Expert Group Meeting on Software for Agroclimatic Data Management

The following paper is the Executive Summary by Ray Motha from the proceedings of an Expert Group Meeting on Software for Agroclimatic Data Management, October 16-20, 2000 held in Washington, D.C. USA.

Introduction

The U.S. Department of Agriculture and the World Meteorological Organization jointly sponsored an Expert Group Meeting on Software for Agroclimatic Data Management on October 16-20, 2000 in Washington, D.C. Three working groups were established during the meeting: a Climate Data Working Group, a Crop and Soils Working Group and a Remotely Sensed/Integrated Packages Working Group. Each group prepared a draft report. This report summarizes the findings and recommendations of all three reports, highlighting the specific issues of each and the common issues of all groups.

Agroclimatic data are used for a highly diverse range of applications. These applications require different sets of data, ranging from hourly observations to long-term climatic records and from point-source data to spatially interpolated products. The results of analysis are used for both short-term tactical decisions and long-term strategic decisions. Traditionally, the user community has relied on time series records of point source data. Both applications and technology have become more sophisticated, however, requiring timely access to additional data sources including automated weather systems, remote sensing platforms and computer-generated products. Software tools for georeferenced data are also becoming a prerequisite for integrated agroclimatic data management.

The diverse user community ranges from farmers to national and international planning agencies. The demand for information by the user community has increased dramatically due in part to the recognition of the importance of agroclimatic information for decision making, and in part to increased economic pressures and environmental concerns. This information is not only needed at the farm level for daily tactical decisions on planting, spraying and harvesting, but also for long-term strategic decisions at the national and international levels concerning trade issues, global variability and risk management (Figure 1).

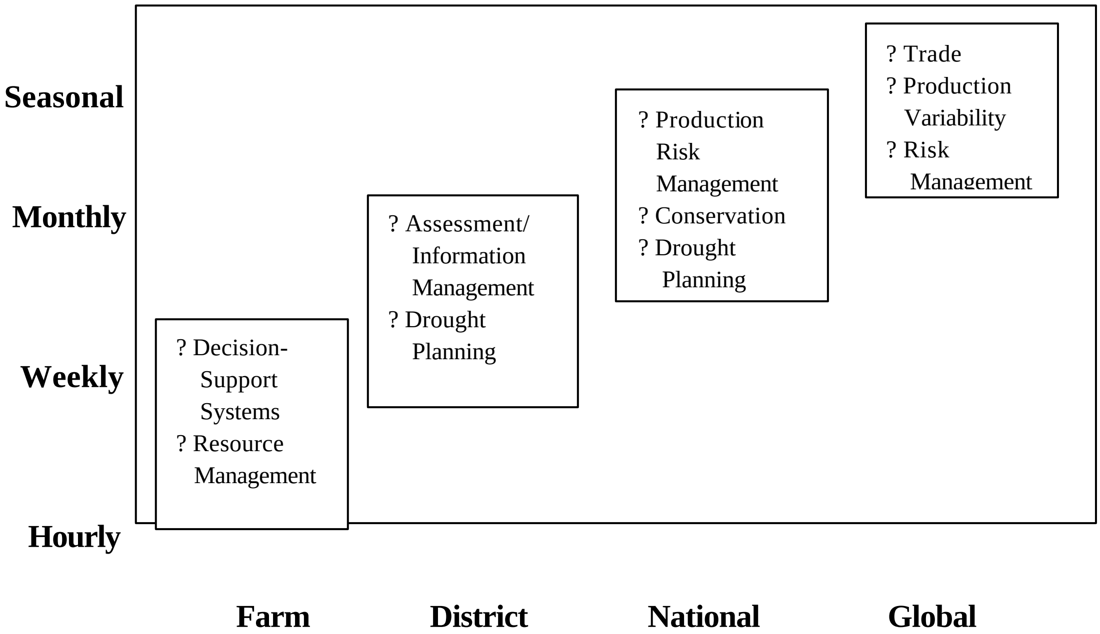

Figure 1. Agroclimatic Decision Making (Tactical to Strategic).

An ideal agroclimatic information system needs the following components:

- An efficient data collection system
- A modern telecommunication system
- An industry-standard data management and processing system
- A modular-based data analysis system
- A technologically-advanced product and information delivery system

Each working group summarized their results according to four specific objectives:

I. Evaluate the current status of software for data management;
II. Identify shortcomings and limitations of current software, and, provide recommendations for future developments;
III. Formulate guidelines for improved management of databases in support of agroclimatic applications specifically to assist training and capacity building; and,
IV. Review and recommend appropriate public domain software packages to be included in CD-Roms for free distribution to all interested parties in agroclimatic data management.

Evaluate the Current Status of Software for Data Management

Each of the working groups noted that while the framework exists for a comprehensive approach to data management, most systems were tailored to satisfy specific needs. These needs range from operational assessments to research applications. Data requirements range from complex hourly values for sophisticated physiological modeling techniques to monthly or seasonal averages for statistical analyses. Analytical tools range from simple time series analyses to georeferenced spatial analyses. Given this broad spectrum of requirements, the experts recognized that the most successful approach to software development will be based on a modular structure with an open architecture. The expert group recognized that there can be greater application of some of the current systems if selected features were integrated into a more comprehensive management systems approach.

A large number of software packages are available for agroclimatic data processing, analysis, and dissemination. Most of these software packages were developed for specific applications. The techniques for weather and climate data management vary widely and are dependent upon the type of data networks, telecommunications, data storage capability and processing power of each system. Similarly, the current status of software for crop models and soil databases is diverse and targets many different applications. Computer models have different specialties and may be appropriate at different scales. Input/output standards as well as model objectives are quite different. The expert group noted that a ranking of the objectives of different models would be useful in trying to assess their applicability. These rankings depend more on model objectives than on model temporal or spatial scales. For example, nutrient balance is not only important for crop productivity at the field level, but also for assessing food security at the national level.

The current status of software for remote sensing and integrated packages is still not focused on agrometeorological applications. It is often necessary to learn and use numerous packages or develop specific applications to provide adequate agricultural analysis, especially for integrated packages. There are numerous viewers for display of raster data and an increasing number of tools for analysis of georeferenced digital data sets. Few commercial software packages are solely directed toward crop modeling, and there is little integration of crop modeling parameters in existing software. Freeware is improving, but is limited to specific data formats and has limited technical support. Commercial software is still too general and is considered primarily as a developmental tool. Furthermore, commercial software packages are often too expensive, but they provide the best technical support and continuous updates. Raster viewing packages are becoming more readily available, but they often require specific hardware and software components. Software packages developed in universities, either freeware or commercial, provide the best opportunity for training and capacity building instructions.

Available Software Packages for Agroclimatic Data Management

A listing of some of the different software packages for data management that were discussed or reviewed in this meeting is presented. This list is by no means complete, nor does it represent an endorsement of any freeware or commercial software. The purpose of this list is to provide an informative listing of currently available software which may serve as a basis for additional inquiry as needed. While there may be some overlap, the listing is presented by the following categories: climate; crop and soils; and remote sensing and integrated packages. Some indication of the use of the software is also presented (COTS refers to "Commercial-Off-The Shelf" software).

A. Climate

1. Programming Languages:

- *C*, *C++*, *Visual Basic* (COTS)
  - Data processing
  - Data import/export
  - Data quality control
  - Modeling (crop stage, PET, soil moisture)
- *SAS* (Subscription, Expensive)
  - Statistical analysis
  - Geographic display
  - Modeling
  - Data base management/development
- *ORACLE*, *SAS* (Subscription, Expensive)
  - Data storage
  - Data base management system
  - Data analysis
- *LOTUS 1-2-3 / MS EXCEL* (COTS)
  - Data time series plots
  - Regression analysis
  - Geographic display
  - Spreadsheet analysis
- *MS ACCESS* (COTS)
  - Application data interfacing
  - Link to spreadsheets
- *FREELANCE GRAPHICS* (COTS)
  - Briefing charts
  - Automated display updates (linked with LOTUS 1-2-3)
  - Graphics
- *ARCVIEW 3.2 / SPATIAL ANALYST / IMAGE ANALYST* (Expensive)
  - Create maps
  - Geographic information system (GIS) map overlay
  - Data analysis
- *CORELDRAW 9* (COTS)
  - Create and edit map graphics
  - Input and edit postscript graphics
- *PAINTSHOP PRO / LVIEW* (COTS)
  - Satellite image editor

- *ADOBE ACROBAT* (COTS)
  - Internet *document dissemination*
- *FRONTPAGE EXPRESS / DREAMWEAVER* (COTS)
  - Create web pages

2. Data Management Software Development

- *Australian Rainman* (Subscription)
  - Seasonal rainfall forecasting information
  - Continuous climate records through "patched point" data interpolation
- *Colchique* (Meteo-France)
  - Software developed for access to automated data networks
  - Fee-based external users via personal computer
- *UCAN-Unified Climate Access Network* (US Regional Climate Center Consortium)
  - Internet-based system
  - Data query and retrieval system for access to diverse climate data sets
  - Plans for a national distribution network for climate products via the Internet
- *FAOCLIM2 (FAO)*
  - Worldwide monthly and time-series climatic database for about 25,000 stations
  - Software for selection of stations and extraction of data
  - Software for selection of stations and visualization of data

B. Crop and Soils

1. Crop

- *WOFOST* (Wageningen University)
  - Limited data required
  - International crop simulation model
  - All key nutrients represented
- *CENTURY* (Colorado State University)
  - Limited data required
  - Considers changes in soil properties over time
- *ALES* (Rossitter)
  - Limited data required
  - Automated land evaluation system
  - Suitable assessments of land for different purposes
- *DSSAT* (University of Florida, University of Hawaii)
  - Comprehensive data required
  - Widely used for a nominal fee
  - Recognized training tool and supports several crops
  - N-balance only included in terms of nutrients
- *STICS* (France-INRA)
  - Multidisciplinary model built with increasing levels of complexity

- Multiple crops

- *GOSSYM* (USDA)
  - Comprehensive data required
  - Single commodity with high management complexity
- *FAOMET* (FAO)
  - Agrometeorological crop forecasting tool
- *FAOINDEX* (FAO)
  - Crop specific water requirements satisfaction index (for 10 crops)
- *ECOCROP* (FAO)
  - Identifies crop environmental response to key climate and soil requirements

2. Soils

- *EPIC* (Texas A&M)
  - Comprehensive data required
  - Soil erosion
- *STATSGO*
  - derived soil root zone available water holding capacity
- *SOTER* (Wageningen University)
  - Global and National Soils and Terrain Digital Database
  - Linked to a relational database management system
  - Used with any commercially available GIS software
- *WOCAT* (University of Berne, Switzerland)
  - World Overview of Conservation Approaches and Technologies
  - Database management system for storage and use of soil and water conservation

C. Remotely Sensed/Integrated Packages

- *WinDisp* (FAO)
  - display and analysis of satellite imagery
  - DOS and Windows versions
- *Addapix* (FAO)
  - pixel-by-pixel classification software (DOS version)
  - university developed
- *Multispec* (Purdue University)
  - Windows classification software
  - university developed
- *LARS* (U.S. Geological Survey)
  - UNIX shareware
- *ArcView/ArcInfo* (COTS)
  - Create maps
  - Geographic information system (GIS) map overlay
  - Data analysis
- *GRASS*
  - Geographic information system (GIS)

    - Data analysis
- *PCI*
    - tailor made, first set-aside agricultural modeling package
- *ERDAS*
    - Robust
- *IDRISI*
    - manual very good
- *ENVI*
    - from research community
- *ERMapper*
- *EPIC* - Texas A&M, freeware
    - Erosion/crop model
    - Multiple crops
- *DSSAT* (University of Florida, University of Hawaii)
    - Multiple crops
    - Relatively low cost
- *CropSyst* (Washington State University)
    - Freeware
- *CropWat* (FAO)

Identify Shortcomings and Limitations in Utilization of Current Software and Provide Recommendations for Future Developments

A number of limitations to the various data management techniques used for both research and operational applications were recognized.

Shortcomings for Climatic Data Management

Specific shortcomings in the use of software include the following:

- Inadequate data exchange standards;
- Diverse and incomplete quality control standards;
- Lack of data continuity over long time periods;
- Inaccessible or difficult to access data sets;
- Poor communication between different systems;
- Cost of systems and data;
- Insufficient or absence of metadata;
- Sparse station coverage in agricultural areas;
- Lack of long-term commitments to sustaining station networks;
- Widely diverse levels of expertise; and,
- A lack of full commitment to exchange necessary data sets at regional, national and international levels.

Shortcomings for Crop and Soils Data Management

With respect to crops and soils, shortcomings and limitations include the following:

- Insufficient treatment of the nutrient balance in some of the key models;
- Crop models are not environmentally sensitive to such factors as salinization or excess nutrients that affect sustainability;
- Pests, including insects, weeds and pathogens, are not generally addressed;
- Hourly step models representing key factors are not readily available (key factors include rainfall intensity and temperature response that strongly affect runoff and growth, respectively);
- Spatial variability is not well addressed in models that assume homogeneity in the micro-environment;
- Mixed cropping models are not commonly available;
- Global change response is limited mainly to climatological models, rather than crop growth models;
- Spatial components for these models are often missing;
- Soils data sets lack a shared structure or standardization;
- Databases tend to offer more information about chemical (fertility) properties than soil physical properties;
- While pH data are available, there is a need for better pedon transfer functions;
- High resolution (field scale) digital soils data are often not available;
- Soil observations (pedon descriptions) are not utilized to estimate soil biota; and,
- Many soil attributes and properties are under utilized by crop growth models.

The expert group meeting recognized that despite the existence of crop models since 1980s, their operational applications have not progressed as expected.

Shortcomings for Remote Sensing and Integrated Modeling Packages

Shortcomings and limitations in the use of current software for remote sensing and integrated modeling packages include:

- Most modelers viewed models as research tools to provide answers to "what if" type questions while others expected the models to provide answers to real-life strategic and tactical tasks;
- Most model development activity tended to focus on research applications and ignored technology transfer, hence, modelers did not establish good contacts with end users that could have facilitated model applications at the field level;
- Majority of the crop models are data intensive and the needed climate/crop/soil data are often not readily available, especially in the developing countries;

- While the modelers' data needs are often quite rigid and they expect the data to come from a single source, in reality data sources and formats are quite variable across countries;
- The problem is further complicated by the fact that climate, soil and crop data are not often collected or available from the same location;
- For most of the public-domain software, user support is often limited; lack of training material and opportunities in remote sensing and modeling applications, especially in the developing countries;
- A good majority of the current software packages that facilitate spatial analysis are not equipped to perform adequate temporal analysis;
- Most process-based models are quite complex and perplex the end users unless they are skilled in the use of the models; and,
- In developing countries there is a growing "digital divide" between senior managers and their younger work force.

Recommendations

Due to these limitations and shortcomings, the working group recommends the following:

- More AWS weather stations are needed in order to provide coverage and support for risk management, crop assessment, crop productivity, fire and rangeland management, and natural resource conservation;
- There is a need for long-term support and commitment (funding, capacity) to sustain reliable station networks;
- Systems should be developed to facilitate data sharing/exchange;
- Continuous data records must be established using standardized methodologies (such as the Patched Point Dataset in Australia);
- Information delivery systems should be Internet based (Internet Data Distribution, IDD);
- Systems should be modular, scalable (able to grow, expandable), open-architecture (run on many different hardware and operating systems) and open-sourced;
- Systems need to be developed with user requirements in mind (both research and operational applications) and should allow rapid response to user requests for specified information. These information needs vary by request and by user;
- Systems should have the capability to handle ad hoc requests from remote users;
- Systems should have the capability to link with other data providers to ensure that the appropriate information is accessible in a timely and usable format;
- Metadata information needs to be developed and should conform to ISO standards;
- All data should be georeferenced and time-stamped for effective integration into GIS software systems;
- Member nations should make available basic geopolitical data sets on the internet. WMO should maintain a web-site with links to these data sets;
- Develop a website with pointers to relevant decision-support software, data and development tools with the capacity for online demonstration of its ability;
- WMO should develop guidelines for the preparation of agro-climatic bulletins;

- There is a need to develop "classical" case studies that motivate applications and exemplify usefulness;
- Soil data management systems should be harmonized and integrated, similar to Soils and Terrain Digital Database (*SOTER*);
- National systems need to be identified as sources of information;
- Since some models and databases are scale independent, an understanding of "data loss" from one scale to another must be acknowledged explicitly;
- Systems need to be developed that simulates the temporal dynamics of soils (*EPIC* provides an approach to erosion but is limited in terms of requirements);
- Time trend analyses are needed for land use databases;
- Efforts must be made to fill the gap between research and operational applications through adoption of appropriate methods and tools;
- Two types of interfaces should be built, one between crop modelers and data providers and the other between modelers and users, to facilitate easy and rapid model applications in decision making;
- National meteorological and hydrological services should be urged to share their knowledge and tools in spatial interpolation with database managers and application developers;
- Special efforts must be made to develop and disseminate uniform formats and data sheets for recording crop and soils information in a format that is compatible for use with crop models and remote sensing applications; and,
- Options for model users need to be further simplified and streamlined to facilitate their wider use by the user community.

Formulate Guidelines for Improved Management of Databases in Support of Agroclimatic Applications Specifically To Assist Training and Capacity Building

Guidelines should be developed to provide a step by step methodology in a generic format for training and capacity building. These guidelines should be in a modular structure to allow sufficient flexibility in a training program to focus on the specific objectives. Recommendations for software development and management should consider the following features:

- Availability of industry standard or open-source software acceptable for a wide variety of agroclimatic applications, ranging from field research to large scale operational applications;
- Data sets must have a minimum metadata base, standard format, standard quality control procedures and adequate continuity of records;
- Procedures to train personnel to recognize inconsistencies of data and the ability to establish appropriate patch-point methods to maintain continuity;
- Software must be compatible with both temporal and spatial data sets to allow for the integration of point source data with georeferenced digital data sets, modeling technology and remotely sensed data;
- New technology in telecommunications should be utilized to bridge the gap between automated data collection systems and web-based information systems;
- GIS metadata are required for appropriate coordinate systems, projections etc.;

- A library of useful applications, case studies, risk analyses and other studies based on archival records must be developed to support crop model applications. This will stimulate the collection, support, maintenance and dissemination of databases;
- In order to make workshops and training seminars more effective for operational applications, an infrastructure for follow-up after the training is needed;
- Establish a "listserve" or other virtual community (The International Soil Reference and Information Centre, ISRIC, Netherlands will develop a website to serve as a "portal for soil information," including metadata information on existing soil databases worldwide);
- Improved data exchange will be fostered by the continued development of geographic frameworks and adoption of standards;
- Cooperation among existing training centers, virtual communities, and international development organizations must be explored to maintain and expand the current capabilities of the agroclimatic community;
- Guidelines need to include new measures and better assessment of soil data reliability given that more robust estimates are obviously linked to data quality and resolution (data resolution is a continuing need previously identified);
- Recommend determining indices of data reliability (quality) vs. error percentages, and, sources of the data (ISO metadata standards). Examples of error tracking and reporting from the agroclimatic community include the IQ index from France and the quality indices from South Africa;
- Guidelines need to emphasize explicit accounting for variability in the soils arena. A "diversity index" (USDA Soil Survey) has been developed to characterize and report soil variability. Similar variability measures are key and they should be associated with input to soil and crop models. Similarly, plant variability (e.g. emergence and growth) needs to be represented especially if this variability is a key to management under a given production scenario;
- Guidelines are needed to report by-country crop yields including sub-national trends and geo-spatial time trends;
- An industry, international standard for agroclimatic data, particularly crop data, should be developed. Minimum metadata, database formats and database content should be resolved, accepted by industry, international centers and academia, and published basic agroclimatic applications;
- Technology transfer of remote sensing and crop modeling software development to users should be improved. This entails supporting international training programs and scholarships and support for applied research;
- More development for telecommunication research which develops new applications using web-based technologies and automated observations should be supported. These developments will help reach more users and encourage standardization;
- Published, professionally accepted standards for content, naming and format convention for applications should be promoted;
- Current measures of standardization are considered mature for meteorological data, poor for crop data and emerging for soils data. Hence, uniform measures of standardization for agroclimatic purposes must be established;
- Spatial interpolation methods for specific applications should be recommended;

- Better curricula and training programs must be encouraged;
- International and regional support for training programs should be improved;
- University-to-university exchange and cross-training programs should be improved;
- International agroclimatic scholarship programs should be approved; and,
- Software should be user-friendly, with good interaction between data, analysis and product creation phases.

Recommended Public Domain Software Packages:
(Some software may be listed under more than one sub-heading)

Database Management Software:

- *PostgreSQL* - relational database management system
- *MySQL* - relational database management system
- *ArcExplorer* - GIS viewing and data query
- *PCGrads* - data plotting, analysis and time series
- *Perl* - practical extraction and reporting language
- *Python* - object-oriented scripting language
- *netCDF* - network common data form: self-describing file format
- *Linux/UNIX* - database management
- *udunits* - library for unit conversion

Soils:

- *SOTER* - Soils and Terrain Digital Database
- *WOCAT* - World Overview of Conservation Approaches and Technologies
- USDA Soil Survey - data aggregation tool (Access)

Crops:

- *DSSAT* - Decision Support Systems for Agrotechnology Transfer
- *CropSyst* - Cropping Systems Simulation Model
- STICS - Multi-crop, multi-disciplinary model
- *GOSSYM* - Cotton simulation software
- *WOFOST* - World Food Studies Crop Modeling Program
- *EPIC* - Erosion Productivity Impact Calculator
- *ALES* - Automated Land Elevation System

GIS:

- *Grass* - Data analysis
- *ArcView 2.x* - Creates maps, data analysis

Georeferenced digitial data sets:

- *WinDisp* - display and analysis of satellite imagery, Windows, FAO
- *Addapix* - classification software, university developed, DOS, FAO
- *Multispec* - classification software, university developed (Purdue),Windows

Spatial image analysis software:

- *ArcExplorer* - image analysis, data query
- *IDRISI* - spatial image analysis

Integration software:

- *EPIC* - Texas A&M, freeware
- *DSSAT* - University of Hawaii consortium initially (ICASA)
- *CropSyst* - Washington State University, freeware
- *CropWat* - FAO
- *Century* - University of Colorado, widely used