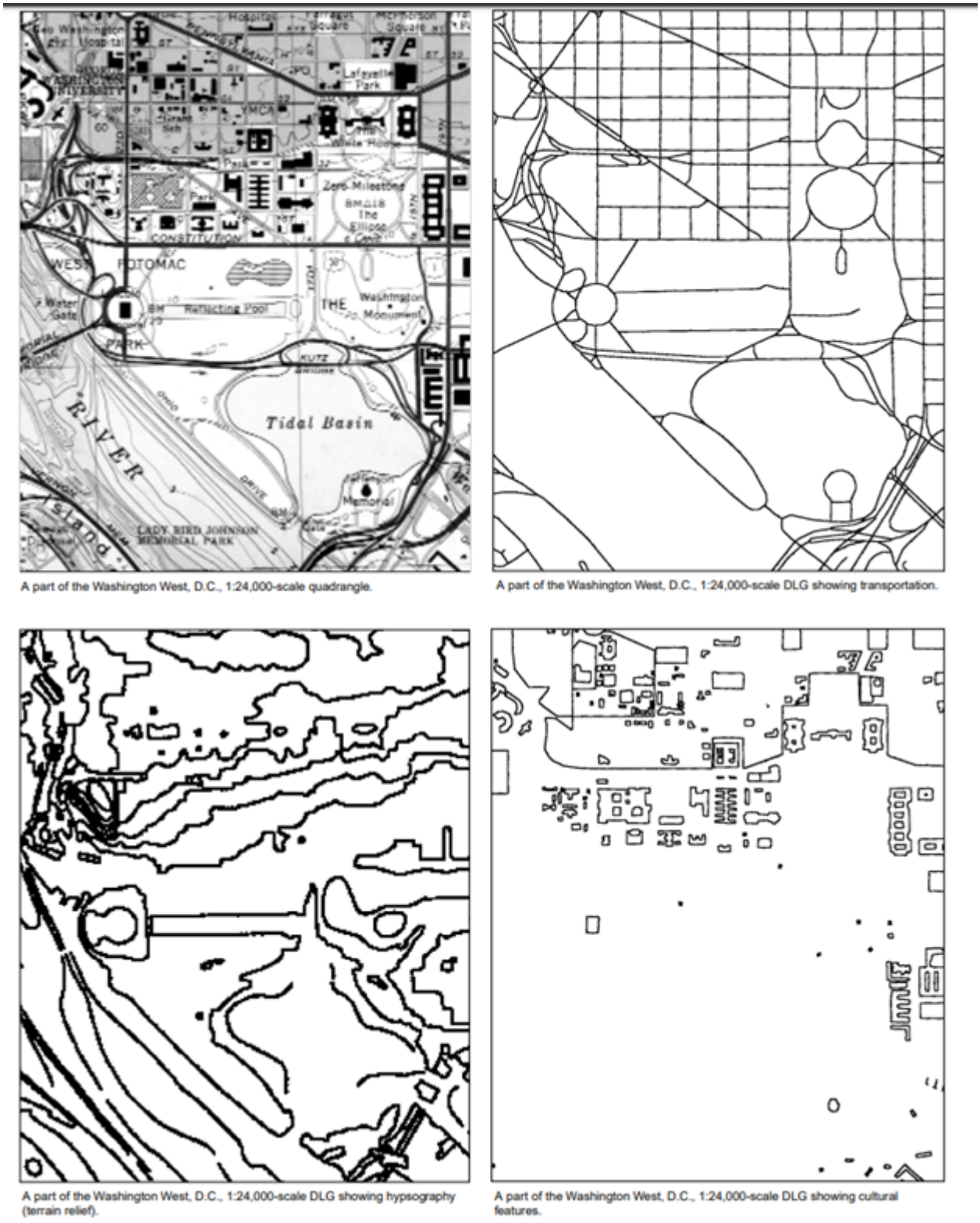


Figure 12. Separation of different themes from the DLG data. Source: USGS.

There were also problems with creating mosaics of features derived from maps with different update cycles. Using laser line following devices and classification algorithms enabled the direct capture of contour lines and linear features. The DLGs that were output from these procedures fueled a major expansion of GIS applications. These vector files were

at different scales: small (1:2,000,000), intermediate (1:100,000), and large (1:24,000) scales (USGS 1996). The 1:2,000,000 set was digitized from the recently updated National Atlas, an atlas of the United States which dates back to 1874 and was suitable as a cartographic base map. The 1:100,000 scale data provided a new consistent nationwide base map that was used to create TIGER for the 1990 decennial Census. These DLG files provided a useful multi-layer base map for numerous applications across the United States. For example, in South Carolina, the Department of Commerce used the 1:100,000 DLG to create a major statewide GIS for economic development. Additional layers such as water and sewer lines were captured by linking manually drawn features to DLG features (Figures 13 and 14).

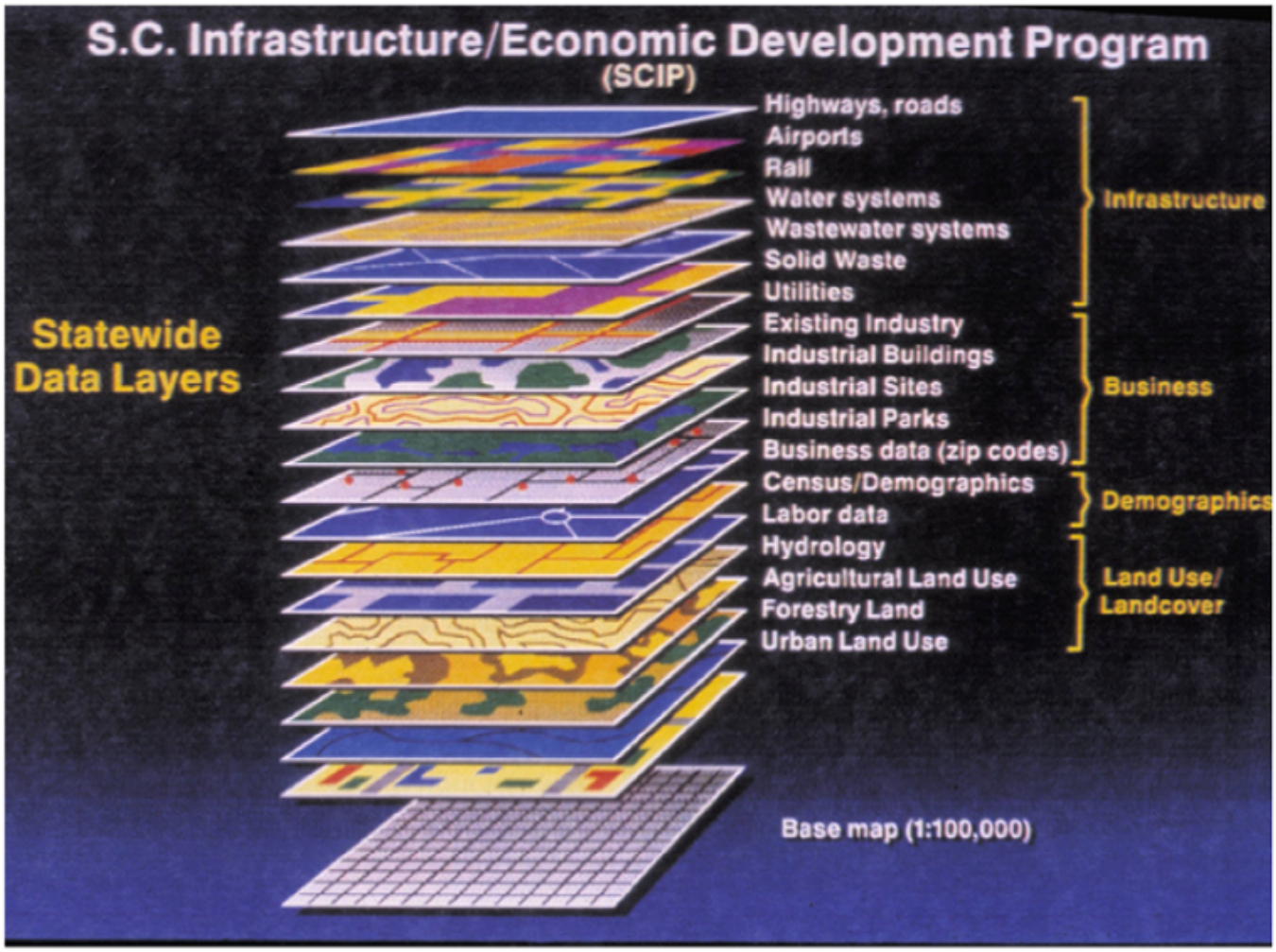


Figure 13. South Carolina Infrastructure and Economic Development Program Layers derived from the 1:100,000 DLG data. Source: authors.

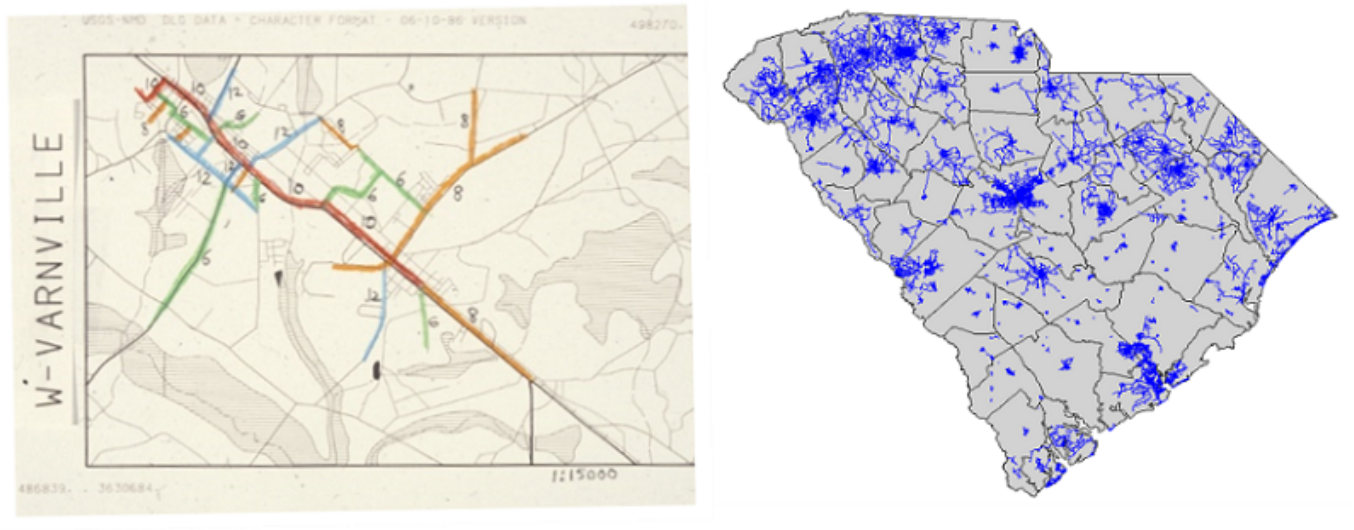


Figure 14. Example of major water lines drawn on 1:100,000 scale digital line graphs used to build a statewide layer of water lines. Source: authors.

In a similar manner, 1;24,000 DLG data became the base for the SC Department of Natural Resources and sixteen quadrangles were integrated to create the Environmental Data Atlas at the Department of Energy's Savannah River Site (Figure 15).

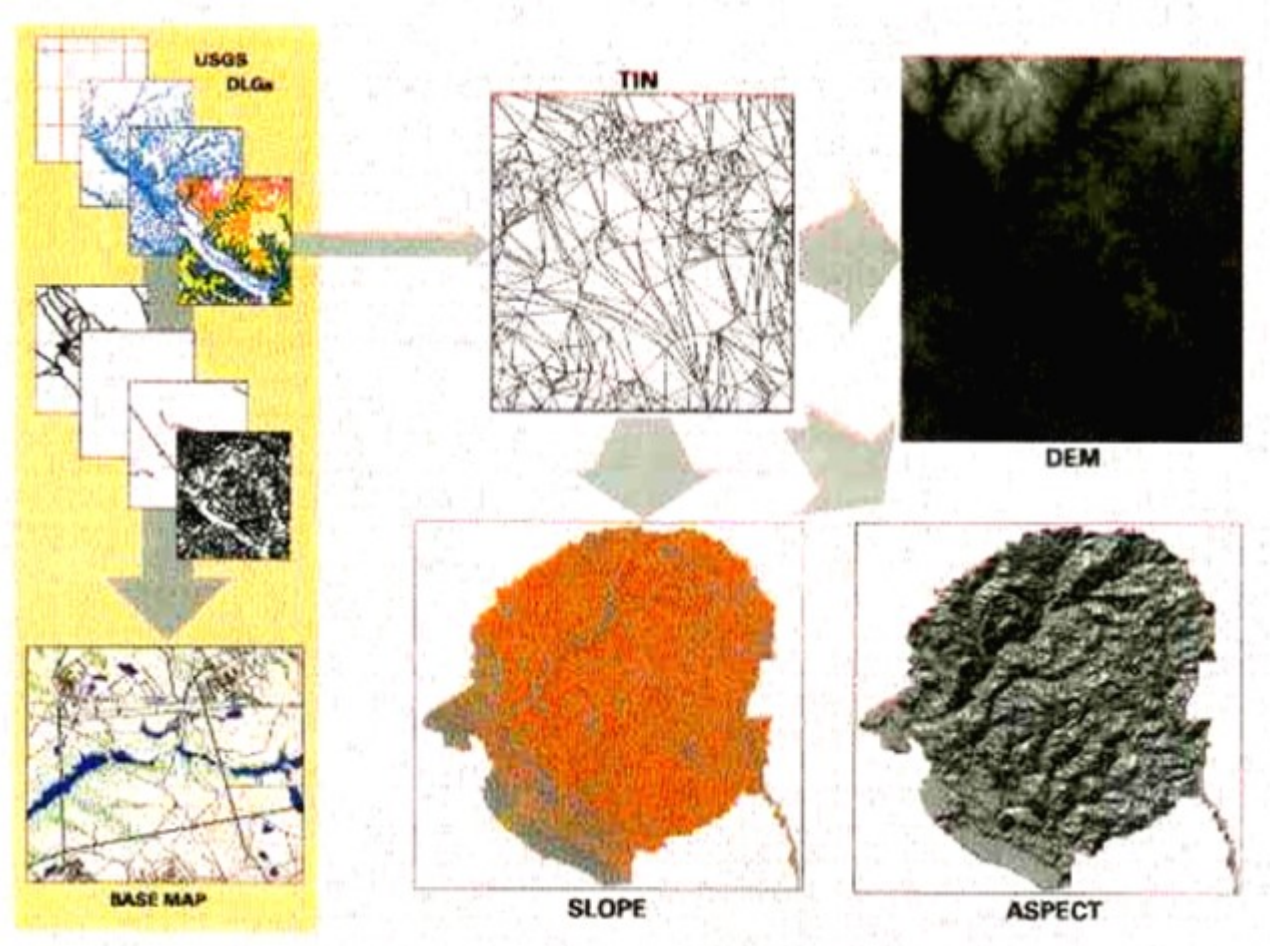


Figure 15. Savanna River Site Environmental Data Atlas Based on USGS 1:24,000 DLG Data.

Image source: ASPRS from Cowen et al. 1995, used with permission.

7.3 Land Use and Land Cover Analysis

In the 1970s the USGS embarked on a program to capture land use and land cover (LULC). This program utilized a new classification system specifically designed for mapping LULC polygons from high altitude photos (Anderson et.al., 1976). The Anderson classification provided a basis for two levels of land use and land cover. Over time, this was expanded to greater detail that required in situ research about specific land uses. LULC program brought into sharp focus issues about interpretation of features from photography. Through its efforts to create a consistent set of complex land cover polygons, the USGS developed the Geographic Information Retrieval and Analysis System (GIRAS) software package (Figure 16), another milestone in the history of GIS. In fact, GIRAS included some of the first routines to create and manage strict topological data structure for a complex set of polygons with embedded islands (Mitchell et al. 1977). The LULC data were used for several regional applications including the Coastal Zone management plan for South Carolina (Figure 17).

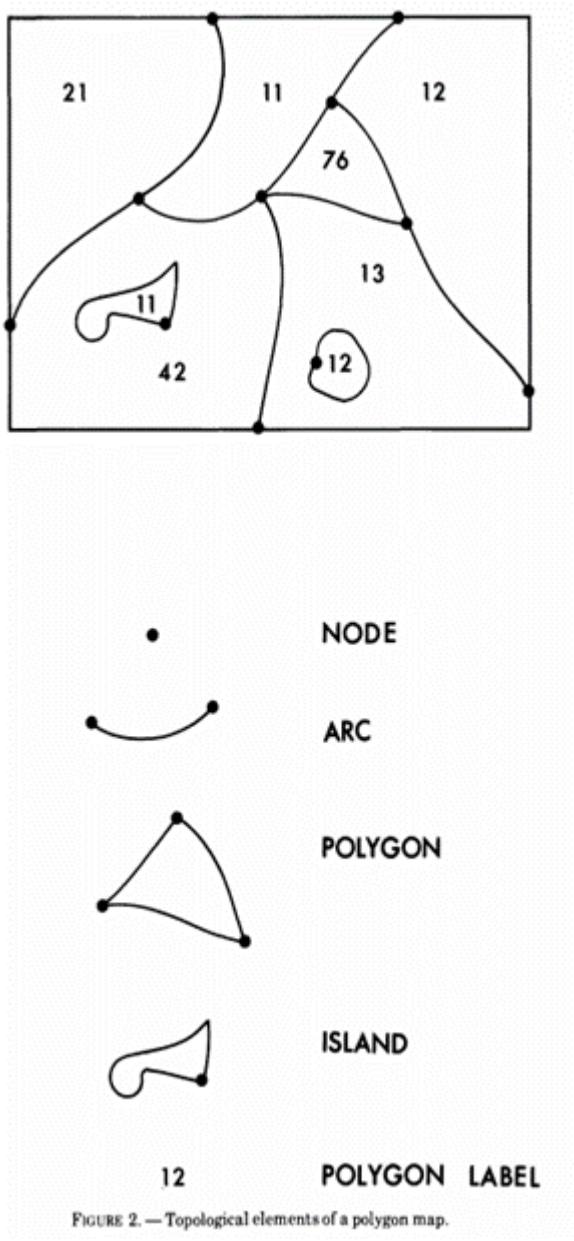


FIGURE 2. — Topological elements of a polygon map.

Figure 16. GIRAS arc node data structure that handled complex embedded islands. Source: Mitchell et. al., 1977 and USGS.

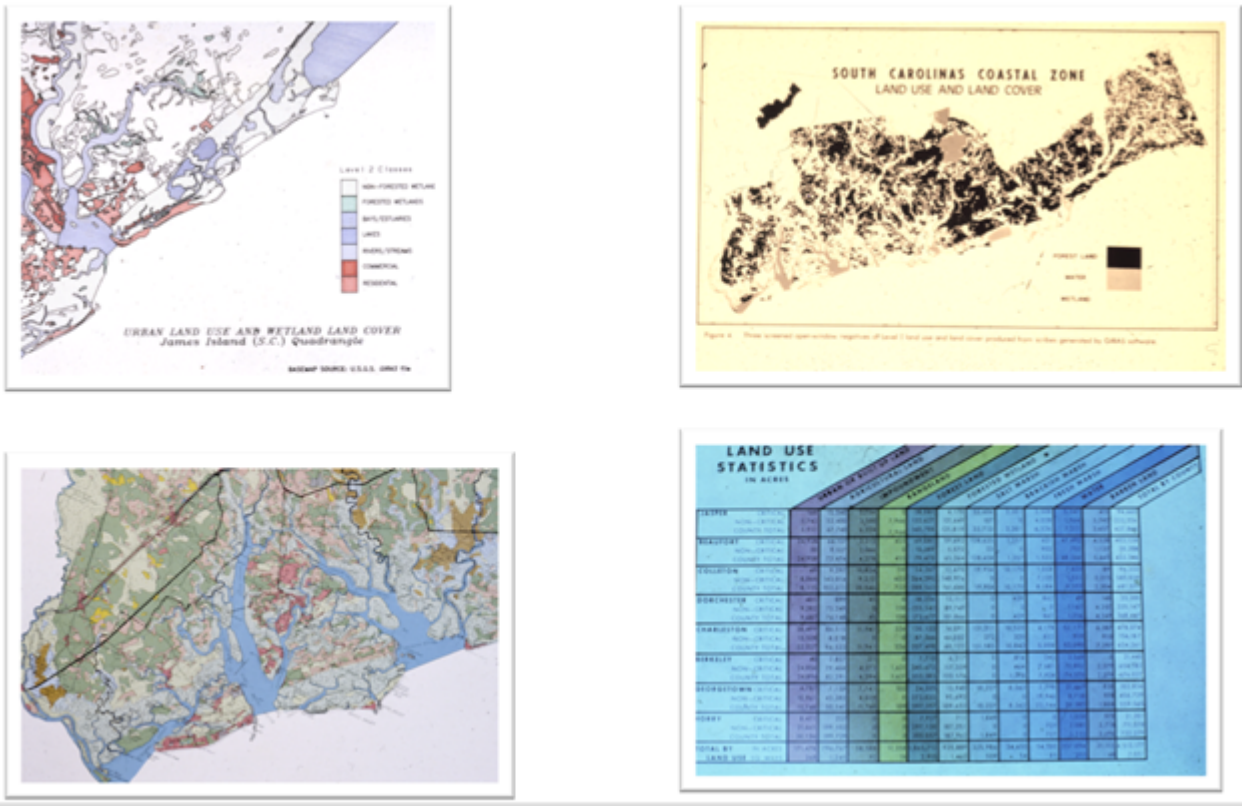


Figure 17. 1978 South Carolina SC coastal zone management maps generated from the USGS 1:250,000 scale USGS land use and land cover (LULC) data. Source: authors.

To create consistent land cover data highlighted the need to have uniform source materials. Assembling this data at the same scale for the same time period across the nation is difficult. Consequently, geospatial data collection often requires the user to make a trade off. The LULC data was interpreted from a set of very high-altitude photos at a scale of 1:250,000. While useful for macro analysis, those photos were not appropriate for the National Wetlands Inventory that was collected at 1:24,000 (Figure 18). A significant issue is the minimum mapping unit to collect. Since the land cover data is a wall-to-wall coverage, small non-mapped polygons are absorbed into larger ones. For example, the Anderson classification for 1:250,000 maps eliminates any urban land smaller than four hectares, and 16 hectares for most non-urban uses. This protocol also meant that only limited access highway features with a minimum width of 400 meters were captured. It is interesting to note that as a nationwide seamless data set the entire interstate system is one polygon!

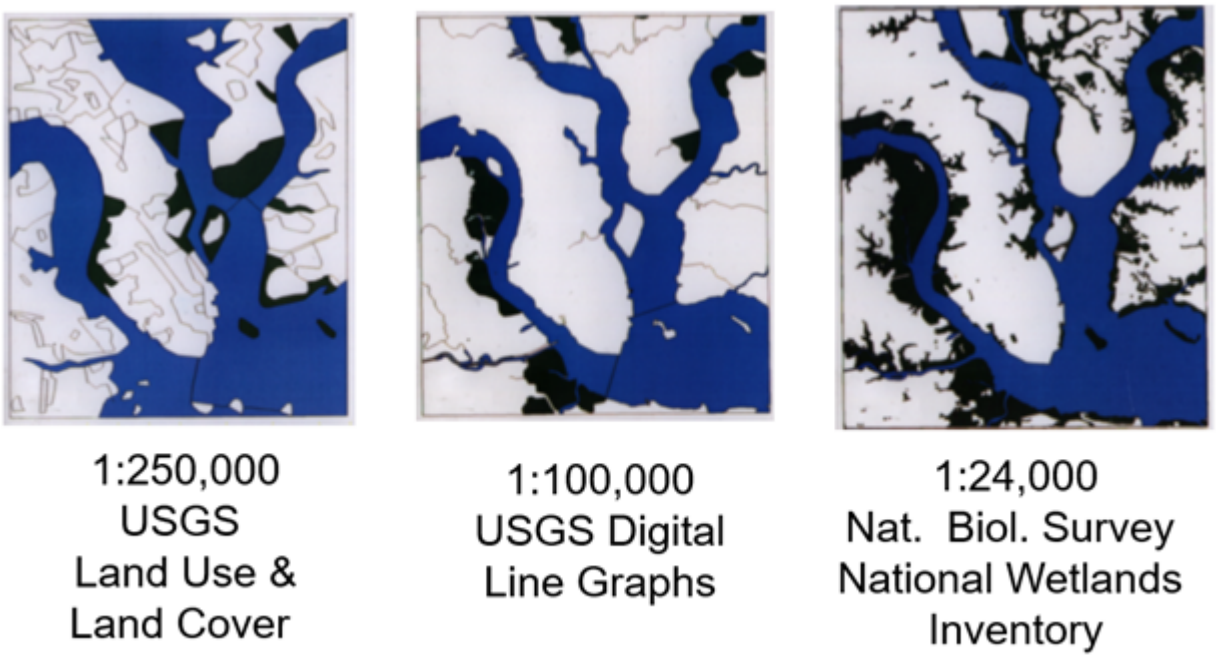


Figure 18. Comparison of wetland delineations from three federal databases. It is important for users to understand the different scale of the source materials and protocols for the collection of the data. Fortunately, this information is now available in metadata. Source: authors.

Clearly the size of the feature to be captured is determined by the scale of the source material. However, in a GIS environment, the user must also be concerned about positional accuracy of the feature. In the 1970s, the existing 7.5-minute quadrangles were considered “large scale” maps that were widely used as base maps, and it took more that 55,000 of them to cover the country. The acceptance of the quadrangle base map imposed a scale 1:24,000 on the capture of features from these maps or other source materials registered to them. According to National Map Accuracy Standards, the horizontal accuracy of any point was a 1/50th of an inch (USGS 1999). That standard means that 90 percent of all points tested must be within 40 feet of the actual location. Clearly this was not acceptable for applications dealing with structures and other small footprints, Furthermore, even Federal agencies digitized features from unstable paper versions of the quadrangles. Now, in the GNSS era, positional accuracy is often measured in a few meters or even centimeters.

7.4 Multispectral Earth Observation Satellites

To continue its efforts to capture land use and landcover data, the USGS has long fostered the use of satellite-based earth observation missions. Effectively, the launch of Landsat 1 in 1972 marked the beginning of a new era of earth observation based on multispectral data captured from digital cameras. In terms of Earth observation, multispectral cameras are passive sensors that capture natural emissions of light reflected from the sun’s rays. Instead of being recorded on film a digital camera assigns values to pixels. A passive sensor can sample various visible and non-visible parts of the electromagnetic spectrum. The resolution of data collected is based on the size of the storage array and the distance to the surface. Today these can range from cameras on low flying UAVs to orbiting

satellites. Remote sensing researchers have devoted their careers to finding the best combination of spectral signals for identifying different types of land cover and conditions on the earth. For example, capturing value in the infrared band helps to identify healthy vegetation. By correlating measurements from training sites on the Earth to values in the satellite data it is possible to generate categories of land cover and other characteristics (Figure 19). Image processing classification procedures are used to extract useful land use/land cover, agriculture, forests, urban sprawl, or other data that can be incorporated as raster data into a GIS environment. The most significant Landsat based program has been The National Land Cover Database (NLCD). This program captures nationwide data on land cover and land cover change at a 30m resolution. The multiple dates of these coverages provide an important resource for monitoring land cover changes.

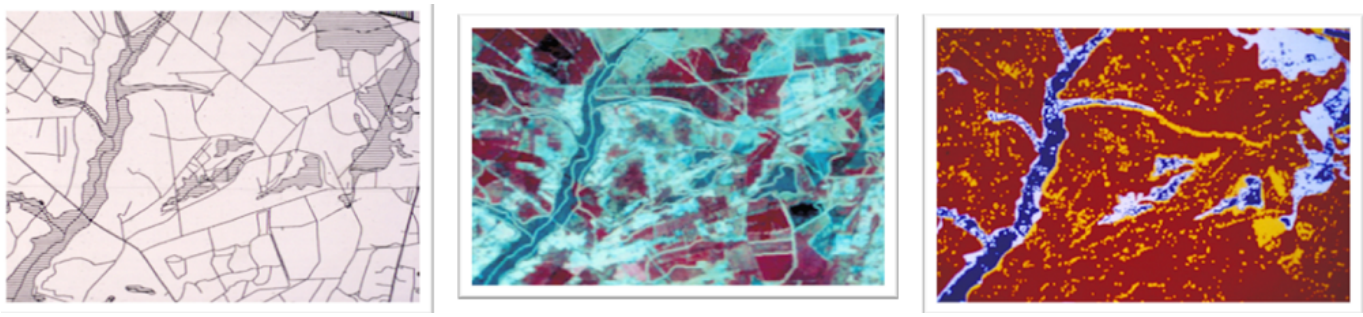


Figure 19. Early example, circa 1988, of vector and raster integration of satellite imagery. USGS 1:100,000 DLG overlaid on satellite image and classification of wetlands categories. This process required a physical linkage between two computers running ArcInfo and Erdas software. Source: authors.

The quality of information derived from remotely sensed data is necessarily limited by the spatial resolution, cloud cover, shadows, classification accuracy and other factors. Nowadays, while the US government continues to acquire 30-meter multispectral data from Landsat 8, other nations and commercial satellite companies have launched a variety of satellites with different spatial, temporal, and spectral characteristics. The current status of satellite-based earth observation is exemplified by Planet Labs' ability to provide 72-centimeter resolution RGB, NIR, and Panchromatic data daily from a fleet of 150 satellites. This level of spatial and temporal resolution has opened the door to several new application areas.

7.5 LiDAR

In 1884, John Wesley Powell persuaded congress to authorize the systematic topographic mapping of the United States. To gain knowledge of the terrain critical for navigation, discovery and settlement. To represent the terrain the USGS created contour lines on more than 55,000 1:24,000 scale maps. Originally, this process required extensive and often difficult field work with surveying instruments to establish horizontal and vertical control points (USGS 2009). Contour lines were manually interpolated from these benchmarks. Usery et.al. provides a useful discussion of the evolution of topographic mapping:

Topographic mapping historically has been approached as a map factory operation through the period 1879-1990. During this time, data were field and photogrammetrically collected; cartographically verified and annotated creating a compilation manuscript; further edited, generalized, symbolized, and produced as a graphic output product using lithography, or more recently, through digital means. Adoption of geographic information systems (GIS) as the primary production process for topographic maps, including digital database preparation (1975-2000) and product generation operations (2001-present), has led to faster and more standardized production in a semi-automated process. (Usery et al. 2018, 87)

Elevation is a continuous surface. In the field, surveyors sample elevation directly at specific points. With photogrammetric tools and stereo photos, it is possible to generate raster dataset of elevation data. For example, when the Gestalt Photo mapper produced raster orthophotos, it also collected a 700,000-point DEM. During the evolution of GIS many tools were developed to handle elevation data. There are routines to generate DTMs from benchmarks or contour lines. There are procedures to generate contours from DTMS and specialized routines to define watersheds and hydrological networks.

While the USGS original efforts provided an important portrait of the terrain of the United States the 10-foot contour lines portrayed on 7.5-minute quadrangles were simply not adequate for critical applications such as disaster planning or development. Consequently, there was an urgent need to find a new way to collect much more granular samples of elevation. In the mid-1980s the USGS enlisted the help of MIT's Draper Laboratory to develop a system to profile terrain based on Light Detection and Ranging (LiDAR) (Hurst 1985). LiDAR and RADAR are active sensors that emit pulses and capture the return values. RADAR devices emit radio waves while LiDAR generated light waves. RADAR is used in applications where detection distance is important but not the exact size and shape of an object. It is a vital tool for meteorological applications that provide real time observations of weather conditions. Synthetic-aperture radar (SAR) is a form of radar that is used to create two-dimensional images or three-dimensional reconstructions of objects, such as landscapes. The Interferometric synthetic aperture radar (InSAR) devices are used to generate DEMs especially in areas with inclement weather such as Alaska. LiDAR instruments emit huge volumes (200,000 pulses per second) of laser signals that provide very precise measurements of features. This revolutionized the direct capture of three-dimensional object. LiDAR devices can be mounted on a variety of vehicles, , and capture elevation and intensity measurements from everything they hit. A cloud of points that have X, Y and Z coordinates as well as intensity values. At high enough frequency, some LiDAR pulses can even penetrate dense tree canopy and create a "bare earth" DTM (Figure 20). LiDAR point clouds provide the input for high resolution DTMs which can generate user defined contours, and these DTMs are critical for hydrologic modeling and applications such as modification of the terrain (cut and fill).

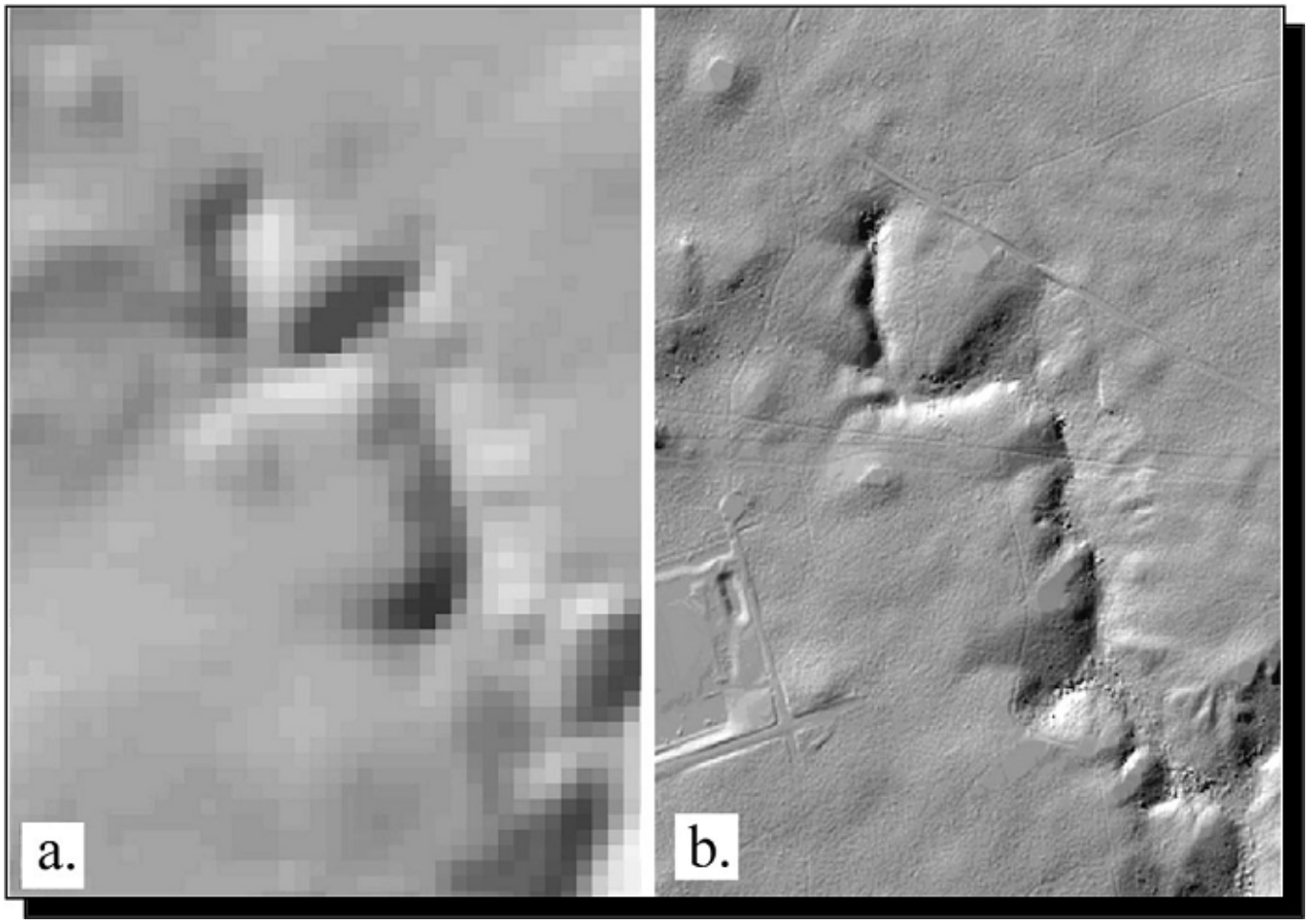


Figure 20. Comparison of Digital Elevation models from NED (a) and LiDAR (b). Source: USGS.

From a GIS perspective LiDAR mounted on aircraft provide a new way to capture the terrain while terrestrial LiDAR is used to capture details about structures to serve the AEC community. LiDAR data has also changed the approach to creating and maintaining hydrological features. The USGS in partnership with states and even local citizen scientists maintains the National Hydrography Dataset (NHD). LiDAR data supplemented by field observations is used to generate the NHDPlus High Resolution data set that aligns water features and watersheds to accurate terrain models.

8. Case Study: US Fish and Wildlife Service, and the National Wetlands Inventory

The National Wetlands Inventory (NWI) provides a valuable example of how imagery was incorporated into a GIS data base environment (U.S. Fish and Wildlife Service 2017). In 1977, the United States Fish and Wildlife Service (FWS) was charged with conducting an inventory of wetlands to support federal conservation efforts. Completion of this task required selection of appropriate source materials, development of protocols for selection of wetlands features and procedures for capturing the digital data. The USDA photography acquired during the growing season provided the best source for delineating wetlands, however, it is much easier to detect other features during leaf off conditions. The FWS

needed to develop a classification scheme for mapping wetlands and aquatic habitats from aerial photos (Cowardin et al. 1979). Protocols were established to help the analyst identify and capture different types of wetlands. The scientists acknowledged that the scale of the source material determined the minimum size of features that could be captured. As Cowardin et al., stated:

There is a limit to the size of a mapping unit that can practically be placed on a map and to the size of a water body or stand of vegetation that can be interpreted from a photograph.... some entire wetland basins and many wetland areas around the margins of basins are not detectable or mappable. (Cowardin et al. 1979).

Analysts manually traced wetland polygons onto acetate overlays with pen and ink. Using an optical zoom transfer scope to adjust scale the acetate overlayers were aligned with a 7.5-minute USGS topographic quadrangle. Cartographers manually transferred the polygons to a mylar overlay attached to the map. These 1:24,000 scale maps became the standard NWI product. To meet the need for a better system in 1980 the FWS contracted for the development of the Wetlands Analytical Mapping Systems (WAMS) for digitizing wetlands and the Map Overlay Statistical System (MOSS) to analyze the data on minicomputers. According to Carl Reed, who has written about the history about GIS:

The Wetlands Analytical Mapping System (WAMS) was an advanced map digitizing and edit package for topologically structured vector data. All work on these projects was completed in early 1980. As far as I know, WAMS was the first interactive digitizing system for capturing and structuring map data as topology in real time. (Reed, n.d.)

With WAMS, analysts were able to delineate wetlands directly on the monitor. In this environment analysts were able to incorporate other layers such as soils and topography to aid in the capture process. Although many areas of the country have not been updated, more than 270,000 unique users accessed the wetlands data last year. through the Fish and Wildlife Wetland Mapper (Figure 21).

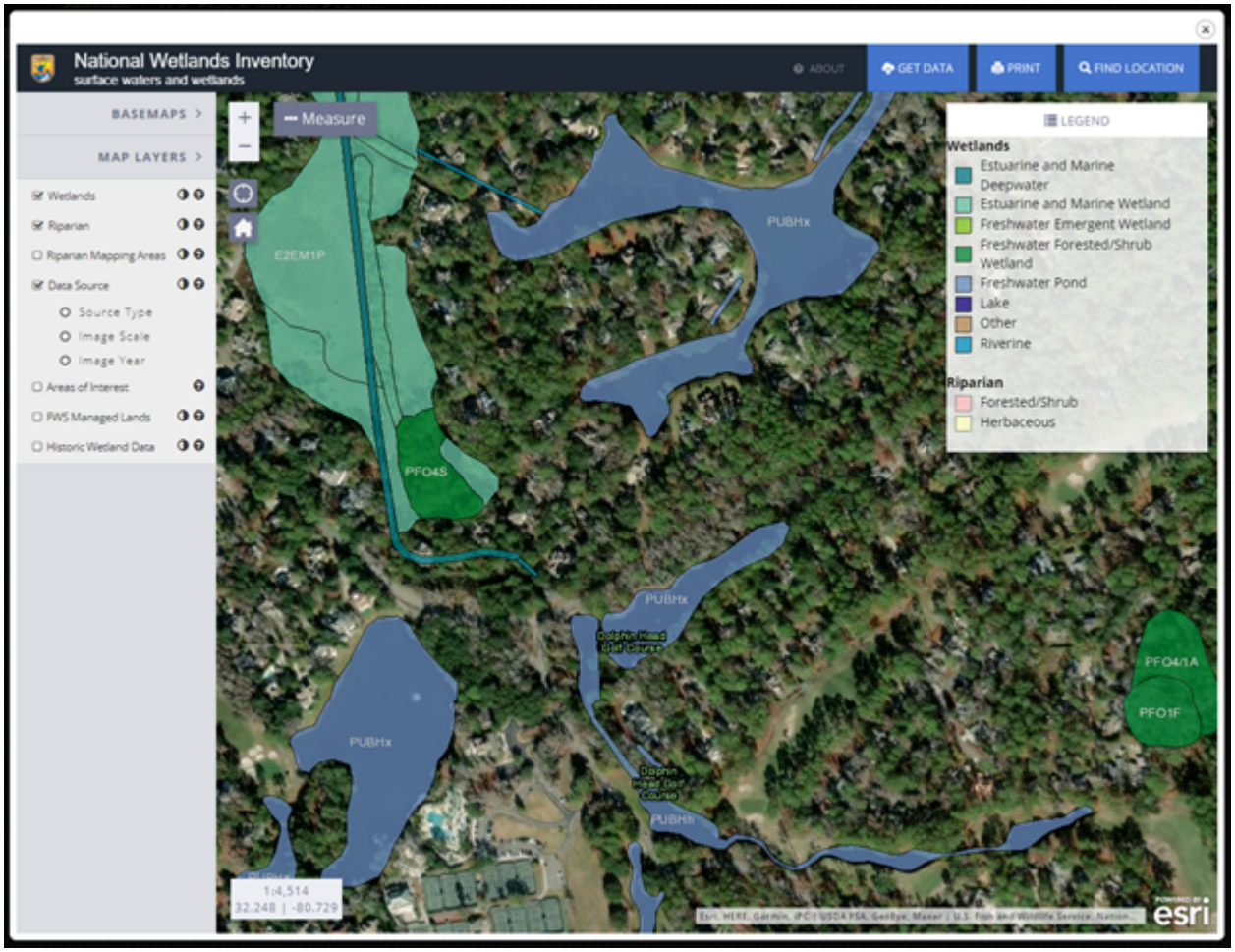


Figure 21. Image from the Wetlands Mapper (https://www.fws.gov/wetlands/data/mapper.html), maintained by the US Fish and Wildlife Service. The wetlands polygons are displayed in relationship to an orthophoto and can be downloaded from the same site. Source: authors.

9. Case Study: the US Department of Agriculture

The United States Department of Agriculture (USDA) began acquiring aerial photographs in the 1930's to inspect soil and crop conditions. In the mid-1950s, it started a systematic program to inventory and monitor agricultural practices. This program conducted by the Farm Service Agency (FSA) utilized large format cameras to acquire 9 x 9 inch panchromatic (B &W), natural color, or color infrared film. The total number of USDA photos produced during that program is unknown. However, the FSA is in the process of scanning more than 10,000,000 film negatives that will be suitable for georeferencing.

Before the digital era, analysts would utilize visible features such as fence lines, roads, and waterways to draw agricultural fields directly on the nine-inch photos. The approximately 35 million Common Land Units (CLU) are the fundamental elements of USDA programs and are not available to the public. Originally, the area of these CLUs was manually calculated by a planimeter. The USDA also utilized imagery to monitor a farmer's compliance with their crop plans. This included the use of 35 mm slides that were manually projected onto the CLU maps. It is estimated that more than a billion of these 35 mm slides are in an archive at the Aerial Photography Field Office (Mathews and Vanderbilt 2008).

In the late 1990s the USDA realized it needed to create a seamless layer of all farms, ranges, and pastures in the nation. To produce and maintain this data, the USDA established 13 digitizing centers. The transformation to a digital data base required analysts to transfer data from hard copy maps to a Digital Orthophoto Quarter Quad mosaic for each county. Old photomaps were scanned for heads up digitizing of CLU polygons aligned to the mosaic DOQ and authoritative control points.

Since 2003, the USDA has used one-meter digital orthophotography from the National Agriculture Imagery Program (NAIP) to monitor compliance. The basic NAIP data provides 1-meter pixel resolution imagery, available through the National Map download client and often used as a base map for many applications. However, finer resolution data is often acquired via partnerships. Predecessors to NAIP include the National High-Altitude Program (NHAP) that operated between 1980 and 1989 as an interagency federal effort coordinated by the USGS. It consists of more than 1.3 million images acquired on 9-inch film and centered over quarters of USGS 7.5-minute quadrangles. The National Aerial Photography Program (NAPP) which operated between 1987 and 2004.was the successor of NHAP and collected photos for the U.S. every 5 years. The archive of images from NHAP, NAPP, and other Forest Service programs are available through the Aerial Photography Field Office.

The federal approach to acquiring new orthophotography is very much in flux. The USDA and the USGS share stewardship of Federal programs for collection of orthoimagery, which is considered a critical part of the framework for the National Spatial Data Infrastructure. Another important source of imagery is NOAA's Coastal Mapping Program which collects orthorectified digital imagery on a regular basis and after natural disasters. According to the recent COGO report card (Coalition of Geospatial Organizations 2018), the Department of Agriculture Farm Service Agency (FSA) is considering licensing a commercial dataset to meet their needs. The demand for high resolution aerial photography has prompted state and local governments to establish their own programs. In fact, some fast-growing local governments acquire 6-inch photographs at frequent intervals to aid in building inspections and other issues (Figure 22). Consequently these public-private partnerships could eliminate NAIP as a public domain imagery source.



Figure 22. Comparison of 6-inch aerial imagery with 1-meter USDA NAIP imagery in Lexington County, South Carolina. Image sources: Lexington County GIS and USGS.

10. Conclusion

Opportunities and responsibilities for geospatial data usage and management have changed dramatically over time. These have affected whole categories of activities, such as those around mapping land use and land cover. Federal agencies involved with data production, maintenance, and distribution have had their workflows significantly altered over time. Final passage of the Geospatial Data Act (2018) was a monumental achievement for this geospatial domain of activities. While many federal programs continue to capture geospatial data, a massive amount of daily data capture is orchestrated by wealthy private sector firms that are constantly working to enhance the variety and quality of information required by billions of users. Most of these users want only to visualize the data to meet their immediate need for spatial search and navigation. At the same time, these users are also sensors themselves. They collect geotagged photos and create traces of their movements. Volunteers help “crowdsource” information about new features and report conditions to authorities. Instead of just using geospatial data, users are active participants in the production and maintenance of the data. Now that these capabilities and practices are in place, they have changed forever the practice of geospatial data capture.

References

COGO (Coalition Of Geospatial Organizations). (2018). Report Card On The U.S. National Spatial Data Infrastructure.

Commonwealth of Massachusetts. (1999). Parcel Mapping Using GIS: A Guide To Digital Parcel Map Development For Massachusetts Local Governments. Executive Office of Environmental Affairs, University of Massachusetts Office of Geographic Information and Analysis for MASSGIS

Cowardin, L. M., Carter, V., Golet, F. C., and LaRoe, E. T. (1979). Classification of wetlands and deepwater habitats of the United States. US Fish and Wildlife Service FWS/OBS 79/31.

Cowen, D. J., Jensen, J. R., Bresnahan, P. J., Ehler, G. B., Graves, D., Huang, X., Wiesner, C., and Mackey, H. E., Jr. (1995). The Design and implementation of an integrated Geographic information System for Environmental Applications. Photogrammetric Engineering & Remote Sensing (PERS) 61(11), 1393-1404.

Esri News. (2009). Governor Martin O'Malley Leads with GIS. ArcNews Summer 2009.

Federal Geographic Data Committee (FGDC). (2000). Improving Federal Agency Geospatial Data Coordination. Sponsored report; MS Word document to download.

Federal Geographic Data Committee (FGDC). (2002). NSDI Framework.

Government Accountability Office (GAO). (2004). Geospatial Information: Better Coordination and Oversight Could Help Reduce Duplicative Investments GAO-04-824T.

Hopkins L. D. (1977). Methods for Generating Land Suitability Maps: A Comparative Evaluation. Journal of the American Institute of Planners, 43(4), 386-400.

Hurst, M. E. (1985). Geography has neither existence nor future. In R. J. Johnston (ed.), The Future of Geography. Methuen.

Jensen, J., et al. (2004). Spatial Data Acquisition and Integration, in McMaster, R.B. and E.L. Usery, (eds.), A Research Agenda for Geographic Information Science, Boca Raton, FL: CRC Press.

Kelly, R. E., Mcconnell, P. R., and Mildenberger, S. J. (1977). The Gestalt Photomapping System. PERS, 43(II) ,1407-1417.

Manning, W. H. (1913). The Billerica Town Plan. Landscape Architecture III(3)108-118.

Mathews, L., and Vanderbilt, B. (2008). Geospatial Data and the APFO: Past, Present and Future. Esri User Conference, 2008.

McHarg, I. L. (1969). Design with Nature. Garden City, N.Y.: Doubleday/ Natural History Press.

Mitchell, W. B., Guptill, S. C., Anderson, K. E., Fegeas, R. G., and Hallam, C. A. (1977). GIRAS: a geographic information retrieval and analysis system for handling land use and land cover data. Professional Paper 1059. Washington, DC: US Geological Survey.

National Geospatial Advisory Committee (NGAC). (2009). NGAC Report: The Changing Geospatial Landscape.

National Research Council (NRC). (1980). Need for a Multipurpose Cadastre. Washington, D.C.: National Academy Press.

National Research Council (NRC). (1993). Toward a Coordinated Spatial Data Infrastructure for the Nation. Washington, D.C.: National Academy Press.

National Research Council (NRC). (1995). A Data Foundation for the National Spatial Data Infrastructure. Washington, D.C.: National Academy Press.

National Research Council (NRC). (2003). Weaving a National Map: A Review of the U.S. Geological Survey Concept of 'The National Map'. Washington, DC: The National Academies Press.

Open Geospatial Consortium (OGC). (2020). What is BIM About?

Reed, C. (n.d.) Short History of the MOSS GIS.

Steinitz, C., Parker, P. and Jordan, l. (1986). Hand-drawn Overlays: Their history and proscriptive uses. Landscape Architecture 66(5) 444-445.

The Congressional Research Service (n.d.).

Tomlinson, R. F. (1967). An Introduction To The Geo-Information System Of The Canada Land Inventory. Ottawa (Canada), Department of Forestry and Rural Development.

United States Census Bureau. (2015). Twenty-fifth Anniversary of TIGER. StoryMap.

United States Department of Agriculture (USDA). (2017). Common Land Unit (CLU) Information Sheet. U.S. Department of Agriculture.

United States Geological Survey (USGS). (1996). GeoData Digital Line Graphs Fact Sheet.

United States Geological Survey (USGS). (1999). Map Accuracy Standards, USGS Fact Sheet 171-99, November 1999. Accessed June 2021.

Usery, E. L., Varanka, D. E., and Finn, M. P. (2009). 125 Years of Topographic Mapping, Part 1, 1884 - 1980, in ESRI ArcNews, vol. 31, no. 3, p. 1.