Journal of Environmental Management 85 (2007) 1120–1131

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Creating affordable Internet map server applications for regional scale applications

Arthur J. Lembo[a,*], Linda P. Wagenet[b], Tania Schusler[b], Stephen D. DeGloria[b]

[a]Henson Science HALL 157H, Salisbury University, 1101 Camden Ave, Salisbury, MD 21801, USA
[b]Crop and Soil Science, Cornell University, 1001 Bradfield Hall, Cornell University, Ithaca, NY 14850, USA

Received 27 March 2006; received in revised form 4 November 2006; accepted 19 November 2006
Available online 17 January 2007

Abstract

This paper presents an overview and process for developing an Internet Map Server (IMS) application for a local volunteer watershed group using an Internal Internet Map Server (IIMS) strategy. The paper illustrates that modern GIS architectures utilizing an internal Internet map server coupled with a spatial SQL command language allow for rapid development of IMS applications. The implication of this approach means that powerful IMS applications can be rapidly and affordably developed for volunteer organizations that lack significant funds or a full time information technology staff.
Published by Elsevier Ltd.

Keywords: Geographic information system; Internet map server

1. Introduction

Geographic information systems (GIS) provide great utility for performing environmental based applications by allowing users to combine different geographic data sets together within a single application. The addition of Internet Map Server (IMS) applications allows users to store, analyze, and communicate the results of environmental analysis to a broad audience. As stated by Peng and Tsou (2003) "Internet GIS is a research and application area that utilizes the Internet and other internetworking systems to facilitate the access, processing, and dissemination of geographic information and spatial analysis". The integration of GIS and IMS allows virtually anyone with an Internet browser to participate in collaborative spatial analysis projects.

An IMS is an Internet based GIS application designed to allow users to interact with geographic data using standard web browsers. Because the application runs inside a web browser (like Internet Explorer or Netscape), IMS applications provide access to spatial data by a wide number of users located throughout the world without the need for more expensive desktop GIS software. However, the capabilities of an IMS application are much more limited than those capabilities provided within a fully functional desktop GIS. Typically, Internet users can perform basic GIS related activities such as zooming and panning around the map or clicking on features to obtain attribute information. Additionally, IMS applications can allow users to perform basic geographic queries such as selecting features based on geographic criteria, or location features based on entering an address.

Traditionally, GIS software costs several thousands of dollars for a single license (OGC, 2003). For some commercial GIS software, off-the-shelf pricing for a suite of products including GIS and IMS can cost upwards of twenty thousand dollars. These costs, coupled with the difficulty to learn complex software, have made adoption of GIS technology prohibitive for volunteer organizations. Other options have been to utilize open source solutions such as PostGIS (http://postgis.refractions.net/), GRASS (http://grass.itc.it/), MySQL (http://www.mysql.com/), and the Minnesota Mapserver

*Corresponding author. Tel.: +1 410 543 6460.
E-mail address: alembo@salisbury.edu (A.J. Lembo).

0301-4797/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.jenvman.2006.11.024

(http://mapserver.gis.umn.edu/). PostGIS, MySQL, and Mapserver are described more fully in Kropla (2005) and Mitchell (2005). These open source products are attractive in that the software is free. Disadvantages in using open source technologies are presented in the literature (Nichols and Twidale, 2003), as are some of the advantages (Lakhani and von Hippel, 2003). Recent web mapping standards (WMS) and web feature standards (WFS) developed by the Open GIS Consortium (http://www.opengis.org/) are a step in the right direction by providing the minimal functionality for IMS applications and GIS data formats.

Upon a review of both Kropla and Mitchell, the complexity of integrating multiple software products to create a functional IMS application is readily apparent especially for public interest volunteer organizations without information technology staff support. In many instances, the open source GIS solutions are somewhat cryptic to just install or use as warned by Kropla (2005, p. 1):

> Building MapServer for the first time can be a challenging experience for someone unfamiliar with Unix build environments. There are many interdependencies between the libraries, and the sequence in which the supporting libraries are created is very important.

Therefore, while the software is free, one can argue that other costs associated with open source solutions make their use prohibitive for some organizations lacking proper technical and sufficient financial resources.

This paper presents an overview and process for developing an IMS application for a local volunteer watershed group using a low cost GIS. The paper illustrates that using a modern GIS and IMS architecture facilitates rapid development while minimizing the cost of implementation. Within the paper we describe the implementation of the IMS application using Manifold GIS, a modern, Windows based GIS software product of moderate price and high functionality. The main objective was to identify a well designed architectural strategy for integrating GIS and IMS together that minimizes the necessary programming requirements to implement an easy-to-use and powerful IMS application.

2. Background

2.1. Watershed monitoring group

Our project was initiated at the request of the Tompkins County Water Resources Council's (WRC) Monitoring Committee. The WRC is an advisory board to the Tompkins County legislature and its Monitoring Committee is responsible for the coordination of monitoring efforts throughout the county. The Six Mile Creek watershed was selected for our initial project development because the IMS application could contribute to a five year project examining sedimentation in the watershed using a readily available set of data collected by a highly participatory group of volunteers.

The Six Mile Creek watershed is a subwatershed of Cayuga Lake, one of the Finger Lakes in upstate New York. The Cayuga Lake Watershed covers seven hundred eighty-five square miles. There are forty-four municipalities and six counties that are all or partially in the watershed. The watershed has been divided into nineteen major subwatersheds (eighteen tributary-based subwatersheds and the remainder in direct drainage) and then further delineated into forty-six minor subwatersheds based on the major tributaries of Cayuga Lake, of which Six Mile Creek is one.

In 2004, the Six Mile Creek volunteer monitoring group organized in the watershed. This program is a partnership of volunteers, the Community Science Institute, Tompkins County Soil and Water Conservation District, and Cornell Cooperative Extension of Tompkins County. Six Mile Creek flows through four municipalities (the City of Ithaca and Towns of Caroline, Dryden, and Ithaca), which contribute funding to the program. This partnership produces management quality data because volunteers are trained in quality assurance and quality control procedures and analyses are conducted by the Community Science Institute's EPA-certified laboratory.

Volunteers who monitor Six Mile Creek assess several water quality parameters, including indicators of minerals, sediment, nutrients and bacteria. Over an eighteen month period, twenty volunteers sampled six times at thirteen locations along the stream.

Recognizing the value of but relative inaccessibility of these data by watershed management agencies, local municipalities, members of the public, and volunteers themselves, the WRC Monitoring Committee expressed interest in reporting these data on a website in order to share water quality trends with others. Neither the WRC Monitoring Committee nor the Six Mile Creek volunteer monitors had the expertise or funding capacity for purchasing and effectively using other traditional commercial GIS software programs. Our project provided them access to a lower cost Internet-based mechanism for such reporting.

2.2. Internet map server architectures

Numerous architectures for Internet based GIS exist (Alameh, 2001). Such architectures include static map publishing, static web mapping, interactive web mapping, and distributed GIS services (Peng and Ming, 2003). In most cases, literature is extensive on the technological methods to implement these architectures (Peng and Ming, 2003; Alameh, 2001; Peters, 2003). Yet in virtually all cases, the technological approach is to migrate functionalities in a stand-alone GIS into interoperable autonomous components/services (Alameh, 2001).

The services approach in traditional IMS technology may actually increase the complexity and cost, forcing users to purchase separate software that is often divorced from any main GIS product. As shown in Fig. 1, traditional IMS technology requires a separate map server that is only loosely connected with the GIS. In addition, rather than having a full suite of geospatial techniques available, only those services exposed within the map server application are available to Internet developers.

Implementation of an IMS application for volunteer organizations would appear to require an easier solution. An alternative strategy is to use a desktop GIS software product with an embedded map server, as shown in Fig. 2. We call this approach Internal Internet Map Server Technology (IIMST).

By using an IIMST architecture, the GIS software itself is the map server and virtually all functionality within the desktop GIS is immediately available for use on the Internet. Also, allowing the desktop GIS to be the actual map server exposes all the data, queries, cartographic display parameters, and the entire GIS object model for customized scripting to Internet users. Having the ability to re-use scripting and high level query code both in the desktop GIS and the IMS application allows for rapid and inexpensive development of customized applications for delivery over the Internet.

To perform this work, we chose Manifold GIS. Manifold is a fully functional GIS (Lembo, 2004a, b) and implements the IIMS technology described in Fig. 2, in addition to being fully compatible with the Microsoft operating system. Some of the benefits of this strategy include:

Ease of use—the desktop GIS is a Microsoft Windows-based system, with an easy to use graphical user interface.
Cost—the desktop GIS costs two hundred forty five dollars, and includes the IMS technology.
Ease of administration—since a separate product is not required, additional software is not necessary to publish the GIS application on the Internet. Installing the desktop GIS on an Internet server with Microsoft Internet Information Services (IIS) establishes the IMS application.
Expanded spatial SQL operators—the desktop GIS software utilizes the structured query language (SQL) in addition to spatial operators built within SQL to perform sophisticated GIS tasks. Therefore, specific spatial queries could be realized using SQL and once written in the desktop system, are immediately available on the Internet.
Internal Internet Map Server technology—with the IMS built directly within the desktop GIS software, all data, cartographic rendering, and queries are immediately available for Internet use, without the need for separate map server software, or reprogramming queries, data, or cartographic rendering within a separate product. Additionally, the IMS application makes use of the same object model used by the desktop GIS. Therefore, more sophisticated scripts written for the desktop can be easily ported over to the IMS application, or, simply referenced from the IMS application.

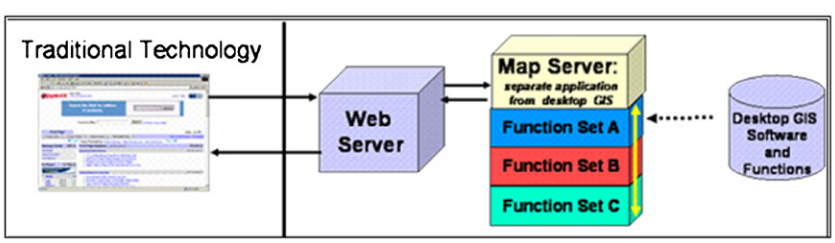


Fig. 1. Traditional approach to IMS technology uses a GIS loosely connected to a separate map server technology. In some instances, the GIS is not connected to the map server at all.

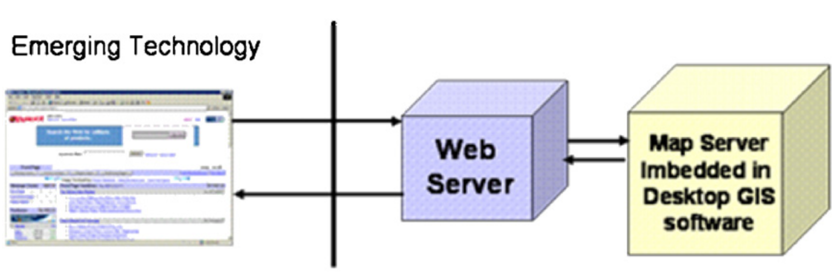


Fig. 2. Emerging technology showing the integration of the Internet map server component (IIMST) within the desktop GIS software. This approach enables the same toolset within the desktop GIS to be available for the Internet developer.

Pack and Ship philosophy—each GIS project file is a self-contained unit, allowing the entire desktop application, including the queries referenced by the IMS application, to easily be copied and placed on another computer system.
Tight Integration with Microsoft tools—Integration with the Microsoft environment allows for seamless integration with Microsoft Office tools. This allows for great flexibility when integrating with data in common Microsoft formats such as Excel and Access.

The IMS application was installed on a 2.0 GHz. Dell Optiplex server with 1 GB of RAM running Manifold GIS (the map server) and Microsoft Internet Information Services (the web server). The application was designed so that all geographic and attribute processing would occur on the server side, thus reducing requirements for the client side (i.e. browser application). Therefore, virtually any standard web browser could be used to operate the application. In this scenario, the client browser would pass requests to the server, such as the command to zoom or pan, using a standard map user interface, and the map server would process the request and send the results back to the browser. For this application, the map server passes an image of the map display for viewing in the browser, and also passes the actual values of the attribute results into an HTML derived table. The client-server nature of this design is ideal for managing many concurrent users. Recent tests using the same hardware and software configuration for an IMS application related to the Indian Ocean tsunami was able to handle over 145,000 requests within a two week period without any noticeable degradation in response.

3. Methods

The data used for this application are from the Six Mile Creek watershed in Tompkins County, NY, a subwatershed of Cayuga Lake. The Cayuga Lake watershed lies within the Allegheny Section of the Northern Appalachian Highland Division of the Hemlock-White Pine Northern Hardwood Region (Braun, 1950). In the southern end of the watershed, where our study occurred, the majority of the land cover is natural forested upland with trees forming greater than twenty-five percent canopy cover (NYSDOS, 1999). We chose this location because it provided our study with historical records from past Cornell University research and offered a range of land cover and physiographic data that affords the opportunity to examine erosion and sedimentation under a variety of conditions (Marks and Gardescu, 1992; Smith et al., 1993). Presently, more than seventy percent of the Six Mile Creek watershed is forested, with ten percent of the watershed devoted to agriculture, and approximately six percent defined as urban. The data layers used in this project are described in Table 1.

The data sets were then integrated together into a multi-layer GIS for further processing as shown in Fig. 3.

One of the most important data sets for this project was the data collected at each monitoring site, collected by the Six Mile Creek volunteer monitoring group. The data were collected at thirteen locations along the stream. Selection of monitoring locations by the volunteer group was based on accessibility, varying land use characteristics along the stream from its forested headwaters to urban outlet of Six Mile Creek at Cayuga Lake, and concerns about specific water quality impacts. Volunteers collected samples under both baseline/low flow and storm event/high flow conditions each spring, summer, and autumn. During each sampling event, volunteers collected water samples from the monitoring sites and transported them to a NYS-certified laboratory at the Community Science Institute for analysis. The data were then provided in Microsoft Excel format, as shown in Fig. 4.

The Six Mile Creek database included historical information, with collections organized by date and sampling period (round). Each sample was a measure of multiple parameters, including alkalinity, total hardness, chloride, total suspended solids, turbidity, total phosphorous, soluble reactive phosphorous, nitrate-+nitrite-nitrogen, total K, nitrogen, total coliform, and *E. coli*, among others.

A number of GIS tools were used to prepare the data. These tools include merging DEM data together, clipping data sets to the watershed boundary, and watershed modeling tools to compute the upstream contributing areas. The process for performing many of these operations, in both Manifold GIS and ArcGIS is found in Lembo (2004b). Discussions with stakeholders identified five basic functions that the IMS application should perform. These functions include zooming to a specified monitoring site, viewing the history of a monitoring site, selecting monitoring sites based upon ad-hoc queries for parameter measurements, having the ability to view daily results of two USGS gauge stations in the watershed, and quantifying both land use and roadway characteristics for each monitoring site's upstream contributing area. Additionally, basic navigation tools (zoom in/out, pan, identify) were noted as important.

Traditionally, such queries would require extensive programming to accomplish in an IMS application. However, the use of spatial SQL and an Internal Internet Map Server Technology (IIMST) allowed for rapid development of each application task. The remainder of this section will provide an overview of how each application task was first developed in the desktop GIS, and then migrated over to the IMS application.

3.1. Zooming to a specified monitoring site

Identifying and zooming to a specific monitoring site was accomplished using a simple SQL parameter query. SQL allows within a single declarative statement, the ability to perform rather sophisticated tasks. The following SQL statement performs the actual query:

```
PARAMETERS [Enter station ID] INTEGER;
SELECT [recent flow data].[geom (i)], [smc database].*
FROM [Recent Flow Data], [smc database]
WHERE [recent flow data].samp_id = [Enter station ID]
AND [smc database].samp_id = [Enter station id]
ORDER BY [date] DESC
```

The above query requires the user to enter a parameter, in this case the monitoring station ID. The query then selects the geometry field [geom (i)] of the monitoring stations drawing ([Recent Flow Data]) in addition to all the historical records from the Six Mile Creek database ([smc database]). The *WHERE* clause selects only those records from the SMC database

Table 1
Base data used in the SMC application

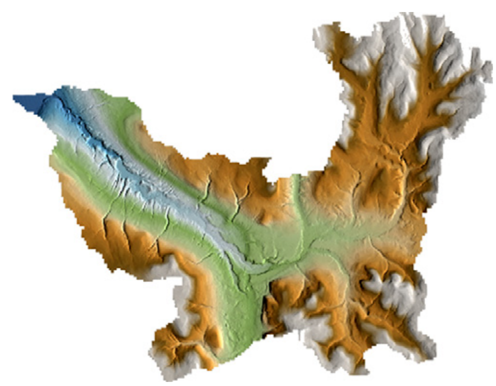
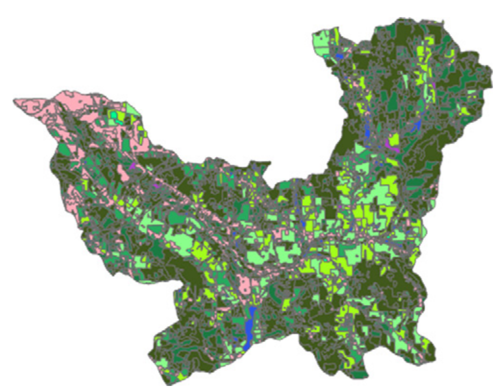
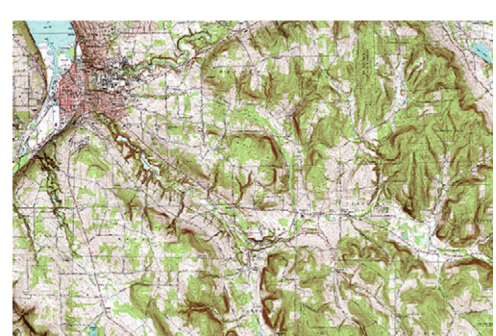
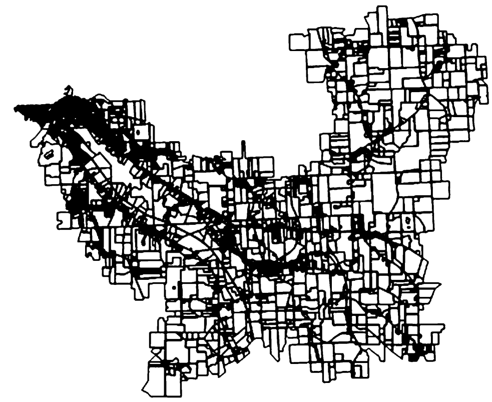
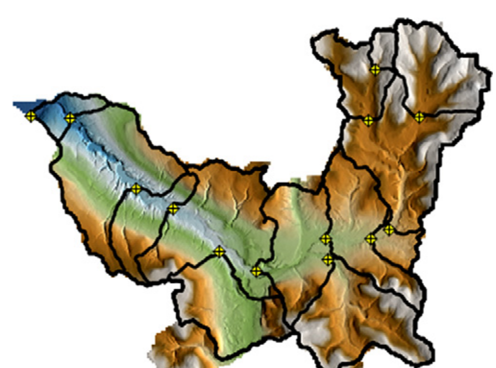
Digital Elevation Model—the digital elevation model of Six Mile Creek covers 127 sq. km, and was assembled from numerous USGS 10 m DEMs. The DEMs were imported into Manifold GIS and extracted out using the clip transformation tool. In each case, the watershed boundary delineated by Tompkins County GIS was used as the clip boundary.	
Land use—the land use map was obtained from the Tompkins County GIS as an ESRI shape file and imported into Manifold GIS. Using the topological *clip* transformation tool in Manifold, the landuse was clipped to the watershed boundary.	
Topographic map—the topographic map of the region was imported from Microsoft Terraservice. Manifold GIS has the ability to dynamically link and import any image stored on Terraservice.	

Table 1 (*continued*)

Property parcels—the property parcels were obtained from Tompkins County GIS as an ESRI shape file and imported into Manifold. Rather than extract the parcels to the watershed boundary, parcels topologically touching the boundary were selected out.	
Roads—The road data was obtained from Tompkins County GIS as an ESRI shape file and imported into Manifold GIS. Roads topologically touching the watershed boundary were selected out.	
Sampling locations—sampling locations were highlighted on USGS quadsheets by the SMC monitoring group. These locations were then manually positioned using heads-up digitizing of marked up USGS quadsheets and a digital topographic map.	
Catchment areas—using the digital elevation model, and the watershed tools in Manifold GIS, flow direction, flow accumulations, and catchment areas were derived for each of the monitoring stations. This provided a mechanism to determine the upslope contributing area for each monitoring site.	

where the station ID [samp_id] matches the station ID entered by the user. The *ORDER BY* clause sorts the data in descending order (*DESC*) and presents the resulting table to the user. The [SMC Database] is separate from the monitoring station geographic layer because the [SMC Database] maintains historical information, and thus creates a one-to-many

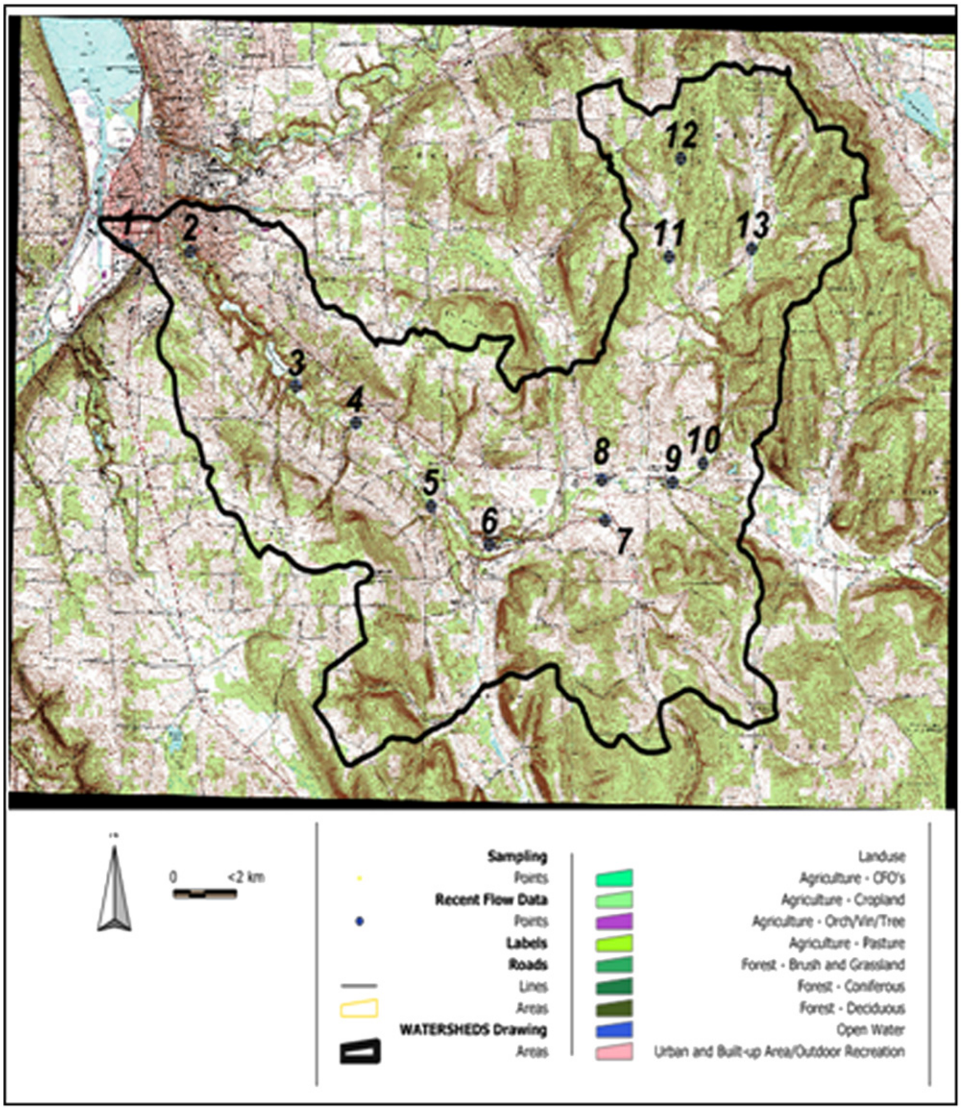


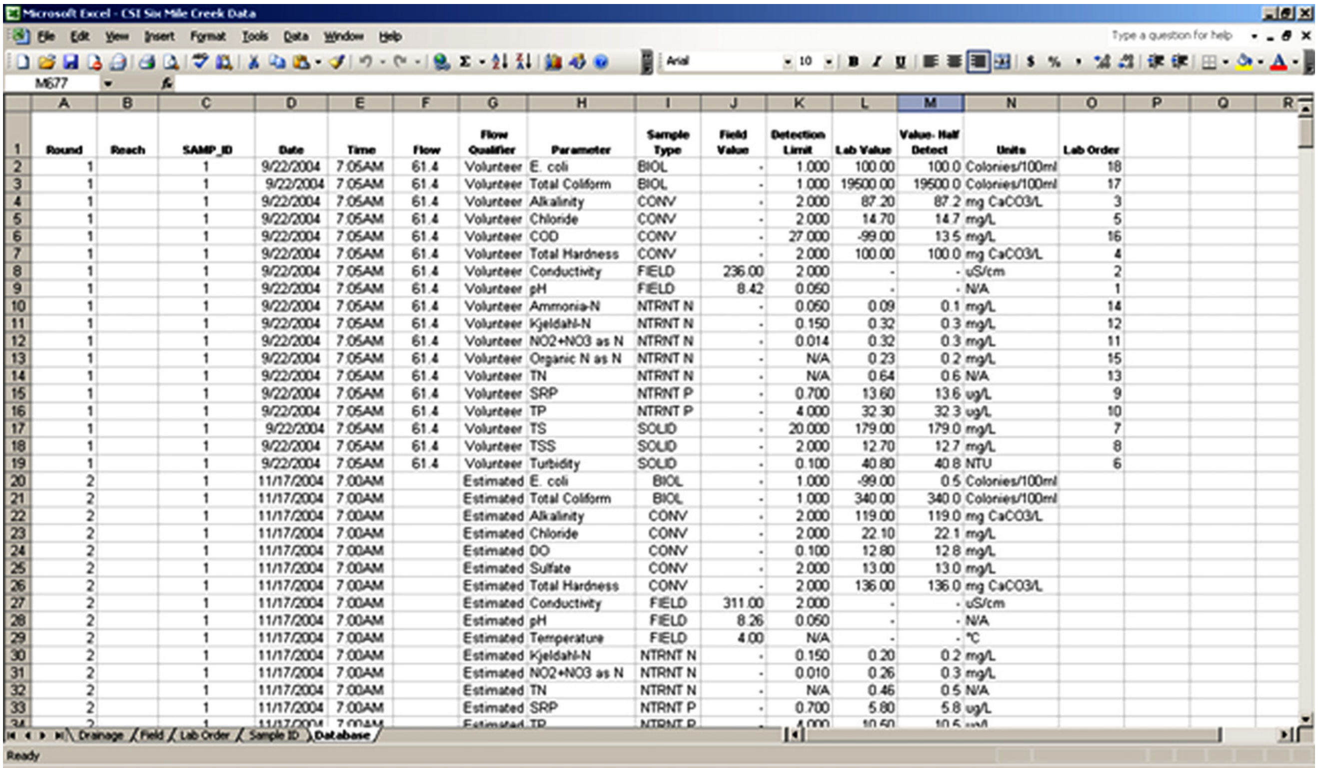
Fig. 3. GIS map of the Six Mile Creek Watershed.

database relationship with the individual point for the station location. But, as can be seen in the SQL query shown above, presenting one-to-many relationship data to the user is simple to accomplish within the SQL statement.

Because the desktop GIS application is also the map server application, this query is immediately available on the web page. Therefore, a simple dropdown field created in HTML allows the user to select a monitoring site number, and the IMS application issues the SQL query, which is received by the GIS application, returning the results. Because a geometry field was selected in the SQL query, the Manifold IMS application has an internal toggle to automatically zoom to the selected feature location.

Therefore, the above SQL query and the following HTML code provides the user with a fully functional drop down query to select and zoom to a specific monitoring site, and display the sampling results sorted by date:

```
⟨select name = ''select2'' size = ''1'' onChange = ''zoomsite(this.options
[this.selectedIndex].value)''⟩
      ⟨option⟩Zoom to Monitoring Site⟨/option⟩
      ⟨option value = ''1''⟩1⟨/option⟩.
      .
      .
      ⟨option value = ''2''⟩2⟨/option⟩
```



Microsoft Excel - CSI Six Mile Creek Data

	Round	Reach	SAMP_ID	Date	Time	Flow	Flow Qualifier	Parameter	Sample Type	Field Value	Detection Limit	Lab Value	Value- Half Detect	Units	Lab Order
2	1		1	9/22/2004	7:05AM	61.4	Volunteer	E. coli	BIOL	-	1.000	100.00	100.0	Colonies/100ml	18
3	1		1	9/22/2004	7:05AM	61.4	Volunteer	Total Coliform	BIOL	-	1.000	19500.00	19500.0	Colonies/100ml	17
4	1		1	9/22/2004	7:05AM	61.4	Volunteer	Alkalinity	CONV	-	2.000	87.20	87.2	mg CaCO3/L	3
5	1		1	9/22/2004	7:05AM	61.4	Volunteer	Chloride	CONV	-	2.000	14.70	14.7	mg/L	5
6	1		1	9/22/2004	7:05AM	61.4	Volunteer	COD	CONV	-	27.000	-99.00	13.5	mg/L	16
7	1		1	9/22/2004	7:05AM	61.4	Volunteer	Total Hardness	CONV	-	2.000	100.00	100.0	mg CaCO3/L	4
8	1		1	9/22/2004	7:05AM	61.4	Volunteer	Conductivity	FIELD	236.00	2.000	-	-	uS/cm	2
9	1		1	9/22/2004	7:05AM	61.4	Volunteer	pH	FIELD	8.42	0.050	-	-	N/A	1
10	1		1	9/22/2004	7:05AM	61.4	Volunteer	Ammonia-N	NTRNT N	-	0.050	0.09	0.1	mg/L	14
11	1		1	9/22/2004	7:05AM	61.4	Volunteer	Kjeldahl-N	NTRNT N	-	0.150	0.32	0.3	mg/L	12
12	1		1	9/22/2004	7:05AM	61.4	Volunteer	NO2+NO3 as N	NTRNT N	-	0.014	0.32	0.3	mg/L	11
13	1		1	9/22/2004	7:05AM	61.4	Volunteer	Organic N as N	NTRNT N	-	N/A	0.23	0.2	mg/L	15
14	1		1	9/22/2004	7:05AM	61.4	Volunteer	TN	NTRNT N	-	N/A	0.64	0.6	N/A	13
15	1		1	9/22/2004	7:05AM	61.4	Volunteer	SRP	NTRNT P	-	0.700	13.60	13.6	ug/L	9
16	1		1	9/22/2004	7:05AM	61.4	Volunteer	TP	NTRNT P	-	4.000	32.30	32.3	ug/L	10
17	1		1	9/22/2004	7:05AM	61.4	Volunteer	TS	SOLID	-	20.000	179.00	179.0	mg/L	7
18	1		1	9/22/2004	7:05AM	61.4	Volunteer	TSS	SOLID	-	2.000	12.70	12.7	mg/L	8
19	1		1	9/22/2004	7:05AM	61.4	Volunteer	Turbidity	SOLID	-	0.100	40.80	40.8	NTU	6
20	2		1	11/17/2004	7:00AM		Estimated	E. coli	BIOL	-	1.000	-99.00	0.5	Colonies/100ml	
21	2		1	11/17/2004	7:00AM		Estimated	Total Coliform	BIOL	-	1.000	340.00	340.0	Colonies/100ml	
22	2		1	11/17/2004	7:00AM		Estimated	Alkalinity	CONV	-	2.000	119.00	119.0	mg CaCO3/L	
23	2		1	11/17/2004	7:00AM		Estimated	Chloride	CONV	-	2.000	22.10	22.1	mg/L	
24	2		1	11/17/2004	7:00AM		Estimated	DO	CONV	-	0.100	12.80	12.8	mg/L	
25	2		1	11/17/2004	7:00AM		Estimated	Sulfate	CONV	-	2.000	13.00	13.0	mg/L	
26	2		1	11/17/2004	7:00AM		Estimated	Total Hardness	CONV	-	2.000	136.00	136.0	mg CaCO3/L	
27	2		1	11/17/2004	7:00AM		Estimated	Conductivity	FIELD	311.00	2.000	-	-	uS/cm	
28	2		1	11/17/2004	7:00AM		Estimated	pH	FIELD	8.26	0.050	-	-	N/A	
29	2		1	11/17/2004	7:00AM		Estimated	Temperature	FIELD	4.00	N/A	-	-	°C	
30	2		1	11/17/2004	7:00AM		Estimated	Kjeldahl-N	NTRNT N	-	0.150	0.20	0.2	mg/L	
31	2		1	11/17/2004	7:00AM		Estimated	NO2+NO3 as N	NTRNT N	-	0.010	0.26	0.3	mg/L	
32	2		1	11/17/2004	7:00AM		Estimated	TN	NTRNT N	-	N/A	0.46	0.5	N/A	
33	2		1	11/17/2004	7:00AM		Estimated	SRP	NTRNT P	-	0.700	5.80	5.8	ug/L	
34	2		1	11/17/2004	7:00AM		Estimated	TP	NTRNT P	-	4.000	10.50	10.5	ug/L	

Fig. 4. Microsoft Excel spreadsheet showing data collected for monitoring locations in the Six Mile Creek Watershed.

3.2. Viewing the history of a monitoring site

Similar to the previous query, viewing the history of a monitoring site is a straightforward SQL query as shown below:

```
PARAMETERS [Enter Station Number] INT, [start date] DATE, [end date] DATE;
SELECT [SMC Database].*, sampling.[geom. (i)]
FROM [SMC Database], sampling
WHERE samp_id = [Enter Station Number]
AND sampling.siteid = [Enter Station Number]
AND [Date]
BETWEEN [Start Date] AND [End Date]
```

In this case, the user enters a starting and ending date for a particular station number. The SQL query then selects all records in [SMC Database] that fall between the two dates, in addition to the geometry field from the monitoring station layer. In this case the query is simple to perform as the GIS supports *Date* fields. And, the *BETWEEN* clause in SQL provides a simple mechanism to select all records falling between the two dates, regardless of the order entered. Similar to the previous query, selection of the geometry field causes the IMS application to automatically zoom to the specified feature.

3.3. Selecting monitoring sites based upon ad-hoc queries for parameter measurements

The third requirement was to allow users to select a parameter (i.e. total solids, *E. coli*, pH) and enter a threshold value. The IMS application would find all historical applications that meet the user defined criteria, present the resulting table, and zoom to the selected monitoring stations.

Once again, the use of SQL facilitates the ability to perform this kind of sophisticated query of two different tables within a few statements as shown below.

```
PARAMETERS [Enter parameter] TEXT, [Enter lab value] REAL;
SELECT [SMC Database].*, sampling.[geom (i)]
FROM [SMC Database], sampling
WHERE [SMC Database].[parameter] = [Enter parameter]
AND [SMC Database].[Lab Value] > [Enter lab value]
```

In this parameter query, the user must supply the parameter of interest and the threshold value. In the IMS application, a small modification of HTML code creates a dropdown list of the parameters so the user does not have to worry about proper spelling, or even remember the names of the parameters.

After selecting the parameter and entering the threshold value, the query once again selects all historical data and the geometry of the monitoring station that meets the criteria of the *WHERE* clause. Again, since the geometry field of the monitoring station is selected the IMS application automatically zooms to the location.

3.4. Ability to view daily results of two USGS gauge stations in the watershed

There are two USGS gauge stations located in the watershed. These stations are continually monitored by the USGS, and their results are posted on the USGS site. The ability to view the most recent data from the gauge station was an important feature for the application to perform. Achieving this was simply adding an HTML link to open a new page on the USGS site for the gauge station, and present the last 24 hours of data.

3.5. Quantifying both the land use and roadway characteristics for each monitoring site's upstream contributing area

This query goes beyond standard SQL and introduces a spatial parameter. In this example the stakeholders were interested in quantifying the total acreage of each land use type and the total miles of roadway within the upstream contributing area for each monitoring site. The first query quantifies all roads within a selected catchment.

```
OPTIONS COORDSYS(''roads'' AS COMPONENT);
PARAMETERS [Enter monitoring site] INTEGER;
SELECT
  SUM(LENGTH((CLIPINTERSECT(roads.id,
    (SELECT id FROM catchments
    WHERE catchments.index =
      [Enter monitoring site]))),''m''))
    FROM roads
```

In this example the user enters the monitoring site number, and the SQL query uses a spatial clause to *Intersect* the roads layer and the catchment layer where the catchment.index represents the upstream area of monitoring site entered by the user. The result of the *ClipIntersect* clause clips out the roads that fall within the catchment area. The query then uses the SQL *sum* and *length* clauses to sum up all the lengths of the roads that were clipped by the catchment area. One should be able to see the power of integrating traditional SQL and spatial SQL within the same declarative statement.

Similarly, stakeholders desired the ability to quantify the land use categories within each catchment area. This too is accomplished through spatialSQL within the GIS, although the code will appear somewhat more complex.

```
OPTIONS COORDSYS(''roads'' AS COMPONENT);
PARAMETERS [Enter monitoring site] INTEGER;
SELECT SUM(AREA(lugeom,''m'')), luclassdes
FROM
  (SELECT luclassdes, lugeom
  FROM (SELECT
       CLIPINTERSECT(landuse.id,
            (SELECT id FROM catchments
            WHERE catchments.index =
            [Enter monitoring site])) AS lugeom,
  luclassdes
              FROM landuse))
GROUP BY luclassdes
```

3.6. Creating an Internet application

The IMS application automatically utilizes the stored queries within the desktop application. However, some HTML customization is necessary to facilitate the site's ease of use. For example, rather than allowing users to enter parameter information such as the monitoring criteria or the site number, a simple drop down box with those variables listed was added to the IMS template. Once the user selected the value from the dropdown box, the data were passed to the mapserver in the form of a query parameter—similar to what a user might enter by hand. In these cases, the results of the selected features are passed to a short javascript function that then sends the information to the mapserver.

Additionally, the placement of the Six Mile Creek specialized tools and the output tables were changed slightly to better improve legibility. Fig. 5 illustrates the general template used within the application.

4. Discussion

Using an IIMST strategy, the IMS application was rapidly developed and required virtually no coding other than the SQL queries presented in this paper. Also, we developed a second IMS site for the Virgil-Fall Creek watershed by re-using the SQL queries and HTML template illustrated earlier. The ability to create functional and robust IMS applications with little or no programming is critical for organizations lacking experienced IT and GIS staff. Typically, these organizations must contract out the effort to build the application, thus increasing the cost of development. Unfortunately, environmental monitoring by volunteer groups is not funded adequately to allow for data collection, monitoring, and web-based system development. Therefore, because the GIS and IMS development could occur so rapidly and inexpensively, the WRC was able to use their existing funds to concentrate on collecting water quality data.

The ease of implementation was augmented by the design goals of the WRC Monitoring Committee. There is a temptation when building an IMS application to recreate a GIS on the Internet. However, the Monitoring Committee understood that virtually all users of the IMS application would have little to no experience with GIS. Therefore, rather than allowing for dozens of spatial analysis capabilities, the Monitoring Committee decided to limit inquiry of the spatial databases to only a handful of tasks as previously described. Drop down boxes and query buttons included leading questions such as 'Zoom to specific monitoring site', 'View monitoring site history', or 'Quantify landcover for a monitoring site'. Therefore, users of the system did not require any training as the queries were self-explanatory.

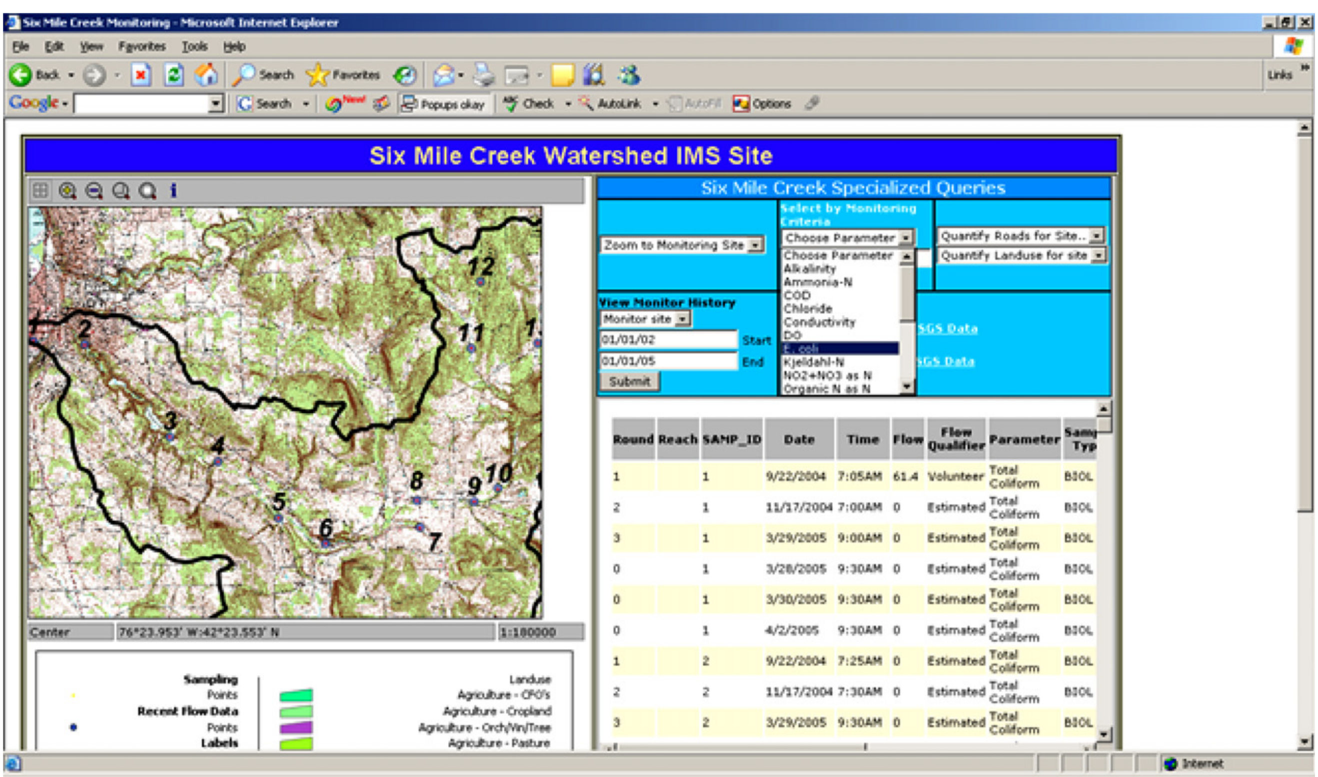


Fig. 5. Internet map server page for Six Mile Creek Watershed, showing display of maps and specialized GIS queries.

4.1. Limitations in system development

Unfortunately, while watersheds do not follow political boundaries, data collection methods do. Development of the application in the Six Mile Creek watershed was simplified by the fact that the entire watershed is located within Tompkins County, New York. Therefore, we could anticipate a consistent representation of GIS data sets. However, the Virgil-Fall Creek watershed is located in both Tompkins County and Cortland County. While both counties maintain GIS applications, their data storage formats, classification schemes, and individual data holdings differ. Therefore, all layers illustrated in Table 1 were clipped at the Tompkins/Cortland border and, including the Cortland County portion of Virgil-Fall Creek was not simply a matter of merging Cortland County's GIS data with Tompkins County data. For instance, the land cover classification system used in Tompkins County is different than the one used in Cortland County. At the moment certain functionality, such as quantifying land cover within each sub-basin, was removed from the Virgil-Fall Creek IMS application until improved data integration between the counties could be achieved.

4.2. Limitations in system use

Certain obstacles were identified during the testing phase. While GIS users are accustomed to understanding icons such as a zoom-in or zoom-out tool, many of the users were still confused by these options. Similarly, the ability to turn layers on or off was difficult for some users to understand. Therefore, the Monitoring Committee developed a short fact sheet to help users understand what each of the systems' functions actually does.

Also, while the system allowed users to select monitoring sites based on the chemical properties of the samples (i.e. levels of pH), many citizens do not understand the meaning of the values, and whether the results are significant or not. One concern of the Monitoring Committee is that users without an understanding of water chemistry may draw erroneous conclusions regarding the quality of the watershed. Therefore, mechanisms to better educate the public about the measurements taken and their meaning must also augment the IMS application. To assist with that effort the Monitoring Committee has asked a Cooperative Extension educator, focused on community based education, to provide a general fact sheet about the purpose of water quality monitoring and how to interpret the measurements taken within the streams.

4.3. A model for other volunteer groups

Perhaps the most valuable lesson from developing the Six Mile Creek IMS application was that of providing a reference framework for other volunteer monitoring groups. The use of an integrated IMS strategy that maximized SQL for performing tasks greatly reduced the development effort, and also reduced the maintenance overhead for the system. At the same time, understanding the limitations of the user base was important in designing a more modest site that was focused on primarily answering the pertinent questions often asked by citizens. This more modest design significantly reduced the complexity of the system (thus making it easier to implement and use), in addition to the complexity often introduced by more broad based IMS applications. The Monitoring Committee believed that developing multiple simplified sites to answer a few targeted questions was more beneficial than attempting to develop an all-encompassing IMS application that mimicked the use of an entire GIS over the Internet. A rule of thumb adopted by the Monitoring Committee was

> If you want to do GIS analysis, buy a GIS. If you want to answer targeted questions and spatially visualize the results, implement a modest IMS application to address those specific questions.

5. Conclusion

While GIS is an attractive technology for solving and communicating environmental issues, cost and complexity often prohibit its use. This is particularly true in the environmental science field where local groups frequently find themselves under-funded and under-staffed to implement leading commercial GIS technology. While innovations in the open source GIS field have produced functional tools for environmental management, implementing the applications is often difficult for users who lack the technical capabilities to integrate multiple software products together, often with little documentation. Unfortunately, many of the easier to use open source GIS products lack the sophisticated functionality of a robust commercial GIS. Enhancements to open source GIS products will hopefully continue and provide users with more user-friendly and robust tools.

This paper presented a case-study of a newer, low-cost GIS architecture that combines robust GIS processing, a user friendly interface often associated with Windows technology, and an IMS strategy that allows for the creation of rapid and highly functional IMS applications with virtually no programming. Also, the presentation of SQL and spatial SQL within the software architecture illustrates the ability to perform sophisticated GIS processing within a single declarative statement. All these benefits together allow for local monitoring groups to competitively participate in the use of GIS for environmental science.

Acknowledgements

This project was partially funded by the United States Department of Agriculture (USDA), Hatch Award NYC 125415. Also, the authors would like to thank Mr. Steve Penningroth for his assistance in compiling the Six Mile Creek database.

References

Alameh, N.S., 2001. Scalable and Extensible Infrastructures for Distributing Interoperable Geographic Information Services on the Internet doctoral dissertation, MIT Libraries, Cambridge, MA.

Braun, E.L., 1950. Deciduous Forests of Eastern North America. Hafner Press, NY.

Kropla, B., 2005. Beginning MapServer: Open Source GIS Development. Apress, Berkeley, CA.

Lakhani, K., von Hippel, E., 2003. How open source software works: "Free" user-to-user assistance. Research Policy 1451 (2002), 1–21.

Lembo (Ed.), 2004a. How do I do that in ArcGIS/Manifold: Illustrating Classing GIS Tasks. Cornell University.

Lembo, 2004b. Software Review: Manifold GIS. Geoplace Magazine.

Marks, P.L., Gardescu, S., 1992. Late Eighteenth Century Vegetation of Central and Western New York State on the Basis of Original Land Survey Records. New York State Museum Bulletin No. 484. The University of the State of New York, Biological Survey. Albany.

Mitchell, T., 2005. Web Mapping Illustrated. O'Reilly Press.

Nichols, D.M., Twidale, M.B., 2003. The Usability of Open Source Software. First Monday 8 (1).

NYSDOS, 1999. Draft Preliminary Characterization of the Cayuga Lake Watershed. New York State Department of State. Albany, New York.

Open GIS Consortium. 2003. Functionality Chart. Available at 〈http://www.geoplace.com/gr/webmapping/Final%20Web%20Charts.pdf〉.

Peng, Z.R., Ming, H.T., 2003. Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks. Wiley, Hoboken, New Jersey.

Peng, Z.R., Tsou, M.S., 2003. Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks. Wiley, New York.

Peters, D., 2003. System Design Strategies: An ESRI White Paper. ESRI Press, Redlands, CA.

Smith, B.E., Marks, P.L., Gardescu, S., 1993. Two hundred years of forest cover changes in Tompkins County, New York. Bulletin of the Torrey Botanical Club 120, 229–247.