ECOHYDROLOGY
Ecohydrol. **2**, 391–398 (2009)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/eco.89

Data management to enhance long-term watershed research capacity: context and STEWARDS case study†

Jean L. Steiner,[1]* E. John Sadler,[2] Jerry L. Hatfield,[3] Greg Wilson,[4] David James,[3] Bruce Vandenberg,[5] John D. Ross,[1] Teri Oster[2] and Kevin Cole[3]

[1] *USDA, ARS, El Reno, OK, USA*
[2] *USDA, ARS, Columbia, MO, USA*
[3] *USDA, ARS, NSTL, Ames, IA, USA*
[4] *USDA, ARS, OIRP, Beltsville, MD, USA*
[5] *USDA, ARS, Ft. Collins, CO, USA*

ABSTRACT

Water resources are under pressure globally due to growing population, human migration into arid regions, and diverse competing needs. In recent decades, progress in the study of information (informatics) and its manipulation via computer-based tools has stimulated development of data systems in many natural resources disciplines. Such informatics systems provide data storage, access, visualization, perhaps with analysis/modelling tools, and download capacity. Application of database technology can overcome problems of fragmentation, inadequate documentation, and cumbersome manipulation of complex data. Data management was a critical requirement for USDA's Conservation Effects Assessment Project (CEAP) which was established to quantify environmental effects of agricultural conservation practices. Although USDA and the Agricultural Research Service (ARS) have conducted watershed research since early 20th century, the data have been managed and disseminated independently from each research location, reducing accessibility and utility of these data for policy-relevant, multi-site analyses. To address these concerns, STEWARDS (Sustaining the Earth's Watersheds—Agricultural Research Data System) was developed to compile, document, and provide access to data from loosely coupled research watersheds. The STEWARDS case study is used to illustrate the role of data management in enhancing ecohydrological research and evolving information technologies available to improve data management from complex ecohydrologic studies. Published in 2009 by John Wiley & Sons, Ltd.

KEY WORDS informatics; database; hydrology; agroecosystems; open source data; STEWARDS; CEAP

Received 29 September 2008; Accepted 27 July 2009

INTRODUCTION

In the face of growing demand for water and projected climate change, uncertainty about precipitation frequency, precipitation intensity, evapotranspiration, runoff, and snowmelt poses severe ecological and societal challenges. Solutions to a multitude of ecological and economic problems will require new approaches to governance, improved information technologies and science capacity, and mobilization and empowering of communities (e.g. RNRF, 2005). Considerable research effort is underway to develop water management tools (e.g. Scholten *et al.*, 2007; Leone and Chen, 2007), but adoption of decision support tools by users outside the development teams lags, partially because of a 'disconnect' between the conceptual models of the development teams and the intended end-users (McIntosh *et al.*, 2007).

Since development of the world wide web, there has been movement toward more open access to information. The Science Commons project (http://sciencecommons.org, accessed 25 June 2009) describes the evolution of a call for open access to information through declarations such as the Bethesda Statement on Open Access Publishing (http://www.earlham.edu/~peters/fos/bethesda.htm, accessed 25 June 2009), the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (http://oa.mpg.de/openaccess-berlin/berlindeclaration.html, accessed 25 June 2009), and the Budapest Open Access Initiative (http://www.soros.org/openaccess/, accessed 25 June 2009), with the latter two advocating open access to data and databases. Klump *et al.* (2006) discussed implications of open access information, including the need for incentives to authors and protection of the intellectual rights of the author while allowing use of data by the scientific community.

The critical role of data in advancing hydrologic scientific understanding was emphasized by the National Research Council, 'Intensifying water scarcity cannot be successfully addressed in the absence of reliable data about the quantity and quality of water over time and at different locations. The end-of-century trend of investing fewer and fewer dollars in data-gathering efforts will need to be reversed if availability is to be adequately characterized' (NRC, 2001). Similarly, Hornberger *et al.* (2001), discussing climate change research, emphasized

* Correspondence to: Jean L. Steiner, USDA, ARS, El Reno, OK, USA. E-mail: jean.steiner@ars.usda.gov

that 'Beyond the need to collect new data, existing long-term records must be archived and preserved carefully, and observations must be continued indefinitely at sites with long high-quality records, so that patterns of temporal variability, including long-term, low-frequency fluctuations, can be identified and studied'.

Kinzig *et al.* (2000) and Kinzig (2001) highlighted the need for interdisciplinary research in the area of communicating scientific information, emphasizing potential benefits of information technology (IT) on flows of scientific information to diverse citizen and stakeholder groups. There is a growing international recognition that use of research data is maximized when the data access, management, and preservation are addressed as an inherent part of the research process (Arzberger *et al.*, 2004).

Considerable effort has been made within the US Department of Agriculture to document and make available data from long-term watershed and ecological research sites. Three watershed teams developed special issues of a peer-reviewed journal to document metadata—Reynolds Creek, Idaho (Marks, 2001); Little river, Georgia (Bosch *et al.*, 2007); and Walnut Gulch, Arizona (Moran *et al.*, 2008). These location-by-location efforts offered peer-reviewed metadata and open access to watershed data. The EcoTrends Project (http://www.ecotrends.info, accessed 25 June 2009) includes centralized data management and download from long-term ecological research sites, including 24 USDA experimental sites.

The objective of this paper is to discuss technological advances in information technologies and data management, identify organizational challenges associated with open access data systems, and present a case study of an ARS data management initiative, STEWARDS (Sustaining the Earth's Watersheds, Agricultural Research Data System) to highlight impacts of changing information technologies on institutional culture relative to data management and access.

TECHNOLOGICAL ADVANCES

The development of the internet and rapid advances in computational capacity and information technologies are transforming how society and the scientific community view and manage information and data. The internet was conceived and developed to provide a collaborative workspace, but the first decade focused primarily on the web as a source of information. More recently, capacities are being developed and adopted that return to the original intent of the internet as a dynamic, interactive communication tool (Butler, 2005). In data management, availability of interactive geospatial platforms is resulting in an evolution toward data systems that are more interactive and dynamic than databases of a decade ago. Cotter *et al.* (2004) summarized key US government policies relating to IT and data management that impact federal agencies, including policies related to geospatial data (with standards developed under the auspices of the Federal Geographic Data Committee, 1999) and that require agencies to disseminate and maximize the usefulness of information to government and the public.

Interdisciplinary environmental research across natural and social sciences to address challenges in water resource management requires comprehensive and long-term data. Application of modern database technology can overcome problems of fragmentation, inadequate documentation, and cumbersome manipulation of agroecological research data. Developing environmental data libraries requires expertise from data information and computing disciplines as well as expertise in environmental sciences (Baker *et al.*, 2000). Futrell *et al.* (2003) identified the need for a 'suite of critical enabling tools for storing, finding, analysing, and synthesizing a diverse array of data' to enable study of complex systems.

More recently, the Consortium of Universities for the Advancement of Hydrologic Science, Incorporated (CUAHSI) initiated the Hydrologic Information System, 'a geographically distributed network of hydrologic data sources and functions that are integrated using web services so that they function as a connected whole.' (Maidment *et al.*, 2005; Horsburgh *et al.*, 2008; http://www.cuahsi.org/his.html, accessed 25 June 2009). The evolution of science databases has included a focus on quality assurance, standards, and interoperability. One essential requirement is documentation about the data, or metadata, to provide adequate descriptive information to enable researchers who weren't involved in collecting or processing the data to understand the details of the data collection and processing (Porter and Brunt, 2001).

ORGANIZATIONAL/INSTITUTIONAL ISSUES

As scientific norms evolve toward more open data systems, institutional constraints are numerous. Challenges include digitization of legacy data, quality assurance, standardization, and lack of documentation. There is disparity across disciplines in the current state of archiving digital data (Esanu *et al.*, 2004). Knapp *et al.* (2007) discuss challenges in a 'rescue' effort in the International Satellite Cloud Climatology Project. The effort was massive, but benefit of preserving the unique data set for future analysis with increased computational capacity was high. This rescue relied on partnerships, as do many large data projects, to share costs and leverage available resources to address a common need. Cotter *et al.* (2004) describe a broad partnership across federal, state, and local agencies, private organizations, and academic partners in establishing the National Biological Information Infrastructure. Their solution resulted in a decentralized system that had regional and thematic structures to address diverse issues, such as invasive species and wildlife diseases. The comprehensive system included focus on infrastructure, content, tools, and services to meet a diverse mix of applications.

Open data policies raise significant privacy issues, leading to the need for restricting access to and/or applications of certain types of data (de Wolf, 2003). This is a major issue in social science, and medical and economic research, as well as for many other situations. In many cases, there are legal restrictions on sharing and application of the data, while in others it is an issue of trust with collaborators. Zinn (1998) synthesized findings from a Privacy and Natural Resources Workshop, that identified the growing value of data, increased interest by private industry in this value, the need to create a climate of trust among agricultural producers, the need for better communication and new partnerships, and the growing importance of information in distinguishing more successful producers from less successful ones. In the area of privacy about natural resource management data (on private land), technologies and expectations are changing rapidly and at an accelerating rate.

Preparation of data for a scientific data system, particularly older data, places a considerable burden on data providers, as the data provider must be involved to ensure that data are accurately represented (Knapp *et al*., 2007). The scientific community as a whole has yet to develop compelling incentives to data providers. In some cases, scientists enter into formal or informal collaborations to share data and information in order to advance their own scientific goals (e.g. Bouma and Jones, 2001; Bostick *et al*., 2004). In addition to collaborative and networking benefits to data providers, the scientific community needs to develop peer review processes for data contributions and recognize provision of data on a comparable basis with other peer-reviewed scientific contributions. Professional credit for provision of open data is also at the discretion of peer scientists who participate in selection panels, tenure review, and award selection.

Legal traditions in the US have not allowed intellectual property rights for data or raw facts. Many proponents of open access propose that the norms of the discipline will determine the citation for use of open data, rather than a legal requirement. However, others propose a Creative Common license, in which scientists could stipulate their rights and credits for reuse of the data (Nature, 2005).

Rodriguez and Solomon (2007) describe the power of innovation in a networked world (http://www.mitpressjournals.org/doi/pdfplus/10.1162/itgg.2007.2.3.3, accessed 25 June 2009) but also discuss ways in which managers in most sectors of our society are ill equipped to effectively manage under this new paradigm. Rewards primarily have focused on innovation and accomplishment by individuals, not recognizing the power of groups of people working in a connected, interactive way that captures energies, ideas, and passions of many toward a common goal. Malone and Klein (2007) describe how such informal networked groups are making major advances toward identifying solutions to thorny issues associated with climate change (http://www.mitpressjournals.org/doi/pdfplus/10.1162/itgg.2007.2.3.15, accessed 25 June 2009).

THE STEWARDS CASE STUDY

Background and overview

The US Government Accountability Office (GAO, 2004) assessed water quality data collection by 15 federal agencies and identified a need for increased coordination in order that water quality data collected become more useful for watershed management decision making. This assessment and other scientific reports (Hornberger *et al*., 2001; NRC, 2001, 2003) have impacted data management planning for many federal agencies, including the Agricultural Research Service (ARS). In 2003, when the USDA Natural Resources Conservation Service entered into partnership with ARS to quantify environmental effects of USDA conservation programs and practices (Conservation Effects Assessment Project, CEAP) a database to retrieve data from watershed studies was identified as a critical requirement. Using field experimental and modelling approaches, CEAP studies require a variety of data that describe hydrology, soils, climate, topography, management practices, and land use to assess the impacts of conservation practices on soil and water quality, and ecological/environmental health (Mausbach and Dedrick, 2004).

Although the USDA and the ARS have conducted watershed research since early 20th century, the data have been managed to address location-specific research needs and disseminated independently from each research location. This is not unexpected in watershed research, which has an inherently local focus, but such practices greatly reduce accessibility and utility of the data for policy-relevant, multi-site analyses. To meet these concerns, the STEWARDS watershed data system was developed to compile, document, and provide access to data from loosely coupled research watersheds across the ARS program (Steiner *et al*., 2008; 2009; Sadler *et al*., 2008).

In an early effort to provide access to this unique data set, about 16 600 station years of rainfall and runoff data from watersheds ranging from 0·2 hectares to 12 400 km^2 were compiled in 1990 by the ARS Hydrology and Remote Sensing Laboratory (Burford *et al*., 1985, Thurman and Roberts, 1995). After 1990, centralized compilation and archiving of ARS watershed data ended until STEWARDS was initiated to develop more comprehensive data sets from these and newer watersheds. A key indicator of widespread interest in these types of data is that the archived pre-1990 dataset (http://hydrolab.arsusda.gov/wdc/arswater.html, accessed 25 June 2009) generated over 4000 site visits and over 500 file downloads in 2007.

The STEWARDS watersheds (Table I) are part of the cropland CEAP Watershed Assessment Studies. These watersheds are a subset of a network of long-term watershed research locations in ARS, plus watersheds established after 2002 to address source water protection or as part of CEAP. The STEWARDS development team was drawn from ARS locations that could provide needed skills and expertise (Figure 1).

Table I. Watersheds in the cropland CEAP watershed assessment studies.

ARS location	CEAP watershed	Date established	Area[a] (km^2)	Regional characteristics/ Research focus
GA, Tifton	Little river	1968	334	Wooded, slow moving watershed/riparian processes, pesticides, nutrients
ID, Kimberly	Upper Snake Rock creek (Twin Falls irrigation tract)	2004	6986 (820)	Western irrigation district/water and salt balance for contrasting irrigation methods
IN, West Lafayette	St. Joseph river	2002	2810	Source water protection/water quality, fish, and wildlife
IA, Ames	North Walnut creek	1990	52	Midwest, corn-soybean/nitrates, drainage
	South fork of the Iowa River	2001	763	Midwest, corn-soybean-CAFOs/nitrates, phosphorus, drainage, soil quality
MD, Beltsville	Choptank river	2003	1756	Tidal estuary, poultry, urban pressure/eutrification, nutrients, ditch drainage, wetlands, cover crops
MS, Oxford	Beasley lake	1995	8·5	Oxbow lake/pesticides, nutrients, ecological impacts
	Goodwin creek	1981	21·3	Loess bluffs/channel processes
	Little Topashaw creek/Yalabousha river	2000	1688	Channelized stream network/erosion, bank failure, channel processes
MO, Columbia	Goodwater creek	1971	72·5	Mixed agricultural land use/pesticides, nutrients, and sediment water quality
	Salt river/Mark Twain lake	2004	6417	Drinking water source/pesticide and nutrient issues
OH, Columbus	Upper Walnut Cr.	2004	492	Source water protection/Atrazine, soil erosion
OK, El Reno	Little Washita river/Upper Washita river	1961	610	Flood retarding structures/climate variability remote sensing, model testing
	Fort Cobb lake/Upper Washita river	2004	788	Reservoir water quality/phosphorus, nitrogen, sediment sources, fate and transport
PA, University Park	Mahantango creek (WE-38)	1965	416 (7·3)	water quality, runoff generation, surface-subsurface transport, frozen soil and snow
TX, Temple	Leon river	2004	6070	CAFOs impacting reservoir/nutrients, *Escherichia coli*, sediment

[a] When detailed research is focused on a sub-watershed of a larger area, the larger area is given.

The STEWARDS system applies ArcGIS (ESRI Corporation, Environmental Systems Research Institute, Redlands, CA) geospatial technologies to provide a flexible approach to visualize, and deliver information to the research community (mention of product name is for information only and does not constitute an endorsement by the US Department of Agriculture). It allows access to the watershed data for internal and external researchers while retaining local control of and responsibility for the data. Data are retained at the watershed level for a period of time for quality assurance and initial analysis and publication. There is no requirement to adopt uniform data management procedures by watershed teams for local use, but delivery to STEWARDS requires standardized parameter names, units, and metadata (Table II). For watersheds established earliest, initial preparation and upload include data collected by staff who are no longer employees, often requiring considerable effort to document the methods. For newer watersheds, initial publications are just being prepared with data delivery to follow. After initial data upload to STEWARDS, updates will be scheduled with each watershed team on an approximately annual basis.

One system requirement was for flexibility to accommodate diverse and dynamic data (Steiner *et al*., 2008, 200_). To accommodate this, data tables are developed around themes (e.g. hydrology, weather, water quality) that may have different content across watersheds. Each

Figure 1. The CEAP Watershed Assessment Study (focusing primarily on rainfed cropland watersheds) was developed from a subset of the ARS watershed research network. The STEWARDS development team was drawn from a subset of the CEAP watersheds (Ames, Iowa; Columbia, Missouri; and El Reno, Oklahoma) along with a programmer from Fort Collins, Colorado, and a systems manager from ARS's Office of the Chief Information Officer in Beltsville, Maryland (dashed lines indicate people from multiple sites working toward the common goal of developing STEWARDS).

Table II. STEWARDS data categories and types.

Data category	Spatial	Measurement	Desired period of record
Minimum[a]	Watershed boundary Digital elevation model Stream network Soils Measurement locations	Stream flow Weather	5+ years
Future minimum	Land use	Water quality, surface	5+ years
Examples of additional open data	Groundwater CAFO locations Irrigation Geomorphology	Water quality, groundwater Soil properties Biological data	n.a.
Proprietary data	Conservation practices	Economic data Social data	n.a.

[a] A minimum data set requires a watershed description, site descriptions, methods descriptions, and other metadata.

data table is paired with a data definition table that defines the content of the data table, and each data table contains a unique SiteID and date/time as the first two columns, providing primary keys within the database structure. Each data table also requires a file to link the SiteID names with geospatial information. Methods used in the data collection are identified by code in the data definition table, with the method code being a primary key to an ARS Methods Catalog, which is modelled on the National Environmental Methods Index (http://www.nemi.gov/, accessed 25 June 2009).

Descriptive information includes watershed descriptions and site descriptions. Keyword searches can be conducted across the entire STEWARDS database to identify where particular types of data are collected or where particular hydro-ecological issues are addressed. Data searches are initiated by selecting a watershed and searching by parameter or site. Data can be viewed as a time series graph or table, and data can be downloaded into a spreadsheet or text file. Associated metadata are downloaded with data download.

Changes in ARS culture toward open data

Development of STEWARDS required and helped foster a change of culture within ARS regarding open data. Flexibility was required to obtain collaboration across internal organizational units that had not worked together in the past. In contrast to past practice, delivery of data to STEWARDS was specified within 5-year research plans, and provision of data was included as a performance requirement for individual scientists and research leaders. While there is no inherent limitation on peer recognition of data provision as a significant scientific contribution, past culture has not valued these contributions highly. It remains to be seen the extent to which scientists will receive adequate credit from their peers and research managers to serve as an incentive for making the effort necessary to develop high-quality data sets for open access in STEWARDS or elsewhere. However, persons who have participated actively in this and other open data efforts are influential within ARS's natural resources research programs.

Data management has received different levels of priorities in different watershed units, and at some locations, improvements are needed. There are several tangible benefits to individuals from participating in the STEWARDS project. Participation in STEWARDS provides learning opportunities for the scientists and staff involved. The structure of the database in STEWARDS provides a template for sites unsure how to organize their data. Additionally, consultation and training are available in various aspects of data management from the STEWARDS operations team or from data managers at other watersheds who may have tackled similar problems.

From an organizational perspective, the process of building the database structure has helped identify commonalties and differences in the data across watersheds. When watershed teams see the STEWARDS structure and prepare their data in that form, they develop a more general understanding of the data—not just the across-watershed data, but also their own. This is a training benefit in database theory and operations. In addition, the recent emphasis on data preparation and delivery to STEWARDS provides a discrete target for completion of data documentation, a task that most researchers consider important, but not necessarily urgent. The focus of peer scientists at other watersheds makes it easier for research managers to raise the priority on data management within a team.

From the outset, there were several desired outcomes from the considerable effort required to develop a watershed data system within ARS. First, expertise within the STEWARDS team and the structured processes developed provided increased opportunity for quality assurance and quality control during data preparation. The comprehensive metadata provide authors, reviewers, and the scientific community information needed to evaluate quality and suitability of data for a given application, improving the credibility of scientific contributions from ARS scientists and their collaborators. An open data system should provide scientists with increased collaboration opportunities, which in turn should increase the productivity and impact of these research teams. In the future, the scientists should be able to shift focus away from time-consuming data compilation and manipulation issues, toward scientific analysis and interpretation. This result was noted by Seber *et al.* (2003) in discussing student response to an interactive geosciences information system developed as a teaching offspring from a research tool. Finally, development of STEWARDS has positioned ARS to better meet stewardship responsibilities to retain and provide comprehensive, long-term watershed data to address future scientific and societal needs.

SUMMARY AND CONCLUSIONS

Interdisciplinary environmental research across natural and social sciences to address challenges in water resource management requires comprehensive and long-term data. The critical role of data in advancing hydrologic scientific understanding has been emphasized in national scientific reports (NRC, 2001; Hornberger *et al.*, 2001). Application of modern database technology can overcome problems of fragmentation, inadequate documentation, and cumbersome manipulation of agroecological research data. Such informatics systems have been created to provide data storage, data access, visualization, analysis/modelling tools and data download.

The development of the internet and rapid advances in information technologies are transforming how society and the scientific community view and manage information and data. As scientific norms evolve toward more open data systems, institutional constraints are numerous. Klump *et al.* (2006) discuss implications of open access information, including the need for incentives to authors and protection of the intellectual rights of the authors while allowing use of data by the scientific community. Open data policies raise privacy issues, leading to the need for restricting access to and/or applications of certain types of data (de Wolf, 2003). In many cases, there are legal restrictions on the application and sharing of the data, while in others it is an issue of trust with collaborators. At an institutional level, when trying to address a large back-log of data, there are tremendous resource constraints. Preparation of data, particularly older data, for a scientific data system places a considerable burden on the data provider, as the data provider must be involved to ensure that the data are accurately represented (Knapp *et al.*, 2007).

The STEWARDS data system represents a move forward for hydrologic and environmental research by providing access to a multitude of data needed to support complex analyses. Key lessons learned, products, and impacts of the development process are summarized in Table III. These data sets represent one of the largest research watershed data collections in the world, with many of the watersheds offering decades of data required to address issues of climate variability and global change. Development of STEWARDS required a balance between accountability to a detailed work plan on one hand, and adaptability to meet evolving requirements, resources, and technologies on the other. Major research efforts that

Table III. Outcomes of development and delivery of STEWARDS.

Lessons learned	• Communication between IT and research personnel requires explicit definition of terms, expectations and roles • Process must continually address balance between flexibility and accountability • Process is inherently iterative and must adjust to rapidly evolving technologies • Data management and IT capacity are highly variable across watershed locations, and require enhancement at some • Team members gained a better understanding of their own data and that from other locations • Team members gained better understanding of agency IT resources and constraints
Products	• Paired data table and data definition table methodology is contribution to the peer-reviewed literature • Meta data about watersheds provides visibility to wider research and management communities • ARS Methods Catalog provides transparency and can serve as a resource for other researchers • Training opportunities for watershed data managers have expanded their professional networks • STEWARDS provides a template for watershed data management • STEWARDS populated with data from multiple watersheds
Impacts	• Improved scientific credibility by documenting QA/QC procedures • Increased collaborative opportunities for individual scientists, watershed teams, and the ARS water resources program • Increased learning opportunities for participants at watersheds • Increased demands on scientists for provision of open data • Better accountability at the agency level for investment in long-term watershed research
Anticipated impacts	• Increased scientific productivity at watersheds • Increased credit to scientists for contribution to open data systems

extend over many years with numerous organizational players will always require such a balance of accountability and adaptability. Challenges remain in data preparation and upload, including initial upload from additional watersheds and periodic updating of data from existing watersheds in STEWARDS. Currently, scientific reward systems offer few incentives to scientists to devote the effort needed to provide data to public data bases. Providing such incentives remains a challenge to be addressed in advancing our capacity to conduct complex, environmental research that is so critical to addressing issues of local, national, and global importance.

REFERENCES

Arzberger P, Schroeder P, Beaulieu A, Bowker G, Casey K, Faaksonen L, Moorman D, Uhlir P, Wouters P. 2004. An international framework to promote access to data. *Science* **303**: 1777–1778.

Baker KS, Benson BJ, Henshaw DL, Blodgett D, Porter JH, Stafford SG. 2000. Evolution of a multisite network information system: the LTER information management paradigm. *Bioscience* **50**: 963–978.

Bosch DD, Sheridan JM, Lowrance RR, Hubbard RK, Strickland TC, Feyereisen GW, Sullivan DG. 2007. Little river experimental watershed database. *Water Resources Research* **43**: W09470. DOI:10.1029/2006WR005844.

Bostick WM, Koo J, Walen VK, Jones JW, Hoogenboom G. 2004. A web-based data exchange system for crop model applications. *Agronomy Journal* **96**: 853–856.

Bouma J, Jones JW. 2001. An international collaborative network for agricultural systems applications (ICASA). *Agricultural Systems* **70**: 355–368. DOI:10.1016/S0308-521X(01)00051-8.

Burford JB, Thurman JL, Roberts RT. 1985. *Hydrologic Data for Experimental Agricultural Watersheds in the United States, 1975*. USDA Misc. Publ. 1446, 580 pp.

Butler D. 2005. Science in the web age: joint efforts. *Nature* **438**: 548–549. DOI:10.1038/438548a.

Cotter G, Frame M, Sepic R. 2004. Integrated science for environmental decision-making: the challenge for biodiversity and ecosystem informatics. *Data Science Journal* **3**: 38–59.

de Wolf V. 2003. Issues in accessing and sharing confidential survey and social science data. *Data Science Journal* **2**: 66–74. DOI:10.2481/dsj.2.66.

Esanu J, Davidson J, Ross J, Anderson W. 2004. Selection appraisal, and retention of digital scientific data: highlights of an ERPANET/CODATA workshop. *Data Science Journal* **3**: 226–232. DOI:10.2481/dsj.3.226.

Federal Geographic Data Committee (FGDC). 1998. *Content Standard for Digital Geospatial Metadata (revised June 1998)*, FGDC-STD-001-1998. Federal Geographic Data Committee: Washington, D.C.

Futrell J. AC-ERE. 2003. *Environmental CyberInfrastructure: Tools for the Study of Complex Environmental Systems*. AC-REE 1, May, 2003. http://www.nsf.gov/geo/ere/ereweb/ac-ere/ac-ere1_ECI.pdf (verified 1/31/08).

Government Accountability Office (GAO). 2004. *Watershed Management: Better Coordination of Data Collection Efforts Needed to Support Key Decisions*. GAO-04-382.

Hornberger GM, Aber JD, Bahr J, Bales RC, Beven K, Foufoula-Georgiou E, Katul G, Kinter JL III, Koster RD, Lettenmaier DP, McKnight D, Miller K, Mitchell K, Roads JL, Scanlon BR, Smith E. 2001. *A Plan for a New Science Initiative on the Global Water Cycle*. US Global Change Research Program, Washington, D.C.

Horsburgh JS, Tarboton DG, Maidment DR, Zaslavsky I. 2008. A relational model for environmental and water resources data. *Water Resources Research* **44**(5): W05406. DOI:10.1029/2007WR006392.

Kinzig AP. 2001. Bridging disciplinary divides to address environmental and intellectual challenges. *Ecosystems* **4**: 709–712.

Kinzig AP, Carpenter S, Dove M, Heal G, Levin SA, Lubchenco J, Schneider S, and Starrett D. 2000. *Nature and Society: An Imperative for Integrated Research*. Executive Summary of a report prepared for the National Science Foundation. http://www.public.asu.edu/~akinzig/nsfes.pdf (verified 1/31/08).

Klump J, Bertelmann R, Brase J, Diepenbroek M, Grobe H, Höck H, Lautenschlager M, Schindler U, Sens I, Wächter J. 2006. Data publication in an open access initiative. *Data Science Journal* **5**: 79–83. DOI:10.2481/dsj.5.79.

Knapp KR, Bases JJ, Barkstron B. 2007. *Scientific Data Stewardship: Lessons Learned from a Satellite-Data Rescue Effort*. BAMS, September, 2007, 1359–1361. DOI:10.1175/BAMS-88-9-1359.

Leone A, Chen D. 2007. Implementation of an object oriented data model in an information system for water catchment management: Java JDO and Db4o Object Database. *Environmental Modelling and Software* **22**: 1805–1810. DOI:10.1016/j.envsoft.2007.05.016.

Maidment DR, Helly J, Piasecki M, Kumar P, Duncan J. 2005. Development of a hydrologic information system for CUAHSI. In *Proceedings of the, Watershed Management Conference—Managing Watersheds for Human and Natural Impacts: Engineering, Ecological, and Economic Challenges*, Moglen GE (ed). July 19–22, 2005, Williamsburg, VA; Environmental and Water Resources Institute (EWRI) of the American Society of Civil Engineers; Reston; 607–613.

Malone TW, Klein M. 2007. Harnessing collective intelligence to address global climate change. *Innovations: Technology, Governance, Globalization* **2**(3): 15–26.

Marks D. 2001. Introduction to special section: Reynolds Creek experimental watershed. *Water Resources Research* **37**: 2817.

Mausbach MJ, Dedrick AR. 2004. The length we go: measuring environmental effects of conservation practices. *Journal of Soil Water Conservation* **59**: 96A–103A.

McIntosh BS, Seaton RAF, Jeffrey P. 2007. Tools to think with? Towards understanding the use of computer-based support tools in policy relevant research. *Environmental Modelling and Software* **22**: 640–648. DOI:10.1016/j.envsoft.2005.12.015.

Moran MS, Emmerich WE, Goodrich DC, Heilman P, Holifield Collins C, Keefer TO, Nearing MA, Nichols MH, Renard KG, Scott RL, Smith JR, Stone JJ, Unkrich CL, Wong JK. 2008. Preface to special section on Fifty Years of Research and Data Collection: U.S. Department of Agriculture Walnut Gulch Experimental Watershed. *Water Resources Research* **44**: W05S01. DOI:10.1029/2007WR006083.

National Research Council (NRC). 2001. *Envisioning the Agenda for Water Resources Research in the Twenty-First Century*. Water Science and Technology Board, National Research Council: Washington, D.C.

National Research Council (NRC). 2003. *Frontiers in Agricultural Research: Food, Health, Environment, and Communities*. Board on Agriculture and Natural Resources, National Academies Press: Washington, D.C.

Nature. 2005. Let data speak to data (editorial). *Nature* **438**: 531. DOI:10.1038/438531a.

Porter J, Brunt JW. 2001. Making data useful: the history and future of LTER metadata. *The Network Newsletter (LTER)* **14**(2): 12–13.

Renewable Natural Resources Foundation (RNRF). 2005. In *Building Capacity for Coastal Solutions*, Colker RM, RD Day (eds). Renewable Natural Resources Foundation: Bethesda.

Rodriguez D, Solomon D. 2007. Leadership and innovation in a networked world. *Innovations: Technology, Governance, Globalization* **2**(3): 3–13.

Sadler EJ, Steiner JL, Chen J-S, Wilson G, James D, Vandenberg B, Ross J, Oster T, Cole K, Hatfield JL. 2008. STEWARDS Watershed Data System: User perspective, operation, and application. *Journal of Soil and Water Conservation* **63**(6): 577–589.

Scholten H, Kassahum A, Refsgaard JC, Kargas T, Gavardinas C, Beulens AJM. 2007. A methodology to support multidisciplinary model-based water management. *Environmental Modelling and Software* **22**: 743–759. DOI:10.1016/j.envsoft.2006.12.025.

Seber D, Moore A, Brindisi C, Danowski D. 2003. Discover our earth: an earth science information system for undergraduate education. *EOS Transactions* **84**(23): 213, 219.

Steiner JL, Sadler EJ, Chen J-S, Wilson G, James D, Vandenberg B, Ross J, Oster TL. 2008. STEWARDS Watershed Data System: Overview. *Journal of Soil and Water Conservation* **63**(6): 569–576.

Steiner JL, Sadler EJ, Hatfield JL, Wilson G, James D, Vandenberg B, Chen J-S, Oster TL, Ross JD, and Cole K. 2009. STEWARDS watershed data system: system design and implementation. *Transactions of the ASABE*.

Thurman JL, Roberts RT. 1995. New strategies for the Water Data Center. (Special Issue: Water Research and Management in Semiarid Environments). *Journal of Soil and Water Conservation* **50**: 530–531.

Zinn JA. 1998. *Data and Information About Natural Resources on Agricultural Land: No Rules, Just Rights*. Privacy and Natural Resources Workshop White Paper. Arbor Day Foundation, Nebraska City, Nebraska. http://www.arborday.org/programs/Papers/PrivacyWpaper.cfm (assessed 6/25/2009).