A WEB-BASED DATA EXCHANGE SYSTEM FOR CROP MODEL APPLICATIONS

W. McNair Bostick,* Jawoo Koo, Valerie K. Walen, James W. Jones, and Gerrit Hoogenboom

Abstract

Development, evaluation and, application of cropping system models (CSMs) require multiple types of data. Cropping system data sets can often be used multiple times by different groups of modelers and experimentalists around the globe. However, for various reasons, data sets are often underutilized while at the same time, lack of data creates a bottleneck for many modeling efforts. In an effort to increase utilization of cropping system data, the International Consortium for Agricultural Systems Applications (ICASA) has developed an internet-based system, the ICASA Data Exchange (ICASA DE; www.icasa.net; verified 15 Feb. 2004), that provides a convenient forum for documenting, archiving, and exchanging cropping system experiment and/or weather data sets. Users of the ICASA DE can enter metadata that describe their data sets, upload their data set files, edit their entries, search for data using specific criteria, browse metadata of data sets in the system, and download data sets from the system. The purpose of this paper is to present the functional design and implementation of this system.

Cropping system models integrate data and knowledge of soil, plant, and atmosphere systems to allow simulation of cropping systems over a wide range of environments and management practices (Larson et al., 1996; Pala et al., 1996; Cavero et al., 1998; Alves and Nortcliff, 2000; Mailhol et al., 2001). This makes them valuable tools for agricultural professionals around the world. However, effective applications of CSMs require a minimum set of weather, soil, and management data (Hunt and Boote, 1998). Development and evaluation of models require all of the aforementioned types of data together with additional data such as time-series data on crop development, soil moisture, and soil nutrients as well as yield and yield components data.

In some modeling efforts, data collection and dissemination systems are in place, and data availability may not be a constraint. Examples of systems include the Georgia Automated Environmental Monitoring Network (Hoogenboom, 1993), which provides weather data from the state of Georgia; MetBroker (Laurenson et al., 2002), which is a *mediator* application to provide access to online weather databases; and EuroSOMNET, which is a soil organic matter experiment database (Franko et al., 2002). However, in many cases, data acquisition requires significant resources, and data availability becomes a constraint.

Data collected in agricultural research can often be used to achieve multiple objectives. For example, data collected by an experimentalist to test a hypothesis may also be usable by a modeler in development or evaluation of a CSM. Examples of cropping system data sets that may be useful for working with CSMs include irrigation and fertilizer response experiments, variety trials, pest management experiments, and long-term crop rotation experiments. In addition, local weather data sets that coincide with these experiments are needed. However, modelers and experimentalists are often unaware of collaboration possibilities and invest significant resources in collecting data solely to meet their specific objective. Furthermore, after a project's objectives have been achieved, the data collected for an experiment, which could be used by multiple researchers working with CSMs, are often filed away with poor documentation and never used again (Hunt et al., 2001; van Evert et al., 1999). We believe that the efficiency of CSM endeavors can be improved by implementing a system for archiving data with adequate documentation (metadata) and providing easy access to data for those who need it.

Internet database systems, such as the Georgia Automated Environmental Monitoring Network, MetBroker, and EuroSOMNET, are an effective way of archiving and disseminating data because anyone with internet access can obtain metadata and/or data from these systems. Moreover, low-cost and public-domain software are now available for developing such systems.

We realize that, for various reasons, not all data can be shared, and some researchers are reluctant to share data at all. On the other hand, data sharing and other exchanges between researchers are often essential to the success of research endeavors. This is one of the reasons that many funding organizations, such as the National Science Foundation (NSF) and the National Aeronautics and Space Administration (NASA), specify that that grantees should make project-supported data available to other researchers and the public (NSF, 2002; NASA, 2003). To provide a forum for data sharing in the cropping system research community, we have developed an internet-based system for researchers to document and exchange cropping system data sets and corresponding weather data sets. The purpose of this paper is to present the functional design and implementation of this system.

W.M. Bostick, J. Koo, V.K. Walen, and J.W. Jones, Dep. of Agric. and Biol. Eng., Univ. of Florida, Gainesville, FL 32611; and G. Hoogenboom, Dep. of Biol. and Agric. Eng., The Univ. of Georgia, Griffin, GA 30223. Florida Agric. Exp. Stn. Journal Ser. no. R-09492. This research was supported in part by the Soil Management CRSP Project, funded by USAID, and by "Carbon from Communities: A Satellite View" funded by NASA. Received 7 May 2003. *Corresponding author (mcnair2@ufl.edu).

Published in Agron. J. 96:853–856 (2004).

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: CSM, cropping system model; ICASA, International Consortium for Agricultural Systems Applications; ICASA DE, International Consortium for Agricultural Systems Applications Data Exchange (system).

System Description

The system is administered by ICASA (www.icasa.net), an open-membership organization that envisions a future characterized by broad acceptance of systems-oriented methodologies in agriculture and natural resources. The goal of ICASA is to advance agricultural systems research through the development and application of methodologies and tools for agricultural systems analysis.

The system is named the ICASA Data Exchange (ICASA DE). Users of the ICASA DE can enter metadata that describe their data sets, upload their data set files, edit their entries, search for data using specific criteria, browse metadata of data sets in the system, and download data sets from the system. Members of ICASA have access to all of the functions of the ICASA DE. To become an ICASA member, one must state that he or she agrees with ICASA vision and goals and will contribute to advancing these goals. By requiring ICASA membership for full access to the system, we feel that we create a forum that users will be comfortable contributing to because it will be accessible by like-minded scientists. In addition, we feel that this requirement will emphasize the importance of voluntary contributions to

OWNERS OF DATASET		
Owner(s)	*Owner (1):	Jose Cavero - Dept. Genetica y Produccion Vegetal, Zaragoza, Spain - e-mail: jcavero@eead.csic.es
EXPERIMENT INFORMATION		
	*Start - End:	1995 - 1996
	*Name (Code):	Spain (ES)
	Single Crop:	Maize
	*Name:	Estacion Experimental de Aula Dei
	City:	Zaragoza
	*Institution:	Agronomic Research Service (SIA)
	*Lat /Lon/Elev:	41.72 / -0.80 / 225m
Soil		Loam
Treatment		Irrigation
Experiment		Multiple Season
DATA		
Climate	Temperature:	5-10 years
	Rainfall:	5-10 years
Daily Weather	Available Data:	Rainfall Temperature (Max/Min) Solar Radiation
	Location Measured:	at site
Crop	Management Inputs:	Planting Date
	Inputs:	Plant Density/Spacing Variety Irrigation Fertilizer
	Measurement(s):	Phenology: Flowering Phenology: Maturity End-of-Season: Biomass End-of-Season: Grain/Seed/Tuber End-of-Season: LAI End-of-Season: Other
FILE & PUBLICATION INFORMATION		
File Information	Data Format:	DSSAT v4.0
	Data Access:	Download
	Verification:	Verified use with DSSAT
	*Contact:	Jose Cavero - Dept. Genetica y Produccion Vegetal, Zaragoza, Spain - e-mail: jcavero@eead.csic.es - fax: 34-976-716145
Publication	Citation:	Cavero, J. et al. Agron J 92:679-690
	Link:	N/A
	Uploaded Publication:	N/A

Fig. 1. Metadata of one of the experiment data sets in the database. * Denotes required fields.

the system and encourage those that download data to contribute data also. If people do not want to become ICASA members, they can still access all of the metadata in the system and contact the data holder if they want to obtain data sets documented in the ICASA DE.

Data Entry and Editing

Metadata and data files are stored in tables of a relational database. For metadata entry, users complete an online form by using either check boxes or pull-down menus for most entries. Examples of metadata for experiment and weather data sets are shown in Fig. 1 and 2, respectively. Metadata fields on the entry form were chosen to provide a complete description of the contents of the experimental and weather data sets. Required fields for both experimental and weather data sets (denoted by an * in Fig. 1 and 2) are owner name and contact information, dates of data collection, site information, and data holder name and contact information. Other fields are optional.

If the users choose to upload data files, they will be prompted to upload after they have completed the metadata entry form. Submission of files that follow the ICASA data standard (www.icasa.net) is strongly encouraged. However, it is not required that file submissions follow this format, as we feel that this could deter scientists that are not familiar with the format from entering data that could be very beneficial to the CSM community. To save space on the data server and to allow for easy and fast file transfers, only compressed .zip files can be uploaded. The ICASA DE data-editing interface provides a user with the option to correct errors found in the metadata or data files after submission or remove the data from the database.

Data Search and Download

When searching for data, users can customize their search by specifying a country, latitude and longitude, and/or crop (for experiment data only). After the search criteria are entered, the system searches the database tables and returns summary information for each data set that meets the criteria. If no criteria are entered, summaries for all data in the database are shown. Users can then select specific data sets to view those data sets' metadata. If data files are available for download, a download link is displayed at the bottom of the metadata (Fig. 1 and 2). If data files are not available, the data holder's contact information is provided so that users can request access to the data or obtain additional information about the data.

File Verification

The only metadata field not entered by a user is the verification status, which is determined by the ICASA

OWNERS OF DATASET		
Owner(s)	*Owner (1):	James Jones - University of Florida - e-mail: jjones@agen.ufl.edu
EXPERIMENT INFORMATION		
	*Start - End:	1970 - 1987
	*Name (Code):	United States (US) Irrigation Park
	City:	Gainesville
	State/Province:	Florida University of Florida
	*Lat/Long/Elev:	29.63 / -82.37 / 10m
DATA		
Level of Detail	Data Measured:	Daily
Measurements	Available Data:	Rainfall Maximum Temperature Minimum Temperature Solar Radiation
Weather Station	Type:	Manual
FILE & PUBLICATION INFORMATION		
File Information	Data Format:	DSSAT v4.0
	Data Access:	Download
	Verification:	Verified use with DSSAT
	*Contact:	James Jones - University of Florida - e-mail: jjones@agen.ufl.edu
	Link:	N/A

Fig. 2. Metadata of one of the weather data sets in the database. * Denotes required fields.

DE administrator. Initially, uploaded files have an unverified status. Uploaded files are checked by ICASA to confirm that they contain data pertaining to cropping systems, and if this is the case, they receive verified status. The main purpose of this first level of verification it to assure that the system is not being used to exchange files unrelated to cropping systems. As an additional level of verification, ICASA may also evaluate whether the data works with a specific application, such as a model. If they do work, this is also specified in the verification status. For example, the verification status in Fig. 1 and 2 indicate that the data files have been verified as operable with the DSSAT CSM. Verification is conducted to confirm the integrity of a data set, but it is not meant to be a rating of the overall scientific quality of the data.

Implementation

The user interface was developed using Active Server Pages scripts (ASP; Microsoft Corp., Redmond, WA) and Hypertext Markup Language (HTML). The relational database was built using the database software MySQL 3.2.53 (MySQL Inc., Seattle, WA). A graphical database management and monitoring application, i.e., MySQL Studio 4.7.2 (PremiumSoft CyberTech Ltd., Central, Hong Kong) was used to allow for easy and visible management of the database. An open database connectivity driver, MyODBC 2.50 (MySQL Inc., Seattle, WA), was used to link the user interface with the database. The server computer used the Windows 2000 Advanced Server (Microsoft Corp., Redmond, WA) operating system and Internet Information Server web server software (Microsoft Corp., Redmond, WA).

MySQL database software was selected because it is commonly used, open-source database software. Active Server Pages scripts were used because they work well with a Windows server without any additions that would be needed with a similar scripting language such as JavaServer Pages (JSP). Active Server Pages scripts are also widely used, and it will be easy to maintain and improve the system.

Discussion

The ICASA DE is convenient and easy to use because it is available for accessing and archiving data at any time from any location with internet access. The system works well with Internet Explorer version 5.0 or later and Netscape version 6.0 or later. We feel that the system will improve data availability and increase multiple uses of data. Increased access to data will result in more opportunities for improvement and applications of CSMs and will increase benefits derived from data collection.

A key improvement that we envision for the ICASA DE involves the metadata entry process. Although the data entry form is clearly laid out and has a simple user interface, manually searching through data files to extract metadata to fill out the forms can be tedious and may deter data holders from entering data. Therefore, a further step of system development would be to automate metadata and data uploading to the database.

Both ICASA as an open-membership organization and the ICASA DE have been in existence for less than one year. As of 8 Oct. 2003, metadata and data files for 52 experiments and two weather stations had been uploaded to the system. One of the main principles of ICASA is cooperation between researchers through sharing data, software applications, and/or other entities. As the ICASA DE becomes more well known in the cropping systems research community, it will be a widely used and important tool to those who are developing, evaluating, and applying CSMs.

References

Alves, H.M.R., and S. Nortcliff. 2000. Assessing potential production of maize using simulation models for land evaluation in Brazil. Soil Use Manage. 16:49–55.

Cavero, J., R.E. Plant, C. Shennan, J.R. Williams, J.R. Kiniry, and V.W. Benson. 1998. Application of Epic Model to nitrogen cycling in irrigated processing tomatoes under different management systems. Agric. Syst. 56:391–414.

Franko, U., G. Schramm, V. Rodionova, M. Korschens, P. Smith, K. Coleman, V. Romanenkov, and L. Shevtsova. 2002. EuroSOMNET—a database for long-term experiments on soil organic matter in Europe. Comput. Electron. Agric. 33:233–239.

Hoogenboom, G. 1993. The Georgia Automated Environmental Monitoring Network. Southeast. Clim. Rev. 4(1):12–18.

Hunt, L.A., and K.J. Boote. 1998. Data for model operation, calibration, and evaluation. p. 9–39. *In* G.Y. Tsuji et al. (ed.) Understanding options for agricultural production. System approaches for sustainable agriculture development. Kluwer Academic Publ., Dordrecht, the Netherlands.

Hunt, L.A., J.W. White, and G. Hoogenboom. 2001. Agronomic data: Advances in documentation and protocols for exchange and use. Agric. Syst. 70:477–492.

Larson, J.A., H.P. Mapp, L.V. Verhalen, and J.C. Band. 1996. Adapting a cotton model for decision analyses: A yield-response evaluation. Agric. Syst. 50:145–167.

Laurenson, M.R., K. Takuji, and S. Ninomiya. 2002. Providing agricultural models with mediated access to heterogeneous weather databases. Appl. Eng. Agric. 18(5):617–625.

Mailhol, J.C., P. Ruelle, and I. Nemeth. 2001. Impact of fertilisation practices on nitrogen leaching under irrigation. Irrig. Sci. 20:139–147.

[NASA] National Aeronautics and Space Administration. 2003. Guidebook for proposers responding to a NASA research announcement (NRA) [Online]. Available at http://www.hq.nasa.gov/office/procurement/nraguidebook/ (accessed 16 Sept. 2003; verified 16 Feb. 2004). NASA, Washington, DC.

[NSF] National Science Foundation. 2002. NSF grant policy manual [Online]. Available at http://www.nsf.gov/pubs/2002/nsf02151/ (accessed 16 Sept. 2003; verified 16 Feb. 2004). NSF, Arlington, VA.

Pala, M., C.O. Stockle, and H.C. Harris. 1996. Simulation of durum wheat (*Triticum turgidum* ssp. Durum) growth under different water and nitrogen regimes in a Mediterranean environment using CropSyst. Agric. Syst. 51:147–163.

van Evert, F.K., E.J.A. Spaans, S.D. Krieger, J.V. Carlis, and J.M. Baker. 1999. A database for agroecological research data: I. Data model. Agron. J. 91:54–62.