GIS DATABASE DESIGN FOR APEX MODEL

Ming-Chih Hung
Department of Geology/Geography
Northwest Missouri State University
Maryville, MO 64468
mhung@nwmissouri.edu

R. Paul Duckworth
Blackland Research & Extension Center
Texas AgriLife Research
Texas A&M University System
Temple, TX 76502
pduckworth@brc.tamus.edu

Yi-Hwa Wu
Department of Geology/Geography
Northwest Missouri State University
Maryville, MO 64468
ywu@nwmissouri.edu

Jamie Patton
Department of Agriculture
Northwest Missouri State University
Maryville, MO 64468
jamiejp@nwmissouri.edu

ABSTRACT

Blackland Research and Extension Center (BREC), Texas AgriLIFE Research, Texas A&M University System maintains a system, the Agricultural Policy/Environmental eXtender (APEX) model, which allows users to input agricultural and environmental data for a particular field or area, store and output that data in a predetermined format. Initial development of this system required text input from users, including spatial data and attribute data. Over the years, improvements were made to create a graphic user interface (GUI) to allow interactive input from users. A recent project focused on developing a database management system (DBMS) for managing spatial data, as well as attribute data, in a GIS environment. The design and implementation of the database was based on two essential components: farmland parcels (fields) and farmers (producers). Each producer is represented as a record containing aspatial data and has his/her own database comprised of field data for parcels of land of which he/she owns or works on. In this database, each parcel of farmland is represented by a unique identification number. Such field data is linked to a process to delineate fields into subfields based on hydrologic and terrain characteristics. The output data from the delineation process is stored in a separate database named after a combination of the producer name and the unique identification of the field.

INTRODUCTION

Since the announcement of the Resources and Conservation Act (RCA) in the 1980s, there has been increased demand for cropping system simulation models to simulate agronomic and environmental impacts of land management decisions. The development of the EPIC (Erosion-Productivity Impact Calculator) model was a response to such needs. EPIC is a field-scale model based upon physical and chemical components of a field, such as erosion rate, nutrient balance, crop growth, etc and is used to simulate the effects of erosion on water quality, as well as examine sediment, nutrient and pesticide transport processes. The APEX (Agricultural Policy/Environmental eXtender) model was developed in the 1990's to allow for the use of the EPIC model at a multi-field level (Blackland Research and Extension Center, 2010). The APEX model incorporates additional hydrologic and erosional parameters, not used in the EPIC model, such as routing of water, stream locations, and watershed outlets, to simulate water flow at the field, farmland, and watershed scale (Gassman et al., 2009).

More recent projects have added increasing flexibility to the APEX model, by including such features as the ability to delineate a field into subfields based upon hydrological features and elevation, incorporating a graphic user interface (GUI) for field data input and a database management system (DBMS) for farmer and farmland entries, and integration with Geographic Information System (GIS). This project focuses on the creation of a DBMS for farmer and farmland entries within a GIS environment.

The central idea of this project was to create a DBMS to manage the input of farmer and farmland data from users, to manage the feed of farmland data into the delineation process (within APEX framework) for identifying subfields, and to manage the reception and storage of resultant subfields from the delineation process.

GIS DATABASE DESIGN

The database design and implementation portion of this project focused on two major components: farmer data and farmland data. Data pertaining to individual farmers in this projectwill be subsequently referred to as producer. Farmland data will be referred to as field, not to be confused with the term field in the database which refers to the space used to store one attribute in a GIS data. An individual farmer can own and/or operate multiple parcels of farmland. On the other hand, although a parcel of farmland could be co-owned or co-worked by multiple farmers, only one farmer is identified as the primary owner and/or worker of any single field.

With such settings and design, a farmer is the identity for storing owner/worker properties, mostly aspatial data not associated with the farmland itself, such as names, address, etc., while farmland data identifies the land's location, such as county, township, range, section, etc. as shown in figure 1. According to ER-modeling under normal circumstances, there should be two individual databases, one for farmers and the other for farmland (Allen et al., 2003; Date and Darwen, 2006; Halpin and Morgan, 2008; Harrington, 2009). Each farmer would be a record in the farmer database. Each parcel of farmland would be a record (and a polygon) in the farmland database. In addition, there would be the need for an ID in the farmland database to serve as the key to establish linkage between farmland and farmer databases. With this design, one farmer may own or work on multiple farmlands.

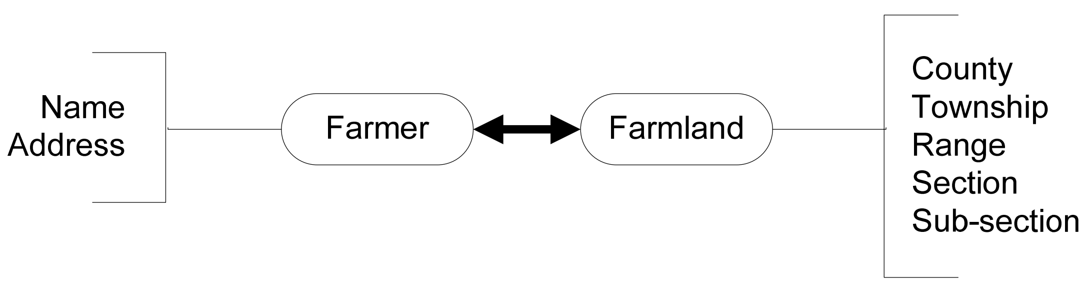

Figure 1. A conceptual diagram to depict the relationship between farmer and farmland, and associated attributes.

However, considering the work associated with delineation of each farmland parcel into subfields, there is the need to leave space for processing spatial features. It was then decided to divide the farmland database into several individual databases, as shown in figure 2. It was also decided to establish a single database for farmers, in which each farmer would represent a record. Once the farmer database is created and farmer data entered (one of the required data is producer ID, which is an arbitrary ID for each farmer), an individual database would be created by the name of the producer ID. This individual farmland database is used to store all farmland data (spatial features) associated with an individual farmer.

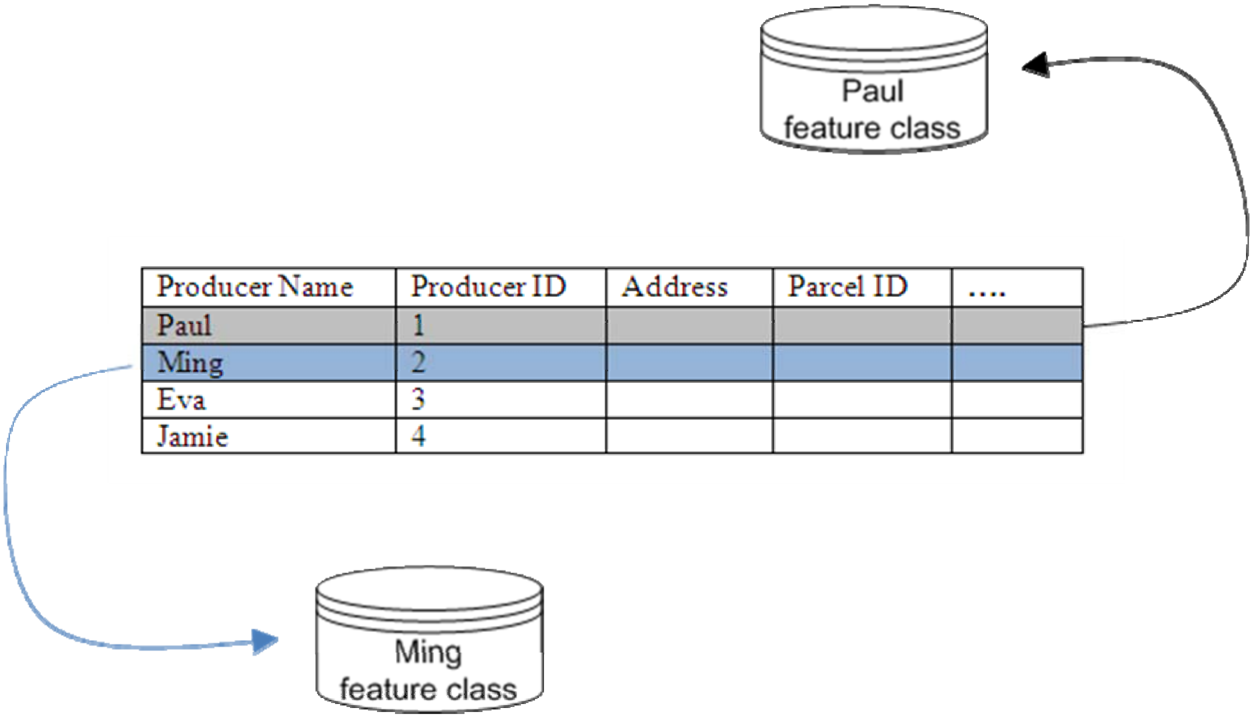

Figure 2. Implementation of farmers and farmland as individual databases.

Farmer/Producer Database

The design and implementation of the farmer database is relatively straight forward. It is a stand-alone table, without any associated spatial features. A graphic user interface (GUI) will be designed and developed for data entry. Table 1 summarizes the necessary attributes for the farmer database.

Table 1. Attribute definitions for the producer table. * indicated required input from users

Attribute	Type	Description
ObjectID	long integer	mandatory attribute, internally created, unique to each producer
IID	string	Internal ID, not used currently, calculated by = ObjectID
*SSN	string	Social Security Number, unique to each producer, required input from user
*ProducerID	string	unique to each producer, required input from user
*FirstName	string	required input from user
*LastName	string	required input from user
Address	string	
City	string	
Zip	string	
Phone	string	

Farmland/Fields Database

The design and implementation of the farmland database is manifold in a hierarchy (Bolstad, 2008; Chang, 2010; Longley, et al., 2005; Worboys and Duckham, 2004) as shown in figure 3. First, after the record for a particular farmer has been created in the farmer database, an empty database (termed feature class by ESRI ArcGIS software) named by the unique farmer ID (Producer ID) will be created. This feature class is used to store the farmland boundaries and associated attributes. A farmer may have multiple pieces of farmland. Each farmland parcel is stored as a polygon (feature) in the feature class databaseand assigned a unique ID, called Parcel ID. When the time comes to delineate a farmland into subfields based on topography, each polygon (feature) will be processed in turn, and the results of such delineation will be stored in a separate and individual feature classes, named by the combination of Producer ID and Parcel ID.

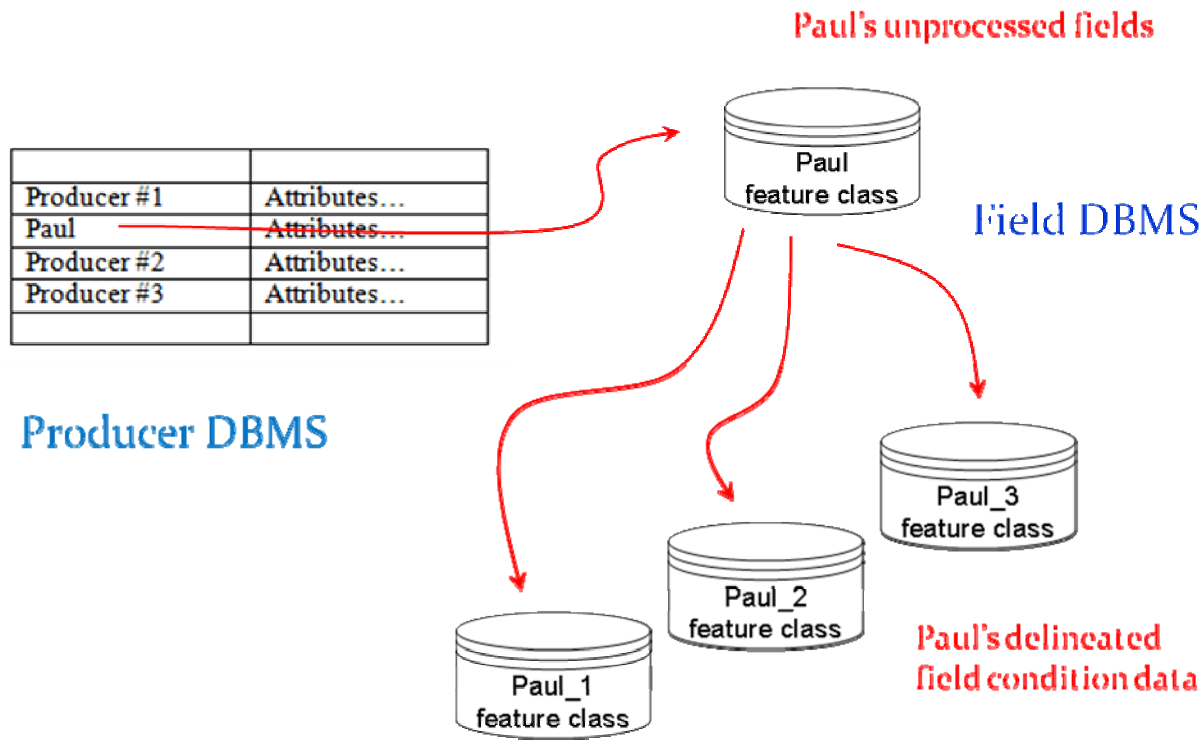

Figure 3. Farmland database to accommodate multiple polygons associated each farmer

File/Folder Structure

The creation of databases was done by VBA (Visual Basic for Application) codes. The file/folder structure of these codes is shown in figure 4. There are two major folders, c:\FMT and c:\fieldhydrotools. Under the c:\FMT folder, there are subfolders: DB and WINAPEX. The subfolder c:\FMT\DB is for databases, such as county boundaries, state NAIP (National Agriculture Imagery Program) images, hydrologic data, elevation data, etc. Also stored in this folder are the farmer database and the farmland database. These two databases are stored in a subfolder c:\FMT\DB\FarmPolygon. The subfolder c:\FMT\WINAPEX is for programs, such as VBA codes developed by this project and APEX programs. The folder c:\fieldhydrotools is for temporary working space for field delineation work, which is done by Python scripts, developed by another project.

CONCLUSION

There are two major data components involved in the APEX model: farmer and farmland. Under normal ER modeling, farmer would be an individual database, while farmland would be another individual database. However, considering the amount of work involved in and amount of storage space needed for further delineation of farmland into subfields based on hydrological and elevation characteristics, it was decided to expand the farmland database into manifold databases, in a hierarchy. The top level is the database for the farmland for a particular farmer. A farmer may own or work on multiple parcels of farmland. All of farmland data is stored in the top level database, named after the farmer's Producer ID. Each piece of farmland is assigned a unique ID, called Parcel ID. During the delineation process, each farmland parcel is processed in turn. These resultant subfields are stored in an individual and independent database, named after the combination of Producer ID and Parcel ID.

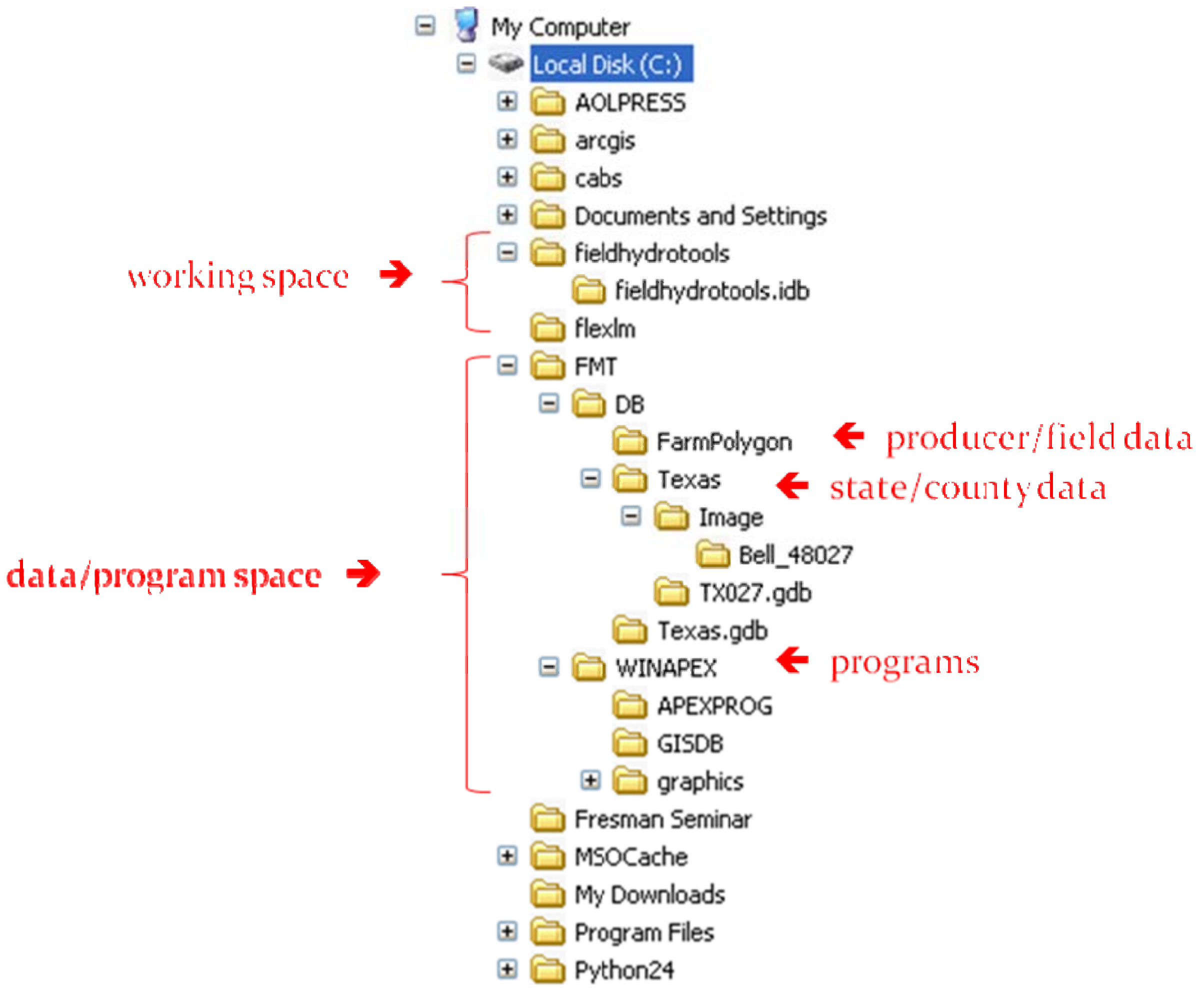

Figure 4. Folder structure as seen in My Computer in Microsoft Windows XP.

REFERENCES

Allen, C. & Creary, C. & Chatwin, S., 2003. *Introduction to Relational Databases*, (New York, NY: McGraw-Hill).

Blackland Research and Extension Center, 2010. EPIC and APEX Models, available online from http://epicapex.brc.tamus.edu/, accessed March, 5, 2010.

Bolstad, P., 2008. *GIS Fundamentals*, (White Bear Lake, MN: Eider Press).

Chang, K.-T., 2010. *Introduction to Geographic Information Systems, Fifth Edition*, (New York, NY: McGraw Hill).

Date, C.J. & Darwen, H., 2006. *Databases, Types and the Relational Model, Third Edition*, (Addison Wesley).

Gassman, P.W. & Williams, J.R. & Wang, X & Saleh, A & Osei, E. & Hauck, L.M. & Izaurralde, R.C. & Flowers, J.D., 2009. The Agricultural Policy Environmental EXtender (APEX) Model: An Emerging Tool for Landscape and Watershed Environmental Analyses, Technical Report 09-TR 49, Center for Agricultural and Rural Development, Iowa State University.

Halpin, T. & Morgan, T., 2008. *Information Modeling and Relational Databases, Second Edition*, (Morgan Kaufmann).

Harrington, J.L., 2009. *Relational Database Design and Implementation, Third Edition*, (Morgan Kaufmann).

Longley, P.A. & Goodchild, M.F. & Maguire, D.J. & Rhind, D.W., 2005. *Geographic Information Systems and Science, Second Edition*, (West Sussex, England: John Wiley & Sons Ltd.).

Worboys, M. & Duckham, M., 2004. *GIS: A Computing Perspective, Second Edition*, (Boca Raton, FL: CRC Press LLC).