

APPLICATIONS OF GIS AND DATABASE TECHNOLOGIES TO MANAGE A KARST FEATURE DATABASE

YONGLI GAO

Department of Physics, Astronomy, and Geology, East Tennessee State University, Johnson City, TN 37614, gaoy@etsu.edu

ROBERT G. TIPPING

Minnesota Geological Survey, 2642 University Ave., St. Paul, MN 55114, tippi001@umn.edu

E. CALVIN ALEXANDER JR.

Department of Geology & Geophysics, University of Minnesota, Minneapolis, MN 55455, alexa001@umn.edu

This paper describes the management of a Karst Feature Database (KFD) in Minnesota. Two sets of applications in both GIS and Database Management System (DBMS) have been developed for the KFD of Minnesota. These applications were used to manage and to enhance the usability of the KFD. Structured Query Language (SQL) was used to manipulate transactions of the database and to facilitate the functionality of the user interfaces. The Database Administrator (DBA) authorized users with different access permissions to enhance the security of the database. Database consistency and recovery are accomplished by creating data logs and maintaining backups on a regular basis. The working database provides guidelines and management tools for future studies of karst features in Minnesota. The methodology of designing this DBMS is applicable to develop GIS-based databases to analyze and manage geomorphic and hydrologic datasets at both regional and local scales. The short-term goal of this research is to develop a regional KFD for the Upper Mississippi Valley Karst and the long-term goal is to expand this database to manage and study karst features at national and global scales.

INTRODUCTION

Several Geographic Information Systems (GIS)-based Karst Feature Databases (KFDs) have been established in many karst regions and countries (Cooper *et al.*, 2001; Florea *et al.*, 2002; Lei *et al.*, 2001). However, the absence of comprehensive database models and standard metadata impeded the data compatibility and management (Gao *et al.*, 2005a). A KFD has been developed in Minnesota in the past few years (Gao *et al.*, 2002; Gao *et al.*, 2005a; Gao *et al.*, 2005b). The development, implementation, and data analyses of the Minnesota KFD are described in detail in the first author's Ph.D. dissertation (Gao, 2002). This paper discusses the last two phases of the process to develop the Minnesota KFD: Building and Implementing Applications and Database Monitoring and Maintenance.

Based on a survey from potential users of the Minnesota KFD, the applications built for the database need to have the following features: easy to use interfaces, compatible with Microsoft Access and ArcView GIS, reliable concurrency control and security, and network and Web accessible. In Minnesota, similar applications already exist in Minnesota Geological Survey's (MGS) County Water-well Index (CWI) Database Management System (DBMS). These applications are used by many users to access and maintain information about Minnesota's millions of water wells. The applications built for the KFD build on the lessons learned from MGS's CWI project.

One of the advantages of a relational database is the capability of data queries. Query languages can be used in a relational DBMS to modify the structures of a database, to update data

in the relations, and to extract useful information needed from the DBMS. The best known query language in relational databases is Structured Query Language (SQL). SQL is also called SEQUEL (Structured English Query Language) and was first designed in a relational database system called System R at IBM (Elmasri and Navathe, 1994). It is now used in most commercial relational DBMSs. Data queries were integrated into the applications to facilitate the functionality of the user interfaces. Stand alone data queries were also used to manipulate transactions of the database by the database administrator (DBA).

The DBA utilized the DBMS applications along with data queries and data logs to maintain the performance, consistency, and security of the KFD. Regular users can also use these applications to monitor the performance of the database.

BUILDING APPLICATIONS

Figure 1 illustrates two sets of applications built for the KFD. The five applications (data entry, data edit, data lookup, data report, and code maintenance) shown on the left side in Figure 1 are built in Microsoft Access. The five applications (digitizing, visualization, location verification, spatial transformation, and map generation) shown on the right side are built in ArcView GIS.

One set of applications, shown as the left portion of Figure 1, was written in Visual Basic programming language and linked to the KFD under Microsoft Access DBMS. These user interfaces are used to process karst feature records from paper-file formats. The other set of interfaces, shown as the right por-

6 correspond to the top-level karst feature index table of the database structure. The tabs (Sinkhole – Remarks) at the bottom section of Figure 6 are linked to the lower level data tables specific for each karst feature and its owner’s address and additional remarks through sub-form interfaces. These sub-form tab controls correspond to the lower level tables in the karst database structure. Figure 6 illustrates information about the spring (shown in Figure 2) entered through the data entry interface.

Figure 7 and Figure 8 are karst feature Data-Edit and Data-Lookup interfaces, respectively. The Data-Edit interface displays attribute values for a specific karst feature record allowing authorized users to modify these attribute values and even delete incorrectly assigned karst feature records. The Data-Lookup interface looks almost identical to the Data-Edit interface. However, Data-Lookup interface is read-only and limited to search and display attribute values of a karst feature record. The Data-Edit, Data-Lookup and Data-Entry interfaces (Figure 6) have very similar layouts which are compatible with the structure of the relational model of the karst feature DBMS.

The search button in both Data-Edit and Data-Lookup interfaces is connected to a smaller interface to search a specific karst feature record to be accessed by the Data-Edit or Data-Lookup applications. Figure 9 is such an interface that searches a karst feature record by asking users to enter its Relate ID or a combination of a feature’s county name, feature type, and numerical ID.

The print-log button in both Data-Edit and Data-Lookup interfaces is connected to a report interface to print out a formatted report sheet of a specific karst feature record. Figure

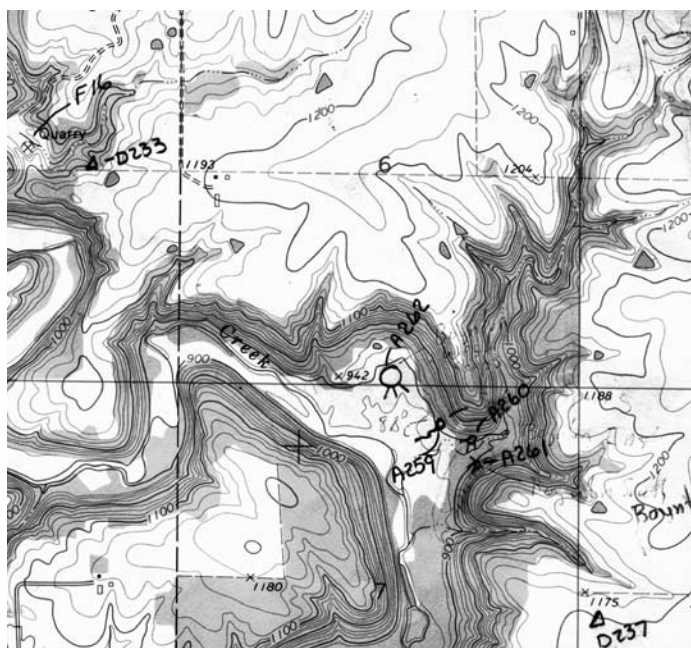


Figure 3. Karst features compiled on USGS 1:24,000 topographic maps. This example is from the Lewiston Quadrangle, Minnesota – Winona Co., 7.5-minute series topographic map, 1974. The area of section 6 is 1 square mile.

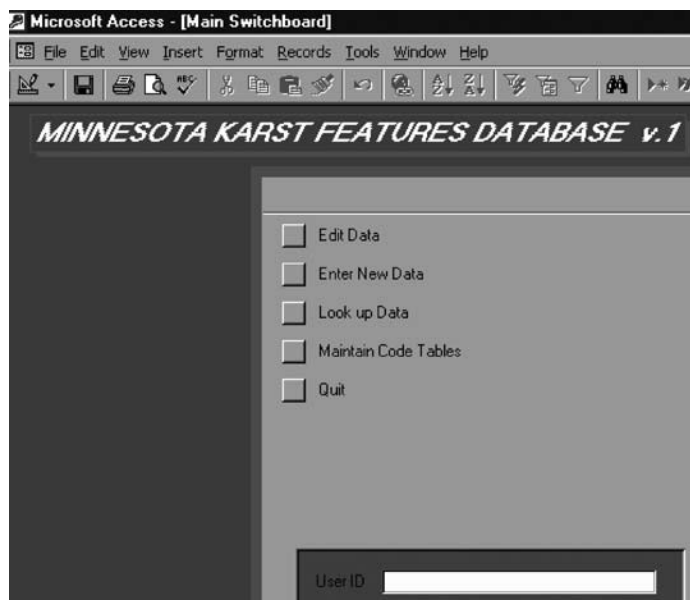


Figure 4. Main interface for the applications built for the karst feature DBMS in Microsoft Access.

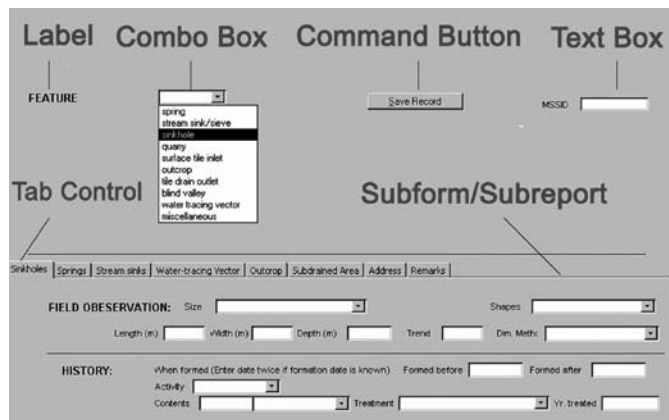


Figure 5. Most commonly used controls in the application interfaces of karst feature DBMS in Microsoft Access.

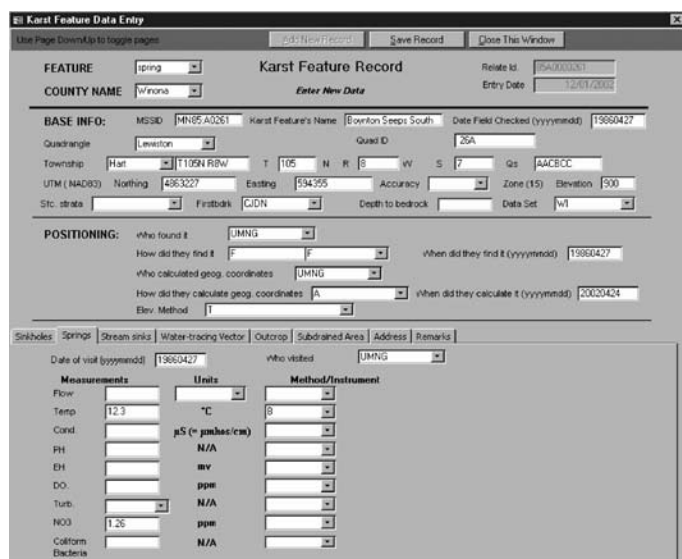


Figure 6. Data-entry interface written in Visual Basic programming language under Microsoft Access.

10 shows a two-pages report for a sinkhole whose Relate ID is 25D0000044. The first page contains information stored in the data tables of karst feature index, remarks, and owner's address. The second page prints out information specific for sinkhole record stored in the sinkhole data table. This sinkhole is also demonstrated in the search interface (Figure 9) and displayed in both Data-Edit (Figure 7) and Data-Lookup (Figure 8) interfaces. Notice that all the codes used in the data tables are converted to their full descriptions to make the report sheet more readable and understandable to the general public.

Figure 11 displays an interface to maintain code tables. Many of the code tables listed in Appendix D of Gao (2002) are subject to change with new information obtained for the karst feature DBMS. Some records in the code tables might get outdated and need to be removed in the future. This interface would allow authorized users to add a new record, delete an outdated record, and modify an existing record for any of the code tables used in the KFD. In Figure 11, records in the GCMCODE table are shown in Code-Table-Maintenance interface.

The Data-Entry, Data-Edit, Data-Lookup, and Code-Table-Maintenance interfaces are directly linked to the main interface. The Search and Report interfaces are accessible through the Data-Edit and Data-Lookup interfaces.

IMPLEMENTING APPLICATIONS IN ARCVIEW GIS

The applications in ArcView can access data in the KFD through Data Sources Open Database Connectivity (ODBC). The main application interface in ArcView GIS is shown in Figure 12. The structure of applications in ArcView GIS is different from the one in Microsoft Access. It consists of two windows. The left window acts as a control interface and the right window is a visualization window to display locations of karst features and maps or images related to locations or some attribute values of individual karst features. These images or maps are used to verify attributes of existing karst features or digitize new karst features and save them into the KFD. The controls in the main interface are similar to those in the interfaces in Microsoft Access. Authorized users utilize these controls to activate specific applications to control what to display in the visualization window.

As shown in Figure 1, the five applications in ArcView GIS are digitizing, visualization, location verification, spatial transformation, and map generation. In addition, applications in ArcView GIS have some comparable functions to those in Microsoft Access. Digitizing is a simplified Data Entry application based on a karst feature's locations on a map instead of its own geographic coordinates. Some basic feature attribute values can be displayed or changed by using controls such as Feature Attributes, Change MSSID #, and those combo boxes as shown on Figure 12 and Figure 13. The control Find Feature functions the same as the search interface in Microsoft Access to find and to display a specific karst feature based on its county number and numerical ID. However, applications in ArcView GIS can only access and modify a limited number of attributes for a karst feature record. These applications are designed for

Figure 7. Data-edit interface written in Visual Basic programming language under Microsoft Access.

Figure 8. A Data-lookup interface written in Visual Basic programming language under Microsoft Access.

Figure 9. Search interface used to search specific karst feature record for Data-Edit and Data-Lookup applications.

Figure 10. Report interface used to print out a karst feature record from Data-Edit and Data-Lookup applications.

CODE	DESC
	Digitized - scale 1:24,000 or larger
AS	Digitized from scanned image- scale 1:24,000 or larger
B	Digitized - scale 1:100,000 to 1:24,000
C	Digitized - scales smaller than 1:100,000
D	Digitized - scale not specified/unknown,
G	GPS (Global Positioning System / satellite);
H	GPS, accuracy < 12 meters (+ 40 feet or worse)
I	GPS, accuracy 3 to 12 meters (+ 5 to 40 feet)
J	GPS, accuracy 1 to 3 meters (+ 3 to 10 feet)
K	GPS, accuracy _ 1 meter (+ 3 feet or better)
M	Manual (estimated from map or description)
P	Visually digitized at parcel centroid from air
S	Surveyed
UK	Unknown

Figure 11. Code Table Maintenance interface written in Visual Basic programming language under Microsoft Access

point features such as sinkholes, springs, stream sinks and they cannot be used for line and polygon features.

The Add Feature and Digitize Location controls in Figure 12 and 13 can be used consecutively to digitize a karst feature based on its location on a USGS topographic map or an aerial photograph. Georeferenced digital raster graphic (DRG) and digital orthophoto quadrangle (DOQ) are used to match karst feature's location and topographic setting. A DRG is a scanned image of a U.S. Geological Survey (USGS) standard series topographic map (U.S. Geological Survey, 2001). The image inside the map neatline is georeferenced to the surface of the earth and fit to the Universal Transverse Mercator (UTM) projection. A DOQ is a computer-generated image of an aerial photograph in which displacements caused by camera orientation and terrain have been removed (U.S. Geological Survey, 1996). If a user clicks the save button on the interface, a karst feature's location information such as geographic coordinates, quadrangle, township, and range will be automatically entered into the database.

There are many visualization controls users can manipulate to decide what to display on the visualization window. Users can select what quadrangle of DRG or DOQ is to be displayed from the combo box Select Quad Name. The control button Find Feature can be used to find a specific karst feature record and its corresponding DRG and DOQ based on its geographic coordinates. Figure 12 demonstrates that users can find the spring entered through Microsoft Access interface (Figure 6) and its corresponding DRG through the main interface in ArcView GIS. Geographic themes such as county, quadrangle, and township are added to the visualization window by default to define and to verify the political boundaries of karst feature records. The visualization window in Figure 13 shows all these themes along with karst features, DRG, and DOQ in Wykoff Quadrangle, Fillmore County. In the main interface of Figure 13, Full-screen Crosshairs and Enable Map Tips at the bottom of the main interface are enabled to display a crosshair and karst feature label when the mouse pointer moves close to a karst feature. In this example, a sinkhole's label D2129 is displayed to identify this feature as a sinkhole in Fillmore County with Relate ID 23D0002129 (Label is composed by the code of a feature type plus the last four digits of the Relate ID).

Users can verify locations of many karst features based on their position on DRG and DOQ. Many karst features were mapped and marked on topographic maps. Many closed depressions on topographic maps and aerial photos can be identified on DRGs and DOQs in ArcView GIS. Once the location of a karst feature is seen to be incorrect, authorized users can move it by using the Adjust Location control in the main interface (Figure 12). Figure 14 illustrates that sinkhole locations in an area of Wykoff Quadrangle were verified and adjusted based on information from a DOQ. The locations of topographically closed depressions on DRG and DOQ are almost identical in this example. The sinkhole locations were placed, whenever possible, where the two maps agreed. The DOQ and DRG can be superimposed on the computer screen. Both maps are important and features visible on one map may be invisible on the other. When the locations of a topographically identified karst

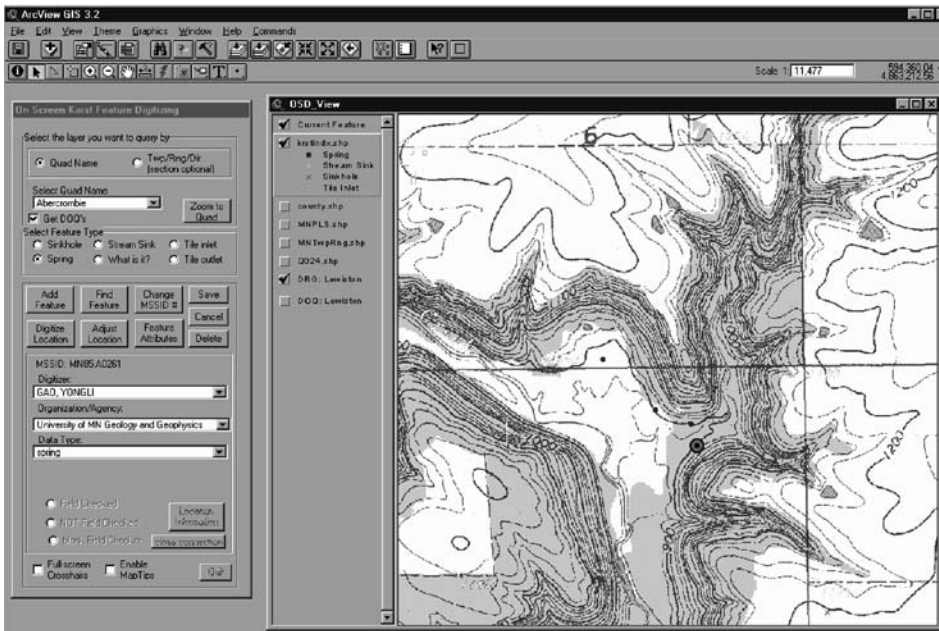
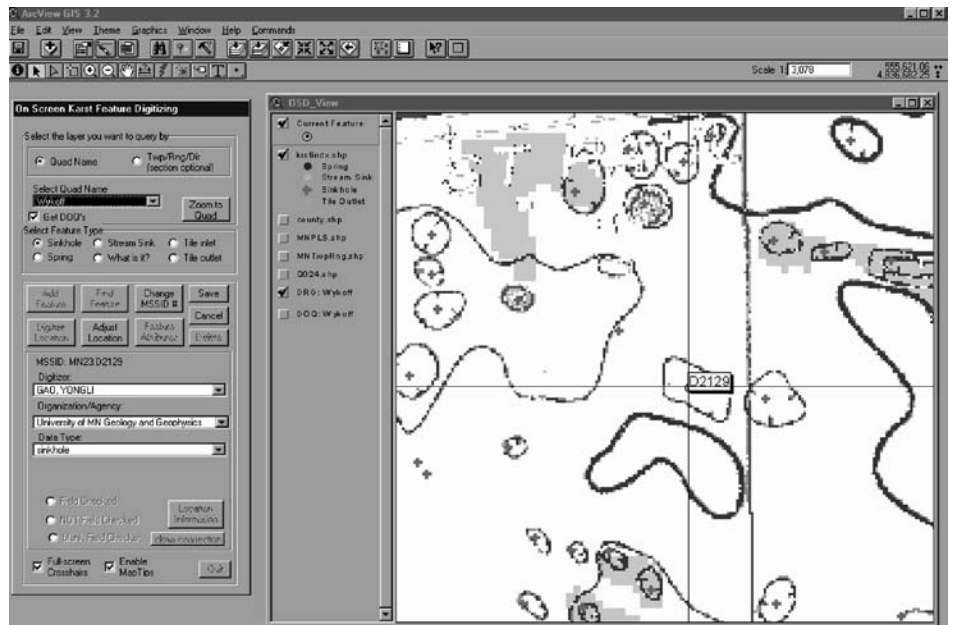


Figure 12. Main interface written in ArcView Avenue programming language under ArcView 3.2 environment. The background DRG in the visualization window is from the USGS Lewiston Quadrangle, Minnesota – Winona Co., 7.5 minute series topographic map, 1974.

Figure 13. Karst feature visualization through application interfaces in ArcView GIS. The background map in the right window is from the USGS Wykoff Quadrangle, Minnesota – Fillmore Co., 7.5 minute series topographic map, 1965.



feature don't match between DOQ and DRG, the DOQ is usually more reliable since this technology is relatively new and the topographic maps used for DRG are much older than the aerial photos used for DOQ. Some of the newly located sinkholes using a Global Positioning System (GPS) unit appear to be in the same locations as the closed depressions shown on DOQs.

In addition to all the tools discussed above, users can use all the available ArcView tools to visualize and analyze karst feature distribution in Minnesota. For instance, bedrock geology and surficial geology themes can be added to the visualization windows to verify or to define some geologic attribute values of any karst feature record.

DATA QUERIES

The karst feature DBMS uses five different kinds of SQL queries: select query, delete query, append query, update query, and make table query.

The select query is the most commonly used query in relational DBMS. The basic form of select query is a SELECT FROM WHERE block (Elmasri and Navathe, 1994). It has the following form:

```
SELECT <attribute list>
FROM <table list>
WHERE <criteria>
```

This form means select a list of attributes from a list of tables based on some criteria to retrieve information from the database. For example, the following query was used to retrieve

A



B



Figure 14. Verified and adjusted sinkhole locations on a DOQ in an area of Wykoff Quadrangle, Fillmore County using Location-Verification and Spatial Transformation tools in ArcView GIS. (A) sinkhole locations before spatial transformation. (B) sinkhole locations after spatial transformation. The cross on the aerial photo reflects sinkhole locations on the land surface.

attributes of all the features whose geographic coordinates do not exist in the karst feature index table. Select kfix.* means to retrieve all the attributes in the table kfix.

```
SELECT kfix.*
FROM kfix
WHERE (kfix.utme = 0 or kfix.utm = 0);
```

This query was used to identify all of the numbered features which had no location in the KFD so that the missing locations could be supplied.

If a select query retrieves information from more than one table, some tables need to be joined together to combine data records from different tables. The following query retrieves RELATEID, UTME, UTMN, and Shapes of all the large sinkholes in Goodhue County.

```
SELECT kfix.RELATEID, kfix.UTME, kfix.UTMN, kfsk.
SHAPES
FROM kfix INNER JOIN kfsk ON kfix.RELATEID =
kfsk.RELATEID
WHERE ((kfix.COUNTY_C="25") AND ((kfsk.
SIZE)="L"));
```

In Microsoft Access, both tables and queries could be joined together. The query above has an inner join statement which joins the karst feature index table and sinkhole table. When it finds matches of RELATEID in both tables, it combines those two records from the two tables and displays them as one record in the query's results. For an inner join statement, if one table or query doesn't have a matching record in the other table or query, neither record appears in the query's results. Other join

types such as left outer join or right outer join can select all the records from one table or query whether or not it has matching records in the other table or query.

The select query also includes functions such as COUNT, SUM, MIN, MAX, and AVG and more clauses such as GROUP BY, HAVING, and ORDER BY to make more complicated selections. The following query retrieves the number of karst features and counties in which more than 100 karst features exist from the karst feature index table. The results of the query are sorted by county code.

```
SELECT Count(kfix.FEATURE) AS NUMBER_OF_
FEATURES, kfix.COUNTY_C
FROM kfix
GROUP BY kfix.COUNTY_C
HAVING COUNT(*) > 100
ORDER BY kfix.COUNTY_C;
```

The delete query removes some records from a table based on certain criteria. The following query would remove any sinkholes in Winona County whose numeric IDs are greater than 1000. Since cascading delete is enforced between karst feature index table and lower level data tables, deleting a record from the index table would delete corresponding records in the related tables as well.

```
Delete Query:
DELETE *
FROM kfix
WHERE RELATEID < "85D0009999" and RELATEID
```

>= "85D0001000";

The append query inserts a set of records into a data table. The following query inserts all records from kfix85 into the table kfix. It loads the most recent Winona County sinkhole data set, which is converted as table kfix85, into the karst feature index table, kfix.

```
INSERT INTO kfix
SELECT kfix85.*
FROM kfix85;
```

The update query is used to modify attribute values of some selected records. The following query modifies all the karst features' labels (FEAT_LABEL) to match their RELATEID. The Mid function is used in this statement to combine the third character and the last four digits from the RELATEID to form a feature's label.

```
UPDATE kfix SET kfix.FEAT_LABEL = Mid([RELATEID],3,1)+Mid([RELATEID],7,4);
```

The make table query is used to convert the results of a query into a table in Microsoft Access. The following query selects all the karst features of Winona County in the karst feature index table and converts it to be an index table in the Winona County DBMS.

```
SELECT kfix.* INTO kf85ix IN 'kf85.mdb'
FROM kfix
WHERE (((kfix.RELATEID)<'85d0009999' And (kfix.RELATEID)>='85d0000001'));
```

In the karst feature DBMS, combinations of different SQL queries provide the fundamental information needed for the applications built on the KFD. The SQL query is also an essential tool for DBMS monitoring and maintenance.

DBMS MONITORING AND MAINTENANCE

After a DBMS is set up, it needs to be monitored closely to maintain its performance, consistency, and security. The KFD of Minnesota was originally put on a Windows NT server and switched to a Windows XP Citrix server in MGS. Authorized users can access the DBMS by connecting to the server as a Citrix client.

Table 1 lists four levels of users to access the KFD. The database administrator (DBA) is responsible for authorizing users to access the database and monitoring its use. The DBA has exclusive access to the DBMS to modify the structure and applications of the DBMS and recover lost or misplaced data due to hardware and software failures and human errors. Power users are authorized to access and update attribute values of karst features and code tables through all the application interfaces in both Microsoft Access and ArcView GIS. Regular users can access the DBMS through all application interfaces except code table maintenance application in Microsoft Access. They are not authorized to modify code tables used in the KFD. Guest users access the DBMS through read-only applications such as the Data-Lookup interface in Microsoft Access and the visualization interface in ArcView GIS. General public can access and download karst feature index files through MGS's website. The GIS-based karst feature index files are generated from the

DBMS on a regular basis to make the most up to date karst feature information accessible to the general public.

Database consistency and recovery are accomplished by creating data logs for the database. Data logs record entry, update, and delete dates for all the karst features in the DBMS. Original attribute values related to location and topographic setting of a karst are also saved in the data logs when it's updated or deleted. When hardware and software failure or human error occurs, the DBA can retrieve critical attribute values and put them back to the KFD. The karst feature DBMS is backed up to magnetic tapes and CD-ROMs at MGS on a regular basis. New and updated karst features are usually printed out as report sheets. Backup files and report sheets will be used for DBMS recovery if the karst DBMS is completely corrupted.

CONCLUSIONS

The KFD of Minnesota is a relational GIS-based DBMS. Microsoft Access and ArcView were used to develop application tools for the working database. Existing county and sub-county karst feature datasets have been assembled into the KFD capable of analyzing the entire data set. Data tables are stored in a Microsoft Access DBMS and linked to corresponding ArcView applications. The current KFD of Minnesota was moved from a Windows NT server to a Windows XP Citrix server accessible to researchers and planners through networked interfaces.

This working database also provides guidelines and management tools for future study of karst features in Minnesota. An inventory of current karst features has been put on MGS's website (<http://mgsnt4.mngs.umn.edu/karst>) to be accessible to the general public. The working database is still incomplete and more data will be put in the database in order to implement more complex statistical and hydrogeological models. More tools need to be built to make the data modules more powerful and interactive. Existing applications will be upgraded to process linear and polygon karst features. The methodology of designing this DBMS is applicable to develop GIS-based databases to analyze and manage geomorphic and hydrologic datasets at both regional and local scales. The KFD of Minnesota is being expanded to manage and study karst features in the Upper Mississippi Valley Karst (Gao *et al.* 2005c).

ACKNOWLEDGEMENTS

This research project was supported with funding from the Minnesota Department of Health (MDH) and conducted through the Minnesota Geological Survey (MGS). We thank Al Epp and Fuliiao Li in MDH for sharing their expertise in building GIS and Database applications. We thank an anonymous reviewer and Adam Kiel for their suggestions to improve the manuscript.

Table 1. Karst feature DBMS users and their authorized privileges.

DBMS User	Authorized Privilege
Database Administrator (DBA)	Grant and revoke privileges to other users and user groups Maintain DBMS consistency and security
Power User	Access the DBMS through all application interfaces
Regular User	Access the DBMS through all application interfaces except code table maintenance application in Microsoft Access
Guest User	Access the DBMS through read-only application interfaces
General Public	Access karst feature index files through MGS's website (http://mgsnt4.mngs.umn.edu/karst/)

REFERENCES

- Cooper, A.H., Farrant, A.R., Adlam, K.A.M., and Walsby, J.C., 2001, The development of a national geographic information system (GIS) for British karst geohazards and risk assessment, *in* Beck, B.F., and Herring, J.G., eds., *Geotechnical and Environmental Applications of Karst Geology and Hydrology*, Proceedings of the Eighth Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karsts: A.A. Balkema, Lisse, p. 125–130.
- Elmasri R., and Navathe, S.B., 1994, *Fundamentals of Database Systems*: Reading, Mass., Addison-Wesley, 873 p.
- Florea, L.J., Paylor, R. L., Simpson, L., and Gulley, J., 2002, Karst GIS advances in Kentucky: *Journal of Cave and Karst Studies*, vol. 64(1), p. 58–62.
- Gao, Y., 2002, Karst feature Distribution in Southeastern Minnesota: Extending GIS-Based Database for Spatial Analysis and Resource Management: [Ph.D. thesis] University of Minnesota.
- Gao, Y., Alexander, E.C., Jr., and Tipping, R.G., 2002, The Development of a karst feature database for Southeastern Minnesota: *Journal of Cave and Karst Studies* vol. 64(1), p. 51–57.
- Gao, Y., Alexander, E.C. Jr., and Tipping, R.G., 2005a, Karst database development, design and data assembly of karst features in Minnesota: *Environmental Geology*, vol. 47(8), p. 1072–1082.
- Gao, Y., Alexander, E.C., Jr., and Barnes, R.J., 2005b, Karst database implementation, analysis of sinkhole distribution in Minnesota: *Environmental Geology*, vol. 47(8), p. 1083–1098.
- Gao, Y., Alexander, E.C., Jr., Bounk, M., and Tipping, R.G., 2005c, Metadata development for a multi-state karst feature database, *in*: Beck, B.F., ed., *Proceedings of the Tenth Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karsts*: ASCE Geotechnical Special Publication, No. 144, p. 629–638.
- Lei, M., Jiang, X., and Li, Y., 2001, New advances of karst collapse research in China, *in* Beck, B.F., and Herring, J.G., eds., *Geotechnical and Environmental Applications of Karst Geology and Hydrology*, Proceedings of the Eighth Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst: A.A. Balkema, Lisse, p. 145–151.
- U.S. Geological Survey, 1996, Standards for digital orthophotos: <http://rockyweb.cr.usgs.gov/nmpstds/doqstds.html>.
- U.S. Geological Survey, 2001, Standards for digital raster graphics: <http://rockyweb.cr.usgs.gov/nmpstds/drgstds.html>.