Information Technology Applications in the ARS Watershed Network

J.L. Steiner, D.C. Goodrich, S. Hardegree, M.R. Burkhart, T.C. Strickland, M.A. Weltz

Abstract

Knowledge gaps relating to water supply, quality, and cycling processes have been identified as critical obstacles to improved water resource management in recent assessments. One critical gap identified was lack of adequate data to evaluate climate and hydrologic processes, particularly variability associated with climate and hydrology that affects management responses. The USDA, Agricultural Research Service operates long-term research watersheds in major physiographic regions of the US. These watersheds provide under-utilized opportunities to evaluate interactions of land use change, management practices, and climate variability in national assessments. A pilot project is underway to increase accessibility and utility of ARS Watershed Network data for enhanced research programs and to support a wide array of stakeholders. Approaches include: 1) evaluate data management models in other ecological and natural resources research networks, 2) develop formats and standards for metadata and data files, involving researchers, end-users, informatics and data base management specialists and 3) develop a common platform to access the data from multiple locations. The overall objective is to provide improved access to the watershed data for internal and external researchers, while retaining local control of and responsibility for the data. New data management systems for the experimental watersheds are expected to reduce delays and costs of developing new research thrusts and partnerships and increase data availability across the entire period of data collection and across different types of data. Users would obtain high quality and timely data, consistent across watersheds. All of this could extend the application of ARS watershed research to ecologic and socioeconomic, as well as agricultural and water resources problem-solving.

Keywords: informatics, hydrology, web-based data products

Steiner is Research Leader, USDA-Agricultural Research Service, El Reno, OK, 73036. E-mail: jsteiner@grl.ars.usda.gov. All authors are with the U.S. Department of Agriculture, Agricultural Research Service. Goodrich is a Research Hydraulic Engineer, Tucson, AZ 85719, Hardegree is Research Leader, Boise, ID 83712, Burkart is Research Leader, Ames, IA 50011, Strickland is Research Leader, Tifton, GA 31793, and Weltz is National Program Leader for Hydrology and Remote Sensing, Beltsville, MD 20705.

Introduction

Water resources face growing pressure globally, and with the prospect of possible future climate change, the water cycle (changes in precipitation frequency and intensity; evapotranspiration, runoff, snowmelt) will likely pose severe societal challenges (Gleik 1998). The critical role of experimental watershed data in the quest for hydrologic scientific understanding was clearly stated in the NRC (2001) report *Envisioning the Agenda for Water Resources Research in the Twenty-First Century*, “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.” In *A Plan for a New Science Initiative on the Global Water Cycle*, Hornberger et al. (2001) emphasized 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 and others (Kinzig et al. 2000, Kinzig 2001) assessed environmental research needs, and recommended increased interdisciplinary research in the area of communicating scientific information, particularly potential benefits of information technology (IT) on flows of scientific information to diverse citizenry and stakeholder groups. Their report explicitly addressed interdisciplinary environmental research across natural and social sciences. In no arena is such interaction of biophysical and social processes greater than in water resource management.

The ecological research community has supported a strong thrust to develop cutting-edge information technologies to transform the research process in support of multi-location, synthetic analyses. The Long Term Ecological Research Network (LTER) (http://www.lternet.edu/data) has provided leadership to develop these technologies (Baker et al. 2000). Similar to genomics, in which scientific advances require collaboration and data sharing among many researchers, new approaches to data sharing and data management have been required in the ecological sciences. The term bioinformatics is most often applied in the realm of genomics, so different terms were adopted for natural resources. Many in the ecological community have adopted the term "ecoinformatics" (http://ecoinformatics.org). In this paper we will use the term "natural resources informatics." A recent effort in hydrology, the Hydrological Data and Information System (HYDIS) (http://hydis.hwr.arizona.edu) also offers insights for multi-location and multi-attribute data synthesis.

Multi-site research efforts and analyses are not new. However, challenges are posed as researchers work across sites, each of which is highly complex, to address questions that span multiple scales of time and space. Not the least of these challenges is how to efficiently compile data sets from dispersed locations and ensure that researchers understand the nuances of the variables from sites where they have not worked. In addition they must ensure that like data from different sites that may have been gathered with different instrumentation or different measurement protocols are comparable in further analyses. A specific consideration for managing hydrologic data in the context of agricultural ecosystems analysis is the need to geographically link soil, crop, and livestock data with physical and chemical hydrologic data. These links are critical if watershed-scale water requirements and water quality responses to agricultural systems are to be understood. Watershed data management systems will be most useful to interdisciplinary studies if they incorporate flexible georeferencing technology.

The LTER Network Information System (Baker et al. 2000) has worked methodically to develop methods to describe and archive data for diverse types of future analyses. One aspect that has required considerable effort is development of metadata systems (Porter and Brunt 2001). The metadata are "data about data" and provide descriptive information to enable researchers who were not involved in collecting or processing the data to understand the details of how the data were collected and processed. Recent and ongoing efforts focus on systems to electronically search data libraries to aid research teams in compiling appropriate data sets to address a particular scientific question. To successfully tackle such problems requires expertise from the data information and computing disciplines as well as expertise in the ecological and natural resources sciences (Baker et al. 2000). The NRC (2003) report *Frontiers in Agricultural Research* identified research in environmental stewardship and integration of leading-edge science concepts and techniques, of which informatics is an example, as an opportunity for USDA research to better address societal needs.

Watershed Research in the ARS

History and overview

The U.S. Department of Agriculture established experimental watersheds as early as the 1930s, when the Soil Conservation Service was first established, and have continued under ARS management since the agency's establishment in the 1954. The ARS watershed facilities serve as stable, high-quality, outdoor laboratories that provide research capacity to conduct basic field research, evaluate management impacts, document effects of global change, and develop new instrument and simulation technologies. Over 100 ARS experimental watersheds ranging in size from a few hectares to over 600 km^2, are currently operated at 14 locations in major physiographic regions of the contiguous United States (Figure 1). The network, including descriptions of individual watersheds, is described at http://www.nwrc.ars.usda.gov/watershed/.

There is no comparable network of experimental watersheds in the world that combines intensive observational infrastructure with a longitudinal knowledge base. This network provides a platform to address complex research questions related to climate variability, atmosphere-earth interactions, and hydrologic processes. Several of these watersheds have served as field sites for large multi-organization remote sensing campaigns. Many of the watersheds are also linked with other national networks to broaden the type of observations made and leverage infrastructure, including: USDA-NRCS Soil Climate Analysis Network, AMERIFLUX, Surface Radiation Network (SURFRAD), ARS Rangeland Carbon Flux Network, and DOE/ARM/CART.

Figure 1. Locations that support experimental watershed research programs within the Agricultural Research Service.

Major contributions from these watershed programs have been made to hydrologic science include development of innovative instrumentation to measure primary components of the water cycle and water quality; development, testing, and application of remote sensing technologies; rainfall frequency analyses from dense gauge networks to modify NOAA National Atlases; improved understanding of spatial and temporal variability of infiltration across a range of hydro-climatic conditions; and development and validation of numerous hydrologic and natural resource models, such as USLE, Curve Number, USDAHL, HYMO, ACTMO, SWRRB, AGNPS, CONCEPTS, CREAMS, GLEAMS, EPIC, KINEROS, SPUR, SWAT, SRM, WEPP, and RUSLE. More recently some of the ARS watersheds have served as core sites and successful examples of integrating science with local policy and decision-making within the UNESCO Hydrology for the Environment, Life, and Policy (HELP) Program (www.unesco.org/water/ihp/help).

Data management at ARS Watersheds

Much of the original instrumentation, installation and data processing procedures for basic rainfall, discharge and meteorological data were guided by Handbook 224 (Brakensiek et al. 1979). However, data collection evolved differently at individual locations to address different research needs and dramatically different watershed response across hydro-climatic regions (e.g. snow, thunderstorm, groundwater dominated watersheds). Instruments, parameters observed, and data reduction procedures vary among watersheds. All locations have information and data about climate, discharge, soils, topography, and land management. Data about channel properties and processes is variable. Some sites collect groundwater and water quality data, and the parameters monitored vary among sites.

Availability of data from the watersheds also varies by location. Until 1990, basic rainfall-runoff data were compiled by the ARS Hydrology and Remote Sensing Laboratory and can be obtained at http://hydrolab.arsusda.gov/wdc/arswater.html. About 16,600 station years of data are stored there from watersheds ranging from 0.2 hectares to 12,400 km^2. After 1990, centralized data compilation and archiving ended. For many of the watershed sites, data are not uniformly accessible across the entire period of data collection or across different types of data. Climate and hydrologic data are generally most easily available, while land management and vegetative cover records are most difficult to obtain in an easily useable form. Many sites are addressing these issues, but have done so independently of one another (e.g., Slaughter et al. 2002). The network as a whole has not implemented many new information technologies, leading to delays and high costs when developing new research thrusts and partnerships. Such issues likely have contributed to under-utilization of ARS watershed data to evaluate interactions of land use change, management practices, and climate variability in national assessments.

Natural Resource Informatics in ARS

Overview

Plans are being formulated to provide computer access to ARS watershed data using modern information technologies. Several locations (Table 1) have agreed to enter into a pilot project to apply new information management principles to the ARS watershed network.

Table 1. Experimental watershed research sites in Pilot Project.

Experimental Watershed	**Location/** *Region*	**Description**[†]	*Research Focus*
Little Washita River Est. 1961	El Reno, OK (Chickasha)/ Great Plains	610 km^2 (236 mi^2)	climate variability, remote sensing, model testing
Walnut Gulch Est. 1953	Tucson, AZ (Tombstone)/ Southwest	150 km^2 (58 mi^2)	semiarid rangelands, downstream water yield, erosion, remote sensing, global change, modeling
Reynolds Creek Est. 1960	Boise, ID/ Pacific Northwest and Great Basin regions	239 km^2 (92 mi^2)	rangelands, snow deposition and melt, riparian processes, model development
Little River Est. 1967	Tifton, GA/ Coastal Plains of the Southeast	334 km^2 (129 mi^2)	low gradient flow, riparian processes, water quality, mixed use watershed, model development
Deep Loess Research Station Est. 1964	Ames, IA (Treynor)/ North Central Corn Belt	30 ha (74 ac) 61 ha (150 ac)	gully and channel erosion, cropping systems, water quality, riparian buffers, model testing
Goodwin Creek Est. 1981	Oxford, MS/ Bluff Hills of lower Mississippi Basin	21.5 km^2 (8.2 mi^2)	remote sensing, riparian corridors, erosion and sedimentation, fluvial geomorphology, model testing
WE-38 Est. 1965	University Park, PA/ Appalachian Valley and Ridge	7.4 km^2 (2.9 mi^2)	runoff generation, groundwater, surface-subsurface interactions, water quality, land use and management, modeling

[†] Most ARS watershed have a nested instrumentation network structure with gauged, internal sub-watersheds with intensive instrumentation.

Approaches include: 1) evaluate successful data management models in other ecological and natural resources research networks, 2) develop formats and standards for metadata and data files, using an interactive, consensus approach (researchers, end-users, informatics and data base management specialists), and 3) develop an operational structure for a common platform or linkages for the network. The overall goal is to develop new technologies to provide improved public access to the watershed data, while retaining local control and responsibility for the data.

Status

Data management will continue to be centered at the individual sites. An information system will be developed to extract, convert, and label data from multiple sites from a shared platform (Figure 2). The system will differ from that supported by the ARS Hydrology and Remote Sensing Laboratory through 1990 in several ways: 1) broader range of data types, e.g., weather, stream flow, topography, soils, land use, management practices, water quality or other parameters, depending on the site, 2) efficient operations that would allow automated updates of data, and 3) more diverse applications because of the more diverse data layers.

There are two primary contacts for each location, one focusing on hydrologic research issues and the other on data management issues. As this project develops, support staff and data managers will require more opportunity to network across sites than they have had historically and increased multisite researcher-data manager interactions. The Agency has committed support for a research associate to work across sites, to help ensure that data are accurately and adequately described. Additionally, this project will be supported by and provide products to the case-study watershed component of the NRCS-ARS Conservation Effects

Assessment Program being conducted under the 2002 Farm Bill.

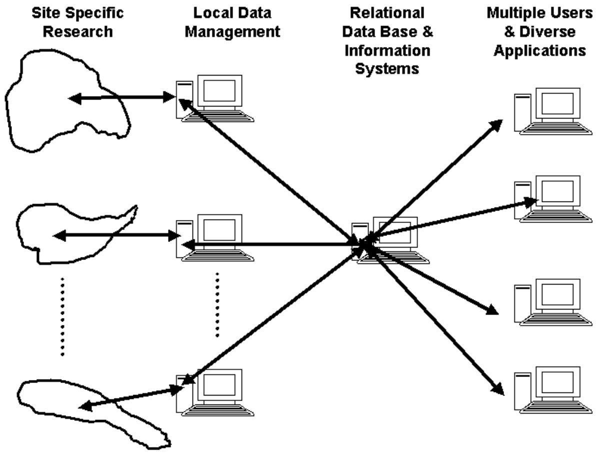

Figure 2. Data system with local control but with access to extract, convert, and label data from multiple sites for diverse problem solving.

Next Steps

As a team, we have identified data management and informatics models (e.g., U.S. Long Term Ecological Research Network http://www.lternet.edu/data/ Oklahoma Mesonet, http://okmesonet.ocs.ou.edu/, USDA UV-B Monitoring and Research Program, http://uvb.nrel.colostate.edu/UVB/uvb_climate_network.html, genomics data networks, and others) that can provide a framework to develop our system. The team will also coordinate with related activities in other USDA agencies, e.g., the Forest Service (FS) and Natural Resources Conservation Service, to ensure what we implement is compatible with other USDA activities. There may be opportunities to pursue a data network collaboratively with FS or other agencies. Three of the FS experimental watersheds are also LTER sites, and they have adopted the LTER system for data management (www.fsl.orst.edu/hydrodb/index.htm).

A critical step is to define formats and standards for metadata and data files. Because of the diverse research approaches and broad partnerships at various watershed sites, we will use an interactive, consensus approach that involves researchers, end-users, data managers, and IT specialists. Even though a pilot activity is planned, researchers and data managers from all the sites will be engaged in this step to facilitate their future incorporation into the data network. It will also be important to explore options for linking an ARS network into existing natural resources informatics networks.

In developing an operational structure, the research team is developing new partnerships within the agency to tap IT expertise. Historically, ARS has maintained a centralized IT staff primarily to support headquarters and administrative functions. The ARS Office of the Chief Information Officer (OCIO) recently developed a new function of support to Digital Research Products. Acceptance of this project within their program would provide IT staff time and some funding for development of the new information system. It could also enhance national/agency level visibility of this product (multi-site natural resource data base). Additionally, the National Agricultural Library has considerable expertise in information management and dissemination and may provide ideas and support to the effort.

This project must operate within Agency data, modeling, and GIS policies. The Agency policy may need to be updated to address critical issues such as standardization, quality control, accessibility, security, (e.g., Office of Science and Technology Policy data policy for Global Climate Research Program http://www.gcrio.org/USGSCRP/DataPolicy.html). In the future, this project could be linked to an ARS-NRCS initiative to develop an Object Modeling System in which simulation modules and appropriate databases could be assembled from a library in order to address specific scientific or natural resources questions.

Conclusions

Recent scientific assessments (NRC 2001, Hornberger et al. 2001) identified critical knowledge gaps relating to water supply, quality, and cycling processes. These reports highlighted the lack of adequate data to evaluate climate and hydrologic processes and variability. The ARS experimental watersheds provide exceptional “outdoor laboratories” where knowledge can be developed to address societal water resource issues in real world settings. These experimental watersheds provide stable and powerful research platforms to support collaborative research to investigate the hydrologic cycle and potential changes to it across a wide range of hydro-climatic conditions and agricultural ecosystems. However, lack of uniformity in data management and availability across sites and within sites impedes such new research collaborations. A pilot project to introduce natural resource informatics within the watershed network is underway to improve the data accessibility and utility. The project should increase productivity within research units, collaboration across units, and multi-partner collaboration. Internal and external researchers should obtain high quality and timely data, consistent across watersheds, extending the application of ARS watershed research to ecologic and socioeconomic, as well as agricultural and water resources problem-solving.

Acknowledgments

The authors gratefully acknowledge the contributions of the many researchers and support staff at the experimental watersheds who have operated and maintained these research programs through the years. Their faithful and dedicated service provides the basis upon which this natural resource informatics project is built.

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.

Brakensiek, D.L., H.B. Osborn, and W.J. Rawls, coordinators. 1979. Field manual for research in agricultural hydrology. U.S. Department of Agriculture, Agriculture Handbook 224.

Gleik, P.H. 1998. The World's Water. The Biennial Report on Freshwater Resources 1998-1999. Island Press, Washington, DC.

Hornberger, G.M., J.D. Aber, J. Bahr, R.C. Bales, K. Beven, E. Foufoula-Georgiou, G. Katul, J.L. Kinter III, R.D. Koster, D.P. Lettenmaier, D. McKnight, K. Miller, K. Mitchell, J.L. Roads, B.R. Scanlon, and E. Smith. 2001. A Plan for a New Science Initiative on the Global Water Cycle. U.S. Global Change Research Program, Washington, DC.

Kinzig, A.P. 2001. Bridging disciplinary divides to address environmental and intellectual challenges. Ecosystems 4:709-712.

Kinzig, A.P., S. Carpenter, M. Dove, G. Heal, S.A. Levin, J. Lubchenco, S. Schneider, and D. Starrett. 2000. Nature and Society: An Imperative for Integrated Research. Executive Summary of a report prepared for the National Science Foundation. http://lsweb.la.asu.edu/akinzig/report.htm.

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, DC.

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

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

Slaughter, C.W., D. Marks, G.N. Flerchinger, S.S. Van Vactor, and M. Burgess. 2001.Thirty-five years of research data collection at the Reynolds Creek Experimental Watershed, Idaho, United States. Water Resources Research 37:2819-2823.