Development of Database and Graphic User Interface for Farmland Delineation in ArcGIS Environment Based on APEX Requirement

Wu, Y.-H.,[1] Hung, M.-C.[1] and Patton, J.[2]
[1]Department of Humanities and Social Sciences, Northwest Missouri State University, 800 University Drive, Maryville, MO 64468, E-mail: ywu@nwmissouri.edu, mhung@nwmissouri.edu
[2]Shawano County, University of WI Extension, Shawano County Courthouse - Room 101, 311 North Main Street, Shawano, WI 54166, E-mail: Jamie.Patton@co.shawano.wi.us

Abstract
A model and a set of tools were previously developed in a separate study. That model is to delineate farmlands into sub-fields based on topography to better simulate hydrological patterns with the assistance from APEX (Agricultural Policy Environmental EXtender) model for further processing. That model requires a lot of manual operations between and within steps, which introduces unnecessary human errors. A GIS database and graphic user interface (GUI) was developed in this study to streamline that model and tools. Because the farmland and delineated sub-fields may contain different attributes, a two-level hierarchical database was proposed. The top level database is used for storing farmlands, with one feature class for each farmer (a farmer may have multiple farmlands). The second level database is used for storing delineated sub-fields, with one feature class for each piece of farmland from which the delineated sub-fields are derived. The GUI provides menus and buttons for user to run the model.

1. Introduction

The Agricultural Policy Environmental EXtender (APEX) Model has been widely used in many agricultural landscape analyses to simulate best management practices or cropping systems (Gassman et al., 2010). Much of this line of work focused on farm or small watershed levels (Texas AgriLife Blackland Research and Extension Center, 2011), in which level the basic unit is farm or larger areas. From a micro management or individual producer perspective, it is essential to know the hydrological pattern in a field, and where erosion or sediment accumulation took place in a field. Knowing such patterns can lead to a more efficient practice on farmland management.

GIS (geographic information system) is a powerful tool for managing spatial data and attribute data (Longly et al., 2010). It has been used in many environmental studies (Clark et al., 2004, Xie et al., 2004, Tuppad et al., 2009, Osward and Treat, 2013 and Zurita-arthos and Mulligan, 2013). One of the major advantages of using GIS is the ability to process attribute data separately (Neményi et al., 2003). Yet at other times, attribute data could be linked to spatial data by "keys" in a relational database model (Loh et al., 1994, Arctur and Zeiler, 2004, Gustavsson et al., 2008 and ESRI, 2010). Because of its unique ability to integrate spatial data and attribute data, it is suitable to process and delineate farmlands into sub-fields, during which process new attributes will be created while some of the existing attributes may be removed.

A previous project developed a model and set of tools to delineate a field into subfields based on topography, such as slope, aspect, etc. (Robert Duckworth, personal communication, 2009). That model consists of 24 steps. The user has to define input data, processing parameters, and output data for most of the steps, if not all. The objective of this project is to design a GIS database and graphic user interface (GUI) in the ESRI (Environmental Systems Research Institute, Inc.) ArcGIS environment to streamline and manage that model, the set of tools, and all of the intermediate or final products. The final products of that model are input data to APEX model for further process. The procedures involved in that model and the methodology used to delineate farmlands is beyond the scope of this study.

2. Initial Database Design for Farmers and Farmlands

The database and GUI design in this project focused on two essential components: farmer (or producer) and farmland (or field, not to be confused with the field in database which refers to a column space used to store an attribute). To keep terminology

ISSN 1686-6576 /

consistent, farmer (instead of producer) and farmland (instead of field) will be used in most places. There are, however, some occasions in which producer (for farmer) and field (for farmland) are used due to actual implementation.

A farmer could have more than one piece of farmland, as well as work on more than one piece of farmland. On the other hand, a piece of farmland could be owned by more than one farmer, and may have more than one farmer working on it. It is a many-to-many relationship between these two components. For the purposes of this project, in situations where one farmland is owned or worked by multiple farmers, only one farmer will be identified as the primary owner or worker. This simplifies the many-to-many relationship to one-to-many relationship, where farmers could have multiple farmlands, but any farmland could be owned by only one farmer. Note, this design does not remove other owners from the official database. Such simplification is for this project only.

There are many attributes associated with farmers. One of them is essential to the design and implementation of the database in this project: Producer ID, which is a unique identification given to each farmer by the system manager. The Producer ID will serve as the primary key to establish the linkage between two databases. There are also many attributes associated with farmlands. Two of them are essential: Producer ID (referring to the primary owner of this farmland) and Field ID (a sequential number given to each farmland for identification).

2.1 Initial E-R Modeling and E-R Diagram

The GIS software chosen for this project is ArcGIS by ESRI. ArcGIS uses a relational database model (Arctur and Zeiler, 2004 and ESRI, 2010). E-R (Entity-Relationship) modeling is a process to reason and identify objects (entity), relationship between objects, and attributes associated with objects. According to the essential components identified above, figure 1 shows the E-R diagram of farmer, farmland, one-to-many relationship, and attributes for farmer and farmland.

In the final stage of E-R modeling is the creation of tables, normally one table for each entity with attributes as columns (Allen et al., 2003, Date and Darwen, 2006, Halpin and Morgan, 2008 and Harrington, 2009). According to figure 1, normal implementation is a table for each entity (Welch et al., 2002 and Gyllencreutz et al., 2007), in this case, a table for farmer and another table for farmland. In the farmer table, there are Producer ID, Address, and Phone attributes. Each farmer will be a record (occupies a row) in the farmer table. In the farmland table, there are Field ID, Producer ID, County, and Township attributes. Each piece of farmland will be a record in the farmland table.

In addition, there is a spatial component associated with farmland, but none for farmer. Farmland database needs to define the boundary of each farmland. Though farmer has an address, which could be considered a spatial component, there is no strong need to show locations of farmer's houses on the map. It was therefore decided that farmer database will be implemented as a stand-alone table without spatial component, while farmland database will be implemented as a feature class (each feature class includes an attribute table and sets of coordinates to define spatial components) within a personal geodatabase (Welch et al., 1995 and Gustavsson et al., 2008). Figure 2 shows the normal implementation of farmer and farmland databases according to the E-R diagram shown in figure 1.

3. New Database Implementation with GIS

Unfortunately, such normal implementation was found difficult to accommodate any further processes to delineate farmland. The farmland delineation processes were developed in a previous project. When each piece of farmland is delineated, many homogeneous sub-fields (hereafter referred to as D-field as delineated field based on topography, such as aspect, slope, water flow direction, etc.) will be created, along with many byproduct polygons

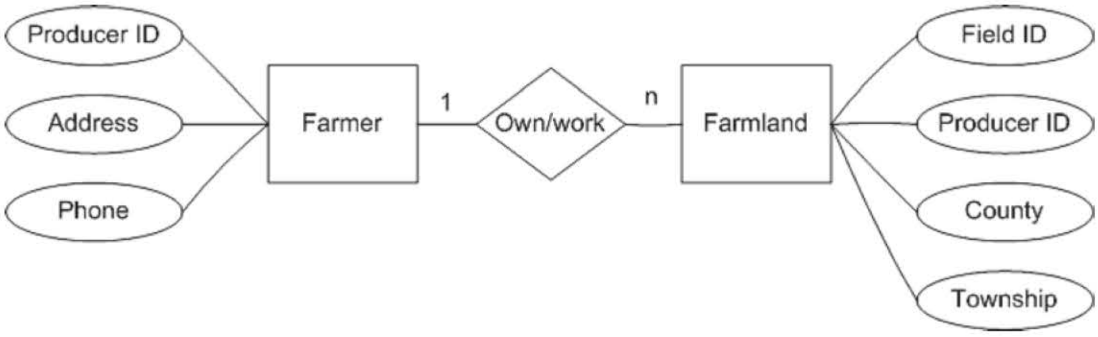
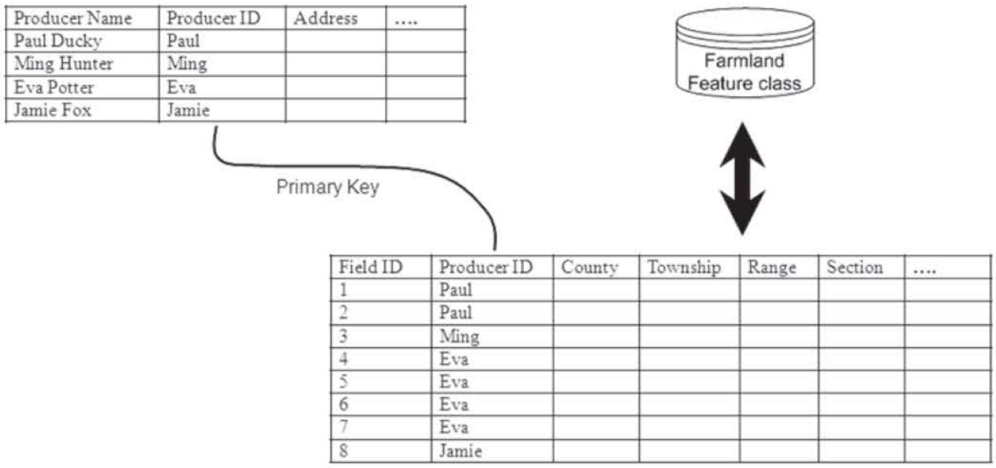


Figure 1: An E-R diagram shows the relationship between farmers and farmlands



Producer Name	Producer ID	Address	....
Paul Ducky	Paul		
Ming Hunter	Ming		
Eva Potter	Eva		
Jamie Fox	Jamie		

Field ID	Producer ID	County	Township	Range	Section	....
1	Paul					
2	Paul					
3	Ming					
4	Eva					
5	Eva					
6	Eva					
7	Eva					
8	Jamie					

Figure 2: Farmer and farmland databases according to E-R diagram in figure 1

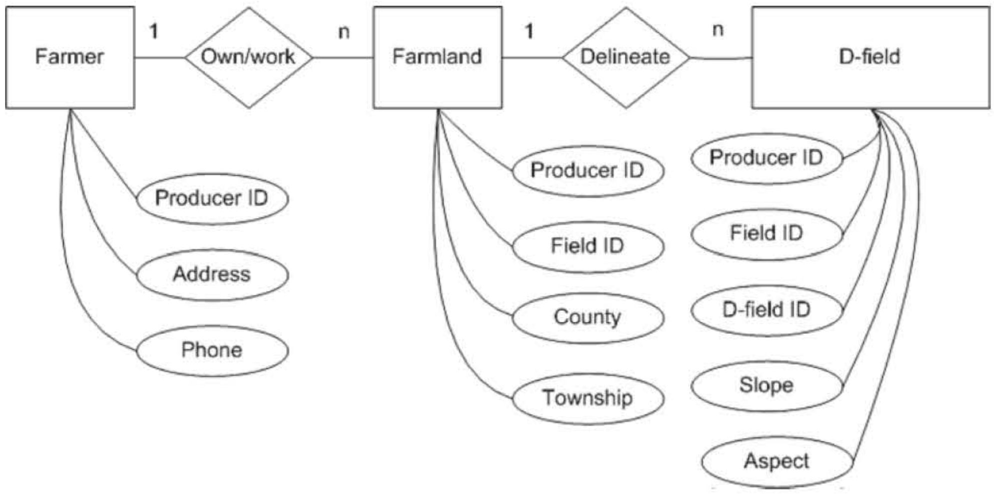


Figure 3: An E-R diagram shows the relationship between farmers, farmlands, and subfields

created by copy, buffer, intersect, union, and other operations. This delineation process causes two major problems for the normal E-R modeling implementation described above. The first problem is the fact that a piece of farmland, after the delineation process, will be divided into multiple pieces of land (D-fields). The second problem is the fact that these D-fields will have a different set of attributes from the original farmland.

One possible solution is to put all farmlands and D-fields in a mega-feature class, which contains all of the possible attributes. Unfortunately, this approach has a lot of waste in terms of space (Harmel et al., 2008). It is much desired to separate D-fields from farmlands and put them in separate feature classes (Sweitzer et al., 1996, Welch et al., 2002 and Gustavsson et al., 2008). However, D-fields should not be totally separated from farmlands because D-fields are sub-fields of farmlands. There is still a spatial linkage between D-fields and farmlands. A new database design and implement was then proposed, described in detail below, to link D-fields feature class to farmland feature class in a hierarchal structure.

3.1 New E-R Diagram

To better describe the relationship between farmers, farmlands, and delineated subfields, the E-R diagram in figure 1 is expanded as a new E-R diagram shown in figure 3. After the delineation process, each D-field contains more properties than farmland. Therefore D-field should be treated as a different object, with its own attributes. According to the new E-R diagram, the implementation of the farmer database as a stand-alone table needs no change. To accomplish the need for better management on farmland and D-fields, it was decided that the farmland database will be implemented as a collection of feature classes in a two-level hierarchical structure.

3.2 Implementation with GIS

On the top level, all of the farmlands belonging to the same farmer will be stored in a feature class, named after the farmer's Producer ID. As shown in figure 4, farmer Paul Ducky has a Producer ID "Paul". This Producer ID is used as the name of the feature class to store all of Paul Ducky's farmlands, which are two pieces of farmlands with Field ID 1 and 2, respectively.

On the second level of the hierarchical database design, after the delineation processes, each individual farmland and associated D-fields will be stored in its own feature class. The name for this feature class is a combination of Producer ID and Field ID. As shown in figure 5, all of the D-fields derived from Paul Ducky's first piece of farmland are stored in a future class named as Paul_1 (Producer ID_Field ID). All of the D-fields derived from Paul Ducky's second piece of farmland are stored in a feature class named as Paul_2.

Figure 6 shows an example of feature classes and farmlands on the top level implementation. As shown in figure 4, Paul Ducky (with Producer ID: Paul) has two pieces of farmlands with Field ID 1 and 2, respectively. These two pieces of farmland will be stored in Paul feature class. Following this pattern, Ming feature class will store one piece of farmland owned by Ming Hunter (with Producer ID: Ming). Eva feature class will store all of the four pieces of farmland owned by Eva Potter (with Producer ID: Eva). Jamie feature class will store one piece of farmland owned by Jamie Fox (with Producer ID: Jamie).

Figure 7 shows an example of feature classes and delineated farmlands (D-field) on the second level implementation. Applying the delineation process onto Paul's first piece of farmland (with Producer ID: Paul and Field ID: 1) results in 43 delineated fields (D-fields). All of these D-fields are stored in Paul_1 feature class. Paul_2 farmland generates 52 D-fields, and all of them are stored in Paul_2 feature class. Ming_1 farmland generates 48 D-fields, and all of them are stored in Ming_1 feature class.

4. Graphic User Interface Design and Implementation

To facilitate a smooth operation on the two-level hierarchical feature classes, a graphic user interface (GUI) has been designed and implemented in ArcGIS as a toolbar, as shown in figure 8. There are six items in the toolbar, representing six groups of functions: Parameters, Zoom, Producers, Fields, AutoProcess, and LinkWinApex. This toolbar controls the workflow in a desired order (Robert Duckworth, personal communication, 2009): the Parameter group controls the global variables, the Zoom group controls the view on the computer screen, the Producers group manages the farmer database, the Fields group manages farmland and D-

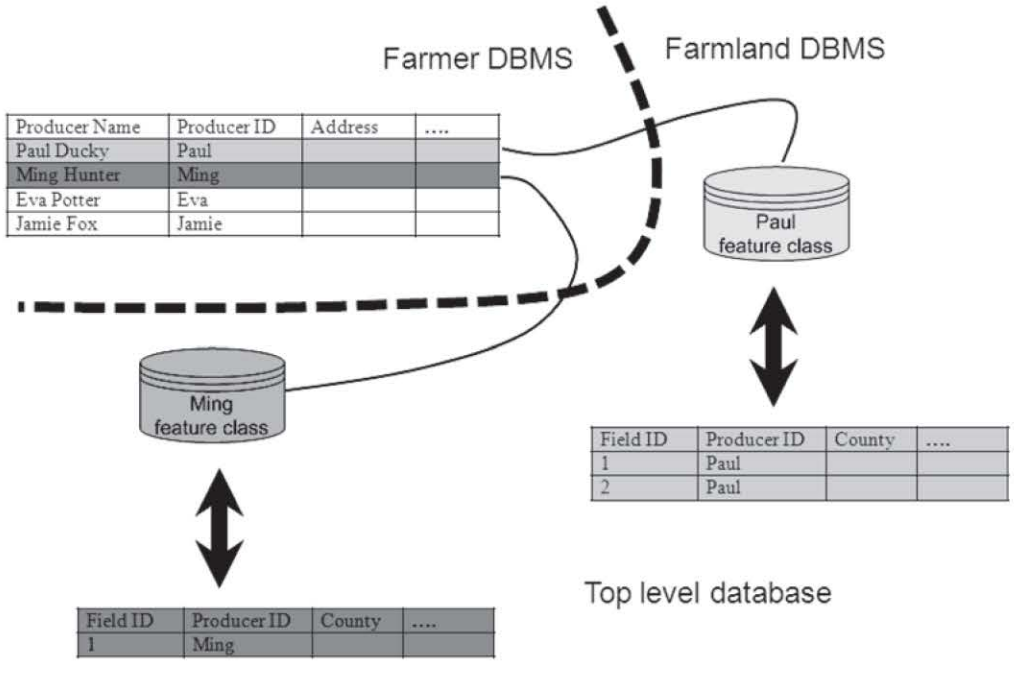
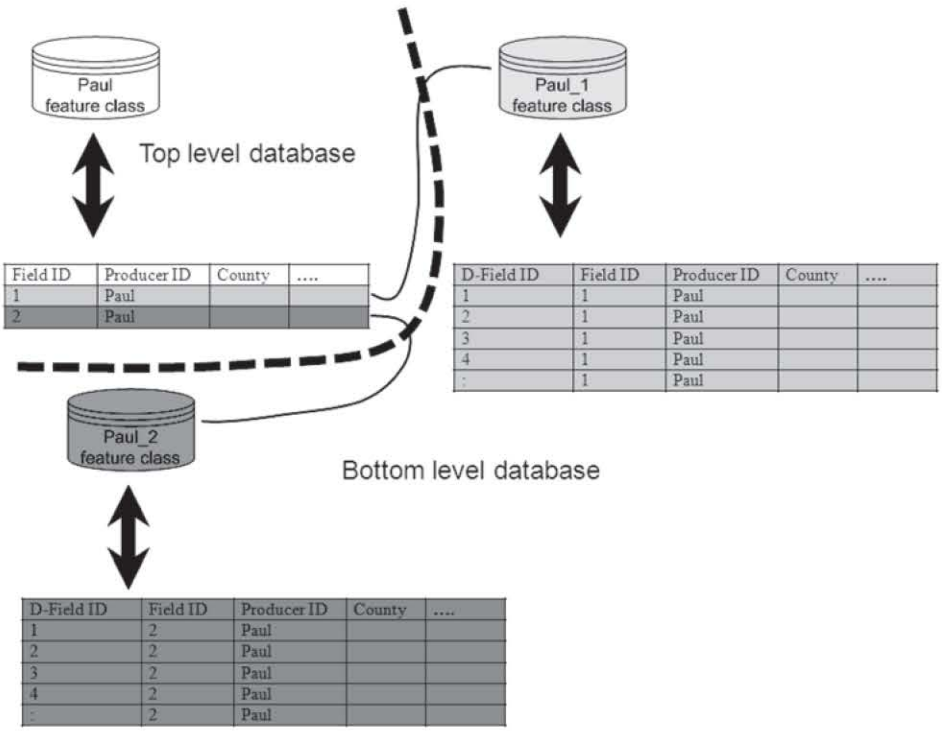


Figure 4: Tables and feature classes for farmers and farmlands on the top level implementation



Field ID	Producer ID	County	....
1	Paul		
2	Paul		

D-Field ID	Field ID	Producer ID	County	....
1	1	Paul		
2	1	Paul		
3	1	Paul		
4	1	Paul		
:	1	Paul		

D-Field ID	Field ID	Producer ID	County	....
1	2	Paul		
2	2	Paul		
3	2	Paul		
4	2	Paul		
:	2	Paul		

Figure 5: Tables and feature classes for delineated farmlands (D-field) on the second level implementation, taking Paul Ducky's farmlands as an example

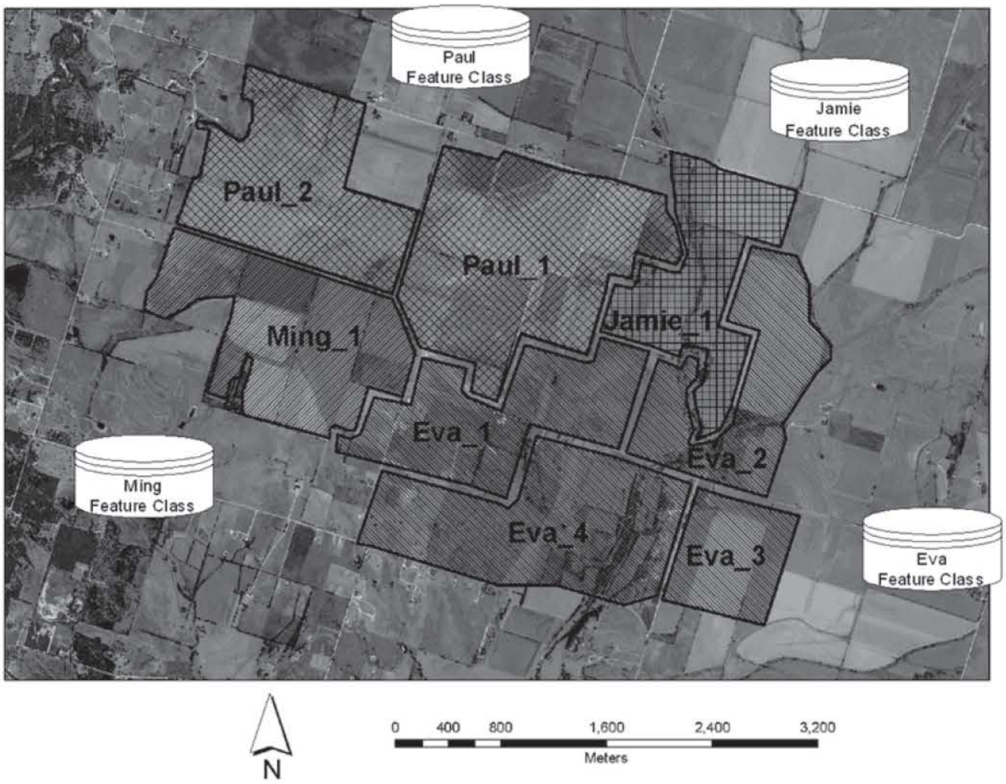


Figure 6: An example of feature classes and farmlands on the top level implementation. In this example, Paul feature class stores two pieces of farmland, Ming feature class one piece, Jamie feature class one piece, and Eva feature class four pieces.

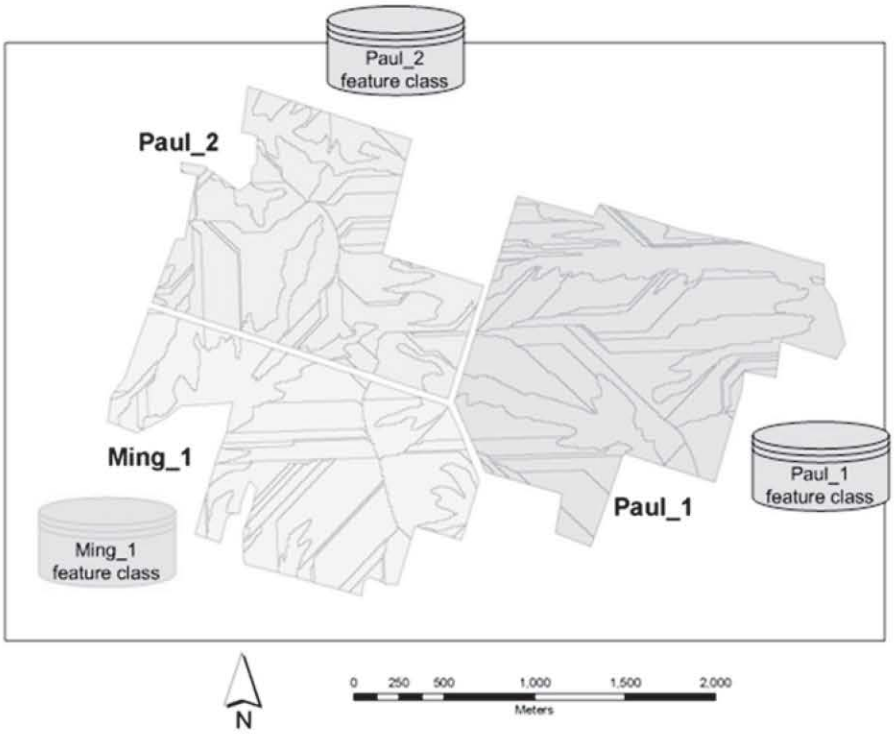


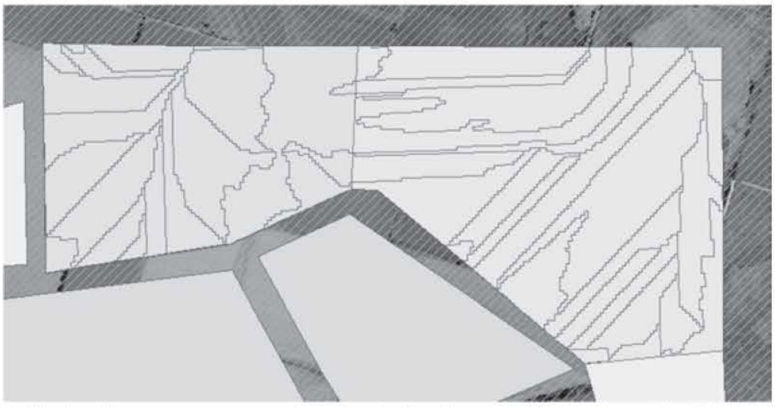
Figure 7: An example of feature classes and delineated farmlands on the second level implementation

Figure 8: The graphic user interface to facilitate a smooth operation in two-level hierarchical feature classes, as implemented in ArcGIS as a toolbar

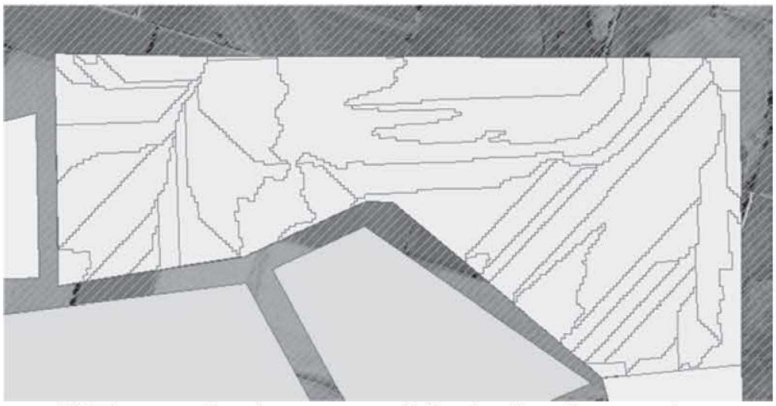
field database, the AutoProcess group controls the execution of the delineation process, and the LinkWinApex group passes the results to APEX model. Each group is explain below in greater detail. The Parameter item will bring up a dialog box allowing users to setup two parameters. The first parameter is to control whether to show or not to show message boxes while performing the AutoProcess to delineate farmlands. The second parameter is to control AutoProcess one farmland at one time and loop through all farmlands or AutoProcess all farmlands at the same time. The effect of the second parameter will be explained later under AutoProcess.

The Zoom item will bring up a dropdown menu allowing users to select Zoom to County or Zoom to Township. Each of the menu items will bring up an associated dialog box allowing users to select the desired area for zooming in.

The Producers item will bring up a dropdown menu allowing users to select New Producer, Select Producer, or Delete Producer. The New Producer menu item will bring up a dialog box allowing users to create a new farmer. During the creation of a new farmer, the user needs to enter the Producer ID, and it will be checked against the database to see if the Producer ID is unique. A duplicate Producer ID will not be allowed. In the meantime, a record for the new producer will be added to the farmer database, and the top level feature class for the new producer. However, the second level feature classes will not be created at this stage. They will be created at a later stage while performing the AutoProcess task. The Select Producer menu item will bring up a



(a) The first option: process one farmland at a time and loop through all farmlands.



(b) The second option: process all farmlands at the same time.

Figure 9: The differences between two options in AutoProcess. Notice the farmland boundary between these two adjacent farmlands in the second option is gone

dialog box allowing users to select or change the current farmer. All further work will be performed on the new or current farmer. The Delete Producer menu item will bring up a dialog box to confirm whether to delete the current farmer. If a farmer is deleted, all of the associated record in the farmer table, the top level features class, and the second level feature classes will be deleted as well.

The Fields item will bring up a dropdown menu allowing users to select Display Field or Select Field. The Display Field menu item will display all of the farmlands (not D-fields) owned or worked by the current farmer. The Select Field menu item will allow users to select an individual farmland, a group of farmlands, or all of the farmlands owned or worked by the current farmer for AutoProcess.

The AutoProcess item will perform the delineation process to the selected farmlands. There are two options to run the delineation process. The first option is to process one farmland at a time and loop through all selected farmlands, and the second option is to process all selected farmlands at the same time. The main difference between these two options is the farmland boundary during the delineation process. With the second option, the boundaries between adjacent farmlands will be resolved during the delineation process, with the resulting D-fields possibly crossing farmland boundaries. This option is particularly useful if a farmer or user is interested in knowing all of the D-fields among all of the farmlands he/she owns. Figure 9 shows the differences between these two options. Notice in figure 9 (b), the boundary between two adjacent farmlands is gone. Furthermore, in the second level implementation, all of the D-fields derived from the same farmland will be stored in an individual feature class, as explained earlier. Due to the potential impacts of D-field crossing farmland boundaries, a special feature class in the second level implementation must be created to accommodate this special AutoProcess. This special feature class will be named as a combination of the Producer ID and a suffix "_m". For example, Paul_m for all Paul Ducky's D-fields derived from the second option of AutoProcess, Ming_m for Ming Hunter's, and so on.

The LinkWinApex item will pass the resultant D-fields to Win APEX model for further analyses, such as crop prediction or watershed environmental assessment (Robert Duckworth, personal communication, 2009). Once data is passed on to Win APEX model, it is a separate process from this project and out of the scope of this paper.

5. Discussion and Summary

A previous project developed a set of procedures to delineate farmlands into smaller units (D-fields) with homogeneous topological conditions, such as slope, aspect, etc. Within these D-fields, basic hydrological characteristics are nearly uniform, such as flow direction, flow speed, etc. Due to such homogeneity, it is essential to store such topological characteristics as attributes for further analyses and query. Because the set of attributes for farmlands and the set of attributes for D-fields are different, it is desirable to store the farmland data and D-fields data in organized and separate databases for further analyses.

The normal approach of database implementation on the relationship between farmers and farmlands may be straight forward with a stand-alone table for farmers (only attributes because spatial component is not necessary) and a feature class for farmland (a database with spatial component to define farmland boundaries). Because of the delineation process and the difference between attributes for farmlands and D-fields, it is necessary to have another feature class for D-fields. Considering the overall relationship between farmers, farmlands, D-fields and intermediate byproducts during the delineation process, it was decided to have a hierarchical structure of two levels of databases, instead of just two feature classes. The first level is for farmlands, with one feature class for each farmer. The second level is for D-fields, with one feature class for each piece of farmland from which the D-fields are derived.

With the proposed approach, it is efficient to manage the databases in terms of inter-person or inter-organization transitions. Feature classes for farmlands and D-fields are named after the farmer who owns them. Some basic database management functions, such as file copy or delete could be performed without additional queries and operations. Though the proposed approach creates a large amount of feature classes, these feature classes are managed by the developed program/tools, and therefore hidden from users. Users do not deal with these feature classes directly. Users only interact with the databases through the developed tools. The perception of users on the program graphic user interface (GUI) / performance will not be affected by the large amount of feature classes.

Acknowledgements

This study was funded by U.S. Department of Agriculture, Natural Resources Conservation Service (USDA-NRCS) subcontract through Blackland Research and Extension Center (BREC), Texas AgriLIFE Research, Texas A&M University System. The authors would like to thank Mr. Robert Duckworth for his work and software testing in Blackland Research Extension Center during this project.

Reference

Allen, C., Creary, C. and Chatwin, S., 2003, Introduction to Relational Databases, New York, NY: McGraw-Hill.

Arctur, D. and Zeiler, M., 2004, Designing Geodatabases: Case Studies in GIS Data Modeling, Redland, CA: ESRI Press.

Clark, C. D., Evans, D. J. A, Khatwa, A., Bradwell, T., Jordan, C. J., Marsh, S. H, Mitchell, W. A. and Bateman, M. D., 2004, Map and GIS Database of Glacial Landforms and Features Related to the Last British Ice Sheet. *Boreas*, 33(4), 359-375.

Date, C. J. and Darwen, H., 2006, Databases, Types and the Relational Model, 3rd Edition, Addison Wesley.

ESRI, 2010, ArcGIS Desktop Help: Geodatabase and ArcSDE, Redland, CA: ESRI.

Gassman, P. W., Williams, J. R., Wang, X., Saleh, A., Osei, E., Hauck, L. M., Izaurralde, R. C. and Flowers, J. D., 2010, The Agricultural Policy Environmental Extender (APEX) model: An Emerging Tool for Landscape and Watershed Environmental Analyses. *Transactions of the American Society of Agricultural and Biological Engineers*, 53(3), 711-740.

Gustavsson, M., Seijmonsbergen, A. C. and Kolstrup, E., 2008, Structure and Contents of a new Geomorphological GIS Database Linked to a Geomorphological Map - with an Example from Liden, Central Sweden. *Geomorphology*, 95(3), 335-349.

Gyllencreutz, R., Mangerud, J., Svendsen, J. I. and Lohne, Ø., 2007, DATED–a GIS-Based Reconstruction and Dating Database of the Eurasian Deglaciation. Applied Quaternary Research in the Central part of Glaciated Terrain. *Geological Survey of Finland*, Special Paper, 46, 113-120.

Halpin, T. and Morgan, T., 2008, Information Modeling and Relational Databases, 2nd Edition, Morgan Kaufmann.

Harmel, D., Qian, S., Reckhow, K. and Casebolt, P., 2008, The Manage Database: Nutrient Load and Site Characteristic Updates and Runoff Concentration Data, *Journal of Environmental Quality*, 37(6), 2403-2406.

Harrington, J. L., 2009, Relational Database Design and Implementation, 3rd Edition, Morgan Kaufmann.

Loh, D. K., Hsieh, Y. T., Choo, Y. K. and Holtfrerich, D. R., 1994, Integration of a Rule-Based Expert System with GIS Through a Relational Database Management System for Forest Resource Management. *Computers and Electronics in Agriculture*, 11(2-3), 215-228.

Neményi, M., Mesterházi, P. Á., Pecze, Z. and Stépán, Z., 2003, The Role of GIS and GPS in Precision Farming. *Computers and Electronics in Agriculture*, 40(1-3), 45–55.

Osward, M. R. and Treat, C., 2013, Identifying Sea Level Rise Vulnerability using GIS: Development of a Transit Inundation Modeling Method. *International Journal of Geoinformatics*, 9(1), 1-10.

Sweitzer, J., Langaas, S. and Folke, C., 1996, Land use and Population Density in the Baltic Sea Drainage Basin: a GIS Database. Ambio, 25, 191-198.

Texas AgriLife Blackland Research and Extension Center, 2011, EPIC and APEX Models, available online from http://epicapex.brc.tamus.edu/ [accessed on Dec. 9,2011].

Tuppad, M., Winchell, F., Wang, X., Srinivasan, R. and Williams, J. R., 2009, ArcAPEX: ArcGIS Interface for Agricultural Policy Environmental Extender (APEX) Hydrology/Water Quality Model. *International Agricultural Engineering Journal*, 18(1-2), 59-71.

Welch, R., Madden, M. and Jordan, T., 2002, Photogrammetric and GIS Techniques for the Development of Vegetation Databases of Mountainous Areas: Great Smoky Mountains National Park. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(1), 53-68.

Welch, R., Remillard, M. and Doren, R. F., 1995, GIS Database Development for South Florida's National Parks and Preserves. *Photogrammetric Engineering and Remote Sensing*, 61(11), 1371-1381.

Xie, H., Zhou, X., Vivoni, E. R., Hendrickx, J. M. H. and Small, E. E., 2004, GIS-Based NEXRAD Stage III Precipitation Database: Automated Approaches for Data Processing and Visualization. *Computers and Geosciences*, 31(1), 65-76.

Zurita-Arthos, L. and Mulligan, M., 2013, Multi-Criteria GIS Analysis and Geo-Visualization of the Overlap of Oil Impacts and Ecosystem Services in the Western Amazon. *International Journal of Geoinformatics*, 9(2), 45-52.