ISEA Discrete Global Grids

By Dan Carr, Ralph Kahn, Kevin Sahr, and Tony Olsen

1. Introduction

This article describes a recently proposed standard, ISEA discrete global grids, for gridding information on the surface of the earth. The acronym ISEA stands for icosahedral Snyder equal area. The grid cells not only have equal areas, they are hexagons when projected onto an icosahedron! Being an advocate of hexagon binning, and corresponding graphics, my (Dan) enthusiasm is such that I want to call attention to this new approach.

Jon Kimerling, a geosciences professor, won the Oregon State University (OSU) Milton Harris Award for his research on the topic. Since I was peripherally involved in the development of the grids, I asked Kevin, Ralph and Tony to help me with portions of this article. They all interacted with Jon in the research and we all shared the desire to promote ISEA grids. Kevin, an admirer of Buckminster Fuller, developed many of the ISEA algorithms and graphics. Along with OSU collaborators Mathew Gregory and Larry Hughes, he developed a web site, `http://bufo.geo.orst.edu/tc/firma/gg/`, that contains foldable figures, descriptions and three reference lists. At the Milton Harris Seminar, Ralph provided an excellent overview on the relevance to summarization and presentation of data from the Earth Observing System (EOS). EOS is a series of NASA satellites designed to detect and monitor global climate change, starting in summer 1998. Material from Ralph's talk is available at the above web site and we include portions here. Tony was the instigator of the whole development with his push to develop a globally consistent environmental sampling methodology. Tony and his EMAP co-workers helped EPA Regions, states and nations develop environmental sampling plans using an EMAP grid that was developed in the early stages of the research.

The structure of this article is as follows. Section 2 traces some of the history behind the ISEA grids. Sections 3 describes the ISEA grid. Section 4 introduces a potential application, storage of summaries derived from Earth Observing System sensors. Section 5 discusses graphics for hexagon grids, Splus algorithms for low resolution ISEA grid smoothing on the globe, and new, resolution 9 (about 200,000 cells) binned map of global elevation data. Section 6 closes with challenges for future research.

2. The Recent Historical Development of ISEA Grids

The impetus for the global grid system came from what many would call an unusual perspective - survey sampling. In 1989, Denis White and Scott Overton, a geographer and a statistician from Oregon State University, held a workshop in Corvallis, Oregon, to discuss the geographic requirements for a general survey design. The survey design would be the foundation for all surveys conducted as part of the Environmental Monitoring and Assessment Program (EMAP) (Messer et al., 1991; Stevens, 1994). Scott Overton, leader of the survey design effort, recommended basing the design on a systematic grid with a random start (Overton et al., 1990). We all know how to accomplish that for planar surfaces but when the design must cover the United States, we are faced with a non-planar surface - the earth. Jon Kimerling, along with the other geographers at the workshop, devised a discrete grid system that satisfied the needs of EMAP at that time (White et al., 1992). The system used a truncated icosahedron model of the earth with a triangular point grid applied to the large hexagon plates. It worked for the contiguous 48 states. However, the initial discrete grid system did not solve all the underlying issues and the embedded triangular grid structure had elements that were arbitrary. As an example, the EMAP team also applied the grid to China, Russia, and Indonesia. The team knew problems would exist for China and Russia, as a single large hexagon plate would not cover either country. Although small in area, Indonesia is stretched out and also is not covered! The initial discrete grid system had problems at the boundaries of the plates.

In 1993, Tony Olsen, faced with these inadequacies, initiated a research effort with Jon Kimerling, Kevin Sahr, and Denis White in the OSU Geosciences department to investigate an alternative discrete global grid system. Tony required the system to be truly global and result in an equal area tessellation. He also had a preference for compact areas, minimal shape distortion, a triangular point grid, and a hierarchical grid structure allowing multiple grid densities. These characteristics would enable global implementation of survey designs for continuous spatial populations (Stevens, 1997).

My (Dan) formal involvement did not start until Oregon State researchers held a workshop on discrete global grids at Santa Barbara in 1994. Others in attendance were Denis White, Jon Kimerling, Michael Goodchild, Waldo Tobler, Tony Olsen, Geoff Dutton, Frank Davis, and David Mark. Many in the group had already developed their own approaches for global grids. Waldo Tobler was already using his methodology to show populations on the global. Geoff Dutton had developed a gridding system that modeled the earth as an octahedron with an appropriate map projection. Kimerling and White presented their icosahedral alternative to the EMAP (truncated) icosahedron model. There are of course additional approaches that work more directly on the globe. (For a recent discussion of distributing points on a sphere, see Saff and Kuijlaars 1997). All methods must deal with the fact that there is no perfect regular partition for the surface of a sphere. One member noted that there is always at least one singularity, as he humorously pointed to the bald spot on his head. Michael Goodchild suggested that the meeting produce a list of desirable properties for gridding systems. The list appears below. Tony knew my objective when I proposed cells being "compact" (having a small dimensionless second central moment - see Conway and Sloane 1988). It was my attempt to promote hexagon cells.

At the Santa Barbara workshop Michael Goodchild proposed a prioritized attribute list for a discrete global grid system. The elements of the list are: the domain is the globe (sphere, spheroid), areas exhaustively cover the domain, areas are equal in size, areas are compact, areas are equal in shape, areas have same number of edges, edges of areas are of equal length, edges of areas are straight on some projection, areas form a hierarchy preserving some properties for $m < n$ areas, each area is associated with only one point, points are maximally central within areas, points are equidistant, points form a hierarchy preserving some properties for $m < n$ points, addresses of points and areas are regular and reflect other properties.

With methods and evaluation criteria at hand, the group planned two sessions at the GIS/LIS94 meeting. To have something to contribute, I, on the spur of the moment, concocted a method based on projecting 3-D lattice points "near" a sphere surface onto the surface. My subsequent attempts with different lattices, packings, and notions of near, did not lead to hexagon patterns over the whole sphere. The redeeming features of my talk at GIS/LIS94 turned out to be the color anaglyph stereo viewgraphs and brevity. The other presentations carried the two sessions.

After GIS/LIS94, work proceeded on the icosahedron model (see Kimerling, Sahr, Song, White, and Iltis, 1995). I called the research to the attention of Ralph Kahn (NASA-JPL) who was looking for better ways to summarize the global data expected from EOS. Tony sought to involve Noel Cressie for dealing with spatial estimation issues.

Jon Kimerling subsequently won the Milton Harris Award. In May 1997, he held the Milton Harris Award Symposium on Global Grids: New Approaches to Global Data Analysis. In addition to presentations by team members Kevin Sahr and Denis White, he invited Ralph Kahn (NASA -JPL), Noel Cressie (Iowa State University), Ross Kiester (USDA-Forest Science Laboratory), Tony Olsen (USEPA-Corvallis) and myself to make presentations. The following sections cover selected topics from the Symposium and Kevin's web site.

3. Icosahedral Snyder Equal Area (ISEA) Grids

The S in ISEA refers to John P. Snyder. He came out of retirement specifically to address projection problems with the original EMAP grid (see Snyder, 1992). He developed the equal area projection that underlies the gridding system. His work at the U.S. Geological Survey on map projections is known by all who spend any time with map projections. John Snyder died this year. By all reports, he was a modest man who would not seek to have procedures named after him. Nonetheless, in honor of his contributions to the field of map projections, those developing the gridding system have desired to use his name.

ISEA grids are simple in concept. Begin with a Snyder Equal Area projection to a regular icosahedron (see the stereo pairs in Figure 1) inscribed in a sphere. In each of the 20 equilateral triangle faces of the icosahedron inscribe a hexagon by dividing each triangle edge into thirds (see the large gray hexagon in Figure 2). Then project the hexagon back onto the sphere using the Inverse Snyder Icosahedral equal area projection. This yields a coarse-resolution equal area grid called the resolution 1 grid. It consists of 20 hexagons on the surface of the sphere and 12 pentagons centered on the 12 vertices of the icosahedron.

To form higher resolution grids, tessellate each equilateral triangle in the planar view with more hexagons and use the inverse projection back to the sphere. The details of the regular tessellation are as follows: Always center a hexagon about the center point of the

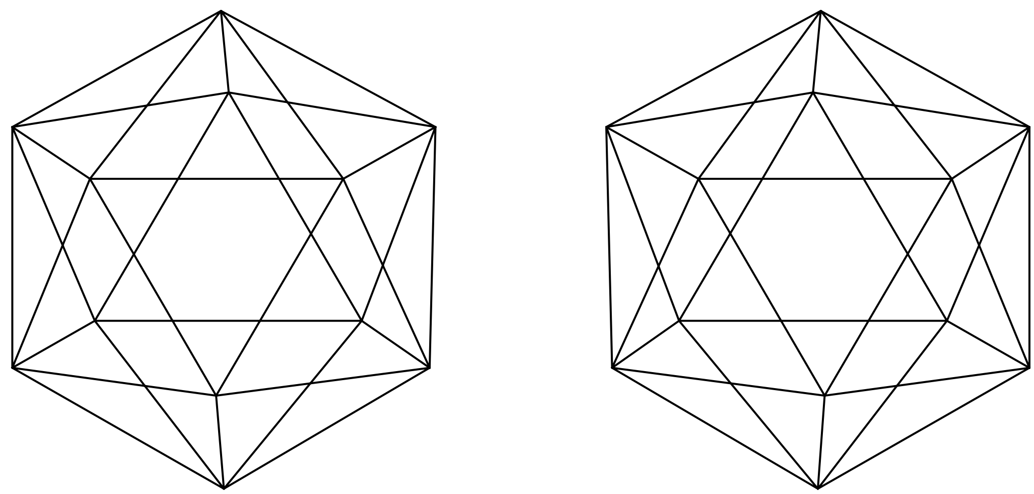

Figure 1. Stereo pairs of a regular icosahedron.

equilateral triangle. For odd resolution grids, orient the hexagon so its base is parallel to the base of the triangle. For even resolution grids orient the hexagon so a vertex points at the base of the triangle. (Figure 2 shows the central hexagons for resolutions 1 and 2 in gray and black, respectively.) Select the edge length of a resolution $r+1$ hexagon so it is $1/\sqrt{3}$ times the edge length of a resolution r hexagon. Thus, the area of a hexagon reduces by a factor of 3 with each increase in resolution. As the resolution increases by 1, the tessellation procedure produces a hexagon centered on each hexagon vertex and center point of the lower resolution tessellation.

As illustrated in Figure 2, the procedure partitions a lower resolution hexagon cell into one central cell and six fractional (1/3) cells. This is not as simple as partitioning a large square into exactly four smaller squares. While the merits of strictly nesting cells within cells depend on the context, one clear merit is aggregation simplicity. The ISEA fractional cells create aggregation and disaggregation problems that are currently under investigation.

The orientation of the icosahedron relative to the globe is an important consideration. The selected orientation for the ISEA grid creates symmetry about the equator. This is desirable for numerical modeling purposes. There are always 12 pentagon cells about the vertices of the icosahedron. The selected orientation places 11 of the pentagon cells over water areas, so that most land mass views will be completely composed of hexagons.

Table 1 on the next page (taken from Kevin's web site) gives the number of cells and characteristic hexagon edge lengths for ISEA grids of increasing resolution.

The advantages of the ISEA grids are (1) they have irregularities (12 pentagon cells) that are minor nuisances rather than being pathological singularities, (2) they are suitable for modeling on all parts of the globe including the poles, (3) they preserve symmetry about the equator, (4) they provide an infinite nesting of equal-area sub-grids, and (5) they provide a basis for uniform global density of sampling for data at all spatial resolutions. The grid facilitates comparisons between high and low latitude data and high and low spatial-resolution data. The grid also improves the isotropy of finite-difference quantities compared to those calculated for rectangular grid schemes. For example Fisch, Hasslacher and Pomeau (1986) note that two-dimensional Navier-Stokes implementations are optimal with hexagons. Finally, no ambiguity exists about nearest neighbors as all nearest neighbor cells share an edge with a reference cell and their distances to the center of a reference cell are nearly equal.

4. EOS and the Potential Application of ISEA Grids

There are many potential applications of ISEA grids. We are particularly mindful of NASA's Earth Observing System and the wealth of global earth science data that it will collect. The EOS AM-1 Platform is scheduled for launch in June 1998. The summarization of this data provides a rapidly approaching opportunity to use ISEA grids.

Resolution	Number of Cells	Length Scale (km)
1	32	4,684.2571
2	92	2,694.2932
3	272	1,553.6212
4	812	896.6139
5	2,432	517.5892
6	7,292	298.8166
7	21,872	172.5192
8	65,612	99.6035
9	196,832	57.5060
10	590,492	33.2011
11	1,771,472	19.1687
12	5,314,412	11.0670
13	15,943,232	6.3896
14	47,829,692	3.6890
15	143,489,072	2.1299
16	430,467,212	1.2297
17	1,291,401,632	0.7100
18	3,874,204,892	0.4099

Table 1. The number of cells and the characteristic hexagon edge lengths for ISEA grids of increasing resolution.

More specifically, ISEA grids are relevant to Level 3 Products in the EOS Data Product Classification. Level 1 Products involve raw radiances with geometric and radiometric calibration. Level 2 Products are geophysical parameters at the highest resolution available. These data sets preserve the non-uniform spatial and temporal sampling of the satellite instruments. Level 3 products are globally and temporally uniform data sets. Level 3 products are needed where large-scale, uniform coverage is required (e.g., global-scale budgets, and problems that depend on data sets from multiple sources). Various tradeoffs will drive the selection of spatial and temporal scales chosen for Level 3 standard products so a multiple-resolution equal-area global grid system is immediately relevant.

The massive amount of data and the resolution issues drive the need for professional algorithms. For example, one instrument on the platform (MISR) will help characterize, on a global basis, atmospheric aerosol type and optical depth, surface bi-directional reflectance properties, and cloud properties. The amount of data to be collected from this one sensor is enormous. With a spatial resolution of 16 values per km^2 and 36 channels, a global description will involve 2.9×1011 basic measurements. The MISR collection rates will be 40 Gbytes/day of raw data, 300 Gbytes/day total data, and

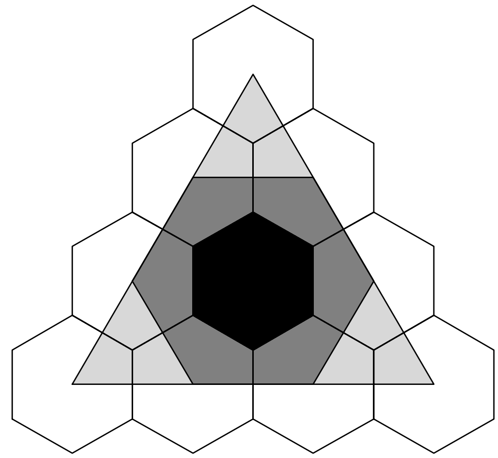

Figure 2. Subdividing the faces of a regular icosahedron: Gray and black regions represent the central hexagons for resolutions 1 and 2, respectively.

15-100 Tbytes/yr for at least 5 years. The computing tools developed for the graphics in this article will not handle such data.

Of course there is the old alternative to handle Level 3 gridding, equal angle grids. The equal angle grid relies on the global latitude-longitude system and uses a cylindrical map projection. It typically has a spatial resolution of 1° (112 km) and sub-grids based on equal-angle divisions at 0.5° (56 km) and 0.25° (28 km). The advantages of the equal angle grid are that the latitude-longitude system is convenient, familiar, and entrenched. Also very important is the fact that the results are easy to represent in 2-D arrays. However, the equal angle approach leads to several issues such as rapidly changing spatial resolution at high latitudes, non-uniform resolution for fine scales, ambiguity of nearest neighbor operations and problems in representing data at multiple scales. The current solutions to the multiple scale problem are discipline-specific variations, for example, specialized grids for polar and for local high-resolution applications. The ISEA approach, among other things, would provide compatible grids across disciplines.

Those seeking additional information on alternative grids and EOS sensors can access Ralph's descriptions at `http://bufo.goe.orst.edu/tc/firma/gg/kahntoc.html`. Of particular interest is an example that shows the huge discrepancies that can result from changing from one grid to another and back. Those seeking more information on Level 2 products or discussion of problems in validating satellite derived parameters can start with Kahn et al. (1991).

5. Graphics for Hexagon Cells, Global Binning and Foldable Figures

Many graphics are available for hexagon cells. Some of these graphics involve spatial smoothing. The figure on page 35 (adapted from Yang and Carr, 1995) shows a breeding bird diversity map based on smoothing to the previous EMAP grid. The brute force smoothing of ten year prevalence data for 615 bird species to the 13000 cell grid involved close to 8 million local logistic regressions! Soon after an article on mortality map smoothing (Carr and Pickle, 1993), Andrew Carr and I created a point and click Splus function (UNIX only) for selecting U.S. cancer mortality rates and smoothing the rates to hexagon grids. The smoother in that context was `loess`. This collection of functions is available as an Splus `data.dump` file, `nchs.dmp,` by anonymous ftp to `galaxy.gmu.edu.` It is located in `pub/dcarr/newsletter/nchs.` While hexagon cell maps are relatively uncommon, the general notion of choropleth maps is, of course, not new.

There are several sources for innovative hexagon graphics. Carr et al. (1987), and Carr (1991) present various density representations and a practical bivariate generalization of box plots. Kevin and Ron Keister (personal communication) have developed an approach for showing the change from cell to cell by coloring triangles within the hexagons. Papers of Carr (1989), Carr (1991), and Carr, Olsen and White (1992) address symbol congestion control with the first showing a stereo regression diagnostic and the last two papers focusing attention on maps. The idea is to partition the map (or plot) using hexagon cells and provide symbols to represent the summary for each cell with data. For example, the angle of a ray glyph can represent a continuous variable, such as a trend estimated from a time series. The rays can point down (below horizontal) for small values or negative trends and up for large values or positive trends. The rays can plot on top of confidence arcs that represent associated confidence bounds. Two rays with common origin, one pointing to the left and one to the right can easily represent two continuous variables on a map.

Splus derivatives of my 2-D lattice functions now facilitate hexagon binning, gray level erosion, smoothing, hexagon plotting and ray plotting. Familiarity and convenience suggested following the conventions in this software when developing binning, smoothing and display procedures for global grids. The result is a set of closely related Splus functions for low resolution grids.

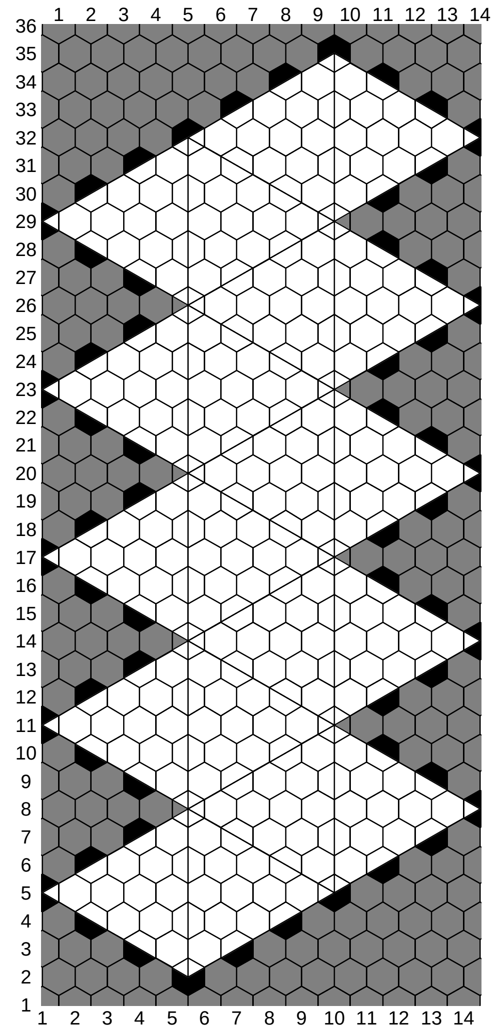

Figure 3. A flattened icosahedron foldable figure for resolution three.

The task addressed here is binning of global ETOPO 5 minute elevation data into an ISEA resolution 9 grid. Conceptually the computation of a cell id for a latitude longitude pair involves four steps. First, the Snyder Equal Area projection produces coordinates in one of the 20 triangles of the icosahedron. Second, an affine transformation (one for each of the twenty triangles) maps the coordinates into a flattened icosahedron foldable figure as shown for resolution three in Figure 3. Next, a hexagon index routine (like `xy2cell` in Splus) produces a planar cell id. The last step uses a look-up table (a vector) to convert planar cell id into a globe cell id. (Globe cell ids are integers ranging from 1 to the number of globe cells.) Given globe cell ids, computation of summary statistics for data falling in the cells is straightforward. The work was in generating the planar cell id to globe cell id conversion vector. The re-indexing omits unused hexagons that cover the Figure 3 rectangle. The re-indexing also accounts for split foldable figure planar cells (for example those containing the five left triangles tips) that are really parts of the same cell on the globe.

Rather than reading the large ETOPO file into Splus, I modified one of Kevin's programs. The program imports the resolution 9 re-indexing vector generated in Splus and the bins on the fly. After reading the binned results into Splus, I used three additional bookkeeping vectors to compute hexagon boundaries and colors for all cells (196832 globe cells and the 832 cloned cells) in the foldable figure. The figure on page 36 shows the average elevation for each cell. Jon Kimerling suggested the basic elevation and depth coloring scheme. A further refinement requiring additional data would be to distinguish land hexagons that are slightly below sea level from ocean floor hexagons. I had problems producing the whole postscript file for the figure on page 36 so I wrote out pieces and connected them using Unix tools. A procedure that reads a value for a location and writes a hexagon directly to a file would be better for graphics output.

My Splus routines for odd resolution ISEA grids are available via anonymous ftp to `galaxy.gmu.edu`. Change directory to `pub/dcarr/newsletter/isea`. There is a README document describing the various functions. For example, one function produces a globe cell near neighbor pointer matrix (for low resolutions). Another function uses this matrix for smoothing values on the globe. (More sophisticated smoothers could restrict domains to land masses, oceans land-ocean boundaries, or address flow constraints.) A script file shows the process of producing a foldable icosahedron. The script starts by randomly generating a vector of 2432 values that implicitly correspond to globe cells in a resolution 5 grid. After smoothing the values and converting them into colors for hexagons, the script plots the hexagons along with tabs for gluing. Creasing along the lines shown in Figure 3 helps in the construction. I have made several figures for the holiday season. Postscript files for different examples and sizes are in the above directory. Kevin's web site contains more examples including one of my favorites. The favorite is an amazing gift from the antiquity of basketry, a six great circle weave.

For stereo presentations on a globe, a simple approach partitions each hexagon into six triangles. The plotting step then renders triangles whose vertices result from the inverse Snyder equal area projection.

6. Additional Challenges and Closing Remarks

This article describes a 1-D indexing system that is viable for modest odd resolution grids. The basic indexing is for hexagon cells that cover a rectangle bounding the planar icosahedron view. A re-indexing vector, whose length is the number of covering cells, removes the unused and redundant indices. After binning with the new indices, pre-computed x and y vectors provide plotting positions for the planar icosahedron view. The binned results correspond to cells on the globe so a short subscript vector extracts values corresponding to split cells in the planar view. The result of concatenating the two vectors corresponds to the planar x and y coordinates. No doubt a similar approach will work for even resolution grids but the bookkeeping to handling unused and split cells will require some work.

When the grid involves billions of cells, the indexing based on the rectangle bounding the foldable icosahedron planar view may be too wasteful. A first challenge is to develop a more efficient indexing system. Quite possibly this will just cover the twenty triangles with hexagons and handle the cells that cross the edges of touching triangles. A second challenge is to move from a demonstration system to professional quality algorithms for high resolution grids.

There are many tasks to be addressed for a collection of algorithms to be complete. Tony is interested in indexing optimized for subsets of the globe such as the continental U.S. Perhaps the most crucial task is to provide fast, conceptually acceptable algorithms for changing resolutions. As indicated earlier, lack of strictly nested cells at different resolutions poses a problem. The equal area projection approach easily adapts to strictly nested triangles, but that would give up some of the merits of hexagon cells.

A second challenge area is to consider the use of spatial models in producing cell summaries. For example, Ralph has noted that the current procedure for producing pixel values for satellite images involves a simple near neighbor averaging process. Noel Cressie addressed some of the spatial modeling possibilities in his talk at the Harris Seminar.

Assuming the computation issues are solved, we will then face the biggest challenge of all, institutional inertia. Proposing a standard is one thing. Getting scientists in different nations and different disciplines to use it is another.

Acknowledgments

Research related to this article was supported by EPA under cooperative agreement No. CR820820-01-0. The article has not be subjected to the review of the EPA and thus does not necessarily reflect the view of the agency and no official endorsement should be inferred.

References

Carr, D. B. (1989), "Discussion of Regression Diagnostics with Dynamic Graphics," *Technometrics*, 31(3), 293-296.

Carr, D. B. (1991), "Looking at Large Data Sets Using Binned Data Plots," *Computing and Graphics in Statistics*, eds. A. Buja and P. Tukey, Springer-Verlag, New York, New York, 7-39.

Carr, D. B., Littlefield, R. J., Nicholson, W. L. and Littlefield, J. S. (1987), "Scatterplot Matrix Techniques For Large N," *Journal of the American Statistical Association*, 82(398), 424-436.

Carr, D. B., Olsen, A. R., and White, D. (1992), "Hexagon Mosaic Maps for Display of Univariate and Bivariate Geographical Data," *Cartography and Geographic Information Systems*, 19(4), 228-236, 271.

Carr, D. B. and Pickle, L. W. (1993), "Plot Production Issues and Details: Smoothed Cancer Rates and Hexagon Mosaic Maps," *Statistical Computing & Graphics Newsletter*, 4(2), 16-20.

Conway, J. H. and Sloane, N. J. A. (1988), "Coverings, Lattices, and Quantizers," *Sphere Packings, Lattices and Groups*, New York. Springer-Verlag, 56-62.

Fisch, U., Hasslacher, B. and Pomeau, Y. (1986), "Lattice-Gas Automata for the Navier Stokes Equation," *Physical Review Letters*, 56(14), 1505-1508.

Kahn, R., Haskins, R. D., Knighton, J. E., Pursch, A. and Granger-Gallegos, S. (1991), "Validating a Large Geophysical Data Set: Experiences with Satellite-Derived Cloud Parameters," *Computing Science and Statistics, Proceedings of the 23rd Symposium on the Interface*, Interface Foundation of North America, Fairfax Station, VA, 133-140.

Messer, J. J., Linthurst, R. A., and Overton, W. S. (1991), "An EPA Program for Monitoring Ecological Status and Trends," *Environmental Monitoring and Assessment*, 17, 67-78.

Overton, W. S., Stevens, D. L. and White, D. (1990), Design Report for EMAP, Environmental Monitoring and Assessment Program, EPA 600/3-91/053, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C.

Saff, E. B. and Kuijlaars, A. B. J. (1997), "Distributing Many Points on a Sphere," *Mathematical Intelligencer*, 19(1), 5-11.

Snyder, J. P. (1992), "An Equal-Area Map Projection for Polyhedral Globes." *Cartographica*, 29(1), 10-21.

Stevens, D. L., Jr. (1994), "Implementation of a National Environmental Monitoring Program," *Journal of Environmental Management*, 42, 1-29.

Stevens, D. L., Jr. (1997), "Variable Density Grid-Based Sampling Designs for Continuous Spatial Populations," *Environmetrics*, 8, 167-95.

White, D., Kimerling, A. J., and Overton, W. S. (1992), "Cartographic and Geometric Components of a Global Sampling Design for Environmental Monitoring," *Cartography and Geographic Information Systems*, 19(1), 5-22.

Yang, K. S., Carr, D. B. and O'Connor, R. J. (1995), "Smoothing of Breeding Bird Survey Data to Produce National Biodiversity Estimates," *Computing Science and Statistics, Proceeding of the 27th Symposium on the Interface*, Vol. 27, M. M. Meyer and J.L. Rosenberger (eds.), Interface Foundation of North America, Fairfax Station, VA, 405-409.

Dan Carr
Institute for Computational Sciences and Informatics
George Mason University
dcarr@voxel.galaxy.gmu.edu

Ralph Kahn
Jet Propulsion Laboratory
ralph.kahn@jpl.nasa.go

Kevin Sahr
University of Oregon
sahrk@cs.uoregon.edu

Tony Olsen
EPA National Health and Environmental Effects Research Laboratory
tolsen@mail.cor.epa.gov