doi:10.2489/jswc.70.5.288

Navigating spatial and temporal complexity in developing a long-term land use database for an agricultural watershed

T.L. Veith, J.E. Richards, S.C. Goslee, A.S. Collick, R.B. Bryant, D.A. Miller, B.W. Bills, A.R. Buda, R.L. Sebring, and P.J.A. Kleinman

Abstract: No comprehensive protocols exist for the collection, standardization, and storage of agronomic management information into a database that preserves privacy, maintains data uncertainty, and translates everyday decisions into quantitative values. This paper describes the development of a relational database intended to meet the agronomic and ecosystem interests of potential users from a long-term experimental watershed located in Pennsylvania, United States' Ridge and Valley physiographic province. We discuss the type and complexity of the data, which has historically been documented in free-form surveys collected through discussion with farmers. We detail the development process of a spatially and temporally explicit land management database and discuss the challenges in standardizing, without generalizing, 13 years of historic free-form data for 13 farms and 315 fields. Finally, we provide examples at field, farm, and watershed scales of how this database serves as a foundation for other data sets and modeling efforts that support research aimed at helping farmers meet long-term production, land stewardship, and water quality goals.

Key words: crop rotations—farm management—spatial—survey—temporal—uncertainty

Long-term, continuous, field- and farm-specific management data are critical in elucidating agricultural watershed processes. Agricultural land use and management vary substantially in both space and time, due primarily to human management interventions, but also to natural processes, both slow (plant community succession in abandoned lands) and sudden (catastrophic flooding). With the growth of long-term watershed databases (e.g., US Geological Survey [USGS]'s Water Quality Portal and USDA'S Sustaining the Earth's Watersheds: Agricultural Research Data System [STEWARDS]) comes the need for spatially explicit land management information that can be used to objectively assess ties in management with watershed trends in water quantity and quality. However, detailed land management data for even a single farm are rarely available to researchers striving to better understand biological processes and interactions in agricultural systems. Collecting and analyzing such data sets require substantial dedication in resources of labor and equipment. Additionally, privacy concerns typically limit availability of the data for research to the original collectors and their project-specific collaborators.

Various classes of spatial and temporal data complement field data collection. Some types of data—for example, elevation maps and soil surveys—are spatially referenced, but generally assumed temporally static at the time scale of field research. Other data sets, such as weather data and aerial photography, require both spatial and temporal reference. Correlation of these data with farm and field boundaries and land covers over time provides critical insight into the effects of management on the agroecological system. To facilitate research on management practices and their role in on-farm soil health and downstream water quality and quantity, ancillary data must be maintained at a relevant spatial resolution and with accurate temporal registration with respect to changing management practices.

The tools available for conducting this type of long-term agroecological research have changed dramatically over the past decades. Originally, farm management information within this experimental watershed was collected by asking farmers to maintain calendars of field management activities, or by interviewing farmers periodically and recording information on paper forms accompanied by hand-drawn maps. These methods introduce considerable uncertainty, especially in details of timing of field activities, rates and forms of amendment application, field conditions, and soil disturbance. For instance, "tillage was carried out in the spring" and "disk tillage followed by cultipacking was carried out on April 23 and 24" were both reported. Over time, aerial photography and then global positioning system (GPS) measurement of farm and field boundaries have been added to the survey process, reducing spatial uncertainty.

Even with improved spatial precision and accuracy, management surveys are inherently difficult to relate to each other. Extracting data from each individual year is complex, time-consuming, inefficient, and error-prone, and requires assumptions to be made that can differ between researchers. Common problems with long-term data collection include lack of metadata, lack of spatial reference, and inadequate documentation of processing and aggregation methods (Volk et al. 2014). Quality assurance, archiving, and especially maintenance of data documentation and metadata are all crucial components of the scientific enterprise (Baker et al. 2000; Le Duc et al. 2007).

Tamie L. Veith is an agricultural engineer, **James E. Richards** is a hydrologic technician, **Sarah C. Goslee** is an ecologist, **Amy S. Collick** is a research hydrologist, and **Ray B. Bryant** is a soil scientist at the USDA Agricultural Research Service (ARS) Pasture Systems and Watershed Management Research Unit, University Park, Pennsylvania. **Douglas A. Miller** is an associate professor of geography and the director of the Center for Environmental Informatics, and **Brian W. Bills** is a senior research assistant and assistant director of the Center for Environmental Informatics at The Pennsylvania State University, University Park, Pennsylvania. **Anthony R. Buda** is a research hydrologist at the USDA ARS Pasture Systems and Watershed Management Research Unit. **Ryan L. Sebring** is a graduate student in agronomy at The Pennsylvania State University. **Peter J.A. Kleinman** is a soil scientist and the research leader at the USDA ARS Pasture Systems and Watershed Management Research Unit, Pennsylvania.

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

Our chosen approach to overcome these difficulties was to develop a spatially referenced relational database to manage complex land use and management information over an entire watershed for many decades. Such databases are increasingly common; data sharing and archiving are increasingly important, both because of the proliferation of large long-term data sets and the emphasis on data sharing and collaborative research (Whitlock 2011; Wolkovich et al. 2012). Good data management tools aid in turning data into knowledge, increase efficiency, and improve our ability to scale up in space and time (Baker et al. 2000). Using a relational database makes data management much more efficient because it facilitates standardized data collection and consistent entry, as well as ensures data security by making it possible to easily back up data and to enforce access controls as desired (Le Duc et al. 2007). Standardized tools facilitate data analysis and visualization (Michener et al. 2011).

The experience of the National Science Foundation (NSF)'s Long-Term Ecological Research (LTER; Hobbie et al. 2003) network has provided valuable insights into developing such a database. Important considerations include (1) ensuring that data management tools are research driven and will meet the needs of participating scientists, (2) providing data documentation and enforcing data quality, and (3) ensuring that tools developed can be scaled to meet future needs (Baker et al. 2000). It is important for us, in turn, to document database development for our complex long-term data set so others can benefit from our experience (Le Duc et al. 2007). More recent efforts include the STEWARDS database for agricultural watershed research, which contains watershed data aggregated into a form suitable for public distribution without privacy concerns, but does not hold the detailed observational data needed for site-specific research and analysis in settings in which data privacy can be assured (Sadler et al. 2008; Steiner et al. 2008).

This paper documents development of a relational database designed to address the spatial and temporal needs described above. The database contains "one-to-many" spatial relationships at the field and farm level. For example, one farmer manages many fields, and each field is used for multiple crops, with many different tillage and fertilizer applications over time. The spatial relationships enable a long-term temporal management history of any particular field to be extracted from the database, as well as snapshots of a field or farm at a given point in time. Data uncertainty indicators are included for all temporal variables. In addition to reporting on the historical data collected, this paper details the process of developing the database structure and illustrates its potential through a variety of simple examples.

Materials and Methods

Data Collection. Field specific management data were collected from farms in the 7.3 km^2 (3.2 mi^2) USDA Agricultural Research Service (ARS) WE38 Watershed, Northumberland County, Pennsylvania, United States. All of the farms are independently owned and operated. Scientists and technicians have conducted long-term stream assessments and shorter-term plot experiments throughout the past 60 years (Bryant et al. 2011; Buda et al. 2011a, 2011b; Church et al. 2011). These efforts have focused on promoting environmentally sound land management by learning from farmers in the non-karst portion of Pennsylvania, United States' Ridge and Valley physiographic province and by better understanding this region's hydrology and nutrient transport processes.

General categories of land use in WE38 have remained largely constant in the past 25 years: 55% cultivated land, 3% continuous pasture, 2% developed, and 40% woodlands (Bryant et al. 2011). The 13 farms in the watershed produce primarily forage crops: corn (*Zea mays* L.) silage, winter cereal, soybeans (*Glycine max*), and alfalfa (*Medicago sativa*) or grass hay. Within the 400 ha (990 ac) of cultivated cropland, the most common agronomic rotation of corn–soybean–alfalfa (2 to 4 years) covers about 120 ha (300 ac). A similar rotation of corn (2 years)–soybean–small grain–hay (2 years) is seen on about 47 ha (116 ac), and continuous corn occurs on about 45 ha (111 ac). Although the general range of practices is narrow, actual crop rotation sequences, timing, and specific management vary constantly throughout the region. In addition, specialty crops can be important. On average over the 13 years of historic survey data, managed Christmas tree fields have occupied 17.2 ha (42.4 ac) of the watershed.

Field sizes across the watershed ranged from less than 0.1 ha (0.3 ac) for vegetable crops to 14 ha (35 ac) for forage crops with an overall average of 1.1 ha (2.6 ac). On 85% of the farms, field areas varied less than 1.2 ha (3.0 ac) over the 13 years. Such small shifts typically represent a cornfield increasing by a few rows and the adjacent soybean field decreasing, a change in setbacks, or perhaps space for a new shed. However, spatial land allocations do vary every few years in this region to accommodate crop changes, field consolidation, and splitting.

Several farmers rent land from other farms within the watershed from time to time, resulting in temporal variation in the operational size of a farm. For example, from 1998 through 2010, the mean farm size was 25.4 ha (62.8 ac) with 25 fields per farm. Three farmers physically owned less than 12 ha (30 ac) each, and two farmers owned more than 40 ha (100 ac), but across the watershed only five operators worked more than 35 ha (86 ac) annually and none worked more than 50 ha (124 ac). Both variation among operational farm sizes and discontinuity between operational land and physical ownership reflect the challenges of small (<50 ha), family-owned farms in responding to economic fluctuations, noncontiguous fields, and a minimal workforce. They also highlight the importance of documenting management and soil characteristics at field and subfield scales so that we can correctly infer causal relationships between land management, soil health, water quality, and ecosystem services.

The on-site ARS technicians, who reside locally, have close working relationships with the agricultural landholders and operators farming the region who are willing and able to share their data, expertise, and observations. Such cooperators manage more than 90% of the farmland within the WE38 Watershed, as well as a substantial portion of the encompassing Mahantango Watershed. For consistency in reporting, a single technician has led the dialogue with watershed farmers annually since the early 1990s to learn about their practices on each field over the prior year and exchange insights in a shared discovery mode. A general survey template, used by the technician each visit to record notes, has provided a basic structure for the visit. Information collected for each field included tillage type and timing; crop type; plant and harvest dates; manure type, application rate, and timing; chemical fertilizer or pesticide type, rate, and timing; field boundaries; and any other relevant information the farmer chose to share.

Also, this same technician observed field operations and crop production throughout

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

the year, during day-to-day work activities, to corroborate survey notes and maintained a comprehensive spatial record of annual changes in land use. This level of observation enabled the technician to be better prepared for farm visits, thus increasing the depth of discussion possible and creating stronger rapport. Likewise, the consistency in personnel strengthened rapport, promoted appropriate confidentiality, and reduced variation in survey notations that could easily result from multiple surveyors.

Database Development Process. The seven-step database development process (table 1) centered on the following five goals: capturing the full content of existing farmer surveys, maintaining temporal and spatial changes for each field, ensuring consistency and quality among entries, identifying clear indicators for the level of uncertainty of variables, and accommodating future watersheds and new research directions. Steps 1 and 2 involved monthly meetings among prospective users of the database to define the research questions that we ultimately wanted the database to help answer. The interdisciplinary nature of the group required focused discussions to agree conceptually on shared controlled terminology and resolve data ambiguities. Step 3 involved determining the attributes of farm management that should be captured in the database and understanding their relationships to each other (figure 1). In Step 4, the attributes were standardized, and data dictionaries were created containing allowable values. This step was vital in ensuring a robust ability to query the database and was tested in the Step 5 build phase. Step 6 was one of the most complex for this data set because the historical data had been, until now, archived in free-form survey notes and the institutional knowledge of the watershed technicians. Finally, Step 7 involved quality control and assessment of the database structure and the entered data.

Conceptually, our data model is organized around the farm field as the smallest discrete unit of agricultural management. Fields, which exist in space and time via spatially referenced boundaries with start and end dates, have one or more land covers (also with temporal bounds). Management attributes directly related to plant growth (i.e., planting, emergence, grazing, and harvest events) are associated with land covers. Other management attributes such as tillage, nutrient and pest management applications, conservation practices, and livestock are directly associated with fields. To support future research, ancillary data such as manure and feed storage infrastructure can be associated with the farms to which fields belong.

Table 1
Brief outline of the database development process.

Seven-step database development process	
1	Identify scientific themes and research goals.
2	Develop conceptual model.
3	Create relationship structure of management attributes.
4	Standardize attribute structures and allowable values.
5	Build and test the system.
6	Determine and document rule interpretation.
7	Quality assessment and control of the database and data.

Each management attribute is represented in the database as either a data table or a lookup table. The data tables contain values that record a particular event. For example, for every farm field, the tillage data table records tillage method, maximum plow depth, spacing, start and end dates, and date uncertainty. The lookup tables, which contain sets of predefined values (e.g., mulch tiller, moldboard plow, generic disc harrow, and chisel plow), are used to populate portions of the data tables (e.g., the tillage method) while minimizing inconsistencies and errors. A complete listing of data and lookup tables along with a brief description of each is provided in table 2.

During design, we discussed information that is applicable to current research interests, but has not been systematically collected, such as information on the livestock components of the farms. We ultimately limited the scope of initial development to include complex structures where sufficient data existed and general structures for accommodating sparse and missing data sets, such as the livestock operations, that are of interest to our current and planned research.

While future possibilities are wide ranging, we maintained a design focus that catered to the basic research interests of the authors and their active collaborators. Otherwise the database could quickly become unwieldy, and the development time could be delayed by addressing components for which the scientists had little experience or current interests. For example, some information on pesticides exists in documentation from farmer interviews, but pesticides have not been a focus of prior research. Thus, some pesticide information was retained and transferred into the database and some efforts were made to clarify and standardize entries, but not to the same degree of effort as was spent on other management practices.

The spatial component of the database includes polygons representing all parcels defined by a uniform land use (including nonagronomic uses such as barnyards, roads, waterways, and fallow areas) over a specified period of time. A unique identification number was assigned to each parcel polygon. When the boundary of a parcel changed, the parcel polygon was redigitized based on GPS measurements and aerial photographs. The identification number for the original parcel was retired on a specified day, and a new number was assigned to the newly digitized parcel from that day forward. For example, two agronomic fields separated by a grassed waterway (three parcels total) might be planted to wheat (*Triticum* spp.), continuously across both fields and the intervening waterway, resulting in a single new parcel with total area equal to the sum of the area of the previous two fields plus the area of the waterway. Although the time that this change occurred may have been recorded as a month or season, a specific day had to be chosen to represent the time of change to ensure that every parcel was represented once and only once during any given time period. Otherwise, parcels would be double counted or excluded, resulting in ambiguity in future research queries and in even the most basic data tallies (e.g., total area of cultivated land in the watershed).

Temporal uncertainties related to the occurrence of farming practices were unavoidable due to a desire to minimize reporting burden for the farmer and the reliance on memories of previous year events during the annual survey collection. Uncertainty levels were recorded throughout the database. For example, an operation occurring on "April 1" was assigned a date uncertainty of "day," and operations occurring in "April" or in "the spring season" were

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

assigned uncertainties of "month" and "season," respectively. However, as discussed for spatial boundaries, temporal ranges of management cannot illogically overlap for a given parcel, and all time must be accounted for. Thus, if the survey data reported a field with corn harvest in October (October 1 to 31) and wheat planted in the fall season (September 1 to November 30), the wheat planting range was adjusted to occur no earlier than the corn harvest (i.e., October 1 to November 30). It is the responsibility of the user to be aware of the uncertainties and implications, as these will vary in importance based on the research question being asked.

The collected information was translated into the database by a single technician, who worked closely with the farm surveyor and other field technicians to decipher notes, reconcile discrepancies, and determine rules for uncertain or unclear data points. A 10% random sample of the input records for the WE38 database was checked for fidelity to the source data. Construction accuracy of the final database to the design structure was then verified by performing a variety of basic summary queries that characterized the range of data values and ensured that the data made real-world sense. The database, while complex in its relational representation of both the spatial and temporal attributes of farm management, is in itself not actually large or overly complex within the context of today's relational database management systems. There are really no practical limits to the number of fields or years of data that can be stored. Thus, expansion of this database to cover a larger area or longer time period is limited only by availability and collection of the input data.

The database structure and contents are housed on a secured server, with regular backups, network security updates, and access monitoring. The database structure is certainly replicable in any Standard Query Language (SQL) database software. However, for privacy reasons, content of the WE38 database is limited to those conducting relevant research and having a current, specific data sharing agreement on file with the location. Products resulting from use of the data are now, and will continue to be, vetted by the location to protect farmer privacy by requiring either spatial aggregation and generality or authorization from relevant land owners and operators.

Results and Discussion

Database Application. Information can be mined from the database at various spatial and temporal scales to explore causal interactions and shift problem solving methods from plan-based to performance-based. The following figures provide representative examples of data mining at three spatial scales: a single field (figure 2), a whole farm (figure 3), and a watershed (figure 4). Please note that the data shown are realistic but not spatially and temporally accurate; these examples are purely illustrative with the intent of better explaining the database. Accordingly, locations and attribute values were modified before extraction from the database to maintain privacy.

By knowing detailed management history for a given field (figure 2), soil scientists and hydrologists can develop more accurate insights on management operations and crop characteristics that impact hillslope variability in nutrient transport and soil stabilization. Ecologists and agronomists can use this information to help inform species competition changes and yield impacts due to climate changes. Land management planners employing nutrient loss risk assessment tools, such as the Pennsylvania Phosphorus index (Weld et al. 2002; Kogelmann et al. 2006), require data collected from a single field or a number of individual fields to generate a particular nutrient loss risk and develop field-based recommendations on manure or fertilizer application rates and timing. Models and risk assessment tools based on detailed practice histories often generate better buy-in with farm stakeholders. Also, such detail in conjunction with water quality and crop yield data improves model process development by providing known inputs and outputs for given management operations.

Land use and management decisions in an agricultural landscape are made at the whole farm scale. Farm-level records that capture temporal management of each field, as well as spatial placement within the whole farm landscape (figure 3), capture and record the effects of these decisions through time. This database is a valuable tool for initializing farm-scale models, such as the Integrated Farm System Model (IFSM; Rotz and Veith 2013; Rotz et al. 2014), which can then be used to simulate the effects of future management decision scenarios. By evaluating which fields and which operators have already employed practices that represent or closely approximate a particular best management practice, we can infer which operators are more likely to be willing to implement that best management practice on additional fields. Additionally, the long-term nature of the database may aid research into the impacts of climate change on agriculture by providing insight on how historical climatic shifts or extreme events have impacted management operations. For example, do long-term shifts in air temperature correspond to shifting trends in planting and harvesting dates? Would shifts in amendment applications have helped avoid excess losses from extreme events during hurricane season, and how can this past information help us formulate practical management guidelines?

Having detailed historical management and monitoring data within a watershed (figure 4) improves our ability to relate management or natural events upstream with historical changes in climate, hydrology, and nutrient loadings to the stream and then to ecosystem health and regional ecosystem services. Long-term trends in climate and hydrology in the WE38 watershed have recently been described (Lu et al. 2014), and we are in the process of relating observed long-term trends on nutrient loadings to the stream to land use and management changes in the context of environmental change. The standardized and complete accounting of land management over time and space enables input tables for watershed-level simulation models, like the Soil and Water Assessment Tool (SWAT; Nietsch et al. 2011), to be developed such that the full complexities of the reported farm operations are maintained. This is particularly important when treating each field as a unique management unit over a multiyear rotation and maintaining the various soil types and topographic distinctions within those fields (Ghebremichael et al. 2008; Veith et al. 2008; Ghebremichael et al. 2010, 2013; Collick et al. 2015). Using the database to perform a combination of farm- and watershed-level analyses, such as those discussed above, one can apply an integrated modeling framework (Ghebremichael et al. 2013) to support further research in how to select and place best management practices that address water quality concerns while maintaining a viable level of production for the farm system.

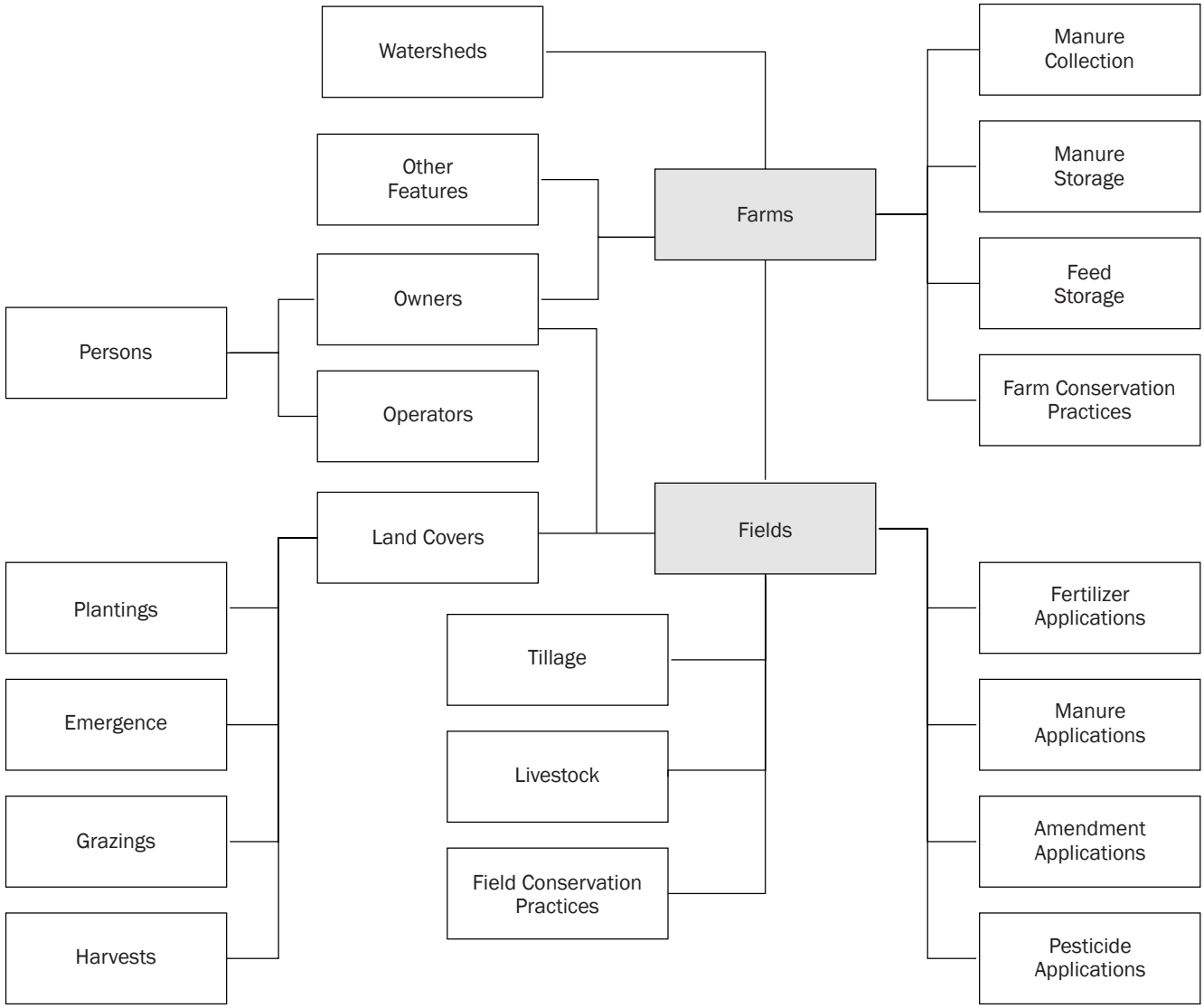
Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

Table 2
Description of all data (D) and lookup (L) tables in the database, ordered by the relationship structure shown in figure 1.

Main database table	Linked tables	Type	Description
(Most tables)	Fields	D	Smallest discrete agricultural management units or parcels
	Units Of Measure	L	Accepted units of measure
Farms	Farm Conservation Practices	D	Farm-specific conservation practices (not tied to a particular field)
	Feed Storage	D	Usage records of feed storage systems
	Manure Collection	D	Usage records of manure collection systems/methods
	Manure Storage	D	Usage records of manure storage systems/methods
	Other Features	D	Records of other features of interest
	Owners	D	Persons with legal title over a farm
	Watersheds	D	Watershed name
Other Features	Other Feature Types	L	Farm-related features of interest
Owners and Operators	Persons	D	Contact information for people associated with farm and field activities (restricted access)
Manure Collection	Manure Collection Types	L	Manure collection method
Manure Storage	Manure Store Capacities	L	Temporal capacity of manure storage systems/methods
	Manure Store Load Types	L	Terms describing how manure is loaded into storage
	Manure Store Types	L	Manure storage systems/methods
Feed Storage	Feed Storage Types	L	Feed storage systems/methods
Farm Conservation Practices	Conservation Practice Groups	L	Groups of conservation practices (see Conservation Practices)
Farms and Fields	Operators	D	Persons who perform work activities in a field
Fields	Farms	D	Fields under common ownership
	Land Covers	D	Crop or land cover records
	Tillage	D	Recorded events where ground was prepared for cultivation
	Livestock	D	Records of livestock kept
	Field Conservation Practices	D	Field-specific conservation practices
	Fertilizer Applications	D	Recorded events of chemical fertilizer application
	Manure Applications	D	Recorded events of animal manure application
	Amendment Applications	D	Recorded events of amendment application
	Pesticide Applications	D	Recorded events of pesticide application
Land Covers	Land Cover Groups	L	Groups of land covers by primary plant type
	Land Cover Names	L	Crops (land covers)
	Plantings	D	Recorded plantings
	Emergence	D	Date of first growth
	Grazings	D	Recorded events where animals are allowed to eat land cover
	Harvests	D	Recorded events where land covers are removed, cut, or killed
Tillage	Tillage Methods	L	Terms that describe preparation of the ground for cultivation
Livestock	Animal Breeds	L	Animal breeds
	Animals	L	Animal genus common names
	Animal Types	L	Animal type based on age, sex, and/or capability of breeding
	Animal Use	L	Primary functions of animals as agricultural commodities
	Livestock Housing	L	Types of livestock shelter
Field Conservation Practices	Conservation Practices	L	Techniques to protect or improve natural resources
Multiple application tables	Application Depth	L	Depth at which a product was applied
	Application Methods	L	How a product was applied
	Application Tank Mix	D	Identifier indicating products applied during same field operation
Fertilizer Applications	Fertilizers	L	Inorganic products used to provide plant nutrition
Manure Applications	Manures	L	Animal manures
Amendment Applications	Amendments	L	Substances added to the soil
Pesticide Applications	Active Ingredient Names	L	Substances classified as biologically active in pesticides
	Active Ingredients	L	Constituent active ingredients in a specific pesticide formulation
	Pesticides	L	Substances applied to protect land covers
	Pesticide Types	L	Groups of pesticides (e.g., herbicide and fungicide)

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

Figure 1
Relational structure of main database tables.



Summary and Conclusions

The USDA ARS Mahantango Long-Term Experimental Research Watershed is a member of the Long-Term Agroecosystem Research (LTAR) Network. The purpose of the LTAR network is to provide regional test-beds where the long-term outcomes of agricultural germplasm technologies, agrochemicals, management strategies, and policies to increase production and/or environmental protection can be evaluated via retrospective (i.e., historical) and prospective (i.e., predictive) research projects. The long-term land management database is a key component of the shared research strategy for this site in that it provides an opportunity to conduct a retrospective evaluation of changes in land use and management and their effects on water quality. Possible scenarios to be tested include the increased adoption of minimum tillage over this period of time, changes in area of perennial crops that provide greater ground cover, and changes in the numbers of livestock and use of manure as a nutrient source.

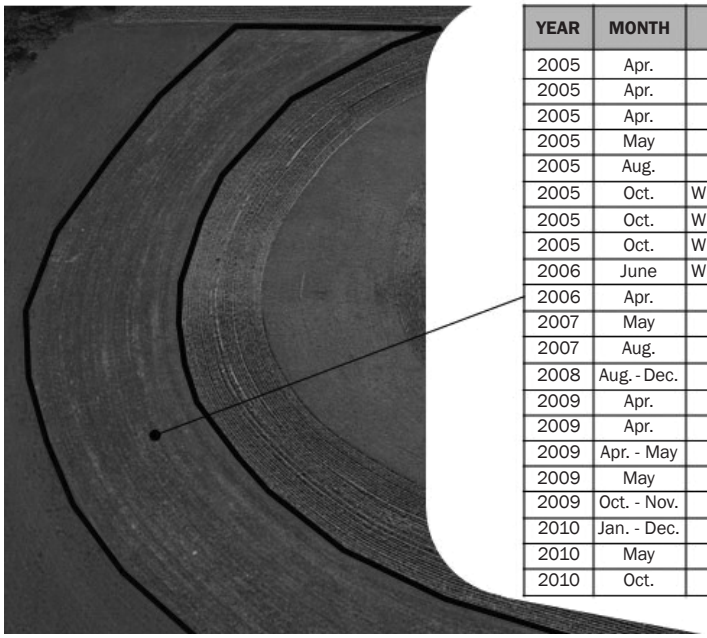
Work is underway to develop a form-based application for collecting future farm surveys digitally using the standardized data model discussed here. This application will enable more efficient collection of the information, maximize consistency through controlled data value choice lists, and improve the likelihood that surveys are entered into the database shortly after collection. According to Michener and Jones (2012), "ecoinformatics is a framework that enables scientists to generate new knowledge through innovative tools and approaches for discovering, managing, integrating, analyzing, visualizing and preserving relevant biological, environmental, and socioeconomic data and information." Employing the principles of ecoinformatics facilitates both long-term research within the Mahantango Experimental Watershed and collaborations across the LTAR network.

The long-term land management database is housed and maintained by scientists at the

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

Figure 2

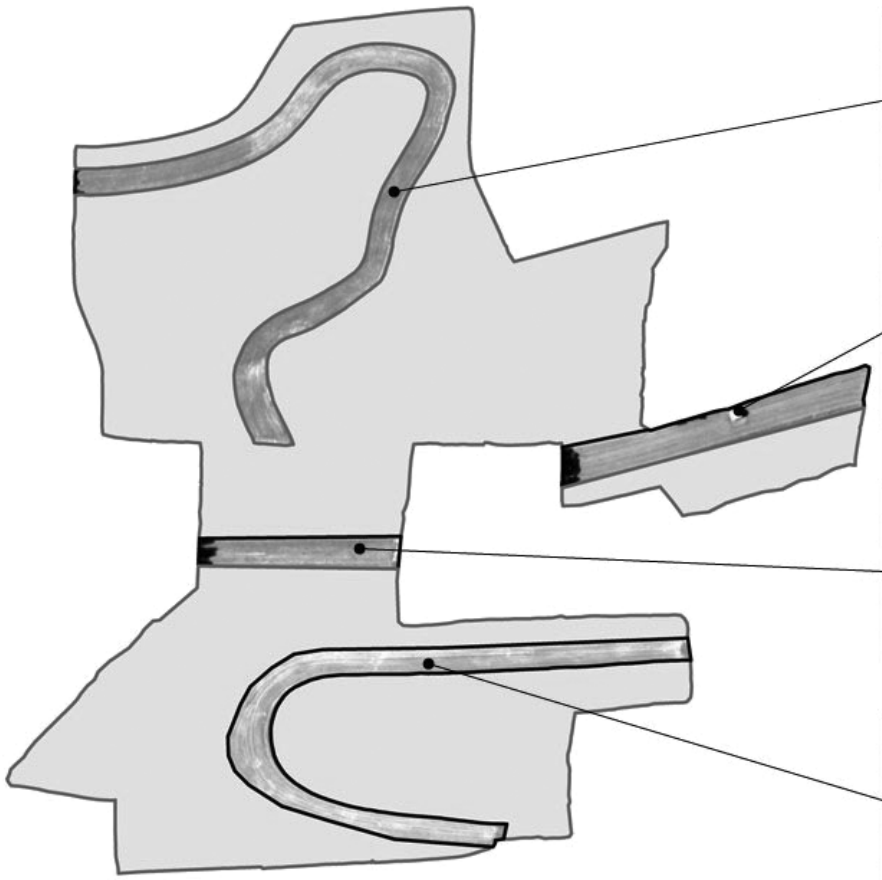
Example of database contents and linkages at the field level.



YEAR	MONTH	CROPS	OPERATION	HARVEST YIELD	TILLAGE METHOD	FERTILIZER	FERTILIZER RATE
2005	Apr.	Oats	Planting				
2005	Apr.	Oats	Tillage		Moldboard		
2005	Apr.	Oats	Tillage		Disc		
2005	May	Oats	Fertilization			10-10-10	10 lb ac^{-1}
2005	Aug.	Oats	Harvest/end	70 bu ac^{-1}			
2005	Oct.	Winter barley	Tillage		Moldboard		
2005	Oct.	Winter barley	Tillage		Disc		
2005	Oct.	Winter barley	Planting				
2006	June	Winter barley	Harvest	90 bu ac^{-1}			
2006	Apr.	Clover	Planting				
2007	May	Clover	Harvest	1.5 tn ac^{-1}			
2007	Aug.	Clover	Harvest/end	0.5 tn ac^{-1}			
2008	Aug. - Dec.		Fertilization			Dairy Manure	5 tn ac^{-1}
2009	Apr.	Corn	Tillage		Chisel		
2009	Apr.	Corn	Tillage		Disc		
2009	Apr. - May	Corn	Planting				
2009	May	Corn	Fertilization			Nitan	30 gal ac^{-1}
2009	Oct. - Nov.	Corn	Harvest/end	150 bu ac^{-1}			
2010	Jan. - Dec.	Soybean	Tillage		No Till		
2010	May	Soybean	Planting				
2010	Oct.	Soybean	Harvest	50 bu ac^{-1}			

Figure 3

Example of database contents and linkages at the farm level.



Field 1

CROP YEAR	PLANTING DATE	HARVEST DATE	CROPS
2007	May 2007	Oct. - Nov. 2007	Corn
2008	May 2008	Oct. 2008	Soybeans
2009	May 2009	Oct. 2009	Corn
2010	May 2010	Oct. 2010	Soybeans
2011	May 2011	Oct. 2011	Corn

Field 2

CROP YEAR	PLANTING DATE	HARVEST DATE	CROPS
2007	May 2007	Oct. - Nov. 2007	Corn
2008	Apr. 2008	July 2008	Oats
2009	Oct. 2008	July 2009	Barley
2010	Aug. - Sept. 2009	June 2010	Timothy
2011	Apr. 2011	July 2011	Oats

Field 3

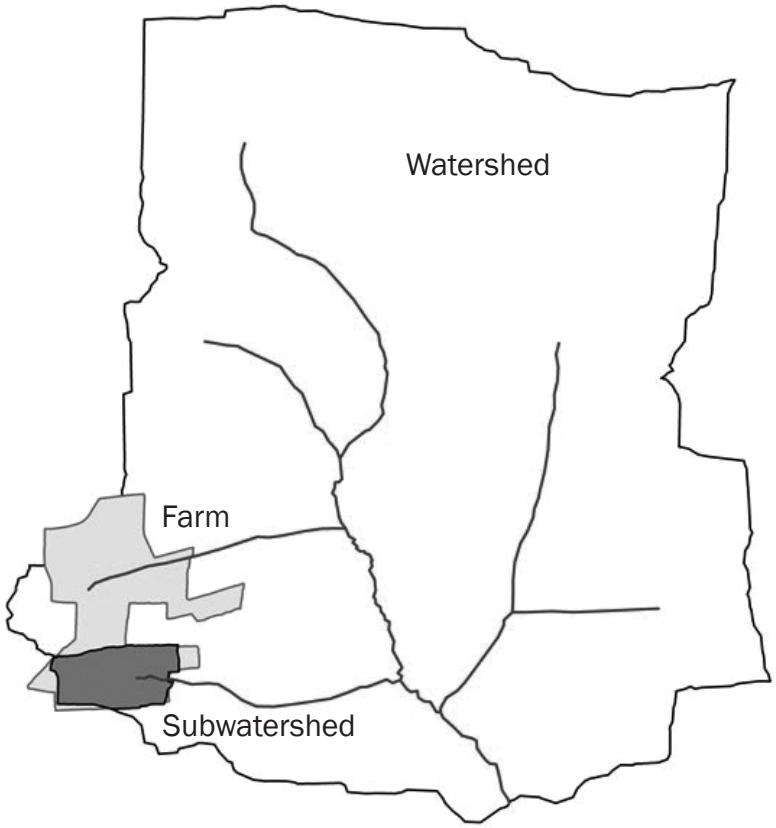
CROP YEAR	PLANTING DATE	HARVEST DATE	CROPS
2007	May 2007	Oct. - Nov. 2007	Corn
2008	May 2008	Sept. - Oct. 2008	Corn
2009	May 2009	Oct. 2009	Soybeans
2010	May 2010	Oct. 2010	Corn
2011	May 2011	Sept. 2011	Corn

Field 4

CROP YEAR	PLANTING DATE	HARVEST DATE	CROPS
2007	June 2007	June 2007	Timothy
2008	May 2008	Sept. - Oct. 2008	Corn
2009	Mar. - Apr. 2009	Aug. 2009	Oats
2010	Sept. - Oct. 2010	June 2010	Barley
2011	May 2011	Sept. - Oct. 2011	Corn

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

Figure 4
Example of database contents and linkages at the watershed level.



YEAR	FERTILIZER NAME	N kg	P kg
2005	Urea	226.9	0.0
2005	Poultry layer manure	228.9	82.1
2005	Dairy manure	20.4	4.3
2005	Beef manure	21.5	5.9
2005	33% Nitrogen	415.0	0.0
2005	0-0-60 Potash w/boron	0.0	0.0
2005	10-10-10	11.2	4.9
2006	Dairy manure	119.8	25.2
2006	Beef manure	31.5	8.7
2006	33% Nitrogen	614.4	0.0
2006	20-10-10	33.6	7.4
2006	0-0-57 w/boron	0.0	0.0
2007	Dairy manure	30.0	6.3
2007	33% Nitrogen	159.0	0.0
2007	10-20-20	100.9	88.8
2007	09-18-09	5.6	4.9
2008	Dairy manure	200.9	42.3
2008	33% Nitrogen	610.3	0.0
2008	10-20-20	78.5	69.0
2009	Urea	103.1	0.0
2009	Dairy manure	122.7	25.8
2009	Beef manure	32.3	8.9
2009	33% Nitrogen	569.6	0.0
2010	Dairy manure	102.2	21.5
2010	33% Nitrogen	671.3	0.0

USDA ARS Pasture Systems and Watershed Management Research Unit (PSWMRU) in University Park, Pennsylvania. Associated geospatial, climate, hydrologic, and water quality databases are available from the USDA STEWARDS Watershed Data System at http://arsagsoftware.ars.usda.gov/stewards/. Users of these data are encouraged to contact ARS scientists at PSWMRU for collaborative assistance with analyzing and interpreting these data.

Disclaimer

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. USDA is an equal opportunity provider and employer.

References

Baker, K.S., B.J. Benson, D.L. Henshaw, D. Blodgett, J.H. Porter, and S.G. Stafford. 2000. Evolution of a multisite network information system: The LTER information management paradigm. BioScience 50:963-978.

Bryant, R.B., T.L. Veith, G.W. Feyereisen, A.R. Buda, G.J. Folmar, C.D. Church, J.P. Schmidt, C.J. Dell, and P.J.A. Kleinman. 2011. USDA-ARS Mahantango Creek Watershed, Pennsylvania, United States: Physiography and history. Water Resources Research 47:10.1029/2010WR010056.

Buda, A.R., G.W. Feyereisen, T.L. Veith, G.J. Folmar, R.B. Bryant, C.D. Church, J.P. Schmidt, C.J. Dell, and P.J.A. Kleinman. 2011a. USDA-ARS Mahantango Creek Watershed, Pennsylvania, United States: Long-term stream discharge database. Water Resources Research 47:10.1029/2010WR010059.

Buda, A.R., T.L. Veith, G.J. Folmar, G.W. Feyereisen, R.B. Bryant, C.D. Church, J.P. Schmidt, C.J. Dell, and P.J.A. Kleinman. 2011b. USDA-ARS Mahantango Creek Watershed, Pennsylvania, United States: Long-term precipitation database. Water Resources Research 47:10.1029/2010WR010058.

Church, C.D., T.L. Veith, G.J. Folmar, A.R. Buda, G.W. Feyereisen, R.B. Bryant, J.P. Schmidt, C.J. Dell, and P.J.A. Kleinman. 2011. USDA-ARS Mahantango Creek Watershed, Pennsylvania, United States: Long-term water quality database. Water Resources Research 47:10.1029/2010WR010060.

Collick, A.S., D.R. Fuka, P.J.A. Kleinman, A.R. Buda, J.L. Weld, M.J. White, T.L. Veith, R.B. Bryant, C.H. Bolster, and Z.M. Easton. 2015. Predicting phosphorus dynamics in complex terrains using a variable source area hydrology model. Hydrological Processes 29:588-601.

Ghebremichael, L.T., T.L. Veith, and J.M. Hamlett. 2013. Integrated watershed- and farm-scale modeling framework for targeting critical source areas while maintaining farm economic viability. Journal of Environmental Management 114:381-394.

Ghebremichael, L.T., T.L. Veith, J.M. Hamlett, and W.J. Gburek. 2008. Precision feeding and forage management effects on phosphorus loss modeled at a watershed scale. Journal of Soil and Water Conservation Society 63(5):280-291, doi:10.2489/jswc.63.5.280.

Ghebremichael, L.T., T.L. Veith, and M.C. Watzin. 2010. Determination of critical source areas for phosphorus loss: Lake Champlain Basin, Vermont. Transactions of the American Society of Agricultural and Biological Engineers 53:1595-1604.

Hobbie, J.E., S.R. Carpenter, N.B. Grimm, J.R. Gosz, and T.R. Seastedt. 2003. The US Long Term Ecological Research program. BioScience 53:21-32.

Kogelmann, W.J., R.B. Bryant, H.S. Lin, D.B. Beegle, and J.L. Weld. 2006. Local assessments of the impact of phosphorus index implementation in Pennsylvania. Journal of Soil and Water Conservation 61(1):20-30.

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org

Le Duc, M., L. Yang, and R. Marrs. 2007. A database application for long-term ecological field experiments. Journal ofVegetation Science 18:509-516.

Michener, W.K., and M.B. Jones. 2012. Ecoinformatics: Supporting ecology as a data-intensive science.Trends in Ecology & Evolution 27:85-93.

Michener, W.K., J. Porter, M. Servilla, and K. Vanderbilt. 2011. Long term ecological research and information management. Ecological Informatics 6:13-24.

Nietsch, S.L., J.G. Arnold, J.R. Kiniry, and J.R. Williams. 2011. Soil and Water Assessment Tool, Theoretical Documentation, Version 2009, Technical Report TR-406 2011. TX: Texas Water Resources Institute: College Station. http://swat.tamu.edu/media/99192/swat2009-theory.pdf.

Rotz, C.A., M.S. Corson, D.S. Chianese, F. Montes, S.D. Hafner, H. Bonifacio, and C.U. Coiner. 2014. The Integrated Farm System Model reference manual, version 4.1. https://www.ars.usda.gov/sp2UserFiles/Place/19020500/Reference%20Manual.pdf.

Rotz, C.A., and T.L. Veith. 2013. Integration of air and water quality issues. *In* Sustainable Animal Agriculture, ed. E. Kebreab, 137-156. Oxfordshire, UK: Centre for Biosciences and Agriculture International.

Sadler, E.J., J.L. Steiner, J.-S. Chen, G. Wilson, D. James, B. Vandenberg, J. Ross, T.L. Oster, K. Cole, and J.L. Hatfield. 2008. Sustaining the Earth's Watersheds-Agricultural Research Watershed Data System: User perspective, operation, and application. Journal of Soil and Water Conservation 63(6):577-589, doi:10.2489/jswc.63.6.577.

Steiner, J.L., E.J. Sadler, J.-S. Chen, G. Wilson, D. James, B. Vandenberg, J. Ross, and T.L. Oster. 2008. Sustaining the Earth's Watersheds-Agricultural Research Watershed Data System: Overview of development and challenges. Journal of Soil and Water Conservation 63(6):569-576, doi:10.2489/jswc.63.6.569.

Veith, T.L., A.N. Sharpley, and J.G. Arnold. 2008. Modeling a small, northeastern watershed with detailed, field-level data. Transactions of the American Society of Agricultural and Biological Engineers 51:471-483.

Volk, C.J., Y. Lucero, and K. Barnas. 2014. Why is data sharing in collaborative natural resource efforts so hard and what can we do to improve it? Environmental Management 53:883-893.

Weld, J.L., R.L. Parsons, D.B. Beegle, A.N. Sharpley, W.J. Gburek, and W.R. Clouser. 2002. Evaluation of phosphorus-based nutrient management strategies in Pennsylvania. Journal of Soil and Water Conservation 57(6):448-454.

Whitlock, M.C. 2011. Data archiving in ecology and evolution: Best practices.Trends in Ecology & Evolution 26:61-65.

Wolkovich, E.M., J. Regetz, and M.I. O'Connor. 2012. Advances in global change research require open science by individual researchers. Global Change Biology 18:2102-2110.

Copyright © 2015 Soil and Water Conservation Society. All rights reserved.
Journal of Soil and Water Conservation 70(5):288-296 www.swcs.org