REIMAGINING THE HISTORY OF GIS

Michael F. Goodchild

Professor Emeritus
Department of Geography
University of California
Santa Barbara, CA 93106-4060, USA

Abstract

Fifty years after the initial efforts that coalesced as GIS, it is possible to look back and ask whether the decisions made then are still viable. Those decisions were constrained by the computing environment of the time, which was extremely primitive compared to today's. The Canada Geographic Information System is used to illustrate those decisions and their consequences. Today it is possible to imagine a very different birth of GIS, based on globes rather than maps, and with positional uncertainty and spatial resolution addressed at the outset. Hierarchical data structures for the globe are introduced, and the advantages of congruent geography are discussed. GIS today still reflects in part the constraints of computing in the mid 1960s.

Introduction

The history of geographic information systems (GIS) as a computer application began with multiple projects in the 1960s, of which perhaps the best known and influential was the Canada Geographic Information System (CGIS), developed by an IBM team under the guidance of Roger Tomlinson. Other projects at that time included efforts by Duane Marble at Northwestern University to support transportation research in the Chicago area; and the work of the UK's Experimental Cartography Unit (Coppock and Rhind, 1991; Foresman, 1997). All of these and other contemporary developments took place in what today would be regarded as a highly constrained and simplistic computing environment. Mainframes were the only computers of that time; they were slow and had almost no mass storage, relying instead on sequentially accessed magnetic tape. There were no high-speed networks, data being exchanged by shipping tapes. Input instructions were provided on punched cards, or perhaps communicated using teletype technology at 300 bits per second.

Technology has advanced enormously in the half century since CGIS was designed and built. Some of the design decisions likely remain correct even today, but others have been superceded, and still others may persist in spite of technical advances. Today we can look back on 50 years of development, and reflect on the ways in which GIS was initially conceived and later elaborated. To what extent is GIS a fast-moving, constantly advancing and improving technology, and to what extent is it limited by the legacy of decisions that no longer make sense, or make less sense than they used to? And to what extent has it missed the opportunities provided by technological advances; opportunities that may not have been available or even imagined fifty years ago? These are the questions that provide the motivation for this paper. The evidence assembled here is mostly well known, and the reader unfamiliar with certain aspects is encouraged to examine the cited references. It would be impossible to review the entire, detailed history of GIS to this point. Instead, this paper is intended to stimulate discussion, and to encourage others to reflect on how GIS has evolved, and whether it might have been different.

The next section uses CGIS as an example of technology at the dawn of GIS, to discuss the approach taken in its design at that time. This is followed by a section on major developments in GIS since then, some of which have replaced the earlier ideas. The next, central sections of the paper contrast maps and globes, identify the advantages and disadvantages of each, and introduce the structures known as discrete global grids (DGGs) in some detail. The penultimate section points to the growing importance of data integration as a motivation for GIS, defines and revisits the history of congruent geometries, and argues that DGGs hold the key to efficient, multiscale data integration. This is followed by a concluding section.

CGIS

In the late 1950s the national and provincial governments of Canada embarked on a comprehensive study of the uses and capabilities of Canadian land, termed the Canada Land Inventory, with an eye to more effective utilization, land conservation, and support for natural resources management. Statistics were promised to the participating organizations on the quantities of land of various characteristics: for example, how much land is currently not in agricultural production, not used for some other purpose, and has soil of sufficient quality to support agriculture? While manual production of such statistics from maps would be extremely labor-intensive, the case was made that computerizing the map analysis would be both cost-effective and fast.

Seven themes were identified, ranging from current land use to soil capability for agriculture. The study area was partitioned into tens of thousands of map sheet areas, each described for each of the seven themes by being divided into irregularly shaped areas that could each be assigned to one of a set of pre-defined classes. A single map sheet of the land-use theme might well include several thousand such areas. Each map was scanned using a fine-resolution custom-built optical scanner, and then vectorized and organized into an arc-node structure so that it could be stored and processed efficiently.

Several design decisions can be identified at this point. First, the vector-based representation made no allowance for uncertainty: the lines drawn on the input maps were taken to represent the truth, that is, real, thin discontinuities, and the areas they enclosed were taken to have perfectly uniform attributes. No allowances were made for any subsequent rethinking of these decisions. Second, vector representation allowed for efficient compression of data on magnetic tape; this, and the decision to omit uncertainty, are in line with the mantra "raster is vaster, but vector is correcter." Third, the arc-node data structure, in which the basic element is a section of boundary separating two neighboring areas, proved to be very powerful, leading to efficient algorithms for area calculation, polygon overlay, and the point-in-polygon task. The structure was widely adopted in the early years of GIS, and implemented by Esri in its "coverage" data model. Fourth, maps were organized in square tiles and stored on tape in "Morton" order, named for Guy Morton, one of IBM's CGIS team, and intended to increase the probability that tiles that were adjacent in space were also adjacent on tape. Morton order is now better known as Z order and is very widely used as a geospatial indexing scheme.

Major technological developments since the 1960s

Many advances have occurred since those early days. Hard disk and vastly increased mass storage have almost entirely replaced magnetic tape. Compute speed has increased by many orders of magnitude: the smart phone or laptop of today has far more power than the IBM 360 of the mid 1960s, and massively parallel machines now offer peta-scale computing.

Major developments have also occurred in data modeling. By the late 1970s the arc-node data structure was being modified to handle road networks, by allowing nodes to be 1-valent and by allowing arcs to cross without intersection. Today, time and the third spatial dimension are widely supported, and object-oriented data modeling allows for inheritance and hierarchical relationships. Today's GIS is able to accommodate both raster and vector structures, together with a wide range of indexing schemes.

Yet many aspects of today's GIS are remarkably unchanged. Little progress has been made in incorporating or addressing concepts of uncertainty, despite the very extensive and widely reported research efforts of the past 30 years (Zhang and Goodchild, 2002). Visualizations of geospatial data rarely if ever communicate any form of uncertainty, and numerical results are still often reported with far more precision than is justified by the accuracy of the source data. To be fair, many of the methods that have been proposed to model uncertainty are complex and hard to calibrate; and maps have long offered a cleaned-up version of reality that leaves out the uncertainty that map readers and decision makers would perhaps rather not acknowledge.

Despite the obvious advances, many of the design features of the 1960s continue to exist in today's GIS. The decision to regard the source map as the truth remains with us, as the previous paragraph made clear. GIS software continues to emphasize the map or layer as the primary organizing element, making it easy to assemble all information about one theme but much more difficult to assemble all themes about one location (this point is the subject of the penultimate section of this paper). We continue to project the Earth onto a flat mapped surface, using a number of popular map projections, and forcing us to accommodate and allow for the various distortions that result.

Maps and globes

In the 1960s two distinct visual methods were available for recording, sharing, and storing information about the variation of geographic themes over the Earth's surface. Globes had been known for centuries, but were difficult and expensive to make, store, copy, and communicate; whereas paper maps were easy and cheap to print in vast numbers, and could be shipped in quantity if appropriately packaged. Moreover maps could be made of parts of the Earth's surface, but there had never been much interest in partial globes. Maps can be made at any scale, though the projections that are required to flatten the Earth so that it can be fitted onto a flat sheet of paper inevitably ensure that scale is never exactly constant. Although globes are scaled models of the Earth and can in principle have constant scale, physical size places an effective limit on the scale of a globe (Eartha, a globe at the headquarters of the DeLorme Mapping Corporation in Yarmouth, Maine, has long been claimed to be the world's largest globe; its diameter of 12.5m gives it a scale of roughly 1:1,000,000).

Over the past 50 years we have seen a steady shift to the digital representation of geographic information. Geospatial databases are now commonplace, and some of them are used every day by the general public. Many forms of geographic information, such as remotely sensed images, are now "born digital", never having gone through a stage of physical representation and subsequent digitization. Map digitizers and scanners, which used to be an essential part of GIS hardware, are now a vanishing species.

In this new digital world the earlier arguments against globes are much less compelling. A digital representation of the information present on a globe is easy to copy, store, share, and ship; as easy, in fact, as a digital representation of a map, as the general public has learned through familiarity with services such as Google Earth. Yet the map remains the primary metaphor for GIS. Many Web services present the world in projected form, often using a Web version of the Mercator projection; and the physical act of overlaying transparent map layers depicting a flattened Earth remains a core concept in GIS education.

The various efforts of the 1960s to computerize geographic information had merged by the mid 1970s into the concept of a GIS, largely through the inspiration of Roger Tomlinson and the two conferences he convened, in 1970 and 1972, under the auspices of the International Geographical Union (Tomlinson, 1970, 1972). Those efforts were all map-based, and so was the technology that emerged. There is no evidence that the globe option was considered, in part because of the limitations of the technology of the time, and perhaps in part because the Earth's curvature was of little concern to these early projects with their small geographic extents. In the ensuing years the idea of a digital globe arose occasionally, but it was not until the mid 1990s, thirty years after the coining of the term GIS, that technology had advanced to the point where a globe-based GIS was realistic; and not until the mid 2000s that it was realistic using everyday consumer technology. Tomlinson's choice of the map as the basis for GIS was clearly the right one in 1965, but is it still the right one today, and how might GIS have been different if the globe rather than the map had been dominant from the outset, and GIS had been conceived as a digital container and processor of globes rather than maps?

The technologies of digital globes

Two distinct technologies have emerged to support digital globes, each with its own advantages and disadvantages. The history of each approach is reviewed in the next two subsections.

Warping an image

In the 1960s computers were very limited in their ability to display results. The line printer, with a spatial resolution of typically 1/10 inch by 1/6 inch, was the primary output medium through the 1970s, and indeed CGIS was originally designed with no graphic output, producing only numeric tables. SYMAP, a cartographic package developed by the Harvard Laboratory for Computer Graphics, used the line printer to make very crude maps by overprinting characters to obtain shades of gray (Chrisman, 2006). Pen plotters, which could produce simple line drawings, were introduced in the early 1970s. But by the early 1990s companies such as Silicon Graphics were marketing fine-resolution display systems with extensive support for 3D. Objects such as spheres could be displayed on the screen, and panned and zoomed by the user in real time. Moreover functions were available to take a 2D image and drape it over a 3D object. Here was the realization of a digital globe: a 2D image of the Earth, draped over a sphere or spheroid, and ready for panning and zooming by the user.

Several groups around the world were quick to realize the potential of this development. The German company ART+COM demonstrated a digital globe at SIGGRAPH 95 (Grueneis, 1995) that they termed TerraVision, while the American company SRI demonstrated a similar system at the same conference and by coincidence used the same name (the ART+COM system was subsequently renamed T_Vision). A few years later Vice-President Al Gore's Digital Earth speech (Gore, 1998) prompted the establishment of a Digital Earth office within NASA to coordinate efforts. Keyhole was funded to develop a Digital Earth prototype, EarthViewer, and was acquired in 2004 by Google, who rebranded the software and launched it as Google Earth in 2005. Initially the high cost of the 3D graphics workstation limited access, but by 2001 the video games industry had stimulated the development of sophisticated graphics accelerators for even the most modest PC, so Google Earth was quickly adoped by hundreds of millions of users worldwide.

Two technical problems had to be solved by these systems. First, in order to zoom rapidly from a global view with perhaps 10km resolution to close-in views with resolutions as small as a meter, some kind of hierarchical structure was needed. The quadtree, which had become popular as a hierarchical version of a raster, provided a suitable solution. The top level in the tree consisted of a coarse representation of the entire Earth as a flattened image; the second level consisted of four tiles, each representing one quarter of the Earth at finer resolution, and so on.

Second, the volumes of data required by a digital globe with fine resolution were enormous, and far beyond the storage capacity of any but the largest systems. 1m resolution, for example, requires at least $5x10^{12}$ data elements to cover the globe. The solution adopted by these systems was to leave the data distributed over a number of remote servers, and to request data for specific areas at specific resolutions as and when needed. The databases on the remote servers were structured as quadtrees, and the user's client was programmed to request specific tiles of data at specific levels in the quadtree. Sophisticated algorithms were developed to anticipate the need for tiles, and to delete tiles from local storage when the user's attention had moved elsewhere.

Several significant advances have been made in these systems since their original introduction in the mid 1990s. The use of multithreading allows the process of requesting new tiles from servers to proceed independently of the process of display. Databases now have 3D elements, including the representation of buildings, so the tiling systems have been generalized to the third spatial dimension. And rather than obtain rasterized, rendered tiles from the server, many systems now obtain tiled data in vector form and render the data in the client. But the basic design is still the same, and its disadvantages are sometimes apparent. The tiles are defined on a flattened Earth, so at the highest latitudes they are in reality curved triangles, defined by two meridians, the Pole, and one parallel (it is often easy to see the effect of this by looking at the digital globe from a point directly above one of the Poles). Moreover a simple task, such as the addition of a straight line, must be executed on the flattened Earth, since the Earth image is only warped to the curved surface at display time. So the result is not a Great Circle arc, as one might expect. If the system's basic image has been created using the Plate Carrée projection, the line will be a linear function of latitude and longitude on the projection, and warped accordingly on the digital globe (Figure 1).

[Figure 1 about here]

A discrete global grid (DGG)

Rather than lay a quadtree on a flattened Earth, DGGs are hierarchical structures on the curved surface of the Earth itself, thus avoiding all of the distortions inherent in map projections. Unfortunately only five ways exist to create a 3D solid using pieces of equal size and shape. These are the five Platonic solids, and have been known since classical times: the four triangles of a tetrahedron, the six squares of a cube, the eight triangles of an octahedron, the 12 pentagons of a dodecahedron, and the 20 triangles of an icosahedron. Instead, DGGs begin with one of the Platonic solids or a simple modification, and then use a hierarchical scheme to subdivide each of its faces. Since triangles are the favored display element of 3D graphics systems, it is desirable that the basic elements at any level in the hierarchy be triangles. Triangles are also convenient since it is desirable that basic elements at each level nest within the corresponding element at the next higher level, and it is easy to create four nesting triangles from a larger triangle by connecting the midpoints of its edges. On the other hand the pentagons of the dodecahedron do not lead to hierarchies with simple and desirable properties.

For example, Dutton (1984, 1989) has proposed a DGG based on the octahedron, first positioning two of the vertices at the Poles and spacing the other four uniformly around the Equator starting at the Prime Meridian. He termed the scheme the Quaternary Triangular Mesh (QTM). Triangle edges can be defined as Great Circles at every level, but it is computationally easier to define them as a mix of parallels and linear functions of latitude and longitude (Goodchild and Yang, 1992). Several schemes have been proposed and implemented based on the icosahedron and truncated icosahedron (Sahr, White, and Kimerling, 2003), and on the cube augmented by four triangles in each polar region (see www.geofusion.com). Criteria have been proposed for evaluating and comparing alternative schemes (Kimerling *et al.*, 1999).

Since no scheme can result in elements that are exactly uniform in size and shape (with the exception of the Platonic solids themselves, and thus only at the top level of the hierarchical structure), much research has gone into minimizing variation. Sahr, White, and Kimerling (2003) have developed a scheme based on the icosahedron that produces elements that are exactly equal in area at a given level, but vary of course in shape and topology. In the QTM, variation of area at a given level is approximately 10% (Goodchild and Yang, 1992), and shape varies from right triangles in the octahedron corners to equilateral in the center of each triangle face.

Any DGG requires an indexing system that uniquely identifies each basic element at each level. In QTM the central triangle can be labeled 0, the apex triangle 1, and the remaining two triangles 2 and 3. At the top level the eight faces of the octahedron are labeled 0 through 7. This creates an index of length $3+2(n-1)$ bits, where $n = 0$ is the entire Earth and $n=1$ is the octahedron level. For example, at $n=9$ the triangles have roughly 1000 sq km resolution.

Given an indexing system, algorithms can be defined for standard functions. Perhaps the most basic is the transformation to and from a DGG index and a position in latitude/longitude (or a standard coordinate system such as UTM). Others include finding the basic elements that are adjacent to a given element, and determining the global distance between the centroids of two basic elements (Goodchild and Yang, 1992).

Several efforts have been made to generalize the concept of a DGG to three dimensions. Recent versions of Google Earth use an Earth-centered, radial octree or "rocktree" (Rohlf and Hancher, 2012) to define and store tiles. Yu and Wu (2009) have proposed a "spheroid degenerated-octree grid" while Wan *et al.* (2013) have proposed a "sphere shell space 3D grid".

Advantages of a DGG

A DGG has several significant advantages over the traditional GIS based on a projected Earth. First, there is no hiatus or break at the Poles or at 180 degrees longitude, as there is for cylindrical projections in Equatorial aspect, such as the Mercator or Plate Carrée. Similar comments apply to conic and azimuthal projections, and in general to all ways of projecting the Earth's surface onto a plane. There is no "top" or "bottom" to the Earth, and no need for Australians to feel isolated "down under". All of these are of course advantages of globes over maps, but the mounting of a physical globe usually serves to limit the user's freedom of perspective (Eartha, for example, is suspended from the North Pole and able to rotate only about the Earth's axis).

Second, the basic elements of a DGG at any level are approximately uniform in size and shape. This property stands in sharp contrast to a raster laid on a flattened Earth. On the widely used Plate Carrée projection, for example, raster cells which are approximately square and 10km on all sides at the Equator become approximately rectangular at latitude 60, measuring 10km north-south but only 5km east-west on the Earth surface, while the rows of cells at the poles are still 10km north-south but only 16m across at the base and

0m across at the apex on the Earth's surface, having lost all but 0.08 sq km of the 100 sq km area that each cell occupied at the Equator.

Third, the spatial resolution of geographic information structured as a DGG is always explicit. This is in contrast to rasters, which inherit the distortions of the projection that was used to flatten the Earth so that a raster could be laid on it. As the previous paragraph demonstrated, the variation in spatial resolution of a raster laid on a Plate Carrée projection is of roughly three orders of magnitude in the east-west direction. QTM's variation in spatial resolution is approximately 10% within most levels of the hierarchy, and zero at Level 1; and the previously cited Sahr, White, and Kimerling (2003) DGG achieves zero variation in area at all levels.

Fourth, the use of a hierarchy based on powers of two lends itself well to the representation of data sets with varying spatial resolution. The quadtree has often been presented as a clever way of compressing geographic information, by using larger basic elements to represent areas of approximate uniformity, and smaller areas where there is greater variation (Samet, 1984). This theme of data integration is addressed in greater detail in the next section.

Fifth, the uniform resolution of a DGG could lend itself well to simulation of Earth-surface processes, and the solution of partial differential equations. Unlike finite-difference methods based on rasters, there are no boundary effects or variation in spatial resolution to deal with. Processes such as those governing the atmosphere can be addressed as easily in the Polar regions as in low latitudes.

If digital globes had been feasible in the 1960s, how might the history of GIS have been different? Would we still need to maintain extensive software libraries to deal with the many ways of flattening the Earth's surface, or would map projections have been rendered obsolete? A typical GIS professional spends many hours mastering the common map projections, but a child of 10 can manipulate Google Earth with no training. Of course the screen of a digital display is still two dimensional, and in principle a perspective orthographic projection must be used to create a screen image of a digital globe. But because the user can manipulate the globe it is perceived as a 3D object. In a related fashion the human eye also receives 2D signals on its retina, and it is the human brain that reassembles and perceives a 3D object.

There is, however, one compelling residual argument for projection, and that is when it is desirable to see the entire Earth at once, since it is never possible to see all of a globe at one time. But this is a matter of visualization rather than representation. It would be a straightforward matter to take the contents of a DGG and render it at display time using a standard projection such as the Plate Carrée. This approach would surely be preferable to projecting the Earth prior to storage, and inheriting all of the distortions of the projection in every application, whether it involve simple display or more sophisticated analysis or modeling.

A related argument can be made about the representation of linear features in a DGG. Like rasters, a DGG represents a linear feature as a collection of basic elements. Although the size of the basic elements may represent uncertainty about the line's position, cartographic sensibilities are offended by the resulting "jaggies". But here, again, would it not be better to use a DGG, with its explicit treatment of positional uncertainty and spatial resolution, as the basis for representation, and to treat the jaggie problem as a cosmetic matter of display? The basic DGG elements of the linear feature could easily be replaced by a smooth single- or double-line vector representation at display time.

Congruent geographies

Because CGIS was organized by layer using an arc-node structure it was straightforward to compute the areas of each polygon, and to obtain totals for each attribute class. Thus for a digital map of land use, for example, the areas of each land-use patch could be computed using a simple algorithm and one pass of the data; and totals could be obtained by adding areas for each class. But CGIS had also promised statistics that could only be obtained by overlaying layers. The example used earlier, "how much land is currently not in agricultural production, not used for some other purpose, and has soil of sufficient quality to support agriculture?", would require the overlaying of two layers: land use and soil capability for agriculture. In the early days of GIS vector overlay proved to be challenging, and was not solved with robust algorithms until the late 1970s. The task was conceptually simple but computationally complex; and because the same real lines tended to appear on many map themes (coastline, for example, would appear on all map themes), and because the different versions of the same line would always deviate due to uncertainties in map drafting, digitizing, and vectorization, early algorithms tended to be overwhelmed by vast numbers of small slivers (Longley *et al.*, 2015, p. 302).

A simple solution was to rasterize the layers, using a common raster geometry. Thus several raster-based GISs emerged in the 1970s, all based on the notion of a data cube: the same raster cells forming the base, and different themes or layers forming the columns. The results of overlay were not as precise as with vector data, since they were estimates subject to the size of the raster cell. But they could be at least as accurate if the raster cell size was no greater than the positional uncertainty of the vector lines. Tomlin (1990) developed what he termed cartographic modeling, a comprehensive set of functions and an associated language, based on the requirement that layers be represented on the same raster geometry.

These raster systems, which filled an important gap between the dawn of GIS and the emergence of robust algorithms based on vector geometry, are examples of a more general class which we might term *congruent geography*: the use of the same basic spatial units, either regular or irregular in shape, to represent every layer or theme. Examples include the Multi-Purpose Cadaster (National Research Council, 1980, 1983), where the basic units are defined by land ownership; Integrated Terrain Units, created by reducing the landscape to pieces with uniform characteristics on all mapped themes; and the Common Land Unit (https://www.fsa.usda.gov/programs-and-services/aerialphotography/imagery-products/common-land-unit-clu/index) of the US Department of Agriculture, a minimum area of land that is homogeneous with respect to ownership and management.

The great advantage of congruent geographies is that they allow the geographic data cube to be examined, analyzed, and visualized either horizontally or vertically. A horizontal slice exposes the spatial variation of one theme for all locations, as does a map or layer. A vertical profile exposes all of the available themes for one location. Several recent projects have resampled all Landsat data to create a simple cube, with time in the third dimension, allowing the user to view a single pixel or collection of pixels through the entire history of Landsat (*e.g.*, http://www.datacube.org.au/about). Yet a sophisticated vector-based GIS today, such as one operated by a city government, would find it enormously difficult and time-consuming to answer a simple query such as "tell me everything you know about my property".

A DGG as a congruent geography

The arguments for DGGs thus far have been concerned with how they address the curved nature of the Earth's surface, avoiding the distortions caused by flattening the Earth. But the argument for DGGs as a congruent geography may ultimately be more powerful. DGGs are hierarchical, with small basic elements that might be used to represent fine-resolution data, nesting within coarser basic elements. A hierarchical indexing scheme such as that described earlier for QTM makes it very easy to identify basic elements at coarser or finer resolution, by either truncating or extending the index. In effect, then, a DGG creates a multiscale congruent geography. It is thus a simple solution to integrating multiple layers of data, and accommodating different levels of spatial resolution or different positional uncertainties (Dutton, 1989).

Data integration or synthesis is presenting a major challenge as we enter the world of Big Data; more challenging, in some ways, than our traditional emphasis on analysis. The literature on data fusion or conflation is small (Li and Goodchild, 2011), and few available GIS provide implementations in the form of appropriate software.

Conclusion

This paper began by posing three questions: to what extent is GIS a fast-moving, constantly advancing and improving technology; to what extent is it limited by the legacy of decisions that make less sense than they used to; and to what extent has it missed significant opportunities provided by technological advances?

The reimagining that forms the central theme of the paper has many dimensions, but at its broadest it imagines a GIS as a computer system for handling digital globes rather than digital maps. The technical ability to create and visualize 3D renderings of the Earth was unavailable in the mid 1960s at the birth of GIS, but it was achieved in the early 1990s, and led directly to Google Earth and its many competitors. Yet the more sophisticated analysis and modeling functions of GIS remain map-based, with all of the distortions and interruptions that flattening the Earth implies. There has not been a wholesale shift to the use of geodetic distance, or of globe-based methods for calculating area. Instead, students continue to devote significant time to learning about map projections, and to becoming sensitive to, and hopefully dealing with the distortions that result. Users of Web mapping services have become used to seeing Greenland and Russia in their vastly exaggerated representations. Passengers on Emirates flights between Dubai and the west coast of North America may be mystified as to why after about eight hours flying the cabin screens suddenly show the aircraft executing an instantaneous left turn, followed by an extremely rapid transit across the screen, and second instantaneous left turn (though alert geospatial professionals may recognize this as the result of using a Plate Carrée projection when flying in a straight line over the North Pole).

Versions of GIS that exploit congruent geographies have existed for decades. The finite size of the basic spatial units, whether raster cells or polygons, has always suggested that these versions of GIS are not as accurate as the dominant vector-based systems. But no set of geospatial data can ever have infinitely fine spatial resolution, or be completely free of positional uncertainty. Thus a GIS based on a congruent geography whose basic spatial units match the known spatial resolution or positional uncertainty of the data must always be as accurate as a vector-based system. A hierarchical system, such as a DGG, has the advantage that the spatial resolution of each theme can be adjusted to the theme's known properties.

Two counter-arguments have been raised: the difficulty of visualizing the entire Earth in a globe-based system, and the unpleasant "jaggies" that result when visualizing raster or DGG data. But these arguments can be addressed at display time, and should not be allowed to determine the method used to represent geographic information in a digital store. Insisting on vector representation because it leads to attractive visualizations ignores the consequences for analysis and modeling: the inability to deal in any way with positional uncertainty or limited spatial resolution, and the spurious slivers that result whenever lines are overlaid that are different in the representation but identical in reality.

It does indeed seem that GIS practice today is to some degree limited by the legacy of the past; and that interesting and viable alternatives exist to some aspects of GIS practice. Instead of a single, authoritative source of data that we could treat as the truth, we would be motivated to find ways of accommodating the many sources of the same geographic facts that are now available, and integrating them in ways that reflect their various properties and uncertainties. If we were to wipe the slate clean and reinvent GIS today, with today's computing power and visualization capabilities, the result might be very different.

References

Chrisman, N.R., 2006. *Charting the Unknown. How Computer Mapping at Harvard Became GIS*. Redlands, CA: Esri Press.

Coppock, J.T. and D.W. Rhind, 1991. The history of GIS. In D.J. Maguire, M.F. Goodchild, and D.W. Rhind, editors, *Geographical Information Systems: Principles and Applications*, Vol. 2, pp. 21-43.

Dutton, G., 1984. Geodesic modeling of planetary relief. *Cartographica* 21(2,3): 188-207.

Dutton, G., 1989. Modeling locational uncertainty via hierarchical tesselation. In M.F. Goodchild and S. Gopal, editors, *Accuracy of Spatial Databases*. New York: Taylor and Francis, pp. 125-140.

Foresman, T.W., editor, 1997. *The History of Geographic Information Systems: Perspectives from the Pioneers*. Upper Saddle River, NJ: Prentice Hall PTR.

Goodchild, M.F. and S. Yang, 1992. A hierarchical spatial data structure for global geographic information systems. *CVGIP: Graphical Models and Image Processing* 54(1): 31-44.

Gore, A., 1998. *The Digital Earth: Understanding Our Planet in the 21st Century*. portal.opengeospatial.org/files/?artifact_id=6210

Grueneis, G., 1995. T_Vision. In SIGGRAPH 1995 Visual Proceedings. ACM SIGGRAPH, New York, p. 134.

Kimerling, A.J., K. Sahr, D. White, and L. Song, 1999. Comparing geometrical properties of global grids. *Cartography and Geographic Information Systems* 26(4): 271-288.

L. Li and M.F. Goodchild, 2011. An optimisation model for linear feature matching in geographical data conflation. *International Journal of Image and Data Fusion* 2(4): 309-328.

Longley, P.A., M.F. Goodchild, D.J. Maguire, and D.W. Rhind, 2015. *Geographic Information Science and Systems*. Hoboken, NJ: Wiley.

National Research Council, 1980. *Need for a Multipurpose Cadastre*. Washington, DC: National Academies Press.

National Research Council, 1983. *Procedures and Standards for a Multipurpose Cadastre*. Washington, DC: National Academies Press.

Rohlf, J.H. and M.D. Hancher, 2012. US Patent 8650220: System and methods for storing and retrieving geospatial data. Washington, DC: US Patent and Trademarks Office.

Sahr, K., D. White, and A.J. Kimerling, 2003. Geodesic discrete global grid systems. *Cartography and Geographic Information Science* 30(2): 121-134.

Samet, H., 1984. The quadtree and related hierarchical data structures. *ACM Computing Surveys* 16(2): 187-260.

Tomlin, C.D., 1990. *Geographic Information Systems and Cartographic Modeling*. Englewood Cliffs, NJ: Prentice Hall.

Tomlinson, R.F., 1971. *Environmental Information Systems: Proceedings of the UNESCO/IGU First Symposium on Geographical Information Systems, Ottawa, September 1970*. Ottawa: International Geographical Union, Commission on Geographical Data Sensing and Processing.

Tomlinson, R.F., 1972. *Geographical Data Handling*. Ottawa: International Geographical Union, Commission on Geographical Data Sensing and Processing.

Wan, G., X.-F. Cao, F. Li, and K. Li, 2013. Sphere space shell 3D grid. In Proceedings, ISPRS WebMGS 2013 & DMGIS 2013, 11-12 November 2013, Xuzhou, Jiangsu, China.

Yu, J.-Q. and L.X. Wu, 2009. Spatial subdivision and coding of a global three-dimensional grid: Spheroid Degenerated-Octree Grid. In IGARSS 2009 (IEEE International Geoscience and Remote Sensing Symposium).

Zhang, J.-X. and M.F. Goodchild, 2002. *Uncertainty in Geographical Information*. New York: Taylor and Francis.