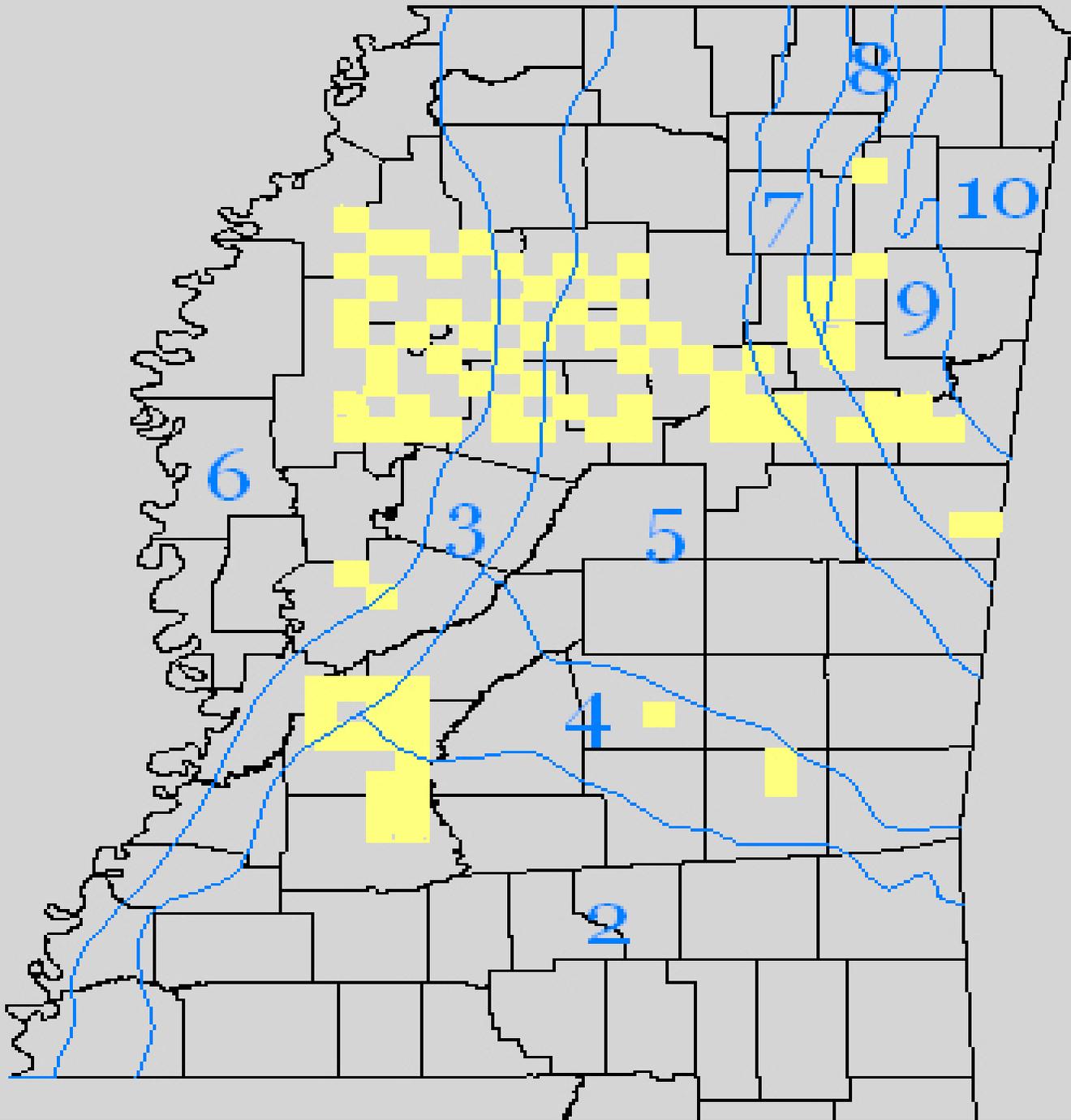




Using the General Land Office Records to Enhance Identification of Cultural Landscapes in Northern Mississippi | 2005-01

Mississippi State University



National Park Service
U.S. Department of the Interior

National Center for Preservation Technology and Training



Using the General Land Office Records to Enhance Identification of Cultural Landscapes in Northern Mississippi

Grant Number: MT-2210-O4-NC-04

Submitted by:

G. Wayne Wilkerson (PI), Assistant Professor

Robert F. Brzuszek, Assistant Professor

Paul Lanning, Graduate Assistant

**Department of Landscape Architecture
Mississippi State University
Mailstop 9725
Miss. State, Ms 39759**

December 9, 2005

Table of Contents

Executive Summary.....	Page 4
Introduction.....	Page 4
Methods.....	Page 8
Results and Discussion.....	Page 14
Conclusions.....	Page 15
Acknowledgements.....	Page 15
References.....	Page 16

List of Figures

Figure 1. GLO Survey districts for Mississippi.....	Page 5
Figure 2. Sample GLO record.....	Page 6
Figure 3. GLO record viewer.....	Page 8
Figure 4. Physiographic regions of Mississippi.....	Page 10
Figure 5. Mississippi townships and ranges.....	Page 11
Figure 6. Translated GLO townships	Page 14

Executive Summary

General Land Office Records (GLO) are the first federal surveys conducted in each state. They are used today to document the earliest verifiable recordings of the historic landscape, early cultural features, and the locations of vegetation types. These surveys have proven to be of immeasurable benefit to resource managers, archaeologists, restoration professionals, and historians. This study focused on converting GLO records for Mississippi, with an emphasis on capturing information about the type and characteristic of the witness trees used during the survey process.

The technical aspect of this study focused on how to best utilize optical character recognition (OCR) and geographic information systems (GIS) to analyze GLO vegetative records. OCR refers to the branch of computer science that involves translation of optically scanned images into forms such as ASCII codes that a computer can manipulate. The primary use of GIS in this project was to locate the converted records in the landscape.

The OCR software tested was unable to accomplish the conversion process as planned. The original GLO records had not been transcribed, and the OCR software tested could not be trained to efficiently convert the hand written files to digital format. All conversion work was accomplished manually. A total of 16,020 records were converted and loaded into an Excel spreadsheet. An attempt was made to link the records to X,Y locations in a GIS. The manner in which the records were entered by the surveyors made this difficult. Instead of X and Y locations, the records are organized by section, township, and range. The translated record set is broken into 89 sub-files, and these have been linked to a WEB page for easy dissemination.

Introduction

The value of using General Land Office records for assessing presettlement landscape in North America are well documented, as are the problems associated with interpreting the data (Bourdo 1956, Bragg 2003, Leahy, 2003, Manies, 1997). The GLO files were produced as part of the original land survey of new states and territories in the early 19th century. Mississippi's land plat records were created primarily during the period of 1832 and 1840 (Muhn, 1996).

In 1803, congress passed "An Act Regulating the Grants of Land, and Providing for the Disposal of the Lands of the United States, South of the State of Tennessee." This act provided for the establishment of two land offices in the Mississippi Territory – one east of the Pearl River and one west of the Pearl River. Mississippi was broken up into the eight cessions and districts (see figure 1). These included the Choctaw purchase of 1805, Choctaw purchase of 1820, Choctaw purchase of 1830, Chickasaw purchase of 1816, Chickasaw purchase of 1832, Spanish Acquisition of 1812, Old Natchez District, and Tombigbee district (Burt, 1973). The records obtained from these offices include field notes and surveys of the districts, tract books, land sales records, official monthly abstracts and various entry records.

Mississippi was surveyed from 1803 to 1855 by 157 deputy surveyors. The survey generals that were in charge of the deputy surveyors include Isaac Briggs from 1803 to 1807, Seth Pease from 1807 to 1810, Thomas Freeman from 1810 to 1821, Wailes, Davis, Turner from 1822 to 1830, Gideon Fitz from 1830 to 1834, and John Bell, Henry

S. Foote, Volney E. Howard, P. H. Fontaine, Benjamin A. Ludlow, Alex Downing, and Charles A. Bradford from 1834 to 1855 (Burt, 1973).

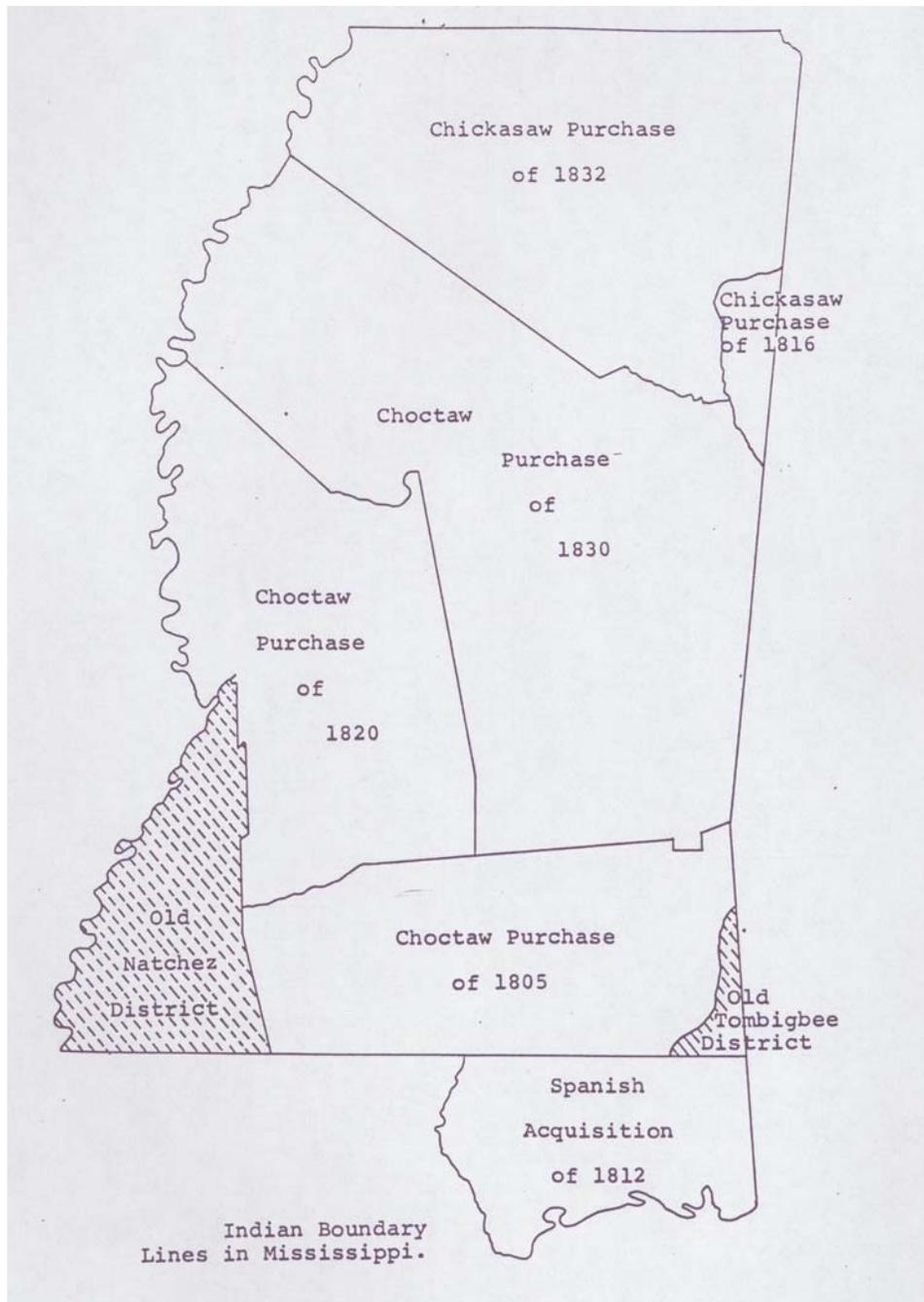


Figure 1. Survey Districts for Mississippi (Burt).

To conduct the land plats, deputy surveyors and their teams established rectangular section lines (one mile square) through the districts. Distances were measured along the section lines, and corner locations were established for each section. In addition to establishing locations, surveyors recorded information by hand in leather-bound field

books. The recorded information listed data about the land and its natural resources; including existing vegetation types, the locations of riparian features and other water bodies, soil types and suitability, various landforms, and evidence of cultural features. The microfilm copies are available for public access and are housed at the Secretary of State's office in Jackson, Mississippi.

Several states have already begun the conversion of the original records from paper and microfilm to digital format and have documented the process (Anderson, 1996). Mississippi's microfilm records were converted to digital format and became available for public access in 2003. Past studies provided valuable insight into the problems encountered for interpreting digital technology and resultant solutions. However, these successful processes are several years old and new technologies are being developed and improved constantly. The technical aspect of this study focused on how to best utilize improvements in optical character recognition (OCR) to convert the files to a usable format. Linking the vegetation records to spatial locations using GIS was also evaluated.

OCR refers to the branch of computer science that involves translation of optically scanned images into forms such as ASCII codes that a computer can manipulate. An OCR system enables the user to take hardcopy or digital data and edit it using computerized tools. Most OCR systems use a combination of hardware and software to recognize characters, although some inexpensive systems do it entirely through software. Advances in OCR technology now allow for the translation of a large variety of font types.

GIS is a combination of hardware and software tools for collecting, storing, analyzing, and outputting spatial data. GIS technology has made significant improvements in the last decade in terms of availability and ease of use. Once only available to computer scientists and programmers, data produced by these systems are readily available to researchers in many non-technical fields. The primary use of GIS in this project is to verify the results of the OCR conversion process. The descriptive locations will be converted to spatial coordinates to verify the accuracy of the records. Environmental Systems Research Institute (ESRI) GIS tools, primarily ArcGIS desktop, was used. One of the secondary goals of this project is to expand data availability through the integration of GIS and the Internet.

Status of Mississippi GLO Records

Initially, it was thought that the GLO records housed at the Mississippi Secretary of State's (SOS) office in Jackson, MS had been transcribed and were in digital format. This was not the case. The records had been scanned, but were still the original hand written note form produced by the surveyors. The data was provided to the Mississippi State University Department of Landscape Architecture via fifteen compact disks (CDs). The disks contained 67,464 sheets in .TIFF files. A sample of one of the original sheet is shown in figure 2.

T. 19 R. 1 West

24 2 Due West Between Sects 23, 24, 25, 26
 3 10 out of cane into a flat slough
 6 Chs wide 30 Chs to high land
 10 16 Set 1/4 Sect post. witness
 Red Elm 4 in Diam N 20° W 1 1/2
 Rose Elder 1 1/2 in Diam S 87° E 1 1/2
 80 12 To Cor. Sects 23, 23, 26 & 27
 See bearings on line Return
 Sects 23 & 24 first 1/2 of this
 illuminated 6 to 10 ft. witness
 high land soil 1/2 acre timber.

Figure 2. Sample GLO record found in the Mississippi archives.

The General Land Office field notes for Mississippi contain many different variables such as a title page, maps, and the surveyor's notes for that particular range. The title page contained the township number and direction, the range number and direction, names of the surveyors who surveyed the east, west, south, and north boundary, as well as the interior section lines. The next page is usually the surveyor's description of when the survey took place, followed by his signature, followed by a map of the land drawn by the surveyors. The majority of the pages were dedicated to the surveyor's field notes. On these pages the surveyor would record the section line that they were surveying, length in chains and feet to important land marks along the line, and an overall description of the land around them. When they were half way between one section corner and the next they would measure distances and bearings to witness or bearing trees and then they would repeat this process when they reached the next section's corner. They would normally measure two trees at the midway point and four trees at the corner. The last page of each township is where you would find an affidavit by which the surveyor swore to have done his work properly and in compliance with the terms of his contract.

Accessing the Original Records

The original records provided by the office of the Mississippi Secretary of State (SOS) were not associated by name or folder to spatial locations such as township and range. Each of the 67,464 files was merely assigned a number. In order to access the files the SOS developed a computer interface that allowed the viewer to select a GLO records using a series of options including district, township and direction, and range and direction. The district option included St. Stephens, Washington, Choctaw, Huntsville, and Chickasaw. An image of the viewer interface is shown in figure 3.

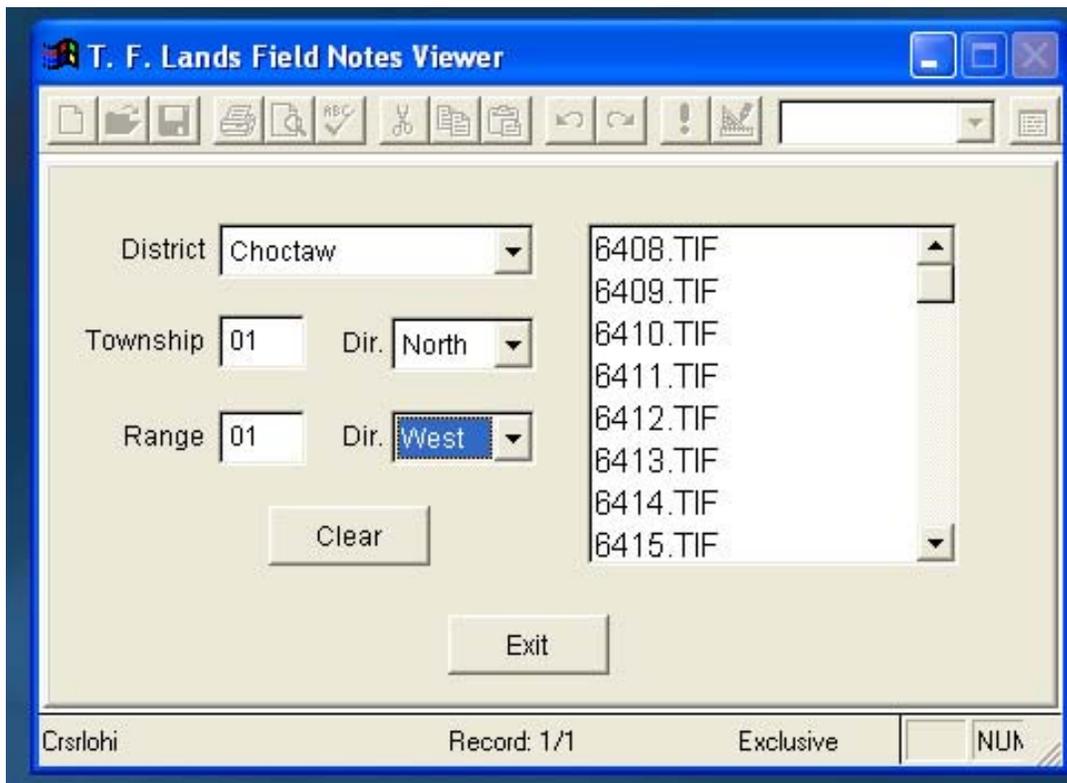


Figure 3. Image of the viewer used to select GLO records.

Methods

Conversion using OCR Software

The scanner software tested for this project was ABBYY FineReader 6.0. It was selected based on cost and product reviews. Several small operational issues were encountered when installing and setting up the software, but these were minor. The major difficulty became apparent when the first attempts were made to convert the GLO records. As the example in the introduction shows, the original surveyor notes were often very difficult to read. The writing method of the day was a very elaborate cursive style. The OCR software tested is best suited to reading text that has been typed. The cursive font styles in the GLO documents require extensive training of the OCR software. If the scanning work is typed, the software recognizes the letters with little coaching. However, this was not the case with the Mississippi GLO records. During the first attempts to read entire documents, the software could not recognize the hand

written fonts, so it read and recognized each document as a mixture of symbols, numbers, and letters. This solution proved to be unproductive.

An attempt was then made to train the program to read a single word. The word chosen was "oak" which occurred 11 times in the initial GLO document. There are two ways to train the program, letter by letter or whole words at a time. Initially, training was tried letter by letter. The way to train the program to read and recognize a word is to put a box around the word, change the box from a recognition box to a text box, and then read the "block" as it is called. After training the program one time, it read 6 out of 11 words correctly. After training the program twice, it read 10 out of 11 correct. To train the program this way you had to put a box around each word you wanted the program to read; therefore, there were 11 boxes on the screen corresponding to the 11 times that "oak" appeared in the document. This is not an efficient means to read the documents. The total amount of time spent on this OCR conversion process was between 4 to 5 days.

A third approach used was to expand the box to include more than one word. The recognition box encompassed eight words, four "oak" and four "white". Recognition by the software went down considerably when using multiple words. Approximately five days was spent using this approach. Little progress had been made in two weeks of training. A different approach was needed.

Conversion using manual methods

It was decided by the PI's that the OCR approach would never produce conversion results on a scale needed to satisfy the initial objectives of the project. A graduate student began to translate the records manually. The data was entered into an Excel spreadsheet. The variables identified to enter into the database, if they existed in the records, included the following:

Survey date: Most of the surveys were dated between 1821 and 1836.

Sheet number: Sheet number find on bottom of GLO record.

Between sections: This identified the location of the section line.

Range: Self explanatory.

Township: Self explanatory.

Survey direction: The direction of the survey party.

Length in chains: Distance to the location of the witness trees.

Tree type: Witness tree type.

Tree diameter: Sometimes in links and sometimes in inches.

Bearing: Bearing from the survey line.

Surveyor: Name of Surveyor.

Purchase location; Survey district such as Choctaw or Chickasaw.

Purchase date: Most of the purchase dates were 1820 or 1830.

Surveyor comments: Observations regarding land quality, vegetative cover, cultural features.

Not every GLO had all the variables identified above. The following is an example of what one row in the database will contain: survey date - **1832**, sheet number - **10**, Map associated with group, between sections – **5,6**, range – **3 E**, township - **19 N**, survey direction - **South**, length in chains – **38.17**, tree type – **Post Oak**, tree diameter – **20 in**, bearing – **S 80, W 25**, surveyor - **Henry**, purchase location - **Choctaw**, purchase date - **1830**, surveyor comments – **Land poor and hilly oaks and pines**. Also, other

variables that were included in the database were the date the notes were transferred, what time the conversion began and ended, and what condition the field notes were in (easy or hard to read). These last two variables were for internal use only, and were not included in the records linked to the WEB page for public dissemination.

Selecting the locations of the translated records

The proposal called for translating GLO records for twenty four counties in northeast Mississippi. This proved to be impossible due to the fact that the dataset provided by the Mississippi Secretary of State's office did not contain records for this portion of the state. Once the project switched focus from using OCR to manual translation, the first records in the dataset were used as a test. These records happened to fall in the central portion of the state around the Jackson area, and were associated with the Choctaw Purchase of 1820. Once the translation process began in earnest, the focus switched back to the northern half of the state. The Choctaw Purchase of 1830 was the district with the records easiest to decipher. At this point it was decided to let the physiographic regions of the state dictate which GLO records were selected. Mississippi is comprised of 10 regions, with the greatest diversity in the northern half of the state (figure 4). The goal was to have GLO records from as many regions as possible.

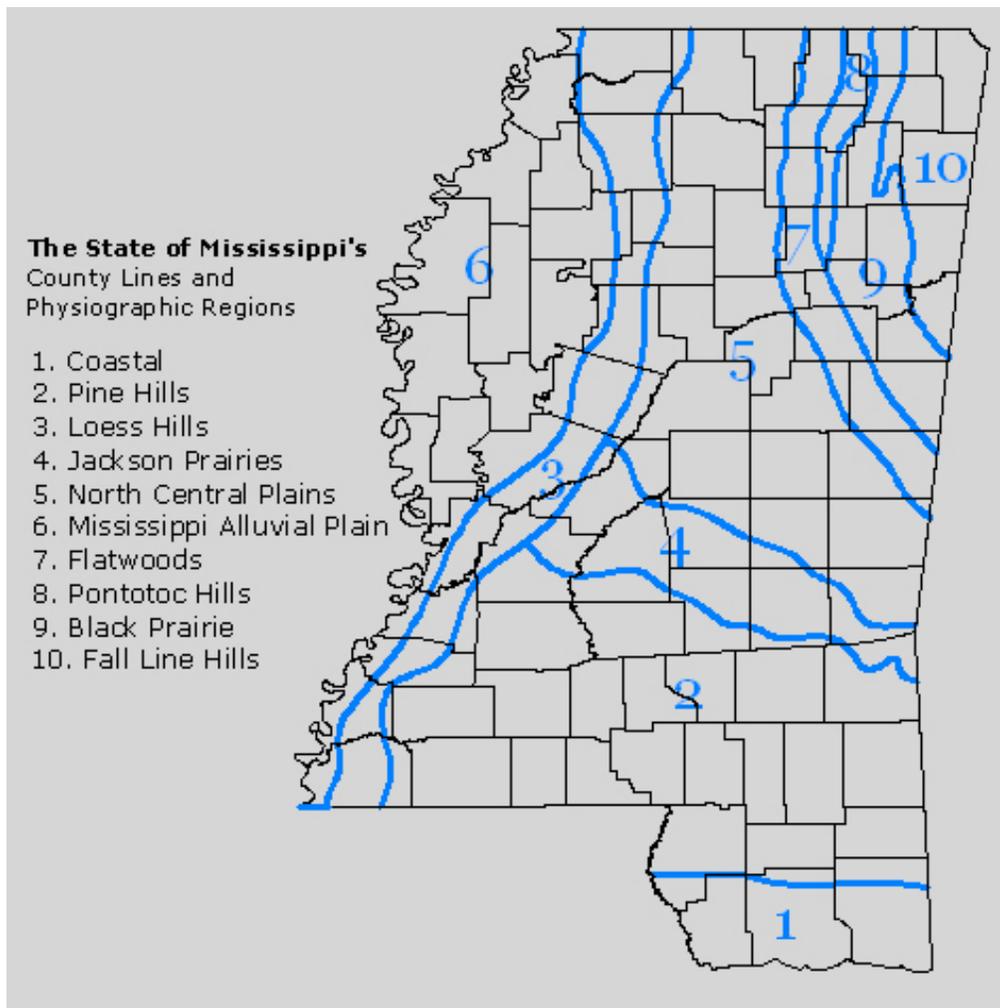


Figure 4. Physiographic regions of Mississippi

Linking records to GIS points

The GLO records contained spatial information that was defined by areas and lines. As described above the original records were tied to section, township, and range, and included the direction and distance along that survey line to the location of the witness trees. However, the records contained no easting/northing or latitude/longitude point locations. This was need to in order to locate the points within a GIS.

The Public Land Survey records for Mississippi are available through the Mississippi Automated Resource Information System (MARIS) Technical Center website. MARIS is a service funded by the Mississippi state government to provide digital natural and cultural data to the public at no cost. One file was downloaded from MARIS, the Sections theme, which included attributes such as Survey District, Section, Township, and Range numbers (figure 5).

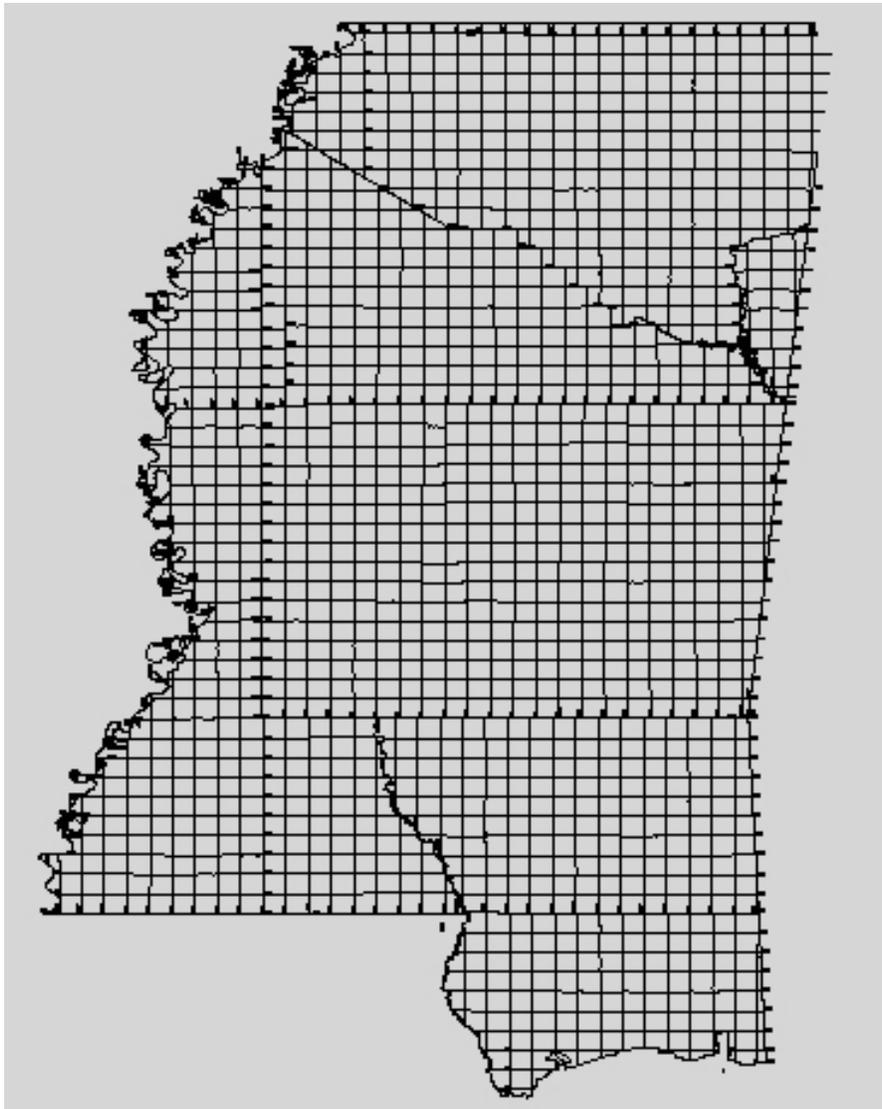


Figure 5. Mississippi Townships and Ranges

Two scripts, Ianko Tchoukanski's Easy Calculate and Shapefile Converter by Robert Scheitlin, were used to establish X, Y coordinate values for the section corners. Both scripts were found on the WEB at ESRI's Support Center Knowledge Base. Finding the X,Y coordinates of the points within ArcMap was extremely important but the primary task was to find a way to locate the witness trees off of these points. Several days were allocated to finding conversion routines on the ESRI site, but without success. It was decided to attempt to solve the problem using customized scripts within the database. The township and range columns for each record were structured with a number and a direction such as Township 1W and Range 7N. These were converted to four columns that looked like Township – "1", Township Direction – "W", Range – "7", and Range Direction – "N".

Next, the "between section line" column in Excel also had to be populated. For example, the combination 35,36 had to be associated with a point for section 35 because the line that separates section 35 and 36 in the GIS starts at point 35 and goes north. This step was repeated for all the records. It was important to convert the between section column into a single point column because this was needed to match the GIS table. Unfortunately we could not find a way to link the GIS table, with the spatial information, to the Excel database with the converted GLO records. Next, the length in chains column was converted into feet. A column was added containing length data in surveyor chains. This column's numbers were found by taking length in chains and multiplying it times 66 (*1 chain = 66 feet*) which converted the distance to feet.

Basic trigonometry was used to locate the points. By using the basic equation $\sin \text{ of an angle} = \text{opposite/hypotenuse}$ we could plug in our angle and the length of the hypotenuse, solve the equation, and find the opposite's line's length. To get the numbers for the Point X column the following formula was used:

```
=IF(K4=$A$233,H4,IF(K4=$A$234,H4,IF(K4=$A$235,SUM(M4,H4),IF(K4=$A$236,M4-H4,"")))).
```

The formula states that if the letter in the survey direction column equals "N", then the X coordinate stays the same, if the letter in the survey direction column equals "S", then the X coordinate stays the same, if the letter in the survey direction column equals "E", then add the number in the "dist in feet" column to the X coordinate, and if the letter in the survey direction column equals "W", then subtract the number in the "dist in feet" column from the X coordinate. To get the numbers for the Point Y the following formula was used:

```
=IF(K4=$A$233,SUM(M4,I4),IF(K4=$A$234,I4-M4,IF(K4=$A$235,I4,IF(K4=$A$236,I4,"")))).
```

The formula states that if the letter in the survey direction column equals "N", then add the number in the "dist in feet" column to the Y coordinate, if the letter in the survey direction column equals "S", then subtract the Y coordinate from the number in the "dist in feet" column, if the letter in the survey direction column equals "E", then the Y coordinate stays the same, and if the letter in the survey direction column equals "W", then the Y coordinate stays the same.

The next step was to take the bearing information that was in one column and break this into three separate columns. Originally the column would read N64, W24 but in three columns it would read Cardinal Direction – “NW”, Bearing – “64”, and Distance in links- “24”. A column for distance in feet was established and took the number in the distance in link column and multiplied it times .66 (*1 link = .66 feet*) to get the numbers into feet with the following formula =R4*0.66. The next step was to find the radians for the angles in the number bearing column by using the radians function in Excel =RADIANS(Q4). This step was required since the radians of the angle was needed for insertion in the sin formula in Excel. The next column, A Distance, was calculated by taking the sin of the radians column times the number in the dist in feet column (=SIN(T4)*S4). This provided the lengths need to input into a later formula. To find the other length the angle in the number bearing column was subtracted from 90 to get the complementary bearing (=90-Q4). This was repeated to find the radians of the new angle by using the formula: (=RADIANS(V4)). Then the sin of the radians was determined and multiplied time the dist in feet with the formula (=SIN(W4)*S4). This formula produced the other length needed to input into the next two formulas. The next two formulas that we would provide the final X,Y coordinates for the witness and bearing trees that were contained in the database.

To get the final X coordinate the following formula was used:

=IF(P4=\$B\$235,SUM(U4,N4),IF(P4=\$B\$236,N4-U4,IF(P4=\$B\$233,SUM(U4,N4),IF(P4=\$B\$234,N4-U4,IF(P4=\$B\$237,SUM(U4,N4),IF(P4=\$B\$238,N4-U4,""))))))

The formula states that if the letters in the Cardinal direction column equal “SE”, then add the distance from column U (A DISTance) to the number in column N (Point X), if the letters in the Cardinal direction equal “SW”, then take the number in column N and subtract it from the number in column U, if the letters in the Cardinal direction equal “NE”, then add the distance from column U to the number in column N, if the letters in the Cardinal direction equal “NW”, then take the number in column N and subtract it from the number in column U, if the letter in the Cardinal direction equals “E”, then add the distance from column U to the number in column N, if the letter in the Cardinal direction equals “W”, then take the number in column N and subtract it from the number in column U.

To get the final Y coordinate the following formula was used:

=IF(P4=\$B\$233,SUM(X4,O4),IF(P4=\$B\$234,SUM(X4,O4),IF(P4=\$B\$235,O4-X4,IF(P4=\$B\$236,O4-X4,IF(P4=\$B\$237,O4,IF(P4=\$B\$238,O4,""))))))

The formula states that if the letters in the Cardinal direction column equal “NE”, then add the distance from column X (A DISTance) to the number in column O (Point Y), if the letters in the Cardinal direction column equal “NW”, then take the number in column X to the number in column O, if the Cardinal direction column equals “SE”, then subtract the number in the Point Y column from the number in the A DISTance column, if the Cardinal direction equals “SW”, then subtract the number in the Point Y column from the number in the A DISTance column, if the Cardinal direction column equals “E”, then take the number in the Point Y column, and if the Cardinal direction column equals “W”, then take the number in the Point Y column.

Results and Discussion

GLO record conversion

Figure 6 below indicates the locations of the 16,020 converted records. A total of 89 townships were converted. More records for the Chickasaw Purchase of 1832 would have been converted, but very few files were included in the SOS data set. As a result the majority of files converted are in the Choctaw Purchase of 1830. They extend across 22 counties and eight of the ten physiographic regions of the state.

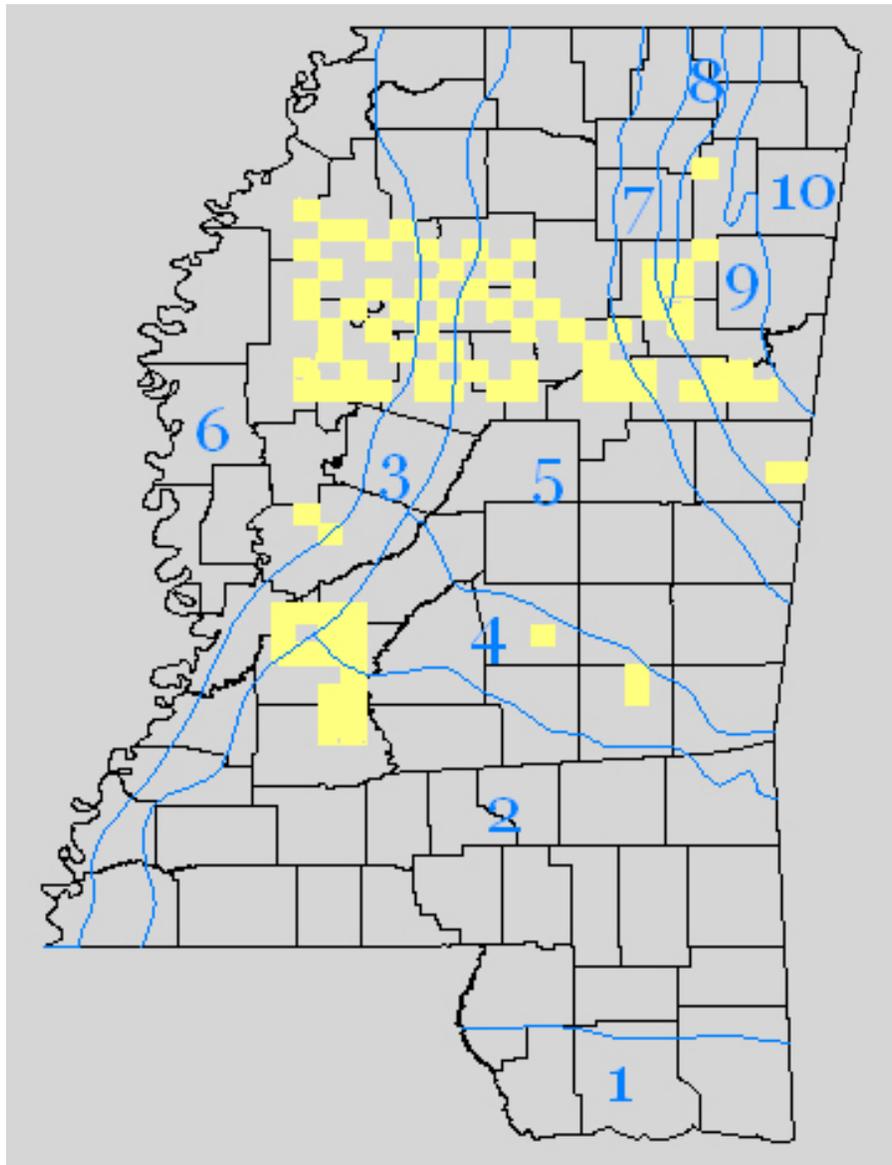


Figure 6. Completed Townships and Ranges by Physiographic Region

OCR conversion software

The software tested was unable to be trained to convert the handwritten records. Multiple factors impacted the unsuccessful conversion attempt using OCR. One is the

nature of the handwritten script used at that time. Another is the quality of the .TIFF files provided. A third is the fact that multiple surveyors were involved, which would have required re-training the software for each new individual. It appears at this time that the current level of OCR software will be of limited use for translating the remaining Mississippi records, or any GLO records for that matter. However, it should be noted that over 25% of the total GLO database was translated by traditional methods. This is a significant contribution to the knowledge base for researchers interested in cultural and biological conditions as they existed in the early 19th century.

Locating the GLO records spatially

An attempt was made to link the records to specific X and Y coordinates. As described above, this may still be possible. However, during the course of the project other research indicated this level of precision was not necessary. A Wisconsin study (Hong, 2000) also used GLO records to construct forest landscapes that predate European settlements. This study aggregates the GLO at the section level for use in creating species indices. The work produced here also aggregates the records at the section level. This should prove sufficient for future cultural and biological research.

Disseminating the translated records

One of the secondary goals of this project was to share the converted GLO information with other researchers and interested parties. The most efficient method was thought the Internet. The records were organized by township and range and stored in zipped Excel™ database files. A total of 89 sub-files were created. They are linked to the website <http://www2.msstate.edu/~gww/glo2/index.html>.

Conclusions

The translated GLO records will be of great use to geographers, anthropologists, and forest managers for their ability to recreate the landscape as it appeared to the first settlers in the 1800's. The ability to place the records using easting/northing values will also be useful, and can probably be accomplished. This research team will continue to work on the problem. Even if the witness trees cannot be placed individually, researchers can still use the aggregated information at the section level. The surveyor comments will also be useful, even at the section level.

Acknowledgements

This research project was developed under a grant from the National Park Service and the National Center for Preservation Technology and Training. Its contents are solely the responsibility of the author and do not necessarily represent the official position or policies of the sponsors listed above

References

Anderson, Paul F. 1996. GIS Research to Digitize Maps of Iowa 1832-1859 Vegetation from General Land Office Townships Plat Maps. Department of Landscape Architecture, Department of Agronomy, Iowa State University. Submitted to Iowa Department of Natural Resources.

Bourdo, E. A. 1956. A review of the General Land Office Survey and of its use in quantitative studies of former forests. *Ecology*, **37**:754–768.

Bragg, Don C. 2003. Natural Presettlement Features of the Ashley County, Arkansas Area. *American Midland Naturalist*, **149**:1-21.

Burt, Roffie. 1973. The Survey of Mississippi's State, Indian, and Township Boundaries, Department of Civil Engineering, Mississippi State University, Starkville.

He, Hong S., Mladenoff, David J., Sickley, Theodore A. and Guntenspergen, Glenn G. 2000. GIS interpolations of witness tree records (1839-1866) for northern Wisconsin at multiple scales. *Journal of Biogeography*, **27**:1031-1042 .

Leahy, Michael J.; Pregitzer, Kurt S. 2003. A Comparison of Presettlement and Present-day Forests in Northeastern Lower Michigan. *American Midland Naturalist*, Vol. **149**:71-90.

Manies, K.L. 1997. Evaluation of general land office survey records for analysis of the northern Great lakes hemlock-hardwood forest, MSc Thesis, University of Wisconsin, Madison. WI.

Muhn, James. 1996. A Brief History of the Disposition and Administration of the Public Domain in Mississippi. U.S. Department of the Interior, Bureau of Land Management.