A COMPARISON OF EQUAL-AREA MAP PROJECTIONS FOR REGIONAL AND GLOBAL RASTER DATA

E. Lynn Usery, Research Geographer
U.S. Geological Survey
1400 Independence Road
Rolla, Missouri, USA 65401

(usery@usgs.gov)

Jeong-Chang Seong, Assistant Professor
Northern Michigan University
1401 Presque Isle Ave.
Marquette, MI 49855

(jcseong@hotmail.com)

Abstract

High resolution regional and global raster databases are currently being generated for a variety of environmental and scientific modeling applications. The projection of those data from geographic coordinates to a plane coordinate system is subject to significant variation and error based on the selected projection, the raster pixel size, and the specific latitude of the location being projected. While equal-area projections are designed to specifically preserve area, research shows that area preservation varies and selection of a projection for the resolution of the data is critical to developing accurate statistics of attributes such as land cover and elevation. In a comparison of four equal area projections, the Lambert azimuthal equal area, the Mollweide, the Goode homolosine, and the equal area cylindrical, results show that total areas of land cover vary with projection type and with raster resolution. While no single projection is best for all resolutions and at all latitudes, and any of the equal area projections tested are accurate with resolutions of eight kilometer pixels or smaller, the Mollweide appeared most accurate at larger pixel sizes. Analysis of the accuracy of raster projection was conducted by two methods. First, a set of twelve one by one degree squares placed at various latitudes were projected at several raster resolutions and compared to a projection of a vector representation of the same squares. Second, several different raster resolutions of land cover data for Asia were projected and the total areas of 21 land cover categories were tabulated and compared. The results indicate a variance in projection accuracy with latitude and among projection types.

Introduction

With the advent of digital computers and their application to map projection problems from the early 1960's, one might think that all projection problems have been solved. It is true that when handling geographic data for small areas at high resolution and large scale, projection effects tend to be small compared to other sources of data error and inaccuracy. Renewed difficulties occurred in the late 1970's and 1980's with the introduction of a datum change in the United States from the North American Datum of 1927 (NAD 27) to the geocentric-based North American Datum of 1983 (NAD 83) (ACSM, 1983). In recent years this datum shift has plagued users of geographic information systems (GIS) and even with the current status of complete ellipsoid, datum, and projection conversions available in most commercial GIS software packages, the knowledge to use such conversions effectively is still lacking in the GIS user community (Welch and Homsey, 1997). Often approximations to projection equations are used resulting in error and comparing the results from various projections is difficult (Snyder, 1985; Tobler, 1986a; 1986b). We are now entering a phase of GIS and digital cartographic use in which large datasets of high resolution are available for global and regional modeling applications. With these large areas and high resolution, data problems of map projections again become significant. In particular, raster datasets suffer accuracy problems directly attributable to projection transformation (Snyder 1983; 1987; Steinwand *et al*.,1995).

Equal-area projections are generally better for raster datasets since preservation of area characteristic yields pixel areas which are more correct and equivalent. The interrupted Goode homolosine projection has been recommended for global-raster GIS databases, particularly for products generated from the National Aeronautics and Space Administration’s (NASA) Advanced Very High Resolution Radiometer (AVHRR) data (Steinwand *et al*.,1995). If a global GIS database is built using the vector data structure, an equal area projection will preserve most of the original information such as the size of areas, but research indicates that even projections designed to preserve areas, *i.e*., equivalent or equal-area projections, may distort original information when the database is built using the raster GIS data structure.

As Steinwand *et al.* (1995) indicate, the loss and distortion of original information occurs during the image warping as well as the reprojection of raster data. In addition, the spatial resolution of a raster pixel can cause an inaccuracy depending on the projection selected. Assuming a projection with minimum area distortion and allowing maximum angular distortion, the projection will be appropriate only when the raster pixel size is small enough not to be significantly affected by the angular distortion (Nyerges and Jankowski, 1989). As pixel sizes are increased, the information for areas are affected significantly due to the distorted shape.

This paper investigates the effect of spatial resolution change in large regional raster GIS databases using four major global projection methods. The next section provides the theoretical basis for the work and an analysis using mathematically constructed datasets. The third section provides an empirical analysis of the problem using regional land cover for Asia. A final section provides some conclusions about raster projection based on pixel size and latitude.

Theoretical Approach

Twelve ground features were designed specifically considering geographic location (Figure 1). Each feature has a rectangular shape covering a one degree by one degree area under the geographic reference system with latitude and longitude coordinates represented in the decimal degree format. The lower left origins of the rectangles are placed at the intersections of the 0, 25, 50 and 75 degree lines of latitude and the 50, 100 and 150 degree lines of longitude.

In order to represent the curvature of projected lines, each polygon is composed of 1,000 line segments. Each arc segment, therefore, spans 0.004 decimal degrees. The length of each arc segment along a meridian is about 445 meters and the length of each arc segment along a parallel

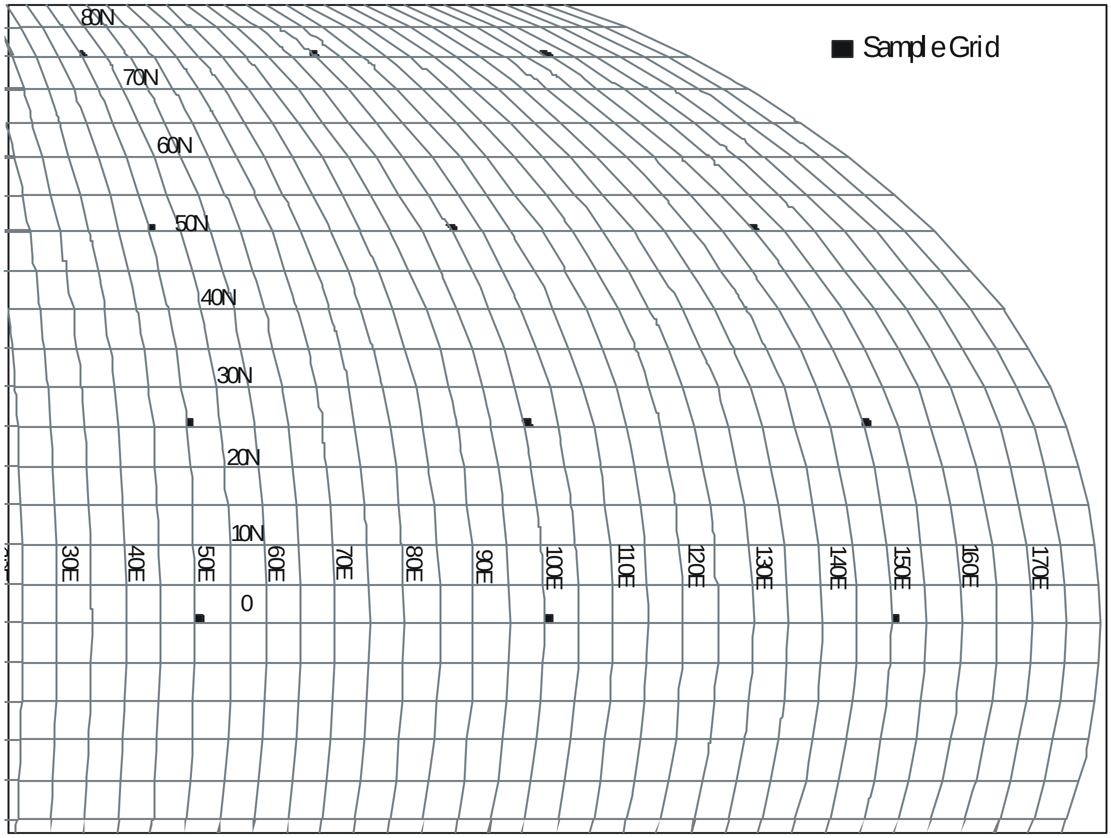


Figure 1: Location of one-by-one degree samples for examining projection effects on area.

of latitude is about 445, 403, 286 and 115 meters at 0, 25, 50 and 75 degrees, respectively (note that distance along a meridian = radius of the earth R 6,371.1 km × difference of latitudes in radians, dn, and distance along a parallel = radius of the earth R 6,371.1 km × cos (latitude in radians) × difference of longitudes in radians, d? (Maling, 1992)). A perfect sphere without flattening is used as a model of the globe. The radius of the earth is considered as 6,371.1 km.

The twelve polygons were imported into the Arc/Info* and ArcView* software systems (ESRI, 2000), and then reprojected to four global projection systems: equal-area cylindrical with the zero-degree central meridian and standard parallel; Mollweide with zero-degree central meridian; Robinson with zero-degree central meridian; and the Goode homolosine interrupted by oceans. As Steinwand *et al*. (1995) suggest, the interrupted Goode homolosine projection was considered first. Because the Goode homolosine projection gives an interrupted look, two other equal area projection methods were selected as alternatives. Mollweide was selected because of its continuous look which is necessary for land-sea integrated global database building. The equal-area cylindrical projection was chosen because of its straight meridians and parallels as well as an entire world look. In addition, a non-equal area projection, the Robinson projection, was selected as an alternative to the equal area projections assuming the attractive look would be useful if it doesn't cause high levels of error when rasterized.

* Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

In terms of the spatial resolution, the raster pixel size used in most coarse-resolution satellite products was considered, because satellite images are very good sources for a global GIS database. Specifically, seven spatial resolutions, 500 m and 1, 4, 8, 16, 25, 50 km, were selected. The 500 m resolution was chosen to match the MODIS sensor system. The resolutions 1, 4, 8 and 16 km were selected representing the products of the Local Area Composite (LAC), Global Area Composite (GAC), and other composite images of the AVHRR. The 25 km resolution was chosen because the longitudinal one-degree distance at 75-degrees latitude is slightly more than 25 km. Also, the 50 km resolution was selected to assess the effect of a pixel size larger than the size of most ground features used in GIS analysis.

Mathematical Basis

The surface area ($S_{a,b}$) covered by a one-by-one degree polygon can be calculated using the integrals of a revolving circle along the x-axis with the radius (R) of 6,371.1 km (Equation 1).

$$S_{a,b} = \int_a^b 2p \cdot f(x) \cdot \sqrt{1 + f'(x)^2}\, dx \qquad \text{(Eq. 1)}$$

where, $S_{a,b}$ = surface area of revolving circle divided by $x = a$ and $x = b$, and

$$f(x) = \sqrt{R^2 - x^2} \qquad \text{(Eq. 2)}$$

$$f'(x) = \frac{-x}{\sqrt{R^2 - x^2}} \qquad \text{(Eq. 3)}$$

$a = R \cdot \sin(n_1)$ and $b = R \cdot \sin(n_2)$, where n_1 is latitude 1 and n_2 is latitude 2.

The surface area, $S_{a,b}$, is 1/360th of the total surface area S, because the distance between longitudes is one degree. Therefore, the size of the one-by-one degree polygon just above the equator is 12,364.072 square kilometers. Using Equation 1, the areas for the rest of the one-degree by one-degree rectangles are 11,160.054; 7,864.816; and 3,095.834 square kilometers at the latitudes of 25, 50 and 75 degrees, respectively.

When the experimental polygons were rasterized using the ‘polygrid’ command in the Arc/Info software, the results were unexpected (note: the polygrid command creates a grid file from the polygon features of an Arc/Info coverage and can take weighting values when two or more polygons overlap in a pixel. In this study, the weighting method was not used, because the 12 sample polygons do not overlap each other in a raster pixel). Figure 2 shows the effects of different projections on the accuracy of the raster area estimation. The X axis denotes pixel sizes from 0.5 km to 50 km and the Y axis shows the percent of area represented by each projection. After the percentage represented in each test polygon was calculated, they were added, and then the sum was divided by the number of test polygons to calculate averages.

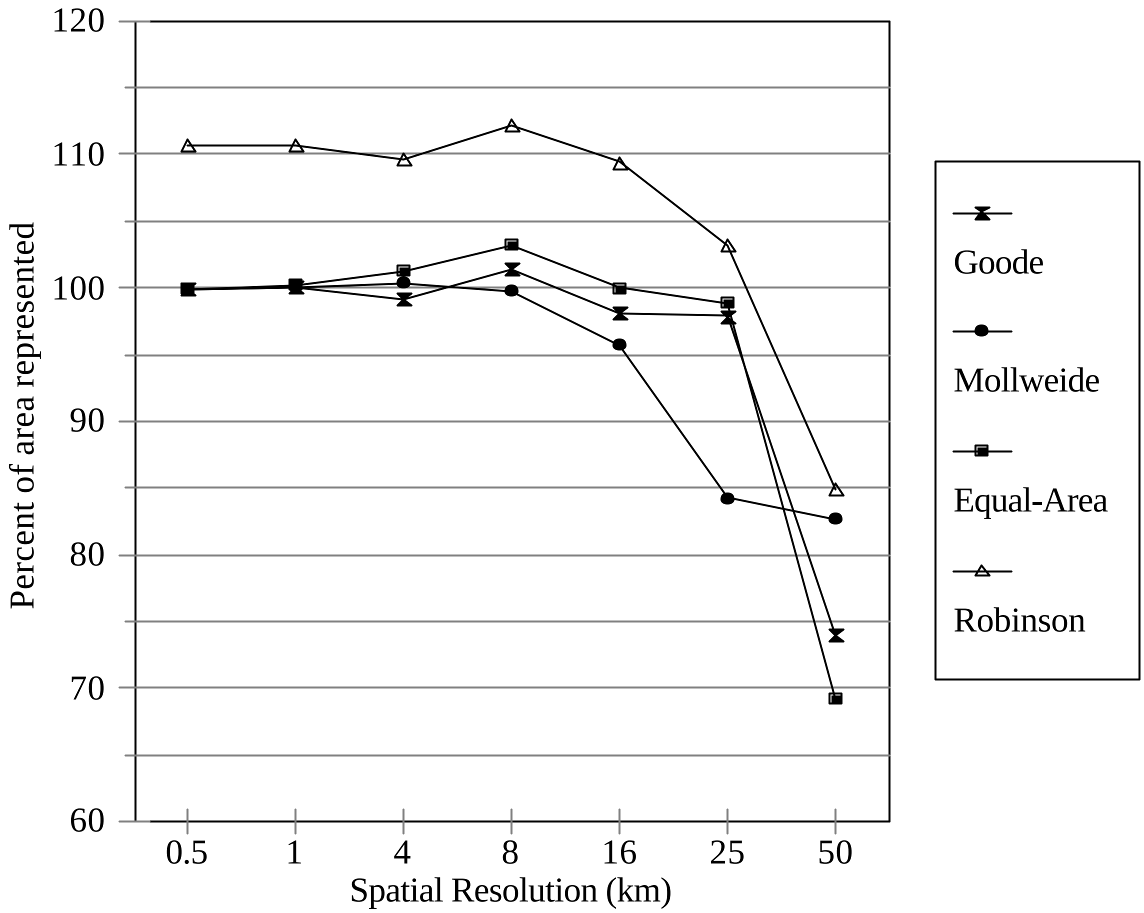


Figure 2. The effect of spatial resolution and global projections on the accuracy of a raster GIS database. The percent of area represented on the Y axis was calculated by averaging the percent values of all the sample polygons, regardless of latitudes and longitudes.

As shown in the figure, the Robinson projection overestimates the area at most pixel sizes except 50 km. At spatial resolutions less than one kilometer, the other three projections show similar high accuracies, which means any equal-area projection can be used for a global GIS database at these spatial resolutions. At a spatial resolution from one kilometer to eight kilometers, the Mollweide projection shows the best representation. The equal-area cylindrical and Goode projections show slight overestimation, but the Mollweide projection shows an almost perfect fit. From 16 km to 25 km, the equal-area cylindrical projection shows the best representation. The Robinson projection shows the best representation at the 50 km spatial resolution.

Figure 3 shows the raster area representation at four latitudes using an eight-km pixel size. The figure shows general over-representation at latitudes of 60 degrees or more. The Mollweide and Goode projections show relatively high accuracy regardless of latitudes. The equal-area cylindrical projection shows slight overestimation at the high latitude. In case of

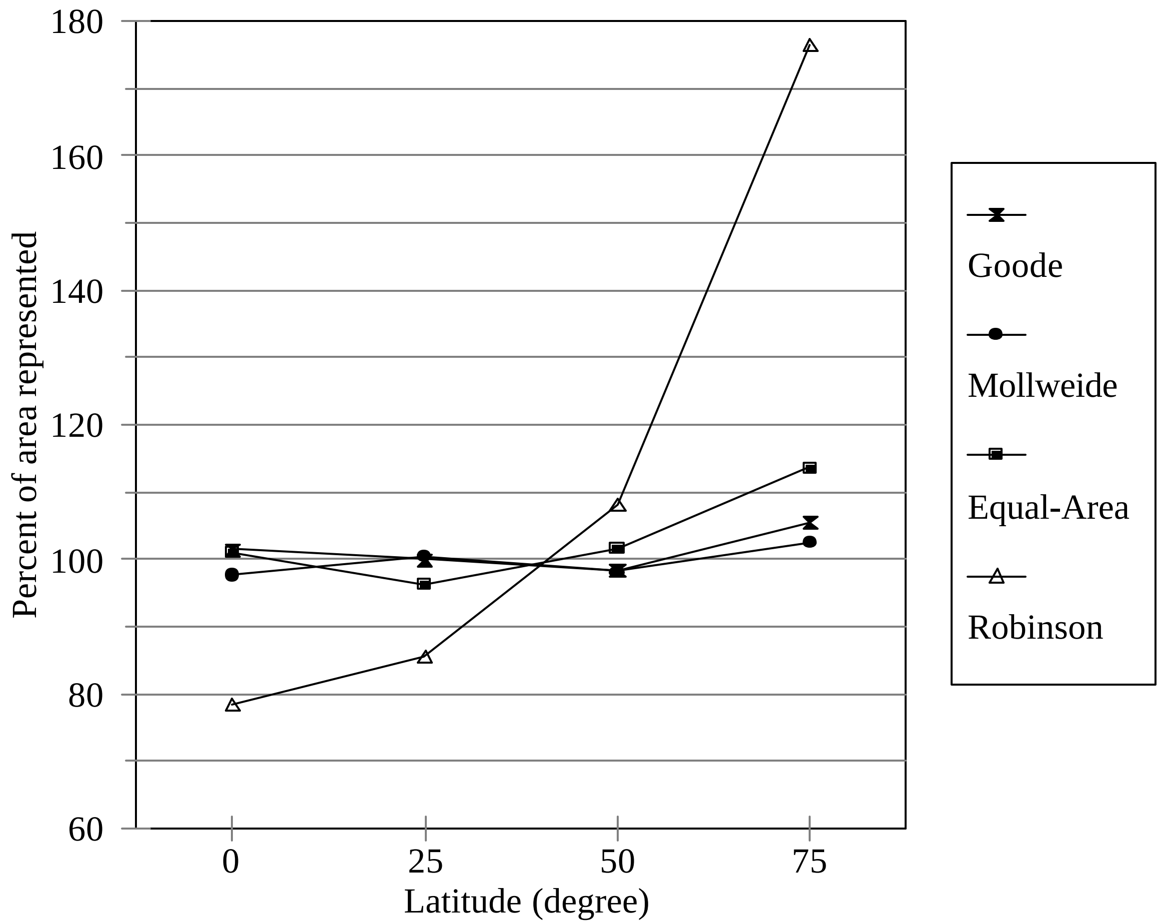


Figure 3. The effects of latitude and projections on the accuracy of the raster GIS database, eight kilometer pixel size. The figure shows general over-representation at the latitudes 60 degrees or more.

the Robinson projection, the represented areas are significantly smaller than the actual areas at low latitudes, while they are significantly larger than the actual area at high latitudes. These findings suggest the Mollweide projection is slightly better than the Goode projection at the spatial resolutions of eight km or less.

Projection Application to Geographical Distributions of Land Cover

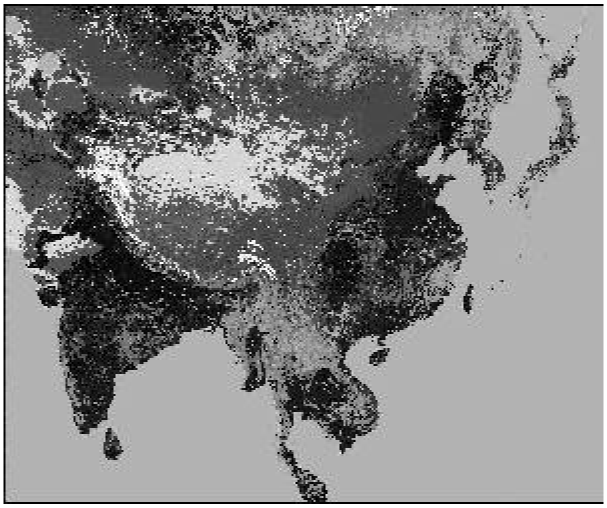
In this section, the problem of establishing area percentages for various land cover categories defined on raster datasets is examined as a function of projection method and data resolution.

The base data for this analysis was downloaded from the EROS Data Center of the U.S. Geological Survey (USGS) (http://edcwww.cr.usgs.gov/landdaac/glcc/glcc.html) and consists of two files of land cover for Eurasia, one in a Goode homolosine projection and the second in a Lambert azimuthal equal area projection. Both were coded in a USGS land cover categorization (Table 1) with 24 categories and the Asia portion extracted with a vector boundary (Figures 4a and 4b). The data from the Lambert projection were projected to geographic coordinates and then reprojected to the Mollweide, Robinson, and equal area cylindrical projections using a rigorous transformation with the ERDAS Imagine 8.4 * software (Figure 4c, 4d, and 4e, respectively). The data in geographic coordinates were projected to the equal area cylindrical projection using ESRI's Arc/Info. The areas for each land cover category were then tabulated as percentages of the total area. Because the extent of the background area and water from the oceans varies from one projection to another depending on the chosen area to be projected, these categories were excluded from the tabulations. Also, since there were no areas of bare ground tundra and dryland/irrigated cropland and pasture in Asia, these two categories were eliminated from the percentage tabulations and do not appear on the maps or in the discussion below.

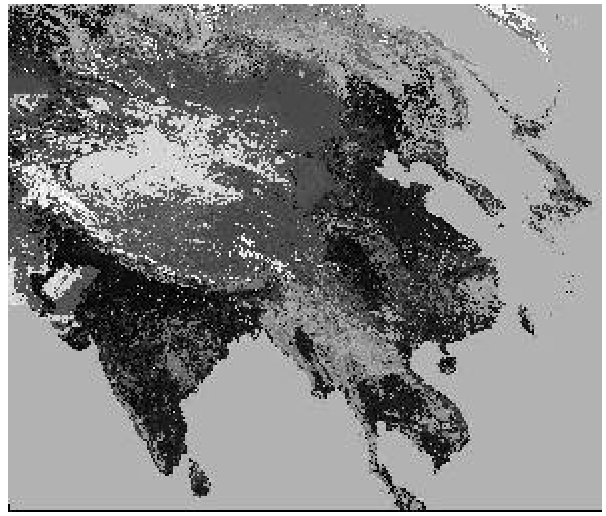
Table 1
Asia Land Cover Percentages by Projection

	16-km Pixels					50-km Pixels				
Land Cover Categories	Lam	Goode	Eq-Cyl	Mw	Rob	Lam	Goode	Eq-Cyl	Mw	Rob
Urban & Built-Up Land	0.16	0.17	0.16	0.16	0.15	0.14	0.17	0.21	0.15	0.09
Dryland Cropland & Pasture	12.36	12.70	12.13	12.47	11.97	12.24	13.16	12.48	12.57	11.76
Irrigated Cropland & Pasture	11.26	12.24	11.20	11.06	10.33	11.60	12.38	11.59	11.07	10.70
Cropland/Grassland Mosaic	5.89	5.78	5.91	5.79	5.83	5.95	5.81	5.66	5.78	5.84
Cropland/Woodland Mosaic	4.24	3.96	4.33	4.32	4.28	3.85	3.79	4.15	4.37	3.97
Grassland	17.12	14.82	17.05	17.00	17.65	17.14	14.55	16.61	16.96	17.57
Shrubland	14.27	11.69	14.45	14.25	14.31	13.94	11.96	14.19	14.44	14.62
Mixed Shrubland/Grassland	2.05	2.39	2.07	2.10	1.96	2.05	2.42	2.24	2.05	2.03
Savanna	4.49	5.23	4.39	4.55	4.58	4.54	4.96	4.64	4.65	5.03
Deciduous Broadleaf Forest	3.20	3.67	3.23	3.16	3.14	3.44	3.44	3.41	2.93	3.02
Deciduous Needleleaf Forest	1.87	2.86	1.92	1.89	2.23	1.88	2.94	1.86	2.04	2.08
Evergreen Broadleaf Forest	2.70	3.09	2.79	2.73	2.40	2.60	3.19	2.72	2.60	2.57
Evergreen Needleleaf Forest	0.83	0.94	0.86	0.85	0.81	0.85	0.76	0.86	0.96	0.85
Mixed Forest	8.56	9.46	8.52	8.64	9.25	8.63	9.53	8.25	8.45	9.01
Herbaceous Wetland	0.14	0.19	0.16	0.14	0.15	0.12	0.18	0.16	0.12	0.11
Wooded Wetland	0.11	0.13	0.11	0.09	0.14	0.17	0.10	0.12	0.11	0.11
Barren or Sparsely	9.20	8.66	9.19	9.24	9.19	9.28	8.65	9.38	9.12	9.02
Herbaceous Tundra	0.16	0.06	0.17	0.16	0.17	0.17	0.07	0.12	0.21	0.17
Wooded Tundra	1.11	1.56	1.07	1.12	1.16	1.23	1.54	1.07	1.10	1.19
Mixed Tundra	0.04	0.11	0.05	0.05	0.07	0.03	0.11	0.04	0.04	0.02
Snow or Ice	0.25	0.28	0.24	0.25	0.25	0.15	0.30	0.24	0.28	0.25

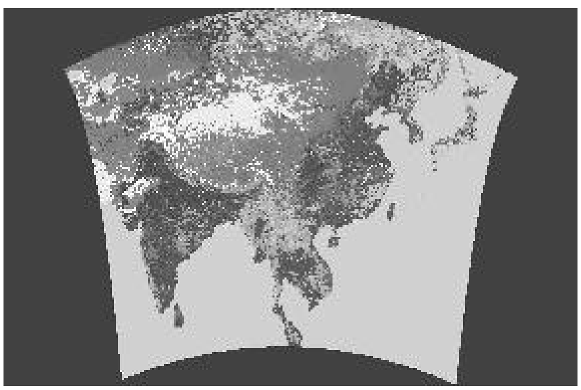
* Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.



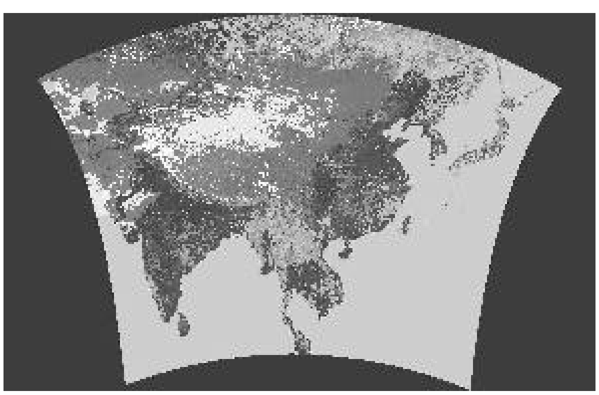
a) Lambert azimuthal equal area



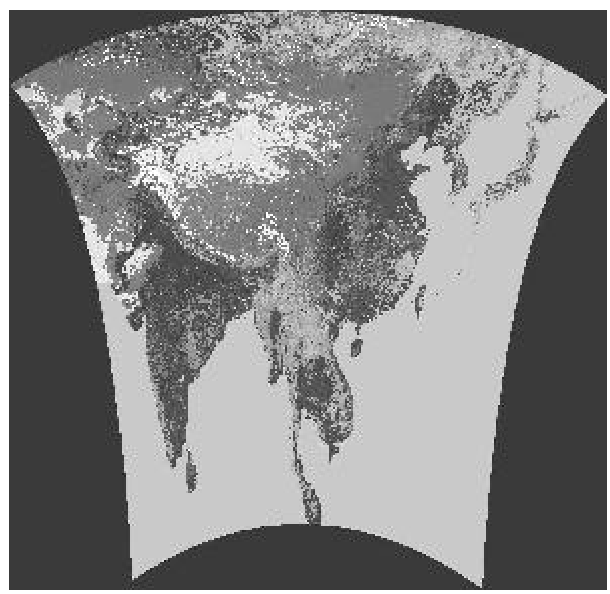
b) Goode homolosine



c) Mollweide



d) Robinson



e) Equal area cylindrical

Figure 4. Asia land cover at 8 km pixels in five different projections.

Analysis of Projection Effects on Land Cover Percentages

The tabulations of total areas for the 21 land cover categories for the five different projections provides a basis for empirically verifying the results of the mathematical analysis above. As can be seen in Figure 2 and Table 1, the Robinson projection significantly overestimates the land compared to the equal area projections. The Lambert, equal-area cylindrical, and Mollweide yield almost identical results at the 16-km resolution, but only the Mollweide retains almost identical percentages at the 50-km resolution. Mollweide and equal-area cylindrical projections provide similar areal percentages. The Goode projection also fails to retain land cover areal percentages between the two resolutions reflecting the mathematical results of Figure 2.

The latitudinal results are most easily verified by examining particular land covers which occur almost exclusively at specific latitudes. For example, the deciduous needleleaf forests occur at high latitudes in Siberia. For the four projections in Figure 3, the Mollweide projected results should have the lowest values, followed by the equal-area cylindrical, the Goode, and the Robinson should show the greatest overestimation. These results are consistent with the tabulated percentages in Table 2, with the exception of the Goode which was projected from a different source and includes a slightly different area (including a portion of the Kamchatka Peninsula not contained in the other source data). The different sources were needed because of the current limitations of commercial software to generate Goode projections. As expected from Figure 3, the Robinson projection overestimates the area at high latitudes, shown by the deciduous needleleaf category, but also shows a reduction of the overestimation at the 50 km resolution. Data projections of the various categories at 1, 4, 8, 16, 25, and 50-km resolutions verify the mathematical results of Figures 2 and 3.

Summary and Conclusions

In this paper, the significant effect of global map projections on the accuracy of tabulated statistical results has been examined. Through integration of raster and vector representations and reprojections of the raster data, accuracy of results has been shown to be dependent on raster resolution and latitudinal position. At resolutions from one to eight kilometers, most equal area projections perform adequately. At resolutions coarser than eight kilometers, variances by projection can be significant with the Mollweide maintaining the greatest consistency over various pixel sizes and over various latitudes.

The result of the effect of spatial resolution and global map projections is significant. According to the results found from the twelve one-by-one degree polygons, the Robinson projection, a non-equal-area projection, showed the poorest estimation in terms of the percentage of areas represented after rasterization, an expected result. Three equal-area projection methods, the interrupted Goode homolosine, Mollweide, and equal-area cylindrical projections, showed little difference in area representation in spatial resolutions of one kilometer or less. However, at the spatial resolutions from one kilometer to eight kilometers, the Mollweide projection showed the best result. At the spatial resolution ranges from 16 km to 25 km, the Goode homolosine and equal-area cylindrical projections showed slightly better results than the Mollweide projection (the Mollweide projection tends to under-represent the original area at this spatial resolution range). The Robinson projection significantly over-represented the original area at the spatial resolution ranges of 16 kilometers or less and the over-representation reached about 10 percent.

At the spatial resolution of eight km, all the global projections used in this study tend to over-represent the original area at latitudes of 60 degrees or higher. The representation is most accurate in the Mollweide projection with Goode homolosine, equal-area cylindrical, and Robinson following in order of accuracy. These findings suggest that the Mollweide projection is a good alternative to the interrupted Goode homolosine projection. Also, the Mollweide projection has an advantage in that it represents the oceans and land masses without any interruption.

References

ACSM, 1983. "The North American Datum of 1983," *American Association for Geodetic Surveying Monograph No. 2*, American Congress on Surveying and Mapping, Falls Church, VA.

ERDAS, 2000. Imagine 8.4, ERDAS, Inc., Atlanta, Georgia.

ESRI, 2000. Arc/Info, ESRI, Inc., Redlands, California.

Maling, D.H., 1992. *Coordinate Systems and Map Projections*, Pergamon Press, Oxford.

Nyerges, T.L. and P. Jankowski, 1989. "A Knowledge-Base for Map Projection Selection," *The American Cartographer*, Vol. 16, No. 1, pp. 29-38.

Snyder, J.P., 1983. "Efficient Transfer of Data Between Maps of Different Projections," *Proceedings, American Congress on Surveying and Mapping 43rd Annual Meeting*, Washington, D.C., pp. 332-340.

Snyder, J.P., 1985. "Computer-Assisted Map Projection Research," *U.S. Geological Survey Bulletin 1629*, U.S. Government Printing Office, Washington, D.C.

Snyder, J.P., 1987. *Map Projections - A Working Manual*, U.S. Geological Survey Professional Paper 1395, U.S. Government Printing Office, Washington, D.C.

Steinwand, D.R., J.A. Hutchinson, and J.P. Snyder, 1995. “Map Projections for Global and Continental Data Sets and an Analysis of Pixel Distortion Caused by Reprojection,” *Photogrammetric Engineering and Remote Sensing*, Vol. 61, No. 12, pp. 1487-1499.

Tobler, W.R., 1986a. "Polycylindric Map Projections," *The American Cartographer*, Vol. 13, No. 2, pp. 117-120.

Tobler, W.R., 1986b. "Measuring the Similarity of Map Projections," *The American Cartographer*, Vol. 13, No. 2, pp. 135-139.

Welch, R. and A. Homsey, 1997. “Datum Shifts and UTM Coordinates,” *Photogrammetric Engineering and Remote Sensing*, Vol. 63, No. 4, pp. 371-375.