Using Digital Spatial Data Sets to Study the Impact of Reservoir Construction on Local Environment and Community

Yong Wang and Scott Wade
Center for Geographic Information Science and Department of Geography
East Carolina University

Digital spatial and demographic data sets have been used to study the impact of the Randleman Reservoir on the local environment and demography of Randolph and Guilford counties, N.C. At surface water heights of 682, 685, 706, and 709 ft. above the mean sea level, the reservoir's capacities were 25,053, 29,242, 92,654, and 106,654 acre-ft., and total surface areas (reservoir plus 200ft. erosion/pollution control buffers) 3,081, 3,516, 7,403, and 8,233 acres, respectively. The capacities and total surface areas were lower than those reported in the environmental impact statement by the U.S. Army Corps of Engineers and those published at the Piedmont Triad Regional Water Authority's Web page. These underestimations were attributed to the inaccurate representation of the digital evaluation model (DEM) data used. After applying a 3 by 3 minimum spatial filter to the DEM data, the recomputed reservoir capacities and areal extents were very close to those reported and published as mentioned in the above. At the surface water heights of 682 and 706 ft., the recalculated capacities were 52,445 and 162,709 acre-ft., and reservoir (only) areas 2,958 and 7,035 acres, respectively.

Introduction

Great effort has been made to create spatial and demographic data sets for documenting and studying the physical and social environments in the United States. These data sets include elevation data, satellite imagery, landuse and land cover types, digital aerial photography, political and statistical boundaries, streets and highways, as well as population and other demographic statistics. Most of the data are available for free or for very little cost (with the exception of satellite imagery) from U.S. government agencies like the U.S. Geological Survey (USGS) and the U.S. Census Bureau, and much of it can be conveniently downloaded from the Internet.

Due to the rapid development in computer technology, varieties of GIS software for desktop/laptop computers are widely available. The leading GIS software includes ArcView, ArcInfo, and recently ArcGIS (Environment System Research Institute, ESRI, California), IDRISI (Graduate School of Geography, Clark University), MapInfo (MapInfo Corporation of Troy, New York), and others (Clarke 2001). The software products have easy-to-use graphical user interfaces. They are reasonably priced and widely used in schools, government agencies, and the private sector. University geography departments use the products to teach geographic information science to students, many of whom decide to pursue undergraduate and graduate degrees in geography. Students from other disciplines (e.g., biology, geology, and business) are also learning to use GIS software in order to enhance their technical research skills.

Having briefly discussed the development and availability of digital spatial data sets and GIS software, we next present an example of how they can be used to study a regional planning problem in North Carolina. After years of planning and preparation of environmental impact studies by governmental agencies and private companies (e.g., Moore and Leonard 1973, Weiss et al. 1973, Black and Veatch 1988, 1990, Lautzenheiser et al. 1997, U.S. Army Corps of Engineers 2000a), the construction of the Randleman dam and reservoir started in Summer 2001. If all goes as planned, the reservoir will be filled by 2004 (http://www.ptrwa.org). The future reservoir will provide water to the Piedmont Triad Regional Water Authority's (PTRWA) six members: Greensboro, High Point, Jamestown, Archdale, Randleman, and Randolph County.

Analysis

The objectives are to: 1) estimate the reservoir capacity produced by the dam project; 2) calculate the surface area of the reservoir and its 200 foot erosion/pollution control buffer; 3) determine the landuse and land cover types inundated by the reservoir; and

4) study the impact of the reservoir on human settlement/resettlement in the area. All of the above will be addressed at two conservation pool and two flood pool surface heights.

Study area

The dam site is situated on the Deep River about 2 miles northwest of the City of Randleman, NC. The reservoir will be mainly along the Deep River (upstream), and also partially along Muddy Creek (Fig. 1). When the surface water height is at 682 ft. (the conservation pool), the reservoir's water will be back up about 13 miles along the Deep River, almost to I-85. When the reservoir is at the surface water height of 706 ft. or its flood pool, the water will reach the City of Jamestown, NC.

Spatial and demographic datasets

a) DEM data. As a part of the National Mapping Program, the USGS led the creation of the DEM data set with coverage for the entire United States and its territories. DEMs are digital elevation data that consist of arrays of elevations in x and y directions and are sampled at regularly spaced intervals (cells). An elevation (z) value of a cell is sampled from elevation values of all locations within the cell, and is measured based on the mean sea level. DEMs are used for presenting and studying the topography of ground

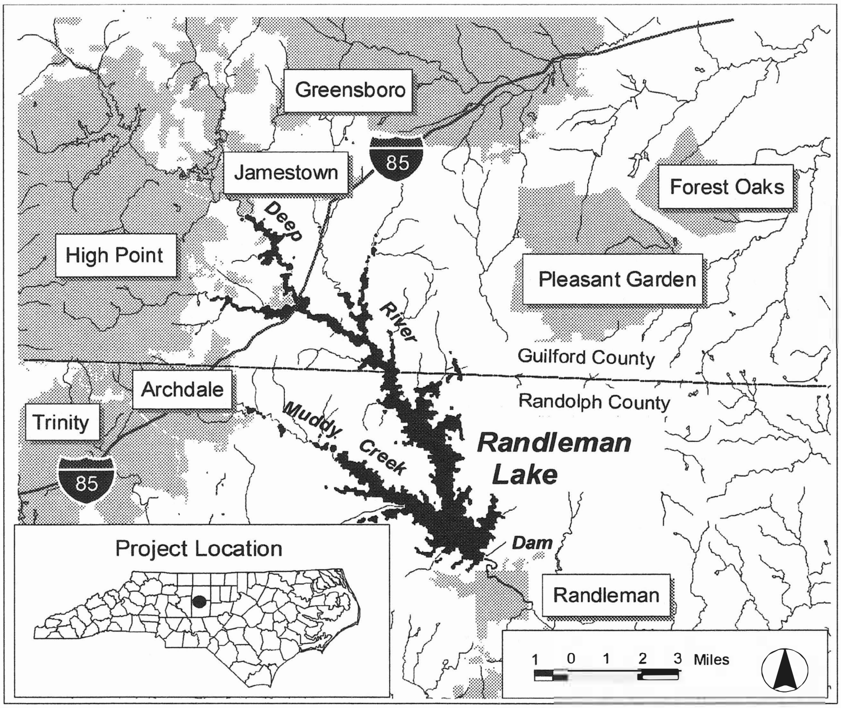

Figure 1. The future Randleman Reservoir and its surroundings in Guilford and Randolph counties, NC.

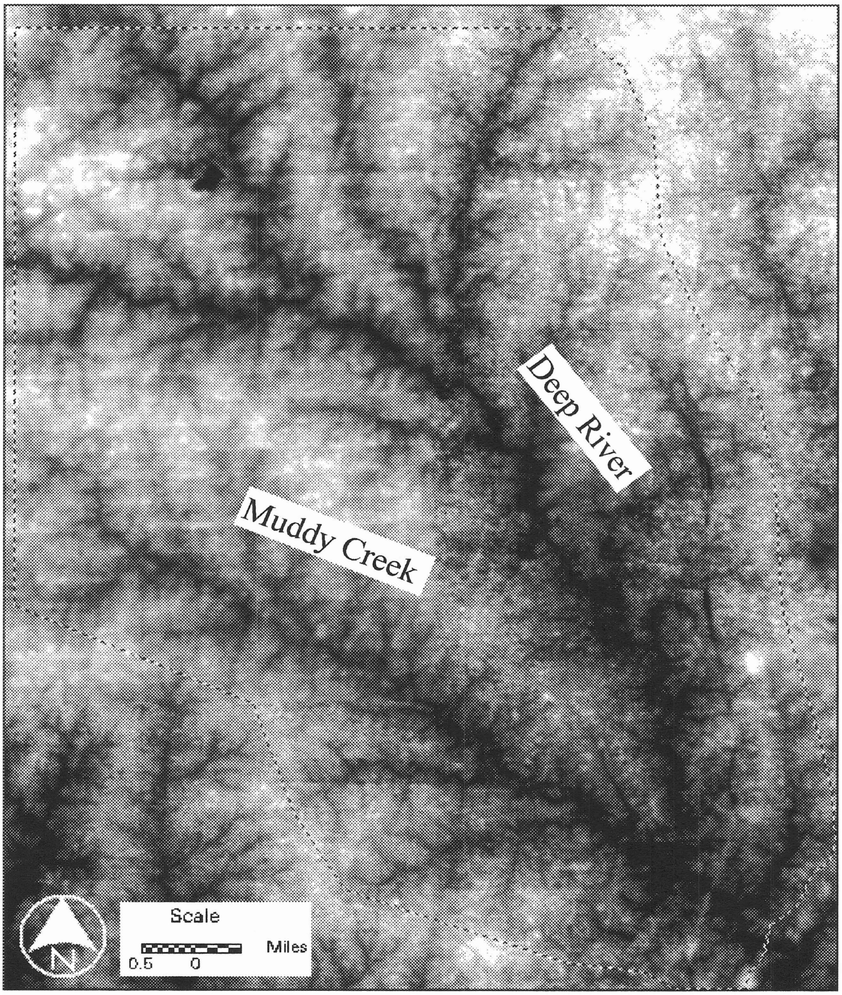

Figure 2. DEM for Randleman Reservoir. The dark areas show low elevation, and bright areas high elevation. The area of interest (AOI) is outlined.

surface, and for hydrological modeling of a local watershed or even an entire river-basin (e.g., U.S. Army Corps of Engineers 1986, 2000b). The DEMs (of the U.S.) are freely downloadable (http://edcwww.cr.usgs.gov/doc/edchome/ndcdb/ndcdb.html). The DEMs used were at 1:24,000 scale with a cell size (x, y) of 30 m by 30 m. The accuracy of the z value is less than or equal to ˚ of the contour lines on the USGS 7.5 minute quadrangle or 5 ft. in the study area. Fig. 2 shows the local topography, with black indicating low elevation and white high elevation. The area of interest (AOI) covering the future reservoir and its surrounding areas is also outlined.

b) *Landsat TM data (the background in figure 4).* Landsat is a series of satellites developed and sent into space by the NASA. The Landsat program dates back to the early 1970s. Landsat 7 launched in 1999 along with Landsat 5 are two Landsat satellites still in operation. Using its onboard remote sensors, Landsat collects information about the earth's surface. One of the sensors is the Thematic Mapper (TM), which measures the reflectance of surface targets on earth illuminated by the solar radiation in the visible and infrared wavelength range or in a portion of electromagnetic spectrum. TM data are distributed as digital images, each of which covers an area of approximately 180 km by 180 km, or 32,400 km^2. The cell (pixel) sizes of TM sensor onboard Landsat 7 are 30 m by 30 m for bands 1-5 and 7, 60 m by 60 m for band 6, and 15 m by 15 m for band 8. The pixel sizes of Landsat 5's TM are 30 m by 30 m for bands 1-5 and 7, and 120 m by 120 m for band 6. The TM data are available from the USGS's EROS data center in South Dakota, EOSAT Company in Maryland, or SpaceImaging Corporation in Colorado. The cost of the Landsat 5 or 7 TM data varies. One Landsat 7 TM image ordered from the USGS costs about $600.00, and a Landsat 5 TM image ordered from the SpaceImaging Corporation ranges from $600 to $1675 depending on levels of processing required by a customer (http://www.spaceimaging.com). TM data as well as remotely sensed data collected by other satellites have been widely used to study the earth's environment (e.g., Verbyla 1995, Jensen, 2000).

c) *Landuse and land cover type data.* The data were derived mainly from the Landsat 5's TM data, coupled with ground observation and other available ancillary information. They were used to quantify each landuse and cover type to be inundated by the construction of the Randleman dam. In the reservoir region, there are 13 landuse and land cover types, ranging from high intensely developed urban area to different types of natural and vegetated surfaces; as well as open water (see Table 2 for other landuse and land cover types). Landuse and land cover types in small areas or patches of small areas might not be identified within the data because the pixel size of the data was 30 by 30 m.

d) *High resolution digital orthophoto quadrangle (DOQ) data.* These are digital photographic images with a resolution of 1 m by 1 m. Panchromatic DOQs created by the USGS in the early 1990s can be freely downloaded from the Microsoft TerraServer (http://terraserver.microsoft.com) at a degraded spectral resolution. Higher quality (panchromatic) originals, as well as false color infrared DOQs created recently (between 1997 and 1999) can be ordered from the USGS ($60.00 per 7.5 minute quadrangle, plus CD charge and shipping). Due to its high resolution, cities, towns, streets, and individual houses can be easily identified (Fig. 3). The DOQs downloaded from the Microsoft TerraServer were used to quantify the number of houses and man-made structures (e.g., barns) that will be inundated by the reservoir or will be within the 200 ft. buffer zone around the reservoir.

e) *Demographic and spatial data.* The U.S. Census Bureau has collected demographic data about this country for over 200 years. The largest single data collection endeavor by the Census Bureau is the decennial census of population and housing, which provides a breakdown of population, housing, and other socioeconomic variables for the national level all the way down to a geographic area equivalent to the

Figure 3. A USGS DOQ, showing individual houses, roads, trees, and vegetation in the study area.

city block. The latest decennial census was taken in 2000. The data collected in the 2000 Census is gradually becoming available, and can be downloaded for free from various Web sites (http://www.census.gov, http://www.geographynetwork.com). The Census Bureau also produces a widely utilized spatial database called TIGER (Topologically Integrated Geographic Encoding and Referencing, http://www.census.gov). TIGER was developed in the 1990s to produce large scale, up-to-date maps which could be used by enumerators in census taking operations. TIGER files contain streets, political boundaries, hydrography and land marks. In addition, TIGER files provide census statistical boundaries (census tracts, block groups, etc.); they are very useful for generating thematic maps of census population data. TIGER files have been converted into native and interchange formats easily read by many GIS software packages. They are distributed by county, and can be freely downloaded from the above Web sites.

Method

a) *Geo-reference the data sets.* The spatial data used were treated as information layers in a GIS. Because these layers were geo-referenced in different coordinate systems, they were reprojected to a common coordinate system before using them together. In this study, the UTM (Universal Transverse Mercator) coordinate system is used as the common coordinate system. The model of the earth's size and shape used for both location (x, y) and elevation (z) is the WGS84 (World Geodetic System – 1984) reference ellipsoid. The distance unit is the meter. (It should be noted that the NC State Plane coordinate system based on NAD83 datum is the standard system for accurate mapping in North Carolina.)

b) *Delineate the area of interest (AOI).* The AOI was delineated such that it contained both the reservoir at its highest flood pool surface height (709 ft. above the mean sea level) and the 200ft. buffer around the reservoir. The dam formed part of the AOI border, effectively excluding downstream areas from reservoir size and capacity calculations.

c) *Calculate reservoir size and capacity.* To determine the extent of the reservoir, we extracted all cells within the AOI where the DEM elevation was less than or equal to a given reservoir water surface height. By summing the area of these extracted cells, the total area of the reservoir was computed for that given water surface height. Then, for each cell within the reservoir, a height difference between the DEM value (or bottom of the reservoir) and surface water height was calculated. The difference was then multiplied by the cell size to compute the volume (of water) in that cell location. By summing all the cell volumes reservoir capacity was estimated.

d) *Determine landuse and land cover types affected by the dam construction.* A simple overlay of the landuse and land cover type layer onto the extent of the reservoir and the buffer around the reservoir provides the information regarding which landuse and land cover types will be inundated by the reservoir, as well as which landuse and land cover types will be within the 200 ft. buffers.

e) *Count houses and other man-made structures to be impacted by the reservoir.* By overlaying the reservoir's areal extent and its buffer zone onto the DOQs, heads-up digitizing was used to identify and count houses and structures. In the future, if parcel boundaries and their corresponding real estate values are available from a county tax office, the total property value impacted by the construction of the reservoir can be calculated. (It should be also noted that census data we had at this time did not contain information about the counts of houses and other man-made structures; only the DOQ was used to count the number of houses and structures.)

f) *Estimate the number of people to be displaced by the reservoir.* Again, by overlaying the reservoir and its buffer onto the census data, the impact on the local demography was assessed. In the analysis, population figures were assumed to be uniformly distributed within each census block, and were proportionally allocated to the block pieces located within the reservoir and buffer boundaries. The estimated population of the block pieces was then summed to provide the estimated total population displaced by the reservoir.

Results

After the dam site and AOI (of the reservoir) have been identified, a model to analyze and to help understand the impact of the reservoir on the local

Table 1. Reservoir capacities and areas at four different surface heights.

Water Height (ft.)	Capacity (acre-ft.)	Reservoir Area (acre)	Buffer Area (acre)	Total Area (acre)
@ 682	25,053	1,605	1,475	3,081
@ 685	29,242	1,895	1,621	3,516
@ 706	92,654	4,465	2,938	7,403
@ 709	106,654	5,043	3,190	8,233

communities was developed. In the model, four surface water heights, 682, 685, 706, and 709 ft. above the mean sea level for the reservoir and its 200 ft. (horizontal) buffer zone were used; four reservoir sizes and capacities were computed (Table 1 and Fig. 4). The first two surface water heights (682 and 685 ft.) could be treated as the conservation pool heights, and last two as flood pool heights. (Due to uncertainty of the z value, ±5 ft. in the DEM data used, the derived reservoir capacities and areal extents at 682 and 685 ft. could be the same, and the capacities and areal extents at 706 and 709 ft. could be the same too.) At surface water heights of 682, 685, 706, and 709 ft., the reservoir's capacities were 25,053, 29,242, 92,654, and 106,654 acre-ft., and total surface areas (reservoir plus 200 ft. erosion/pollution control buffers) 3,081, 3,516, 7,403, and 8,233 acres, respectively. Also, benefit/cost trade-offs of building the dam at lower or higher heights were evaluated. For instance, the ratios of reservoir's capacity to total affected areas (reservoir and buffers) at surface water heights of 682, 685, 706, and 709 ft. were 8.1, 8.3, 12.5, and 13.0 (acre-ft./acre), respectively. This ratio could be used as one possible trade-off indicator.

The location and areal extent of each landuse and land cover type within the reservoir and its (200 ft.) buffer zone were identified and estimated (Table 2). The most affected landuse and land cover types were mixed upland hardwoods, managed herbaceous cover, and cultivated lands.

Using high resolution DOQs, the number of houses and man-made structures within the reservoir and its buffer zone at four surface water heights of the reservoir were counted; 82 to 321 houses and man-made structures would be affected depending on reservoir surface water height (682 ft. to 709 ft., Table 3). The ratios of reservoir capacity to the total number of houses and man-made structures within the reservoir and its buffer at surface water heights of 682, 685, 706, and 709 ft. were 305.5, 278.5, 338.2, and 332.3 (acre-ft. per number of houses and structures), respectively.

By overlaying the reservoir and its buffer onto the 2000 census block data (Fig. 5), the number of people to be displaced (Table 3) was estimated. The number of people to be affected ranged from 399 to 1376 for the surface water heights of the reservoir between 682 and 709 ft. At surface water heights of 682, 685, 706, and 709 ft., the ratios of the reservoir's capacity to the number of people to be relocated were 62.8, 62.3, 80.8, and 77.5 (acre-ft. per person),

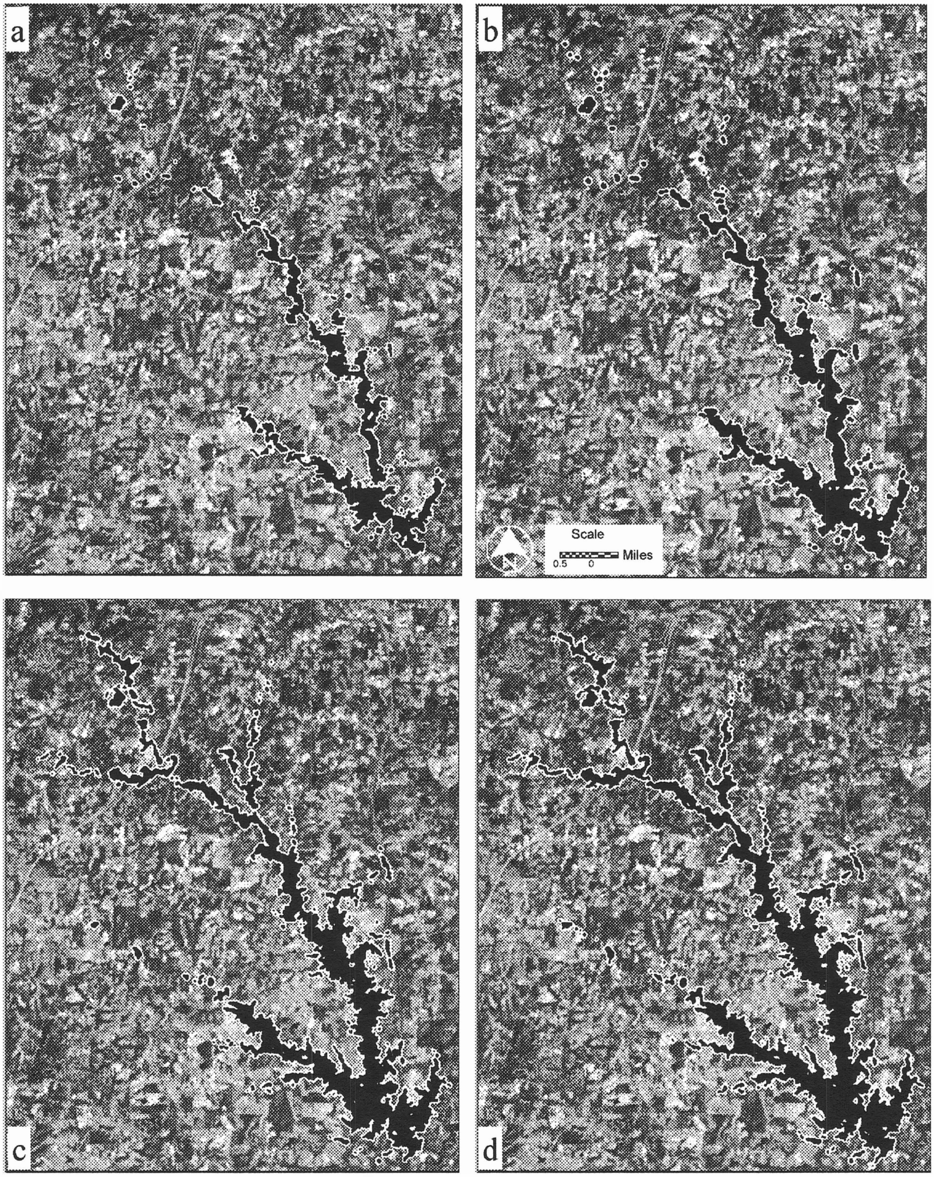

Figure 4. Randleman Reservoir (black) and its 200 ft. buffer zone (white) at water surface heights of 682 (a), 685 (b), 704 (c), and 709 (d) ft., respectively. The background for this figure was the TM image.

Table 2. Areas (acre) of each landuse and land cover type within the reservoir and erosion/pollution control buffer zone at two surface water heights.

	Water height	@682 ft.			@709 ft.	
Land use and land cover type	Reservoir	Buffer	Total	Reservoir	Buffer	Total
High intensely developed urban area	29.7	29.7	59.3	93.9	81.5	175.4
Low intensely developed urban area	9.9	12.4	22.2	39.5	17.3	56.8
Cultivated land	86.5	49.4	135.9	180.4	79.1	259.5
Managed hervaceous cover	432.4	378.1	810.5	1,341.7	798.1	2,139.9
Unmanaged hervaceous cover-upland	14.8	4.9	19.8	24.7	12.4	37.1
Evergreen shrubland	12.4	4.9	17.3	17.3	4.9	22.2
Deciduous shrubland	42.0	42.0	84.0	116.1	44.5	160.6
Mixed upland hardwoods	914.3	901.9	1,816.2	3,031.9	1,989.1	5,021.0
Southern yellow pine	32.1	34.6	66.7	98.8	121.1	219.9
Other needle leaf evergreen forest	0.0	0.0	0.0	0.0	2.5	2.5
Mixed hardwoods/confiers	32.1	14.8	46.9	71.7	22.2	93.9
Water bodies	0.0	2.5	2.5	19.8	14.8	34.6
Unconsolidated sediment	0.0	0.0	0.0	2.5	2.5	4.9
Total	*1,606.1*	*1,475.2*	*3,081.3*	*5,038.3*	*3,190.0*	*8,228.3*

Table 3. Estimated numbers of houses and man-made structures, and people within the reservoir and buffer zone at four different reservoir surface heights.

	# of houses & structures			# of people
Surface Water Height (ft)	**Within Reservoir**	**Within Buffer**	**Total**	**Within Reservoir & Buffer**
@ 682	27	55	82	399
@ 685	37	68	68	469
@ 706	124	150	150	1,146
@ 709	139	182	182	1,376

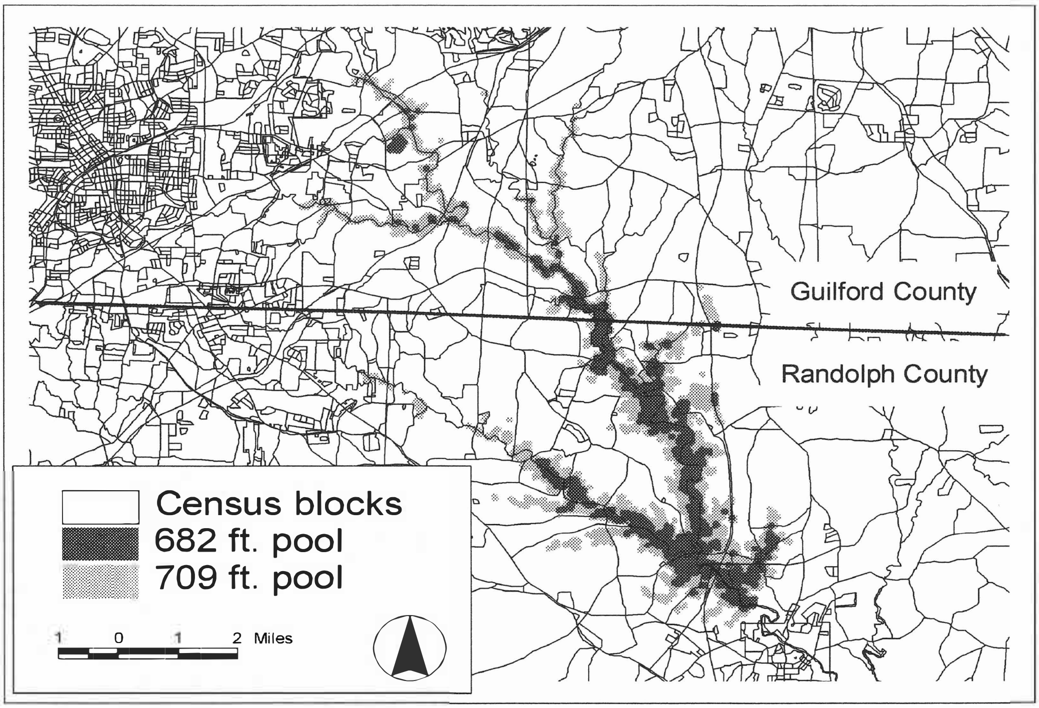

Figure 5. Reservoir and its buffers at surface water heights of 682 ft. and 709 ft. were overlaid over the Census 2000 block data; the blocks need to be redrawn before the 2010 census.

respectively. For the latter two ratios, higher values indicated that a greater volume of reservoir water was produced for each inundated house and man-made structure and for each displaced person. If the two types of ratios were used to determine the surface water height, 706 ft. would be the optimal height.

Discussion

While low cost and ease of access make digital spatial databases attractive for use in studies such as this one, they are not without limitations. For example, note the isolated reservoir polygons in Fig. 4. These polygons resulted from the coarse spatial resolution (30 m by 30 m) and uncertain vertical resolution of the DEM data. The isolated polygons were separated from each other and from the main reservoir polygon by DEM cells with sampled elevations that were higher than the respective reservoir surface water elevations. While at least a portion the ground covered by these "elevated" cells contained lower elevation streambeds, the sampled elevation for these cells was derived from the streambeds and higher ground beyond the banks of the streams when the DEM data were created (Fig. 6). Not only did this create the visually inaccurate "ponding" in Fig. 4, it was at least partly responsible for the discrepancy between the initial reservoir area and capacity estimates presented in this study and those presented in the final environmental impact statement (EIS) (U.S. Army Corps. of Engineers, 2000a) and those at the PTRWA Web page (http://www.ptrwa.org) (see Tables 1 and 4). Area estimates in the EIS were calculated manually using a planimeter to trace an interpolated reservoir boundary on a topographic map. Capacity estimates in the EIS were calculated from contours on topographic maps using the *average end area method*, a technique that is commonly used to calculate volumes in engineering applications.

One method tested to remove the "ponding" and improve area and capacity estimates was to apply a minimum spatial filter to the DEM data. Reservoir surface area and capacity estimates derived from the DEM after applying a 3 by 3 minimum filter (Table 4) were remarkably improved - much closer to those stated in the final EIS and on the PTRWA Web page than the original estimates. To really understand the filtering effects and to more accurately estimate the reservoir's area and capacity, better DEM data are required. Fortunately, better DEM data may be available soon. The USGS is currently creating higher resolution DEM data with a 10 by 10 m (x, y) resolution for the mountain and piedmont regions of North Carolina (http://mcmcweb.er.usgs.gov/status/mac/nc/nc_dem10.html). Furthermore, the State of North Carolina is creating high resolution DEM data derived

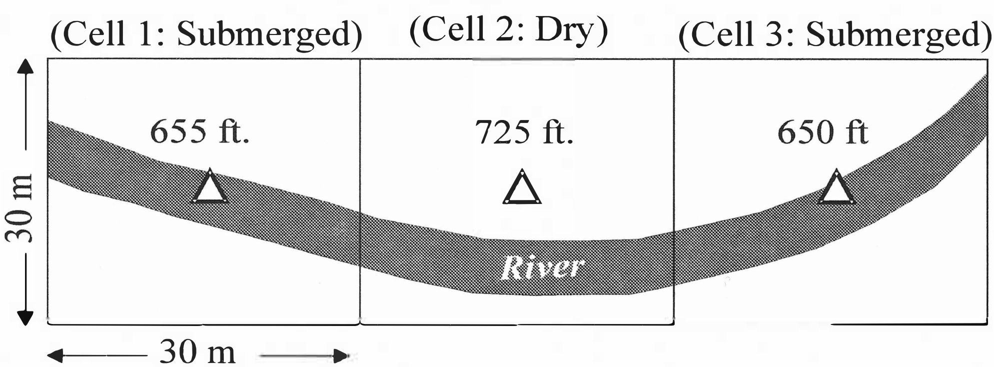

Figure 6. At a reservoir surface height of 709 ft., DEM cells 1 and 3 would be classified as submerged, but cell 2 would not be because its elevation sampled was higher than 709 ft.

Table 4. Reservoir capacities (acre-ft.) and reservoir sizes (acre) derived after applying a 3 by 3 minimum spatial filtering operation to the DEM data, reported in the environmental impact statement (EIS), and published at the Web page by the Piedmont Triad Regional Water Authority.

	After spatial filtering operation		In the EIS		At the Web site	
Surface Water Height (ft.)	**Capacity**	**Size**	**Capacity**	**Size**	**Capacity**	**Size**
@ 682	52,445	2,958	62,000	3,200	--	3,007
@ 706	162,709	7,035	160,000	6,200	--	--

from LIDAR (Light Detection and Ranging) data (Dorman, 2000).

Another potential problem involved the estimation of displaced population using census block population. The proportional allocation method used assumed that the population was evenly distributed throughout the census block. If the population was unevenly distributed in reality, then the estimates of the total displaced population could be inaccurate. While it was beyond the scope of this study, the only way to accurately estimate displaced population was to conduct a field enumeration of the population within the proposed reservoir boundary.

Concluding remarks

Digital spatial and demographic data sets have been used to study the impact of the Randleman Reservoir on the local environment and demography of Randolph and Guilford counties, N.C. At surface water heights of 682, 685, 706, and 709 ft. above the mean sea level, the reservoir's capacities were 25,053, 29,242, 92,654, and 106,654 acre-ft., and total surface areas (reservoir plus 200ft. erosion/pollution control buffers) 3,081, 3,516, 7,403, and 8,233 acres, respectively. The capacities and total surface areas were lower than those reported in the final environmental impact statement by the U.S. Army Corps of Engineers and those published at the Piedmont Triad Regional Water Authority's Web page. These underestimations were attributed to the inaccurate representation of the digital evaluation model (DEM) data used. After applying a 3 by 3 minimum spatial filter to the DEM data, the recomputed reservoir capacities and areal extents were very close to those reported and published as mentioned in the above. At the surface water heights of 682 and 706 ft., the recalculated capacities were 52,445 and 162,709 acre-ft., and reservoir (only) areas were 2,958 and 7,035 acres, respectively.

The most affected landuse and land cover types due to the construction of the reservoir were mixed upland hardwoods, managed herbaceous cover, and cultivated lands.

This study has demonstrated the potential of using geo-referenced spatial and demographic datasets once they have been integrated into a GIS. Also, these datasets are available to the public at little or no cost. With a general background in geographic information science and training in the use of remote sensing/GIS software, many users can carry out studies incorporating the spatial and demographic data sets relevant to their sub-fields, and can generate many eye-opening applications in the near future.

Acknowledgement

The authors wish to thank two anonymous reviewers whose comments have greatly improved the quality of the paper. This research was supported by the Center for Geographic Information Science, East Carolina University.

References

Black and **Veatch.** (1988). *Randleman Lake Safe Yield Analysis*, prepared for the Piedmont Triad Regional Water Authority.

——. (1990). *Water Quality and Quantity Studies to Support Randleman Lake Environmental Impact Statement*, prepared for the Piedmont Triad Regional Water Authority.

Clarke, K. C. (2001). *Getting Started with Geographic Information Systems*, 3rd edition, Prentice Hall, Upper Saddle River, New Jersey 07458.

Dorman, J. (2000). North Carolina floodplain mapping initiative, in *the Symposium Proceedings of Research in Support of Hazard Mitigation: Science in Service of Society*, North Carolina Emergency Management, held on October 30-31, 2000 in Raleigh, North Carolina. 179-182.

Jensen, J. R. (2000). *Remote Sensing of the Environment, an Earth Resource Perspective*, Prentice Hall, Upper Saddle River, NJ.

Lautzenheiser, L., Eastman, J., Seramur, K., and **Holm, M.** (1997). *Archaeological Survey, Proposed Randleman Reservior, Randolph and Guiford Counties, North Carolina*, prepared for the Piedmont Triad Regional Water Authority.

Moore, J. H. and **Leonard, S. W.** (1973). *Preimpoundment Studies; Randleman Project, A survey of the Vascular Plants*, Department of Environmental Sciences and Engineering, School of Public Health, University of North Carolina at Chapel Hill, ESE No. 323, 112 p.

U.S. Army Corps of Engineers. (1986). Accuracy of Computed Water Surface Profiles, Research 26, Hydrological Engineering Center, Davis, California.

U.S. Army Corps of Engineers at Wilmington District. (2000a). *Final Environmental Impact Statement, Randleman Lake, Guilford and Randolph Counties, NC*, 178 p.

U.S. Army Corps of Engineers. (2000b). HEC-GeoRAS An Extension for Support of HEC-RAS using ArcView, User's Manual Version 3.0. Hydrological Engineering Center, Davis, California.

Verbyla, D. L. (1995). *Satellite Remote Sensing of Natural Resources*, Lewis Publishers, New York, NY, 198 p.

Weiss C. M., Francisco, D. E., and **Lenat.** (1973). *Preimpoundment Studies; Randleman Project, Water quality characteristics*, Department of Environmental Sciences and Engineering and the UNC Wastewater Research Center, School of Public Health, University of North Carolina at Chapel Hill, ESE No. 327, 174 p.