GIS Data Layer Development

Final Report

June 2004

Project Manager: David D. Diamond
Director
Missouri Resource Assessment Partnership (MoRAP)
4200 New Haven Road
Columbia, MO 65201
Phone: 573-876-1862
Fax: 573-876-1863
David_Diamond@usgs.gov

Contacts: Kathryn Riesenberg
GIS Technician
573-441-2987
kriesenberg@usgs.gov

Melissa Lanclos
GIS Technician
573-441-2795
mlanclos@usgs.gov

Diane True
GIS Coordinator
573-441-2794
truecd@missouri.edu

Michael Morey
Database Manager
573-441-2790
Moreyme@missouri.edu

MoDOT Contact: George Kopp
CADD Support Center
200 Harrison Street
P.O. Box 270
Jefferson City, MO 65102
Phone: 573-751-7886
Fax: 573-526-4535
koppg@mail.modot.state.mo.us

Table of Contents

List of Figures

List of Tables

Abstract

The projects discussed in this report are a result of coordination between the Missouri Department of Transportation (MoDOT) and the Missouri Resource Assessment Partnership (MoRAP) staff. Three very distinct projects were defined and created based on agency needs. These projects were (1) development of an archaeological database prototype for the state of Missouri, (2) an update of metropolitan public lands data for major cities in the state, and (3) improving the SEMA data layer MoDOT currently has in-house.

The Missouri Department of Transportation (MoDOT), State Historical Preservation Office, Department of Natural Resources (SHPO-DNR), and (MoRAP), are working together to create a geodatabase which will represent known areas of historic significance in the state of Missouri. This project is the first of its kind in the state; therefore the early phases were focused on methodology development. Initially MoRAP upgraded and appended the SHPO Access database and used this format for capturing archaeology survey attributes. Surveys are areas that have been surveyed for archaelogical sites, generally in locations where human infrastructure development is slated to occur. Sites are identified locations of historical or cultural significance. ArcMap was used to create two shapefiles, one of surveys and one of sites, taken from photocopies of 1:24,000 topo quads with hand-drawn survey and site polygons. We have since decided to capture survey and site data using a geodatabase using SDE and an SQL relational database server. This way digitizing of polygons and attributing of data can occur nearly simultaneously, and the product is portable and supports multiple users. The end goal is to have digital layers of archaeology sites and surveys statewide, including attributes of the hard copy survey reports and site forms. At this time surveys are being attributed with fields determined by MoDOT and SHPO within St. Charles County. Our prioritization of counties to be digitized states that first the St. Louis area will be completed, followed by the Kansas City area, then the I-70 corridor, and on from this mid-state area.

Metropolitan areas in the state of Missouri are growing rapidly. Urbanization trends indicate that population density is decreasing in the cores of cities and increasing in the outlying areas. This growth has led to the need for more infrastructures within the metropolitan areas. MoDOT and MoRAP are working together to determine the location of the publicly owned areas within the cities of St. Louis, Kansas City, and Springfield, MO. These were the initial metropolitan areas selected based on their size and the fact that they are currently the fastest growing cities in the state. Data were collected for each area and merged together to create a seamless coverage of the lands for each. Further work can be done in the future to collect and reconcile public lands data for other cities within the state of Missouri.

MoDOT currently has in-house SEMA buy-out lands for the state of Missouri. These data have been put into shapefile format and given spatial relevance. However, some of the addresses for the data points were not accurate enough to give it an exact, unique location; instead a general point was assigned. These points were run through a

geocoding function in ArcMap in order to attempt to find a more exact spatial location for each. Many of the addresses were found in the TIGER road files and assigned a more exact placement in the shapefile.

Introduction and Goals

MoDOT has been a member of MoRAP from inception in 1995, and has been a key state agency supporter. The overarching goals behind MoDOT's participation include the production of needed GIS and remote sensing information at low cost and the facilitation of coordination and cooperation among key MoDOT partners. Data layers produced cooperatively have wide buy-in from multiple agencies, which provided increased credibility, and funds are leveraged so that costs are shared among partners, therefore saving MoDOT money.

The proposed activities listed here are based on discussions among MoDOT and MoRAP staff beginning in 2002, including at least two formal meetings involving multiple staff on October 25, 2002, and on April 16, 2003.

The goals of this project are to:

1) Produce an Archaeological database based on the Cultural Resources data housed at the State Historic Preservation Office (SHPO). This will include a Geodatabase of the information and data capture of the survey and site polygons for each archaeology report.
2) Create an updated public lands data layer for several major metropolitan areas in the state of Missouri.
3) Improve the spatial accuracy of the existing SEMA data currently in-house at MoDOT using the Geocoding function of ArcMap and the TIGER road files.

A summary of activities on these three goals follows in three separate sections.

Archaeology Database

MoDOT Contact: Bob Reeder
MoRAP Contact: Kathryn Riesenberg

Background and Goals

For a number of years MoDOT, DNR, and the University of Missouri–Columbia have been discussing the usability and necessity of a digital representation of archaeological surveys and sites as an aid to cultural resource management and planning. After some months spent designing a methodology for data capture that is a best fit for end-users and a quick, streamlined, production oriented procedure, three counties have now been completely entered and checked for QAQC in a relational geodatabase. Our original

goals were to form an interagency committee to guide database development and to design the database. Not only has a relational database been created, but data from three counties have been captured, production gets faster and more efficient every day as workers become more familiar with the software and hard copy survey reports and site forms from which data are collected. The end goal of this project is to have a manageable and up-to-date digital representation of all archaeological surveys and site locations that have been and are being registered at SHPO. At the completion of this year's work, (due date June 30, 2004), the following counties have been completed and checked for QAQC:

Atchison
DeKalb
Jackson (in progress)
Knox (in progress)
Newton (in progress)
Oregon (in progress)
Osage
St. Charles
St. Clair
St. Louis (in progress)

Counties have been ranked in order of priority, with highest priority being on the St. Louis and Kansas City areas, followed by the Interstate 70 corridor, then areas of municipality, and finally ending with the more rural parts of the state (Figure 1). Some counties with a low priority ranking have been completed in the first year of this project because SHPO has allowed datasets for these counties to leave their building for short time periods, which has allowed us to capture data on-site at the MoRAP offices.

General Methods and Procedures

This is an unusual project for MoRAP in that data capture has been carried out by employees at the SHPO building in Jefferson City. Two employees have been traveling to SHPO an average of two days a week since February, 2004. In addition, student workers have been working on selected rural counties at the MoRAP offices at least 24 hours per week since May. We have also purchased two new high-end laptop computers to ensure hardware availability for workers who travel to SHPO to collect data.

The data capture process is time consuming due to the complexity and variability in survey reports. Each report must have a topo map with survey and site polygons superimposed to be included in the database, then the narrative, or text, section of the reports must be perused for the information required to attribute each survey polygon.

Before data capture can begin, the database must be “checked out” to the MoRAP laptops while connected to the SDE server at the USGS facility in Columbia. Upon arrival in Jefferson City, workers select a box of survey reports and start collecting data. The first step is to digitize a survey polygon in the correct spatial location using an ArcMap

project specially built to help us locate the surveys and sites. Layers include but are not limited to digital raster graphics, roads, municipalities, stream networks, railroads, and county and quadrangle boundaries (Figure 2). After the polygon is digitized, the related tables (survey statistics, secondary author, methodology, project type, and related sites), can be checked for accuracy and filled in by culling through the narrative included in the survey report (Figure 3). These attributes are then linked to the survey polygon and related site polygons are digitized. The initial focus of the project team has been to move forward to capture survey and site polygon information and to attribute survey polygons. The team has not yet settled on what site attributes to capture, so only site polygons and numbers are currently captured.

The Cultural Resource Management Geodatabase

The geographic information system (GIS) and tabular data for the Cultural Resource Management Project are stored and managed in a geodatabase using a spatial database engine (SDE) on a structured query language (SQL) server. Relationship classes are used to relate (link) tables using a common identifier.

The geodatabase consists of features, data tables, relationship classes and domains. Survey and site boundaries are digitized features. They are related (linked) to tables. Currently, there are tables for reporting secondary authors, project types, project methodologies and general survey information. Relationship classes determine how tables are linked to the digitized features. Domains are lists of items, such as project types or methodologies. Domains offer the person entering data a list of choices. Clicking on an item enters it into the database.

Tables

Surveys Table

The Surveys table is used to hold information concerning a survey area and project report. It is the main table to which most other tables are linked. Currently this table is maintained independently from the Survey Polygons. This table contains legacy data from a Microsoft Access database that had no spatial coordinates. Once all the survey reports for the State have been digitized, this table will be permanently attached to the Survey Polygons.

Field Name: ObjectID
Data Type: Integer
Definition: A unique number used by the geodatabase system to identify a record.

Field Name: New_Edit
Data Type: text
Definition: New = new survey record; Edit = existing survey record that has been edited.

Field Name: SHPO_ID
Data Type: text
Definition: A unique identifier given to a proposed project by the SHPO staff.

Field Name: Survey_ID
Data Type: text
Definition: A unique identifier given to a survey.

Field Name: Date
Data Type: integer
Definition: Year the survey was conducted.

Field Name: Arch_Site
Data Type: text
Definition: Yes = a site of cultural or historical significance is present in the survey area; No = a site is not present.

Field Name: Author
Data Type: text
Definition: The primary author of the report.

Field Name: Title
Data Type: text
Definition: The title of the report.

Field Name: Cndctd_By
Data Type: text
Definition: The agency, business or organization that conducted the survey.

Field Name: Cndct_For
Data Type: text
Definition: The agency for which the survey was conducted.

Field Name: Comments
Data Type: text
Definition: A field used for miscellaneous information.

Survey Polygons Table

Most of the fields in the Survey Polygons table are used by the geodatabase system to define and draw the outline of the survey area. This is the attribute table of the survey polygon features.

Field Name: ObjectID
Data Type: Integer

Definition: A unique number used by the geodatabase system to identify a record.
Field Name: Survey_ID
Data Type: text
Definition: A unique identifier given to a survey. This field links the survey polygon table to the Surveys table.

Methodology Table

The Methodology table is used to identify the methods used to explore a survey area. It is a separate table because it has a one-to-many relationship with the Surveys table.

Field Name: ObjectID
Data Type: Integer
Definition: A unique number used by the geodatabase system to identify a record.

Field Name: Survey_ID
Data Type: text
Definition: A unique identifier given to a survey. This ID links the Methodology table to the Surveys table.

Field Name: Method_ID
Data Type: integer
Definition: This ID links the Methodology table to Methodology domain. It identifies the method(s) used to look for culturally or historically significant sites in a survey area.

Project Types Table

The Project Types table is used to identify the purpose(s) of the survey. It is a separate table because it has a one-to-many relationship with the Surveys table.

Field Name: ObjectID
Data Type: Integer
Definition: A unique number used by the geodatabase system to identify a record.

Field Name: Survey_ID
Data Type: text
Definition: A unique identifier given to a survey. This ID links the Project Types table to the Surveys table.

Field Name: Prjct_Type_ID
Data Type: integer
Definition: This ID links the Project Types table to the Project Types domain. It identifies the type or purpose of the survey.

Secondary Authors Table

The Secondary Authors table is used to give the names of any additional authors of a report. It is a separate table because it has a one-to-many relationship with the Surveys table.

Field Name: ObjectID
Data Type: Integer
Definition: A unique number used by the geodatabase system to identify a record.

Field Name: Survey_ID
Data Type: text
Definition: A unique identifier given to a survey.

Field Name: Authors_Other
Data Type: text
Definition: The name(s) of additional authors of the report.

Site2Survey Table

The Site2Survey table is used to list any sites of cultural or historical significance to the survey area. It is a separate table because it has a one-to-many relationship with the Surveys table.

Field Name: ObjectID
Data Type: Integer
Definition: A unique number used by the geodatabase system to identify a record.

Field Name: Survey_ID
Data Type: text
Definition: A unique identifier given to a survey.

Field Name: Site_ID
Data Type: text
Definition: A unique identifier given to a site of cultural or historical significance.

Sites Polygon Table

Most of the fields in the Sites Polygons table are used by the geodatabase system to define and draw the outline of the sites area. This is the attribute table of the sites polygon features.

Field Name: ObjectID
Data Type: Integer
Definition: A unique number used by the geodatabase system to identify a record.

Field Name: New_Edit
Data Type: text
Definition: New = new site record; Edit = existing site record that has been edited.

Field Name: Site_ID
Data Type: text
Definition: A unique identifier given to a site of cultural or historical significance.

Domains

Methodology Domain

The Methodology domain is used by the Methodology table to identify the method(s) used to explore a survey area. It appears as a drop-down list when using the geodatabase. Possible values are auger probe; excavation; interviews; literature search; mechanical stripping; pedestrian; photographic analysis; plowed; raking; shovel probe; surface collection; test pit; transects; trenching; and screening.

Project Type Domain

The Project Type domain is used by the Project Type table to identify the purpose(s) of the survey. It appears as a drop-down list when using the geodatabase. Possible values are archaeology; architecture; event location; phase II; and phase III.

New_Edit Domain

The New_Edit domain is used by the Surveys and Sites Polygons tables. It appears as a drop-down list when using the geodatabase. Possible values are new and edited.

Yes_No Domain

The Yes_No domain is used the Surveys table. It appears as a drop-down list when using the geodatabase. Possible values are yes and no.

Table Relationships

The Surveys table has a one-to-many relationship with the Survey Polygons, Methodology, Project Types, Secondary Authors and Site2Survey tables. This means there is one record per survey in the Surveys table, but there can be one or more records related to the survey in each of the other tables. The tables are linked via the Survey_ID field.

Results

Five counties have been completely entered into the geodatabase and checked for QAQC and five more are in progress at the time of this writing. The team has completed more than was envisioned based on goals outlined in the original proposal, thanks to diligent and energetic efforts of all of the members. The geodatabase will be a useful tool for storing data, accessing data, providing information for researcher, and quickly updating data sets in the future.

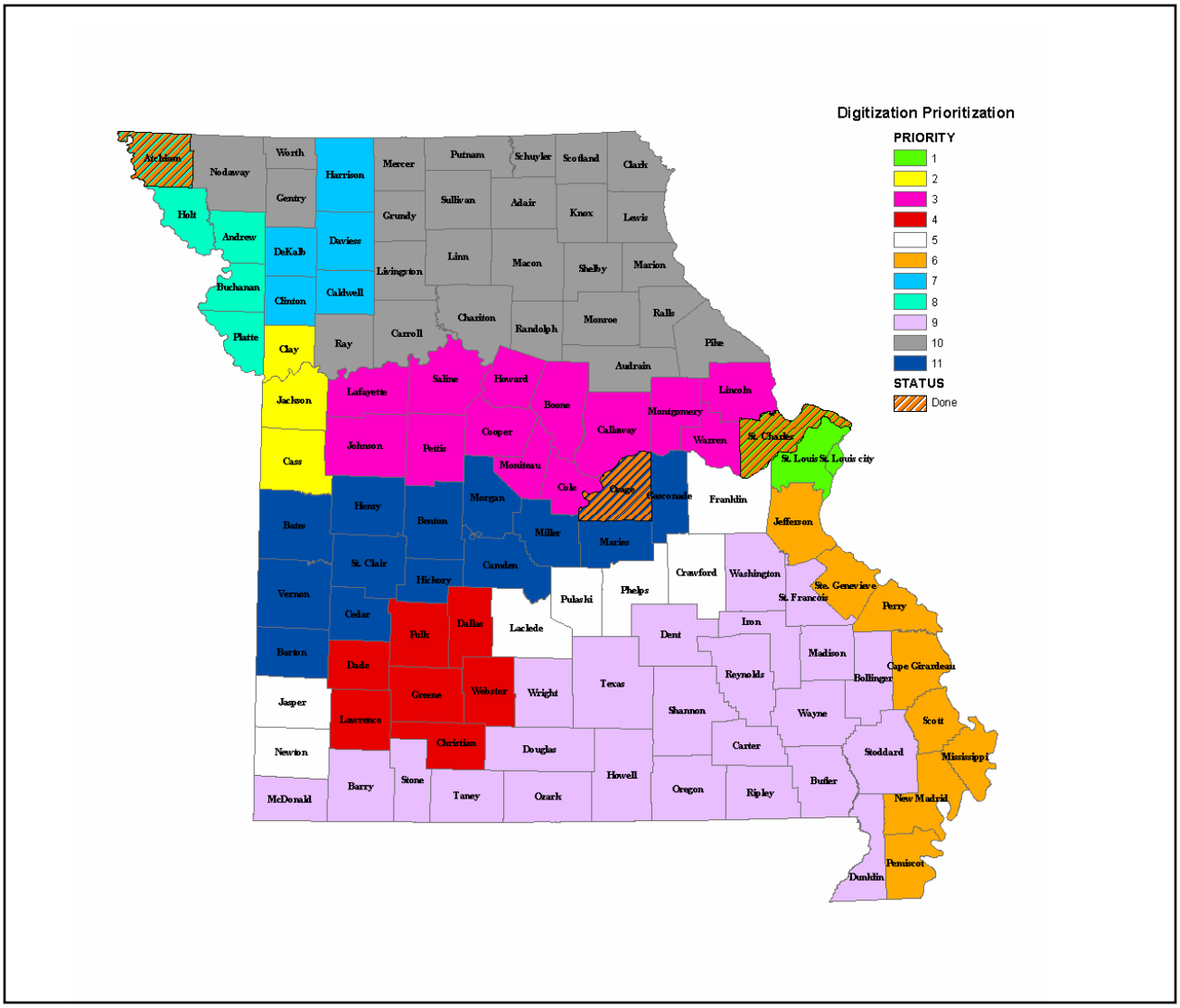

Figure 1. Prioritization of counties to be digitized and completed counties.

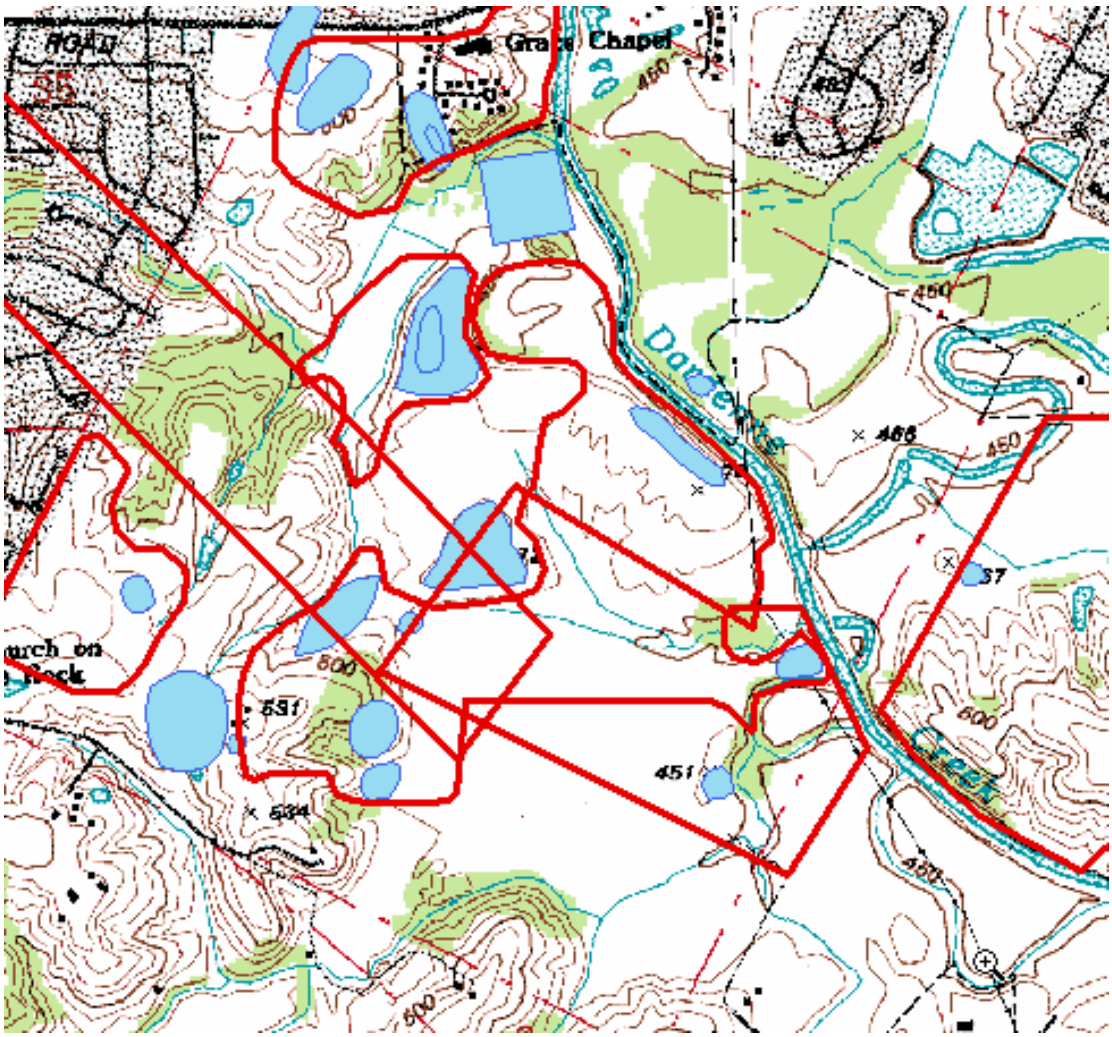

Figure 2. Red polygon outlines represent archaeological surveys; filled blue polygons are sites of cultural significance.

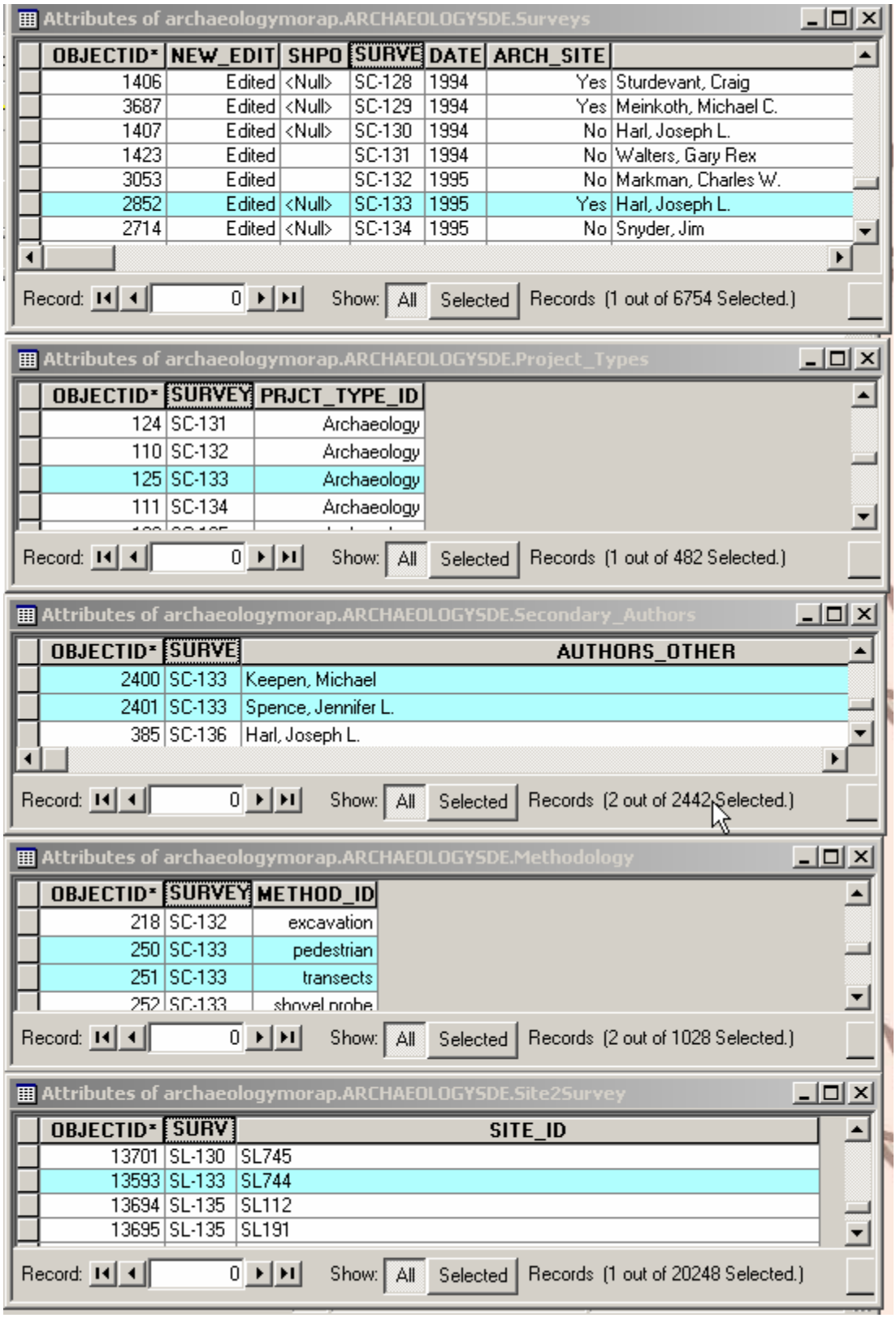

Figure 3. Related tables within the Archaeology geodatabase.

Metropolitan Public Lands

MoDOT Contact: Terri Wren
MoRAP Contact: Melissa Lanclos

Background and Goals

As urban areas within the state of Missouri are growing and their boundaries are expanding, the amount of infrastructure is also changing. It would be useful to MoDOT to know where public lands are located within metropolitan areas in order to assist with road construction and other transportation related projects. This project came about as a way to locate the most recent data representing the public lands for the largest and fastest growing metropolitan areas in Missouri. These cities are St. Louis, Kansas City, and Springfield.

General Methods and Procedures

Existing Data

A public lands data coverage currently exists for the entire state of Missouri. This coverage was originally created by MoRAP in the mid-1990s using data from the early 1990s. This data layer included the following land-holders:

- Audubon Society of Missouri
- City of Excelsior Springs
- City of Joplin
- United States Army Corps of Engineers
- Missouri Department of Natural Resources
- Department of Defense
- Jackson County Parks and Recreation
- Missouri Department of Conservation
- United States Forest Service
- State of Missouri
- The Nature Conservancy
- University of Missouri
- United States Fish and Wildlife Service

An update has recently been created (Fall 2003) using newer data for the major land-holding agencies in the state. The newer data layer consists of the public lands used in the original dataset with updates from the Missouri Department of Natural Resources, the Missouri Department of Conservation, the United States Forest Service, and the United States Fish and Wildlife Service. This dataset acted as the base layer upon which the remainder of the project was built.

It is then necessary to clean up any overlap between datasets. There are some areas where the boundaries from two or more layers overlap slightly (this is due to the fact that the various layers were created by different agencies and individuals, resulting in slight discrepancies in where one boundary ends and the next begins). The area where the layers overlap is known as a sliver polygon and these must be removed for data analysis and integrity.

The overlap is corrected by snapping the borders of the individual layers together where overlap exists. Snapping simply means that where two lines are in close proximity they can be snapped to each other, therefore becoming one single line. The sliver polygons in general are extremely small and therefore snapping the edges of two layers together does not have any major impacts on the area of the individual polygons (Figure 4).

Data Collection

Once the base data layer was created, individuals from each metropolitan area were contacted in order to obtain city-specific data for St. Louis, Kansas City, and Springfield. Many individuals were contacted in each city in an effort to procure the data. A list of these individuals is included at the end of this report (Tables 1-4). Contacts were attempted by either telephone or email, depending on what contact information was available. Individuals were contacted repeatedly until they referred us to someone else or it was determined that they did not have data that would benefit this project. In some instances, contact was never established with an individual. Generally, approximately four or five attempts were made over a month to a month and a half time period to reach the individuals before considering them an unreachable data source.

St. Louis. - For the St. Louis metropolitan area, 12 individuals / agencies were contacted in an effort to obtain data. Most of those contacted referred us to another source. In the end, all data that was collected for the St. Louis area came from one source. Gary Mook of East-West Gateway Council of Governments based in St. Louis provided a data layer of "Open Lands." This data includes city-owned, private, county-owned, and state agency owned lands for the greater St. Louis area. It is the data set that they have collected from various sources and put together to create a single data layer. This is an extremely complete data layer and includes data for both the Missouri and Illinois sides of the Mississippi River.

Data layers for such things as walking trails and bike paths would be a beneficial addition to the project data.

Kansas City. - For the Kansas City metropolitan area, 10 individuals / agencies were contacted in an effort to obtain data. Most of those contacted referred us to another source. Data sets were collected from two separate places. Data were purchased from the City of Kansas City at a price of $100 per megabyte. The total cost for the data was $122. The contact at the GIS shop in Kansas City was Kristin Collins. The data that were obtained included city-owned lands, areas owned by various businesses, Kansas City Parks and Recreation lands, and some cemeteries. Data were also downloaded from

the Automated Information Mapping System (AIMS), the Geographic Data & Service Provider for Johnson County, Kansas. The data set obtained from this source was parks, bike paths, trails, and golf courses for the Kansas side of the metropolitan area.

Data layers for bike paths and walking trails on the Missouri side of the metropolitan area would be a beneficial addition to the project data.

Springfield. - For the Springfield metropolitan area, four individuals / agencies were contacted in an effort to obtain data. Data were collected from Jim Vandiver of the Greene County Government and Mark Styles of the City of Springfield. Jim Vandiver provided data layers for the whole of Greene County. The data layer contained all of the parks, golf courses, and cemeteries for the county. Mark Styles provided a data layer containing data for lands owned by the City of Springfield, church-owned lands, fire stations, the school districts, and many other various owners.

Merging of Data Sets

The statewide public lands layer, which is acting as the base data for this project, is clipped to each of the three metropolitan areas. Each of the data layers collected for this project come from different sources. Due to that fact it is difficult and impractical to attempt to merge the data layers into one layer. Several polygons in an area may fall completely on top of one another (such as a golf course in the middle of a park), and merging these datasets in the same manner as the statewide public lands layer would lose valuable attribution for those layers. The best option therefore is to put all of the collected data layers into a personal geodatabase for delivery. This will allow all the data to be viewed and stored together, but will maintain the integrity of each of the individual layers.

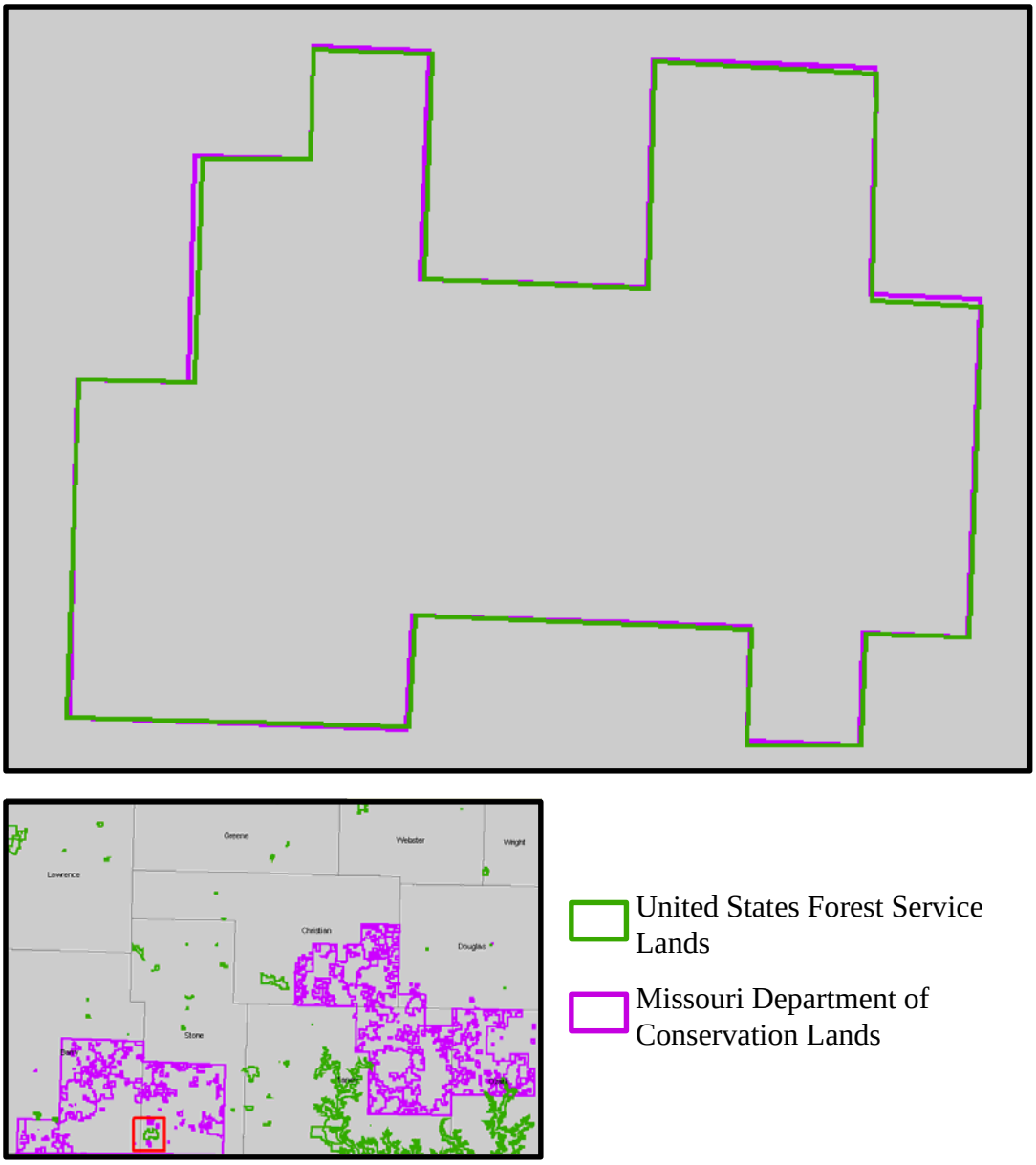

Figure 4. Example of overlap that occurs between datasets obtained from different agencies. These are the sliver polygons that must be corrected by snapping the boundary lines of each polygon to one another.

Results

The resulting data consists of a detailed public lands layer for each of the three metropolitan areas. The St. Louis and Kansas City data layers include some of the public lands in their neighboring states of Illinois and Kansas respectively. The public lands data for each of the metropolitan areas has been integrated into the statewide public lands dataset currently in existence for Missouri. This results in a much more detailed data layer for the three metropolitan areas (Figure 5).

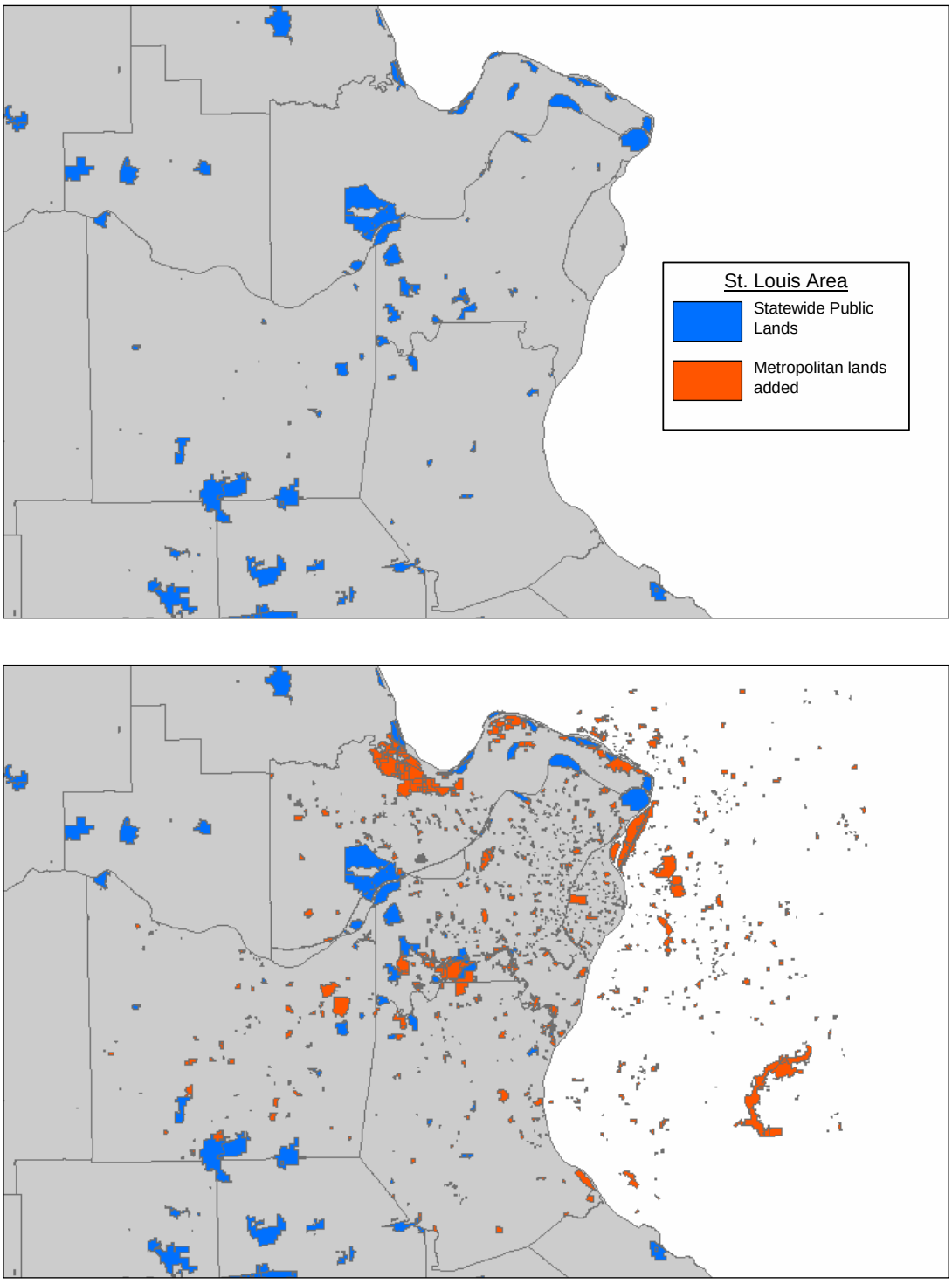

Figure 5. Example of the metropolitan public lands. This shows the St. Louis area with the statewide data at the top and the added metropolitan lands at the bottom.

	Contacted	Result
St. Louis	Don Brannon	Referred to Another Source
	Mike Duncan	Referred to Another Source
	Franklin Co	No Response
	Gary Hunt	Referred to Another Source
	Chuck Kindleburger	Referred to Another Source
	Melisa McLean	Referred to Another Source
	Gary Mook	Data Sent
	Hilary Perkins	No Response
	Chad Quin	Referred to Another Source
	Sonny Sanders	Referred to Another Source
	St. Louis GIS Users Grp	Referred to Another Source
	Monica Wilbur	No Response
Kansas City	Kristin Collins	Data Sent
	Christian Cooley	No Response
	Dawn Hilderbrand	Referred to Another Source
	Kansas City GIS Users Grp	No Response
	Kevin Kuzma	Referred to Another Source
	Steve Marsh	Referred to Another Source
	Tony Perkins	No Response
	Shannon Porter	Referred to Another Source
	Steve Rhodes	No Response
	Peter Veenstra	No Response
Springfield	Mike Fonner	Referred to Another Source
	Mark Styles	Data Sent
	Jim Vandiver	Data Sent
	Bill Weaver	Referred to Another Source
All Areas	Army Corps	No Response
	GeoComm	No Response

Table 1. Individuals / Agencies contacted in an effort to obtain data for the metropolitan public lands

Metros:	Area Counties:	Contacted:	
Lake St. Louis	St. Louis City	Sonny Sanders	ESRI
St. Charles	St. Louis	Don Brannam	St. Charles County GIS Director
St. Louis City	St. Charles	Melisa McLean	St. Louis County
	Franklin	Chuck Kindleburger	St. Louis County
	Jefferson	Gary Mook	East West Gateway Council Inc
		St. Louis Arc Users Group	
		Mike Duncan	St. Louis County Planning
		Hilary Perkins	Jacobs Engineering
		Franklin County	
		Gary Hunt	City of St. Louis
		Chad Quin	City of St. Louis
		Army Corps	

Data Provider	Data	Data Type	Area	Contact Name	Contact Number / Email	Availability	Cost	Status
MO Dept. of Conservation	Public lands	Arc Coverage	Statewide	Tony Spicci	573-522-4115 x3259/Tony.Spicci@mdc.mo.gov	Upon request	$0	In House
MO Dept. of Natural Resources	Public lands	.shp file	Statewide	Ron Curry	573-526-2005 / nrcurrr@dnr.state.mo.us	Upon request	$0	In House
United States Forest Service	Public lands	Arc Coverage	Statewide	Mike Schanta	573-341-7447 / mschanta@fs.fed.us	Upon request	$0	In House
United States Fish & Wildlife	Public lands	.shp files	Statewide			Upon request	$0	In House
St. Louis County Dept. of Planning	Parcel info	.shp files	St.L County	Mike Duncan / Melisa McLean	Mike_Duncan@stlouisco.com / Melisa_McLean@stlouisco.com	Upon request	$0	In House
MO Dept. of Transportation / SEMA	Buyout lands	.shp files	Statewide			Upon request	$0	In House
East West Gateway Council Inc.	"Open lands"	.shp files	8 County	Gary Mook	314-421-4220 / Gary.Mook@ewgateway.org	Upon request	$0	In House

Table 2. Data collection results and information for the St. Louis metropolitan area

Metros:	Belton
	Blue Springs
	Grandview
	Indepence
	Lee's Summit
	Liberty
	Peculiar
	Pleasant Hill
	Raytown

Area Counties:	Jackson
	Clay
	Cass
	Platte

Contacted:	GIS in KC	City of Kansas City
	Steve Rhodes	Mid-America Regional Council (MARC)
	Christian Cooley	Unified Government of KC
	Shannon Porter	AIMS
	Tony Perkins	City of Lenexa
	Steve Marsh	Jackson County
	Peter Veenstra	M.J. Harden
	Dawn Hilderbrand	City of Kansas City
	Kevin Kuzma	Platte County
	Kristin Collins	City of Kansas City

Data Provider	Data	Data Type	Area	Contact Name	Contact Number / Email	Availability	Cost	Status
MO Dept. of Conservation	Public lands	Arc Coverage	Statewide	Tony Spicci	573-522-4115 x3259/Tony.Spicci@mdc.mo.gov	Upon request	$0	In House
MO Dept. of Natural Resources	Public lands	.shp file	Statewide	Ron Curry	573-526-2005 / nrcurrr@dnr.state.mo.us	Upon request	$0	In House
United States Forest Service	Public lands	Arc Coverage	Statewide	Mike Schanta	573-341-7447 / mschanta@fs.fed.us	Upon request	$0	In House
United States Fish & Wildlife	Public lands	.shp files	Statewide			Upon request	$0	In House
MO Dept. of Transportation	Parks	.shp files	KC			Upon request	$0	In House
MO Dept. of Transportation / SEMA	Buyout lands	.shp files	Statewide			Upon request	$0	In House
City of Kansas City	City Owned Lands		KC	Kristin Collins	816-513-2812 / Kristin_Collins@kcmo.org	Upon request	$100/mb	In House
AIMS	Parks; bike/walking trails; golf	.shp files	KC	Web Site	http://aims.jocogov.org/	Download	$0	In House

Table 3. Data collection results and information for the Kansas City metropolitan area

Metros:	
	Battlefield
	Brookline
	Freemont Hills
	Nixa
	Ozark
	Republic
	Strafford
	Willard

Area Counties:	
	Greene
	Christian

Contacted:		
	Bill Weaver	City of Springfield
	Mike Fonner	City of Springfield
	Jim Vandiver	Greene County
	Mark Styles	City of Springfield

Data Provider	Data	Data Type	Area	Contact Name	Contact Number / Email	Availability	Cost	Status
MO Dept. of Conservation	Public lands	Arc Coverage	Statewide	Tony Spicci	573-522-4115 x3259/Tony.Spicci@mdc.mo.gov	Upon request	$0	In House
MO Dept. of Natural Resources	Public lands	.shp file	Statewide	Ron Curry	573-526-2005 / nrcurrr@dnr.state.mo.us	Upon request	$0	In House
United States Forest Service	Public lands	Arc Coverage	Statewide	Mike Schanta	573-341-7447 / mschanta@fs.fed.us	Upon request	$0	In House
United States Fish & Wildlife	Public lands	.shp files	Statewide			Upon request	$0	In House
MO Dept. of Transportation / SEMA	Buyout lands	.shp files	Statewide			Upon request	$0	In House
Green County	County Lands		Greene Co	Jim Vandiver	417-868-4005	Upon request	$0	In House
City of Springfield	Government Property			Mark Styles	417-864-1969	Upon request	$0	In House

Table 4. Data collection results and information for the Springfield metropolitan area

SEMA Data Update

MoDOT Contact: Bill Graham
MoRAP Contact: Melissa Lanclos

General Methods and Procedures

Existing Data

The project began with MoDOT acquiring all of the SEMA buy-out lands for the State of Missouri. Neil Meredith, a student employee of MoDOT, went through all of the data during the fall/winter of 2002 and 2003, entered it into a database, and assigned a spatial location to each address associated with a buy-out land. Some of the SEMA data had exact addresses and was given an exact spatial location. However, some of the points had a more generic address with less specific information, such as only a street name with no numerical addresses associated with it. These were given the accuracy of "Geocoded". The result was several SEMA data points all occurring directly on top of one another in a general location.

MoRAP was included in this project due to the addresses within the SEMA buy-out land data that were given a general address rather than an exact (those referred to as "Geocoded" in the accuracy field of the table). It was agreed that MoRAP would use the TIGER line files to assign a more accurate location to as many of the data points with "Geocoded" accuracy as possible, in order to improve on the already existing data.

Address Updating

The SEMA buy-out lands data set included a total of 1,479 records. Of these records, 593 of them received the accuracy of "Geocoded." These 593 records were separated from the original dataset in order to attempt to assign a more specific spatial location to each.

Most of the 593 records did in fact have an address which included a street number. However, for various reasons an exact location had not been assigned. These reasons might include: the street did not exist in the data layer MoDOT was using to assign points, the street name was one that existed in several different cities / towns in the state, or the street name was misspelled or missing a portion of the name (such as St., Dr., N., etc.).

A city field was added to the table in order to narrow down the location of the street for each of the 593 records. This acted to eliminate the issue of the street name occurring in several different places, and aided in determining where each point was located.

Geocoding Procedures

ESRI products have a Geocoding function which can be used to address match. This function can be described as:

> A feature is an object that has geometry. In most cases, this geometry is captured by digitizing or scanning paper maps. In many cases, however, geographic data exists that indirectly captures geometry by describing locations such as street addresses, city names, or even telephone numbers. While humans understand what these descriptions mean and how they relate to locations on the earth's surface, computers do not. In order to display these locations on a map and perform analyses with them, a computer must be given geometric representations (such as point features) of these locations.
>
> Geocoding (also known as address matching) is the process of creating geometric representations for descriptions of locations. A geocoding service defines a process for converting nonspatial descriptions of places into spatial descriptions.

Geocoding was performed on the data set using the TIGER road files. The initial run through had fairly poor results – only 23 of the 593 records were given new points. It was determined that the issue was that many of the streets were missing a portion of their names (St., Dr., Ave., etc). Each of the 593 records was manually searched for and located using the TIGER files and additions / subtractions were made as needed (it should be noted that some of the street names did not exist in the TIGER data). Once the addresses had been updated, Geocoding was run again.

Results

The Geocoding function assigns one of three options to each record in the table when analysis is performed. These are "M" for matched and indicates that an exact match has occurred, "T" for tied which indicates that there was more than one possible match found in the reference dataset, and "U" for unmatched which indicates that no match was found. The final analysis resulted in 214 of the 593 records receiving "M's" or exact address matches, 136 records received a "T", and 243 records remained unmatched. The accuracy for each of the 593 records was changed accordingly. The records that were coded as "M" received an accuracy of Street, the "T's" received an accuracy of Multiple (a new accuracy option referring to the fact that there are several possibilities for location), and those records which received a "U" kept the Geocoded accuracy. These updated addresses were merged back into the original data layer.

In the original data layer, many of the points with an accuracy of Geocode occurred directly on top of one another. When the Geocoding function was performed in ArcMap each of those points was moved. For the points where an exact address was found, they were moved to the correct spatial location. For the points where no exact address was determined, the points were simply moved slightly apart so that they no longer occur in the exact same location.

There are several suggestions as to what to do with the remaining addresses with an accuracy of Geocode. The TIGER line files are the most accurate road data MoRAP has

available, and some of the most detailed road network data available. However, it does not contain every single road in existence for all areas. This is why not all of the "Geocoded" addresses were matched to the road data. Several suggestions could be made in order to assign spatial locations to the remaining data points. 1) An individual could visit each of the data points and take GPS readings of the location. These data could then be incorporated into the file with the rest of the address points. 2) A roads data layer could be created for individual areas using higher resolution data where address points need better accuracy. This would involve obtaining road data for the areas at a fairly large scale and creating a new road layer. It is unknown how well this would work, largely based on whether or not such data exists or is available.

Another factor to consider is that some of the addresses listed in the database might be incorrect. Spelling mistakes, inaccurate housing numbers, and streets that have been renamed or no longer exist could result in the addresses not being found in a roads data layer.