ICASA Version 1.0 Data Standards for Agricultural Research and Decision Support

L.A. Hunt[1], G. Hoogenboom[2], J.W. Jones[3], and J.W. White[4,*]

[1]University of Guelph, Guelph, Canada, [2]University of Georgia, Griffin, GA, USA, [3]University of Florida, Gainesville, FL, USA, [4]ALARC-USDA, Maricopa, AZ, USA

* Corresponding author, email: jwhite@uswcl.ars.ag.gov

International Consortium for Agricultural System Applications, P.O. Box 527, Honolulu, Hawaii 96822

10 August 2006

ver ICASA1.06

Abstract

The use of standard data formats for documenting experiments and modeling crop growth and development can greatly facilitate exchange of information and software, allowing researchers to focus on science rather than on re-formatting data. The standards developed by the International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) project and used in the original Decision Support System for Agrotechnology Transfer (DSSAT) software package were of considerable value in this context, but they contained ambiguities that made them unsuitable for certain uses. The International Consortium for Agricultural Systems Applications (ICASA) revised the standards in the year 2000. Specific goals were to provide more complete documentation of experiments and to allow use of the standards in a wide range of decision support situations. Further modifications proved necessary, both to permit use with crops not previously considered and to place the standards in a more general framework, necessary as software systems evolve and researchers move towards more generic mechanisms for data interchange. The ICASA standards have thus been reformulated with greater emphasis on standardizing vocabularies and clearly specifying relations among variables. This change moves the standards closer to data ontologies being developed in other branches of plant science. The modified standards, which may be implemented in an ASCII file format, in a relational database, in eXtensible Markup Language (XML) format, or in other formats, are quite similar to those documented previously. They organize data into a two-level hierarchy, using indices or keys to link information both across and within hierarchies. An overview of the ICASA Version 1.0 standards is presented in this document, and the ASCII implementation is described in some detail. Implementations being developed for relational databases and XML are also discussed briefly. The ICASA Version 1.0 Data Standards will be implemented in the next version of DSSAT.

Introduction

Efficient interchange of data among researchers, especially for use in simulation models and other decision support tools, requires use of a common vocabulary and strategy for organizing data. The agricultural research community increasingly encounters research problems that require interdisciplinary collaboration. Physiologists and molecular biologists work together to develop a better understanding of the genetic control of productivity-related traits. Agronomists, soil scientists and irrigation specialists combine efforts in order to increase the efficiency of crop water use. In such collaborations, ready data interchange is essential. Genomic data are widely available through publicly accessible databases (Blanchard, 2004). Daily weather records and soil profile data are increasingly available through the Internet. The International Research Institute for Climate Prediction recently developed two daily weather data download options in ICASA format that can be accessed from the ICASA web site (www.icasa.net/weather_data). Efforts are also underway to make the "World Inventory of Soil Emission Potentials" (WISE) database developed by the International Soil Reference and Information Centre in The Netherlands available for crop model applications. Field research data, however, are seldom available through public databases. Although there have been various initiatives to develop systems for reporting and storing data from field research, e.g., van Evert et al. (1999a; 1999b); Bostick et al. (2004), to date, no system is widely accepted as a standard.

Among the largest sustained effort to promote the use of standards in relation to field research has been that of the International Consortium for Agricultural Systems Analysis (ICASA), and one of its predecessors, the International Benchmark Soils Network for Agrotechnology Transfer (IBSNAT). As early as 1983, the IBSNAT project developed data standards to be used both to document experiments and to provide input for models or other software tools (Uehara and Tsuji, 1998). The standards and their implementation in ASCII files facilitated interactions among experimenters and modellers (Hunt et al., 1994; Jones et al., 1994). They were extensively used by experimenters and modellers using the DSSAT system (Tsuji et al., 1994) and were adopted by the Global Change and Terrestrial Ecosystem project (GCTE) for use in documenting experiments and regional yield investigations (GCTE, 1996). Within GCTE, the standards greatly assisted model comparisons (Goudriaan, 1996; Jamieson et al., 1998; White et al., 2003) that led to model improvements. To promote actual interchange of data, ICASA has also developed the ICASA Data Exchange (IDE; www.icasa.net/data_exchange/) where ICASA members can store and share experimental data.

Experience with the IBSNAT standards and files showed that they contained ambiguities and lacked fields to characterize certain crops and management practices. Members of ICASA and other organizations thus defined a revised set of standards that were intended to be unambiguous, easily processed by a broad range of software tools, and more inclusive of crops and production practices (Hunt et al., 2001). Use of this initial draft of the ICASA standards and associated ASCII files highlighted further issues that needed attention before wider use, as for example in the DSSAT software, could be promoted with confidence. Furthermore, diverse stakeholders requested that the standards be defined not only for the "flat" ASCII file format but also for formats such as relational databases and eXtensible Markup Language (XML). Thus, the ICASA standards have been reformulated, with greater emphasis on standardizing vocabularies and establishing relations among variables. This somewhat parallels efforts to develop data ontologies in other branches of plant and agricultural sciences (The Plant Ontology Consortium, 2002; FAO, 2004).

The goal of the ICASA standards, following the thrust of the earlier IBSNAT standards, is to provide a reliable, portable, and flexible structure both for documenting field experiments (or their equivalents in greenhouses or growth chambers) and for specifying realistic conditions for running dynamic simulation models. For documentation, it is desirable to have sufficient detail on weather, soil, crop cultivars, weeds, diseases, pests, and crop management to permit a full reproduction of the experiment and to conserve any information measured or noted during or at the end of the experiment. Often, the variables defined within the standards are taken as indicating the required degree of detail, the "minimum dataset" (Nix, 1984), for experimental documentation. We emphasize, however, that because of the diversity of experiments, the list of variables documented within the current standards should not be interpreted as a "minimum dataset". Indeed, since the standards can be extended to accommodate new variables, and the implementation files contracted by omitting variables, the standards are better viewed as describing a "flexible dataset" concept that should be moulded to the needs of specific lines of research or decision support.

This document provides an overview of the most recent revision, which represents the ICASA Version 1.0 standards. It presents various examples as implemented in ASCII file format and briefly discusses progress in implementing the standards in databases and XML. A number of ASCII implementations of the new standards have been developed and will be used in the next official release of the DSSAT software (Jones et al., 2003; Hoogenboom et al., 2004). The standards are expected to evolve over time, mainly through addition of new variables. Updates to the definitions will be posted at the ICASA web site, www.icasa.net/standards/.

Standards

General

The basic organization of the data is defined for experiments that may include multiple sites and years as well as various crop and weed species, managements, and initial conditions at one site in one year. Various subsets of site, species, managements and initial conditions data are referenced in a central subset of data items identified as “Treatments.” Additional subsets of descriptive data and results are linked to “Treatments” through level indicators that are defined in the treatment subset. These linked subsets describe the genotypes, the fields, crop management and other features of the experiment (Fig. 1). Data from experiments are recorded in two separate subsets, one dealing with measurements or observations made at one or a few times during the course of the experiment, the other with those made at intervals throughout the experiment. Field data typically include crop developmental stages, yield and yield components, and growth analysis data such as leaf area index (LAI), stem, leaf, aboveground and grain biomass, but they can include measurements of soil water content, soil nutrient levels, pest damage or any other variables deemed relevant. Weather and general descriptions of soil profiles are managed separately since a single set of data may apply to multiple experiments.

Data items

The basic unit is a data item that contains one or more values, which may be numeric, identifiers, codes, or descriptive text, plus a name. The names are character strings with no distinction between upper and lower case. Information in a data item can be either:

- Variables - information pertinent to the experiment/situation documented.

  Examples:

`2070`	Value in kg ha^{-1} for stem dry weight on a given date.
`1982`	The year an experiment was planted.
`3.2`	The number of seed per pod from a soybean treatment.
`-5.1`	The minimum temperature in °C on a given day.
`BAT 477`	The name of a common bean cultivar.
`UFGA8701`	The name of a weather data set.
`Early planting`	The name of a specific planting date treatment.

- Level indicators - character strings or numbers that link with data reported elsewhere.

  Examples:

`7`	Treatment number 7 in an experiment.
`6`	Irrigation regime number 6 in an experiment.
`IB0488`	A cultivar identifier that links to a table of cultivars.
`KSAS0401WH`	An experiment identifier.

Variables can be numbers (decimal or integer), character strings, or text. Variables associated with a single name must be of the same data type. Units for numeric variables largely follow the International System of Units (SI), but “cm” and “ha” (hectare) are permitted in order to conform to dominant practices in agricultural research, e.g., by the American Society of Agronomy (Anon., 1998). Times of events such as planting, fertilization, or anthesis are recorded using year (four- or two-digit, depending on context) and day of year. Further examples of variables are presented in Table 1, and the full set of standard variables is listed at the ICASA website.

Codes are provided for non-numeric variables where some degree of standardization is convenient or required, such as for describing fertilizer types, irrigation methods, or planting methods. Examples are presented in Table 2, and the complete list of codes is provided at the ICASA website.

Data Sets and Subsets

Data items are organized into a hierarchy with two levels, i.e., sets, and subsets. Sets are the highest level of aggregation. They allow connected but not necessarily related data to be kept together. Three types are currently recognized (Table 3):

- Experiments. A description of treatments, initial conditions and field measurements for a single experiment, which could be multi-location or multi-season.
- Weather. Daily weather data from one or more recording stations.
- Soils. Collections of soil information, usually from a single geographic region or data source.

Data subsets are comprised of closely related data items stemming from one set of measurements or field operations, one soil profile characterization, or one weather station. For experiments, most subsets correspond to specific management activities, e.g. planting or irrigation, or to field measurements (Table 3).

Names and Identifiers - General

Experience from managing data from large numbers of experiments has demonstrated the need for datasets and subsets to be identified in a consistent manner. Furthermore, a compact name is valuable for manipulating data electronically, i.e., in spreadsheets or statistical packages. Thus, a consistent and compact naming system has been defined.

Nonetheless, it is recognized that some flexibility is necessary in naming datasets and subsets, to accommodate both user preferences and established local practice. When a user introduces an identifier that deviates from the standard convention, however, it is important that an effort is made to ensure that such introductions are consistent, especially in linking weather or soil data to experiment descriptions.

Names and Identifiers - Datasets

Datasets are identified by one of three names, EXPERIMENT, SOIL or WEATHER (Table 3), associated with a specific identifier constructed to provide information on the contents of the set. These are constructed differently for experiments, soils, and weather. For experiments, the set of identifiers are constructed by combining:

An institute or region code (two characters, e.g., "UF" for "University of Florida", "CA" for "Canada"),

A code for the site or set of sites (two characters, e.g., "GA" for "Gainesville"),

A year code (two characters representing the year in which the experiment was initiated, the year in which it was finally harvested, or another year of significance to the principal investigator or coordinating Institute),

An experiment number or code (two characters), and

A crop, multi-crop (for mixed cropping or crops with weed populations) or sequence (for rotation experiments) code (two characters).

Thus, the third experiment (03) conducted by the University of Florida (UF) at Gainesville (GA) in 2006 (06) with soybeans would yield a specific dataset identifier of UFGA0603SB.

For weather, dataset identifiers can be constructed from Institute and Site codes plus, if desired, four digits to indicate the starting year (e.g., UFGA2006). Optionally, a twelve-character code may be used, where the first two additional characters indicate the number of years of data and the last two characters can be used to identify other characteristics of the set. Thus, "UFGA196825R1" might indicate a 25-year series from Gainesville, Florida that started in 1968 and that used a method "R1" for estimating daily solar radiation.

For soil data, specific set identifiers can be constructed using a two- character code for the institute or region, plus a two-character code for the site or collection of sites. Alternatively, a longer name (but staying within an eight character limit) can be used to provide significant information on the contents of the set. Thus, "ARIZONA" could be used as a general name for a dataset containing soil profile descriptions from diverse sites in Arizona

Names and Identifiers - Subsets

For experiments, in which each set is restricted to data from one experiment, general names (Table 3) provide an unambiguous identification for each subset, e.g., TREATMENTS, PLANTING, but for weather and soil files, which may contain information from different weather stations or soil profiles, the single name may not guarantee unique identification. Thus, a specific subset identifier may need to be appended to the general name to provide a unique reference for the data items. For example, for weather subsets that contain only a part of an overall dataset, e.g., a single year or portion of a year, the subset should be identified with the general name plus some specific information, as shown below:

```
WEATHER_STATION:UFGA2004S1
```

In which the ending "S1" might indicate "Season 1".

For soil subsets, the specific identifiers are of ten-character length, with the Institute and Site codes in the leading four positions, the year the profile was described in the field in the next four positions, and a specific profile identifier in the remaining two positions, e.g., UFGA198501. A soil subset identifier thus may have the appearance shown below:

```
SOIL_PROFILE:UFGA200401
```

Names and Identifiers - Variables

Many variable names in DSSAT were limited to four or five characters to permit displaying a name as a label over a column of data that contained no more than five digits. Recognizing, however, that the expansion of the standards to other electronic formats reduces the need to limit the size of variable names, the ICASA standards now include two name formats, the long variable name and the abbreviated name (Table 1). The long name is generally 12 to 24 characters long and uses complete words as much as possible. Words are separated by the underscore ("_") character. The abbreviated name usually corresponds to the previous DSSAT name, although some variables have been re-named to correct inconsistencies. Abbreviated names for most data are no longer than five characters to permit their use as compact column headings. Allowing five spaces for numeric data permits displaying at least four significant figures, a level of precision greater than achieved in most agronomic, weather and soil measurements.

To facilitate interpretation of the abbreviated names, a major effort has been made to use a consistent naming strategy. Thus, for observed data relating to plant tissue masses or nutrient content, the first character indicates the tissue type, e.g., “L” for leaf, the second character describes the quantity being measured, e.g., “W” for dry weight, “N” for nitrogen, the third and sometimes the fourth character(s), the measurement reference, e.g., “A” for area, “PC” for percentage, and the final character indicates the time or frequency of measurement or observation, e.g., “D” for time series data referred to specific sampling or observation dates, “H” for data recorded at harvest.

ASCII Implementation

General

When the standards are implemented in ASCII, a single file generally corresponds to a dataset, occasionally to a subset. Data items are column-formatted and arranged in columns headed by the abbreviated variable name. These abbreviated names are presented above the variables or level indicators on a header line that begins with a '`@`' symbol. Multiple names can occur on one line. Text is usually specified by a string of dots after the header to indicate the full width of the field, e.g., "`TRT_NAME...............`" under Treatments in Fig. 1, and the text value must be presented below the header plus dots. For data items at the end of a row, however, text can be used without the dots that are necessary elsewhere. Blank lines can be inserted anywhere in the file.

Examples of ASCII file implementations of the standards are given in Figs. 2 to 6. Fig. 2 contains the documentation for a single experiment, and Figs. 3 and 4, the associated field measurements, separated as summary (Fig. 3) and time series (Fig. 4) data. Fig. 5 is for weather data, and Fig. 6, for soil profile descriptions.

File organization and naming

As outlined above, the usual file organization is for data from a given set or subset to reside within a single file. The file is named using a two-part convention. File prefixes equate to the specific identifiers for the set or subset, whilst the file extension distinguishes among the type of information. The prefix plus, if necessary, part of the extension provides a specific identifier. File extensions are constructed as follows:

- `ccX` Experimental details, where "cc" indicates a crop, multi-crop, or rotational sequence code. Examples are "WH" for wheat, "PN" for peanut, "MC" for multi-crop, and "SQ" for a rotational sequence or experiment,
- `SOL` Soil profiles,

WTH Weather.

For convenience, it is often desirable to place the summary and time-course subsets from an experiment in separate files. To allow this, two additional file types are used. Their respective extensions are:

ccA Experiment – summary results,
ccT Experiment – time-course results.

Further, large weather datasets may need to be subdivided for easy file handling. In such a case, the subset names indicate the time period covered, and these are used as the file names. Thus, a dataset with the name UFGA197520, having 20 years of daily data starting in 1975, could be broken into two ten-year subsets, with the files containing the subsets identified as UFGA197510.WTH and UFGA198510.WTH.

Some example file names are given below:

```
UCEA0501.SBX
UCEA0501.SQX
UCEA0501.MCX
UFGA.WTH
UFGA2005.WTH
UFGALCOR.WTH
UC.SOL
ARIZONA.SOL
```

File symbols

Within individual files, specific symbols are used to guide machine reading of data and to facilitate visual inspection of data. Datasets are identified with a dollar sign ($) in the first column of the first line. This symbol is followed by the general name and the specific identifier, as shown below:

```
$EXPERIMENT:UCEA9601SB
$WEATHER:UFGA2001
$SOIL:UC
$SOIL:ARIZONA
```

Data subsets are identified by placing an asterisk (*) in the first character position before the identifier, as shown below:

```
*TREATMENTS
*IRRIGATIONS
*WEATHER_STATION:UFGA2004S1
*SOIL_PROFILE:UFGA200401
```

Linking data across datasets and subsets

Numeric level identifiers and character strings are used to link information across datasets and subsets much as keys are used to link tables in a relational database. Links between different subsets of data require multiple identifiers in order to accommodate combinations of treatments, replicates, rotations and crop or weed associations as specified in the `TREATMENTS` subset (Table 4; Fig. 2). Treatment levels and the links are identified with integers, preferably numbered from 1 within each experiment. A value of 0 for the replicate implies that means across replicates are given.

The `TREATMENTS` subset of Fig. 2 illustrates cases of single replicates and means and of different plant species components (wheat and weeds). Note that while data for individual replicates are given for the wheat, the weed component is only specified at the level of means of the two replicates.

Special syntax and forming rules

The ASCII implementation could be used without restrictions on line length or number of characters allowed in identifiers, variable name abbreviations, or character variables. However, to simplify the development of associated software, and to facilitate viewing and manual checking of the files, the following maximum line and name lengths are used:

1.	Lines	254 characters
2.	Variable name abbreviation	31 characters
3.	Variables	31 characters (but 80 for text)
4.	Level indicators	31 characters
5.	Dataset and Subset Identifiers	31 characters

Specific rules for comments, missing data, non-applicable data, data flags, sub-samples, and end-of-file markers are summarized in Table 5. Missing numeric data are identified by –99, and missing character strings or text, by the string "-99". In some cases, specific codes can be used to indicate that an application method or fertilizer type is unknown.

Dates for growth stages are presented using two digits for the year and three digits for the day ("yr-day of year" format such as "81253" for day 253 of 1981). This allows a compact presentation, and ambiguities that could arise with the two-digit year format are easily resolved by reference to the planting date as given in the `PLANTING` subset.

File Additions and Modifications

To incorporate new data items, additional abbreviations are defined, and the corresponding columns of data items added within an existing or new subset. Adding additional data items at the end of existing rows of data items is discouraged. Two exceptions are when the variable is needed as a link to data elsewhere in the file, e.g., additional factors in the treatment subset, or when adding extra rows of data items would disrupt the overall configuration of the file, e.g., for a new daily weather variable.

File Validation

A valid file has no headers repeated within a subset, and it uses unique subset names within the dataset. For experiments, all levels defined in the TREATMENTS subset must link to levels in other subsets. In addition, for a crop grown in the field, a file should contain subsets dealing with GENOTYPES, FIELDS, PLANTING, and INITIAL_CONDITIONS, as a minimum. For a fallow, subsets dealing with FIELDS, INITIAL_CONDITIONS and HARVESTS are required. In the current implementation, the HARVESTS subset is used to specify the end of the fallow.

Relational Database and XML Implementations

A major change in revising the ICASA standards was the conclusion that they must work with diverse formats for electronically storing and manipulating the data. The foremost requirement is that the dictionary of variable names, identifiers and codes, which includes definitions and units of measurement, is comprehensive enough to embrace experiments conducted with many different objectives and is completely enough defined to allow use in different implementations. Thus, in revising the standards, a major effort was made to develop a consistent master list of variables that includes not only variable names, definitions, and units of measurement but information on the data type and validation criteria. The second requirement is to define the relations among variables so that treatment structures, sampling regimes (e.g., dates or positions within a soil profile) are correctly represented.

Implementing the standards for flat ASCII files required the definition of a structure that both facilitated visual inspection of data and ensured that software could read files in a consistent manner. Implementation in a relational database schema requires the definition of tables and fields within tables, and the establishment of links between tables. For this purpose, the sets and subsets correspond to major tables, and individual data items are represented by fields within tables. In the case of subsets containing data with a two-dimensional structure, such as time series of irrigation applications or soil profile descriptions that vary with depth, two tables may be required per subset. The first provides general descriptors or metadata, e.g., name and general characteristics of an irrigation treatment, and the second describes the activities for each treatment on individual dates, e.g., irrigation methods and amounts, or the soil measurements at different depths. Lookup tables are also required for the various codes. Relations among the tables are defined through database keys. Since relationships serve the same function as the various level indicators used in the ASCII implementation, links among keys mainly correspond to the links among level indicators.

Work is underway to adapt the XBuild software tool of DSSAT4 (Uryasev et al., 2004) to the ICASA standard. XBuild currently creates a temporary relational database for a single experiment using the DSSAT4 formats. The new tool, AgroBuild, extends this approach to the ICASA standards and allows managing multiple experiments. The initial emphasis is on managing data for experiment descriptions (FileX) and measured/observed data (FileA and FileT), but the proposed relational structure includes daily weather and soil profile data. Figure 1 presents a simplified representation of the tables and relations.

XML is widely used as a data interchange format (XML Core Working Group, 2004). A key feature is that by conforming to basic syntax rules and providing information to document the structure of a given data file, any XML-compatible software can read the data, perform basic validation checks, and correctly interpret the data for subsequent processing (XML Core Working Group, 2004). Typically, discipline groups develop an XML standard for the subject matter of interest. This work basically consists of defining the file structure through a document type definition (DTD) or an XML schema (XML Core Working Group, 2004). Thus, implementing the ICASA standards in XML requires re-expressing the standards through a DTD or XML schema. Efforts are underway to determine what XML standards have been proposed by other groups involved in agricultural research or industry, and to develop and test an XML prototype based on the ICASA master list of variables.

Standards Documentation and Revision

Complete definitions of variable names, variable codes, dataset and subset names, and relations are maintained as a set of tables accessible at the ICASA web site (www.icasa.net/standards). The standards allow for new datasets, subsets, and variables to be defined according to the needs of specific users, but the utility of the standards depends upon use of a uniform vocabulary. Users are encouraged to report new variable names and codes, as well as dataset and subset names, together with concerns over inconsistencies in the overall standards.

A recent example of the revision process is from a project by the University of Florida to develop a data management system for the Florida Department of Agriculture and Consumer Services (FDAC). In reviewing data to be reported by researchers, it was found that mulching is an important management practice, e.g., plastic mulches are widely used for vegetable production. Thus, a new subset was defined as MULCH_MANAGEMENT and linked to a new treatment category. Since mulches vary in type, thickness, portion of ground cover provided, and other characteristics, new variables were defined for use within the mulching subset. New codes were also defined for mulch types and application methods. We anticipate, however, that these variable list and codes will be modified as actual field data are processed.

Suggestions for additions or modifications will be reviewed periodically and updates to the standards provided as necessary. Backward compatibility will be maintained by retaining synonyms and presenting these in the master list of standard variable names, which is downloadable from the ICASA site. Comments or suggestions on the ICASA standards may be submitted through the ICASA Forum (www.ICASAForum.net) or to the corresponding author of this document (J.W.White: j.white@uswcl.ars.ag.gov).

References

Anonymous. 1998. Publications handbook and style manual. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, WI.

Blanchard, J.L. 2004. Bioinformatics and Systems Biology, rapidly evolving tools for interpreting plant response to global change. Field Crops Res. 90:117-131.

Bostick, W.M., J. Koo, V.K. Walen, J.W. Jones, and G. Hoogenboom. 2004. A web-based data exchange system for crop model applications. Agron. J. 96:853-856.

FAO. 2004. Agricultural Ontology Service Project (AOS) [Online]. Available at http://www.fao.org/aims/aos.jsp (verified 18 July 2006).

GCTE. 2006. Global Change and Terrestrial Ecosystems Report No. 2, Second Edition. GCTE Focus 3 Wheat Network: 1996 Model and Experimental Metadata. GCTE Focus 3 Office, Wallingford, UK, 1996.

Goudriaan, J. 1996. Predicting crop yields under global change. p. 260-274. *In* B.H. Walker and B.H. Steffen (ed.). Global change and terrestrial ecosystems. International Geosphere-Biosphere Programme Book Series #2, Cambridge University Press, Cambridge, UK.

Hoogenboom, G., J.W. Jones, P.W. Wilkens, C.H. Porter, W.D. Batchelor, L.A. Hunt, K.J. Boote, U. Singh, O. Uryasev, W.T. Bowen, A. Gijsman, A. du Toit, J.W. White, and G.Y. Tsuji. 2004. Decision Support System for Agrotechnology Transfer Version 4.0. Ver. 4.0 [CD-ROM publication]. University of Hawaii, Honolulu, HI.

Hunt, L.A., J.W. Jones, G. Hoogenboom, D.C. Godwin, U. Singh, N.B. Pickering, P.K. Thornton, K.J. Boote, and J.T. Ritchie. 1994. Input and output file structures for crop simulation models. p. 35-72. *In* P.F. Uhlir and G.C. Carter (ed.) Crop modelling and related environmental data. A focus on applications for arid and semiarid regions in developing countries. CODATA, International Council of Scientific Unions, Paris.

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.

Jamieson, P.D., J.R. Porter, J. Goudriaan, J.T. Ritchie, H. van Keulen and W. Stol. 1998. A comparison of the models AFRCWHEAT2, CERES-Wheat, Sirius, SUCROS2 and SWHEAT with measurements from wheat grown under drought. Field Crops Res. 55:23-44.

Jones, J.W., L.A. Hunt, G. Hoogenboom, D.C. Godwin, U. Singh, G.Y. Tsuji, N.B. Pickering, P.K. Thornton, W.T. Bowen, K.J. Boote and J.T. Ritchie. 1994. Input and output files. p. 1-94. *In* G.Y. Tsuji, G. Uehara, and S. Balas (ed.). DSSAT v3.Vol 2-1. University of Hawaii, Honolulu, HI.

Jones, J.W., G. Hoogenboom, C.H. Porter, K.J. Boote, W.D. Batchelor, L.A. Hunt, P.W. Wilkens, U. Singh, A.J. Gijsman and J.T. Ritchie. 2003. The DSSAT cropping system model. Eur. J. Agron. 18:235-265.

Nix, H.A. 1984. Minimum data sets for agrotechnology transfer. Proceedings of the International Symposium on Minimum Data Sets for Agrotechnology Transfer, 21-26 March 1983, ICRISAT Center, India. ICRISAT, Patancheru, India.

The Plant Ontology Consortium. 2002. Conference report: The Plant Ontology Consortium and plant ontologies. Comp. Funct. Genomics 3:137-142.

Tsuji, G.Y., G. Uehara and S. Balas. 1994. DSSAT version 3 Volumes 1, 2, and 3. International Benchmark Sites Network for Agrotechnology Transfer, University of Hawaii, Honolulu, HI.

Uehara, G., and G.Y. Tsuji. 1998. Overview of IBSNAT. p. 1-7. *In* G.Y. Tsuji, G. Hoogenboom and P.K. Thornton (ed.). Understanding Options for Agricultural Production, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Uryasev, O., J.W. Jones, G. Hoogenboom, and C. Porter. 2004. XBuild: Creating crop management files for documenting experiments and simulating crop growths and yield for DSSAT v4. *In* Decision Support System for Agrotechnology Transfer Version 4.0. Vol. 2. DSSAT v4: Data Management and Analysis Tools, Vol. 2, University of Hawaii. Honolulu, HI.

van Evert, F.K., E.J.A. Spaans, S.D. Krieger, J.V. Carlis, J.M. Baker. 1999a. A database for agroecological research data. 1: Data model. Agron. J. 91:54–62.

van Evert, F.K., E.J.A. Spaans, S.D. Krieger, J.V. Carlis, J.M. Baker. 1999b. A database for agroecological research data. II. A relational implementation. Agron. J. 91:62–71.

White, J.W., S.S. Dhillon, L.A. Hunt and P.D. Jamieson. 2003. A Comparison of Approaches to Modeling Phenology as Applied to Genotype by sowing Date Interactions in Wheat. p. 35-41. *In* J.W. White (ed.). 2003. Modeling temperature response in wheat and maize: proceedings of a workshop, CIMMYT, El Batán, Mexico, 23-25 April 2001. NRG-GIS Series 03-01. CIMMYT, México, D.F.

XML Core Working Group, 2004. Extensible Markup Language (XML) [Online]. Available at http://www.w3.org/XML/ (verified 18 July 2006).

Table 1. Examples of variables used to describe inputs, crops, management practices, environment, and other aspects of an experiment as specified in the master list of variables for the ICASA standards. Date types (i.e., "doy" and "yeardoy") apply only to ASCII implementation.

Long Name	Abbreviated Name	Description	Unit or Type
cultivar_name	CUL_NAME	Cultivar name	text
field_elevation	FLELE	Elevation of field site	m
field_name	FLNAME	Field name	text
fertilizer_level	FE	Fertilizer level in treatment structure	number
fertilizer_level_name	FE_NAME	Fertilizer level name	text
fertilizer_applic_depth	FEDEP	Fertilizer application/mixing depth	cm
fertilizer_applic_code	FEACD	Fertilizer application, code for method	code
fertilizer_applic_day	FEDAY	Fertilizer application date, doy or dap	day
fertilizer_applic_year	FEYR	Fertilizer application year	year
irrigation_level	IR	Irrigation level in treatment structure	number
bund_height	ABUND	Bund height (e.g., for flooded rice)	mm
irrigation_applic_depth	IRADP	Irrigation application depth	cm
irrigation_day	IRDAY	Irrigation application date, doy or dap	day
irrigation_operation	IROP	Irrigation operation	code
irrigation_year	IRYR	Irrigation or water management year	year
planting_doy	PLDOY	Planting day of year	doy[1]
plant_pop_at_planting	PLPOP	Plant population at planting	number/m^2
planting_material_age	PLAGE	Planting material age (also dormancy)	day
planting_material	PLMA	Planting material	code
anthesis_days_after_pl	ADAP	Time to anthesis as days after planting	dap
anthesis_date	ADAT	Anthesis date	yrdoy[2]
anthesis_day_of_year	ADOY	Anthesis day of year	doy
zadoks_21_growth_stage	Z21D	Zadoks 21 growth stage date	yrdoy
grain_N_area_maturity	GNAM	Grain N at maturity	kg/ha
harvest_index_maturity	HIAM	Harvest index (tops only) at maturity	kg/kg
grain_dry_wt_area_maturity	GWAM	Grain dry weight at maturity	kg/ha
leaf_area_index_maximum	LAIX	Leaf area index, maximum	m^2/m^2
leaf_number_per_stem	LNOSD	Leaf number per stem on a day	number
leaf_area_index	LAID	Leaf area index on a day	m^2/m^2
tops_dry_wt_area	CWAD	Tops dry weight on a day	kg/ha
grain_number_area	GNOAD	Grain number on a day	number/m^2
grain_dry_wt_area	GWAD	Grain dry weight on a day	kg/ha
grain_dry_wt_per_grain	GWGD	Grain unit dry weight on a day	mg
harvest_index	HIAD	Harvest index on a day	kg/kg
panicle_dry_wt_area	PNWAD	Panicle dry weight (grain+structural) on a day	kg/ha
pod_dry_wt_area	PWAD	Pod dry weight on a day	kg/ha
stem_dry_wt_area	SWAD	Stem dry weight on a day	kg/ha

[1] doy = day of year (e.g., 196).
[2] yrdoy = two digits of year + day of year (e.g. 81245). Used only in ASCII implementations.

tiller_number_area	TNOAD	Tiller number (area basis) on a day	number/m^2
tuber_dry_wt_area	UWAD	Tuber dry weight on a day	kg/ha
grain_N_area	GNAD	Grain N on area basis on a day	kg/ha
grain_N_percent	GNPCD	Grain N concentration as percentage on a day	%
plant_P_area	PLPAD	P content on a given day	kg/ha
soil_NO3_content_layer	NO3WD	Soil layer NO3 N content (dry basis) on a day	ug/g
soil_NO3_concentration	NO3CD	Soil solution NO3 N conc at a depth on a day	ug/ml
soil_CO2_emission	SCO2D	Soil CO2 (as C) emission, daily	g/m^2 d
temperature_ maximum	TMAX	Temperature of air, daily maximum	^{o}C
temperature_ minimum	TMIN	Temperature of air, daily minimum	^{o}C
rain_snow_fall	RAIN	Daily total precipitation, including snow	mm
wind_speed_daily	WIND	Wind speed (run), daily	km
ambient_CO2_conc	ACO2	Ambient CO_2 concentration, daily average	vpm
soil_runoff_curve_NRCS	SLRO	Runoff curve, Nat. Resources Conservation Service	number
soil_organic_C_percent	SLOC	Organic carbon as 100 x g/g dry soil	%
soil_pH_in_water	SLPHW	pH of soil in water, from layer in profile	number
soil_root_growth_factor	SRGF	Root growth factor, soil only (0 to 1)	number
soil_bulk_dens_moist	SLBDM	Soil bulk density when moist	g/cm3
soil_water_lower_limit	SLLL	Soil water, lower limit for extraction	cm^3/cm^3

Table 2. Examples of codes (variable codes) used to represent specific inputs, crops, management practices, implements or other aspect necessary to characterize an experiment.

Category	Code	Description
Chemicals	CH001	Alachlor (Lasso), Metolachlor (Dual) [Herbicide]
	CH022	Malathion, Mercaptothion [Insecticide]
	CH051	Captan [Fungicide]
Crop	AL	Alfalfa/Lucerne
	MZ	Maize
	SC	Sugarcane
Pest organism	CEW	Corn earworm (*Heliothis zea*)
	VBC	Velvetbean caterpillar (*Anticarsia gemmatalis*)
	SBL	Soybean looper (*Pseudoplusia includens*)
Application methods	AP001	Broadcast, not incorporated
	AP004	Banded beneath surface
	AP006	Foliar spray
Drainage	DR000	No drainage
	DR001	Ditches
	DR002	Sub-surface tiles
Fertilizers	FE001	Ammonium nitrate
	FE005	Urea
	FE006	Diammonium phosphate
	FE021	Rock phosphate
	FE022	Calcitic limestone
Irrigation methods	IR001	Furrow, mm
	IR003	Flood, mm
	IR005	Drip or trickle, mm
Organic materials	OM001	Crop residue
	OM002	Green manure
	OM005	Compost
Planting material	S	Dry seed
	T	Transplants
	R	Ratoon
Plant distribution	R	Rows
	H	Hills
	U	Uniform/Broadcast
	RB	Rows on beds
Soil P analysis methods	SA001	Olsen
	SA003	Bray No. 2
	SA004	Mehlich I (double acid, 1:5)
Tillage implements	TI002	Tandem disk
	TI005	Moldboard plow
	TI006	Chisel plow

Table 3. General names and descriptions of datasets and subsets.

General Name	**Description**
<u>Data Sets</u>	
EXPERIMENT	Complete description of management and initial conditions for a real or synthetic experiment (or very closely linked set of experiments). Data measured during or at the end of the experiment. The information presented should be sufficient to allow both for an intellectual interpretation of the results and for simulation of the experiment with a computer model.
SOIL	Soil profile data for pedons from one or more sites in a region.
WEATHER	Daily weather data from one or more stations.
Data Subsets	
CHEMICALS	Herbicide, pesticide, or growth regulator application details.
ENV_MODIFICATIONS	Adjustments to weather variable such as those that could be made during growth cabinet, CO_2 enrichment, or rainout shelter studies.
FERTILIZERS(INORGANIC)	Fertilizer amounts, types, and dates of applications.
FIELDS	Field description including links to weather station and soil profile information.
GENERAL	Names and addresses of people involved, location(s) and general background details.
GENOTYPES	Crop and cultivar identifiers, names, and background.
HARVESTS	Harvest dates, components harvested, and percentages removed.
INITIAL_CONDITIONS	Starting conditions including above- and below-ground residues, and water and nitrogen in the complete profile.
IRRIGATION	Irrigation dates and amounts, flood and water table depths, thresholds for automatic applications.

ORGANIC_MATERIALS	Details of straw, manure and other organic material applications.
PLANTING	Planting date, seed and initial plant populations, seeding depth, and plant distribution (row spacing, etc.).
SOIL_ANALYSES	Details of classical soil surface layer analyses.
SOIL_PROFILE	Surface and layer characteristics for a specific soil.
SUMMARY	Summary data (may include data from means over replicates or from individual plots) that are collected during the course of or at the end of an experiment.
TIME_COURSE	Time-course data (may include data from means over replicates or from individual plots) collected during the course of an experiment.
TREATMENTS	Treatment names and level codes for rotation and crop components and experimental factors.
TILLAGE	Dates and types of tillage operations.
WEATHER_STATION	Daily weather data typically for a single experiment or season, but preferably for a complete year or longer.

Table 4. Indicators required to link data in different datasets and subsets.

A. Indicators required to link subsets in different datasets		
Set 1	Set 2	Links Required
Experiments	Weather	Weather set and subset identifiers
Experiments	Soil	Soil set and subset (soil profile) identifiers
B. Indicators required to link subsets in the same (Experiment) dataset		
Main Subset	Subsidiary Subset	Links Required
Treatments	Genotypes	Genotype level no.
Treatments	Fields	Field level no.
Treatments	Soil analysis	Soil analysis level no.
Treatments	Initial conditions	Initial conditions level no.
Treatments	Plantings	Planting level no.
Treatments	Irrigations	Irrigation level no.
Treatments	Fertilizers	Fertilizers level no.
Treatments	Organic materials	Organic material level no.
Treatments	Chemicals	Chemicals level no.
Treatments	Tillage	Tillage level no.
Treatments	Env_Modifications	Env_Modifications level no.
Treatments	Harvests	Harvests level no.
Treatments	Summary	Treatment no., Replicate no., Sequence no., Option no., Component no.
Treatments	Time-course	Treatment no., Replicate no., Sequence no., Option no., Component no.

Table 5. Special syntax and formatting rules for the ICASA standards.

Item	Rule
Comments	Information on data quality, problems with treatments, or aspects of an experiment that are difficult to quantify should normally be recorded under ‘Notes’ in the ‘General’ subset of the file or using the note fields in the corresponding subset. However, comments can also be inserted immediately after a dataset or subset identifier, or after a line of data using ‘!’ as the first character in the line. (Used only in the ASCII implementation.)
Missing data	Indicated by a value of ‘-99’.
Non-applicable data	Indicated by a value of ‘-99’. For example, row width and spacing for a crop that is broadcast sown.
Data ‘flags’	For specific data values that require annotation (e.g., due to quality concerns or to indicate an estimated value), data flags may be defined for summary, time series and weather data. The allowed values of the flag and their respective definition are defined in the General section of the dataset, and the actual flag values are reported in the respective file under the variable FLAGA, FLAGT or FLAGW. (See examples in Figs 3, 5 and 6).
Sub-samples	Data obtained from sub-samples can be entered under the optional header 'SS' in any subset for any of the datasets. These subset data items can be related to individual replicates or to overall averages (level 0 in the experiment datasets).
End-of-file	The symbol ‘=’ can be entered as the only character on the last line of a file to indicate the end of a file. Its use is recommended to indicate whether a file has unintentionally been truncated. (Used only in the ASCII implementation.)

Figure 1. Overview of the ICASA standards representing relations among subsets in one experiment dataset. Most boxes represent subsets, which correspond to tables in a relational database or to XML elements. Note that several subsets contain data items that involve time series or multiple soil depths, so a complete representation of the relations may require additional database tables or hierarchies of XML elements. Weather and soil profile data are stored in separate datasets.

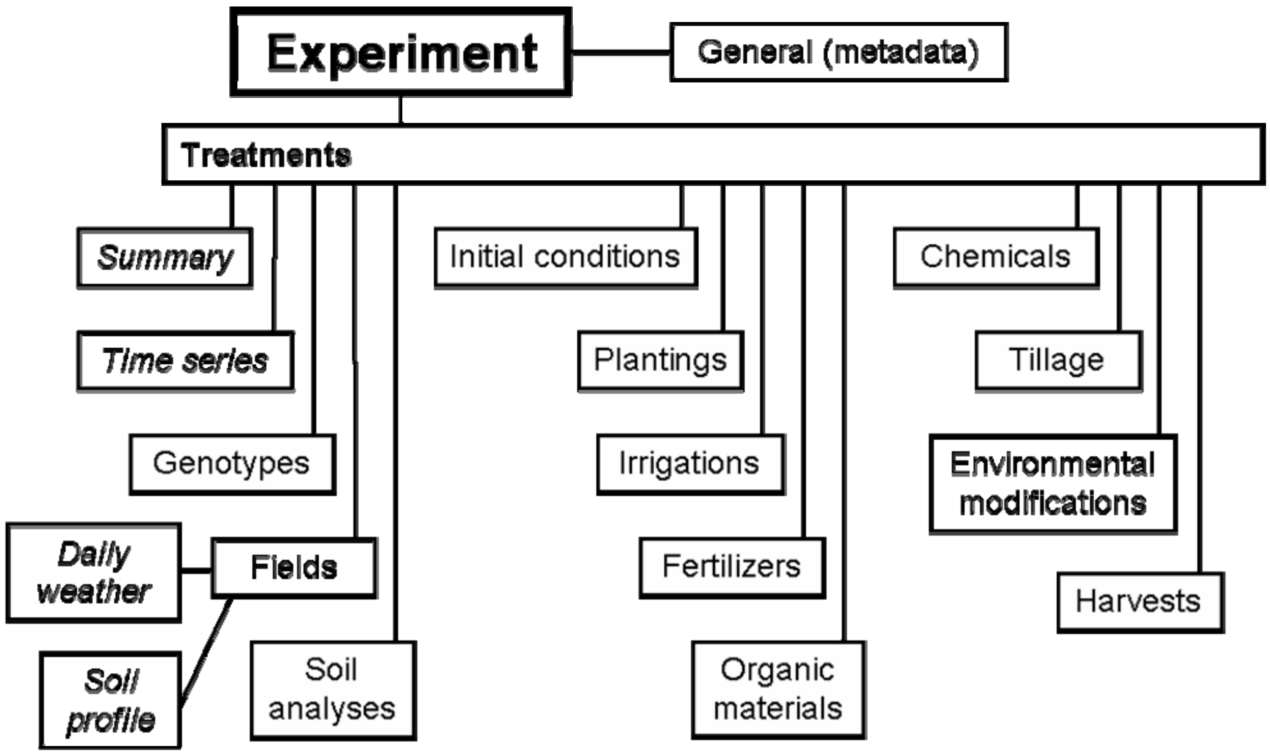

Figure 2. Example of an ICASA experimental details file. Note that lines that start with "!" are for comments.

```
$EXPERIMENT:KSAS0401WH

!Some general information is desirable in all files.

*GENERAL
@ NAME
  Fictitious experiment based on KSAS8101
@ LOCAL_NAME MAIN_FACTOR  FACTORS
  Godwin     FE(N)        4GE,4PL(Date),3IR,3FE(N)
@ PEOPLE
  Wagger,M.G. Kissel,D.
@ INSTITUTES
  Kansas State Univ.
@ CONTACTS
  Hunt,L.A. thunt@uoguelph.ca
@ NOTES
  Here should note any constraints and complications.
  Constraints could relate to: Soil physical conditions
  (surface compaction, duripan); Moisture deficit; Moisture
  excess; Nitrogen; Phosphorous; Potassium; Salinity; pH;
  Aluminum/Manganese; Boron; Other micro-nutrients; Above-ground
  diseases; Below-ground diseases; Nematodes; Aphids; Insects;
  Weeds; Other
  Complications could deal with: Stand loss (e.g.,winter kill);
  Wind damage (->leaf shredding); Sandblasting; Hail; Lodging;
  Bird feeding; Shattering; Sprouting; Combine settings;
  Seed mixing; Animal grazing; Other.
@ PUBLICATIONS
  Wagger,M.G., ????  N cycling in the plant-soil system. Ph.D.,Kansas State
@ DISTRIBUTION
  Use at will but acknowledge source. No secondary distribution
@ VERSION
  ICASA1.0 10-08-2006(GH,JW,LAH;email)

!With additional background information being added if thought
!necessary for comprehensive documentation.
@ OBJECTIVES
  To determine the degree of interaction between N and water supply
@ DESIGN
  Assumed to be a randomized complete block.
@ LAYOUT
  Not known.
@ PLTA PLTR# PLTLN PLTOR PLTOD PLTSP PLTHA PLTH# PLTHL PLTHM
  11.5 16.00 4.000 90.00 90.00 18.00  0.72     4   1.0 Hand
!If plot organization cannot be documented as above, describe in 'Notes'
@ METHODS
  Standard field-plot technique using cone seeder.
  Standard breeding observations for phenology. Hege combine for yield.
  MDAT2 obtained using dye absorption technique (MDAT loss of green)
  Small sub-plots for TWAD.
  Gravimetric procedures for SW1D
@ FLAG FLAG_DETAILS
     0 Data ok
     1 Processing delay possibly spoiled sample.
     2 Sampled spilled. Some grain possibly lost.
```

```
!And the following specific information is necessary in most cases. Data
!subsets can be omitted, however, if the particular aspect (e.g., irrigation)
!is not a factor in the experiment.

*TREATMENTS                                -----------FACTOR LEVELS-----------
@TRTNO RP SQ OP CO TRT_NAME............... GE FL SA IC PL IR FE OM CH TI EM HA
     1  1  1  1  1 0N,dryland+weeds Rep1    1  1  1  1  1  0  0  1  1  1  1  1
     1  2  1  1  1 0N,dryland+weeds Rep2    1  1  1  1  1  0  0  2  1  1  1  1
     1  0  1  1  2 0N,dryland+weeds         2  1  1  1  2  0  2  1  1  1  1  1
     2  0  1  1  1 180N,dryland             1  1  1  1  1  0  3  1  1  1  1  1
     3  0  1  1  1 0N,irrigated             1  1  1  1  1  1  0  1  1  1  1  1
     4  0  1  1  1 60N,irrigated            1  1  1  1  1  1  2  1  1  1  1  1
     5  0  1  1  1 180N,irrigated           1  1  1  1  1  1  3  1  1  1  1  1
     6  0  1  1  1 180N,irrig(dap)          1  1  1  1  1  2  3  1  1  1  1  1
     7  0  1  1  1 180N,irrig(automatic)    1  1  1  1  1  3  3  1  1  1  1  1
     7  0  2  1  1 RedClover Post-harvest   3  1  0  0  3  0  0  0  0  0  0  0
     7  0  2  2  1 RedClover2 Postharvest   4  1  0  0  4  0  0  0  0  0  0  0

*GENOTYPES
@ GE   CR CUL_ID   CUL_NAME
   1   WH IB0488   Newton  Hard red winter,Central Plains
   2   GW IB0001   Garytown  Central Gt.Plains quackgrass
   3   RC IB0001   Bigyield  General double cut
   4   RC IB0002   Earlywonder  Single cut

*FIELDS
@ FL    FL_COUNTRY            FL_REGION         FL_LOCATION
   1    USA                   Kansas            Research Park, Ashland
@ FL    FL_LAT     FL_LONG    FLELE  FLSL FLSLL FLSLA FAREA FLLWR
   1      37.1167  -90.4521     -99    20   200    90   226   -99
@ FL    FL_PLOWD   FL_PLOWDC  FL_DRNTYPE FLDRD FLDRS
   1          25         0.4  FLD00          0     0
@ FL    WTH_SS     SOIL_SS    WTH_DS   SOIL_DS
   1    KSAS2004   KSAS200401 KSAS     KS
! The weather station identifier can also be presented as shown below:
! FL    WTH_SS     SOIL_SS    WTH_DS   SOIL_DS
!  1    KSAS0401   KSAS200401 KSAS     KS

*SOIL_ANALYSES
@ SA    SA_NAME
   1    Science Associates Lab
@ SA    SAYR SADAY SAMHB SAMPX SAMKE
   1    2004   260   -99   -99   -99
@ SA    SABL SABDM  SAOC  SANI SAPHW SAPHB  SAPX  SAKE
   1      15   -99   -99   -99  5.85  5.21   -99   -99

*INITIAL_CONDITIONS
@ IC    IC_NAME
   1    10 days before planting
@ IC   ICRYR ICRDY ICPCR ICRDP ICRIP ICRAG  ICRN  ICRP  ICRK ICRLI  ICRT  ICND
   1    2004   260    WH    10    50  2000  1.00  1.00   -99  10.0   100   -99
@ IC    ICYR ICDAY  ICSW  ICWT  ICIN ICRZC ICRZE
   1    2004   260   300   100   100   -99   -99
@ IC    ICBL ICH2O ICNH4 ICNO3
   1       5 0.205   3.4   9.8
   1      15 0.205   3.4   9.8
   1      30 0.170   3.2   7.3
   1      60 0.092   2.5   5.1
   1      90 0.065   2.2   4.7
```

```
   1     120 0.066   2.7   4.3
   1     150 0.066   2.7   4.3
   1     180  0.66   2.7   4.3

*PLANTINGS
@ PL    PL_NAME
   1    Early planting
   2    Late planting
   3    Planting 25 days after harvesting the previous crop
   4    Planting 45 days after harvesting the previous crop
@ PL    PLYR PLDOY PLDOE PLPOP PLPOE  PLMA  PLDS  PLRS  PLRD  PLPH  PLDP PLMWT
   1    2004   270   -99   162   162     S     R    17    90   1.0   5.5   120
   2    2004   289   -99    40    40     S     R    17    90   1.0   3.5    20
   3     -99    25   -99   400   400     S     B   -99   -99   1.0   0.0     5
   4     -99    45   -99   400   400     S     B   -99   -99   1.0   0.0    10
@ PL   PLAGE PLENV PLSPL PLGPC PLMSOURCE
   1       0   -99   -99    97 Univ. breeder
   2       0   -99   -99    90 Univ. weed scientist
   3       0   -99   -99    95 Univ. breeder
   4       0   -99   -99    95 Univ. breeder

*IRRIGATIONS
@ IR    IR_NAME
   1    Irrigation on specified dates
   2    Irrigation on specified days after planting
   3    Automatic irrigation
@ IR   IOEFF IOADP IONPC
   1     1.0   -99  0.03
   2     1.0   -99   -99
   3     1.0   -99   -99
@ IR    IRYR IRDAY  IROP IRVAL IRADP IRNPC
   1    2005    96 IR001    65     0   -99
   1    2005   110 IR001    78    20   -99
   1    2005   117 IR001    70    20   -99
   2     -99   170 IR001    65    20  0.03
   2     -99   200 IR001    78    20  0.05
   2     -99   210 IR001    70    20  0.06
@ IR   IIRYR IIDAY IISTG IIROP IIMDP IITHR IIEPT IIVAL IINPC
   3    2005   200   -99 IRA01    60    60    85   -99  0.01
   3    2005   250   -99 IRA00    60    60    85   -99  0.05

*FERTILIZERS(INORGANIC)
@ FE    FE_NAME
   2    Single application at planting
   3    Applications at planting and in spring
@ FE    FEYR FEDAY  FECD FEACD FEDEP FEAMN FEAMP FEAMK FEAMC FEAMO FEOCD
   2    2004   289 FE001 AP001    15    60   -99   -99   -99   -99   -99
   3    2004   289 FE001 AP001    15    90   -99   -99   -99   -99   -99
   3    2005    56 FE001 AP001     1    90   -99   -99   -99   -99   -99

*ORGANIC_MATERIALS
@ OM    OM_NAME
   1    Low lignin
   2    High lignin
@ OM    OMYR OMDAY  OMCD OMACD OMDEP OMINP OMAMT
   1    2004   280 OM000 AP002    10    80  1000
   2    2004   280 OM000 AP002    10    80  1500
@ OM    OMYR OMDAY OMH20 OMCPC OMNPC OMPPC OMKPC OMLPC
   1    2004   280   0.1   0.1   0.1   -99   -99    10
   2    2004   280   0.1   0.1   0.1   -99   -99    20
```

```
*CHEMICALS
@ CH    CH_NAME
   1    Pre-plant weed control
@ CH    CH_NOTES
   1    Poor control of leafy spurge due to dry conditions
@ CH    CHYR CHDAY  CHCD CHACD CHDEP CHAMT CH_TARGETS
   1    2004   260 CH001 AP001   1.3   2.2 Broad-leaf weeds

*TILLAGE
@ TI    TI_NAME
   1    Standard preparation for wheat
@ TI    TIYR TIDAY TIIMP TIDEP TIMIX TI_NOTES
   1    2004   250 TI005  15.0    30 Lea type plough with long moldboard

*ENV_MODIFICATIONS
@ EM    EM_NAME
   1    CO2 supplementation early
@ EM    EMYR EMDAY ECDYL EMDYL ECRAD EMRAD ECMAX EMMAX ECMIN EMMIN
   1    2005   180     0     0     0     0     0     0     0     0
   1    2005   220     0     0     0     0     0     0     0     0
@ EM    EMYR EMDAY ECRAI EMRAI ECCO2 EMCO2 ECDEW EMDEW ECWND EMWND
   1    2005   180     0     0     R   600     0     0     0     0
   1    2005   220     0     0     R   360     0     0     0     0

*HARVESTS
@ HA    HA_NAME
   1    Harvest with no screening
@ HA    HAYR HADAY HASTG HACOM HASIZ  HAPC HABPC
   1    2005   273 GS090 HAC0P HAS0A   100    50
=
```

Figure 3. Example of an ICASA summary data file.

```
$EXPERIMENT:KSAS0401WH

*SUMMARY
!The following are presented to indicates the general layout.
!Actual data items will vary from experiment to experiment.In
!most cases, some 'keys' (e.g.SQ,OP,CO) will not be necessary.
!Mean values (RP=0) can be placed together at the top or bottom
!of a data subset, or as individual rows after the corresponding
!replicate data.
@TRTNO RP SQ OP CO ADAT  MDAT  GWAM  CWAM  FLAGA
     1  0  1  1  1 05162 05221  5700 12300     0
     1  0  1  1  2 05162   -99   -99  1020     0
     2  0  1  1  1 05158 05218  5500 12800     0
     1  1  1  1  1 05162 05221  5700 10100     0
     1  2  1  1  1 05162 05221  5700  9900     0
     1  3  1  1  1 05162 05221  5700 10300     0
     1  1  1  1  2 05162   -99   -99  1020     0
     1  2  1  1  2 05162   -99   -99   810     0
     1  3  1  1  2 05162   -99   -99   980     0
     2  1  1  1  1 05158 05218  5500 13200     2
     2  2  1  1  1 05158 05218  5500 13560     0
     2  3  1  1  1 05188 06002  5500 12800     0
=
```

Figure 4. Example of an ICASA time-course data file.

```
$EXPERIMENT:KSAS0401WH

*TIME_COURSE
!The following are presented to indicates the general layout.
!Actual data items will vary from experiment to experiment.In
!most cases, some 'keys' (e.g.SQ,OP,CO) will not be necessary.
!Mean values (RP=0) can be placed together at the top or bottom
!of a data subset, or as individual rows after the corresponding
!replicate data.
@TRTNO RP SQ OP CO  YEAR   DOY  TWAD FLAGT
     1  0  1  1  1  2005   063  1280     0
     1  0  1  1  2  2005   063   980     0
     1  0  1  1  1  2005   092  4440     0
     1  0  1  1  2  2005   092  1440     0
     1  0  1  1  1  2005   120 10080     1
     1  0  1  1  2  2005   120  4080     0
     2  0  1  1  1  2005   063  1280     0
     2  0  1  1  1  2005   092  4440     0
     2  0  1  1  1  2005   120 10080     0
     1  1  1  1  1  2005   063  1010     0
     1  2  1  1  1  2005   063  1180     0
     1  3  1  1  1  2005   063  1240     0
     1  4  1  1  1  2005   063  1280     0
     1  1  1  1  2  2005   063   090     0
     1  2  1  1  2  2005   063   380     0
     1  3  1  1  2  2005   063   640     0
     1  4  1  1  2  2005   063   880     0
     1  1  1  1  1  2005   092  4640     0
     1  2  1  1  1  2005   092  4320     0
     1  3  1  1  1  2005   092  4530     0
     1  4  1  1  1  2005   092  4440     0
     1  1  1  1  2  2005   092  1640     0
     1  2  1  1  2  2005   092  1320     0
     1  3  1  1  2  2005   092  1530     0
     1  4  1  1  2  2005   092  1440     0

[data series truncated]

@TRTNO RP SQ OP CO  YEAR   DOY  SW1D
     1  0  1  1  1  2005   063  0.33
     1  0  1  1  1  2005   092  0.23
     1  0  1  1  1  2005   120  0.33
     1  1  1  1  1  2005   063  0.33
     1  2  1  1  1  2005   063  0.35
     1  3  1  1  1  2005   063  0.31
     1  4  1  1  1  2005   063  0.33
     1  1  1  1  1  2005   092  0.23
     1  2  1  1  1  2005   092  0.25
     1  3  1  1  1  2005   092  0.21
     1  4  1  1  1  2005   092  0.23
     1  1  1  1  1  2005   120  0.33
     1  2  1  1  1  2005   120  0.35
     1  3  1  1  1  2005   120  0.31
     1  4  1  1  1  2005   120  0.33
=
```

Figure 5. Example of an ICASA Weather File

```
$WEATHER:KSAS

!The following data subset is desirable.

*GENERAL
@ PEOPLE
  Wagger,M.G. Kissel,D.
@ INSTITUTES
  Kansas State Univ.
@ CONTACTS
  Hunt,L.A. thunt@uoguelph.ca
@ NOTES
  Wind and dewpoint included; wind in m/s as average over 24h
  Data from DSSAT 3.0 file
@ DISTRIBUTION
  Use at will but acknowledge source. No secondary distribution
@ VERSION
  ICASA1.0 10-08-2006 (GH,JW,LAH;email)
! And additional data items can be added for comprehensive documentation:
@ METHODS
  Standard Met Station instruments
@ PUBLICATIONS
  None of direct application
@ FLAG FLAG_DETAILS
     0 All data ok
     1 SRAD estimated from sun hours

!And data subsets similar to the following are necessary for all stations.

*WEATHER_STATION:KSAS2004
@ NAME
  Example weather dataset
@ COUNTRY     REGION           LOCATION
  USA         Kansas           Ashland
@ LAT         LONG        ELEV   TAV  TAMP TEMHT WNDHT   CO2  CO2A
     37.1137    -90.4567    81   8.5  18.9   2.0   2.0   370   1.1
@ YEAR   DOY  SRAD  TMAX  TMIN  RAIN  FLAGW
  2004     1  11.5   1.4  -1.4   3.5      1
  2004     2  11.4   0.0  -2.1   2.5      0

[data series truncated]

!Additional data items can be added as illustrated below.
@ YEAR   DOY  SRAD  TMAX  TMIN  RAIN  WIND  TDEW
  2004     1  11.5   1.4  -1.4   3.5   4.2   0.6
  2004     2  11.4   0.0  -2.1   2.5   5.4   0.5

[data series truncated]
=
```

Figure 6. Example of an ICASA Soils File.

```
$SOIL:KS

!Some general information is desirable in all files.

*GENERAL
@ PEOPLE
  Wagger,M.G. Kissel,D.
@ INSTITUTES
  Kansas State Univ.
@ CONTACTS
  Hunt,L.A. thunt@uoguelph.ca
@ NOTES
  Put together to illustrate the ICASA standard files. Fictitious
@ DISTRIBUTION
  Use at will but acknowledge source. No secondary distribution
@ VERSION
  ICASA1.0 10-08-2006 (GH,JW,LAH;email)
@ METHODS
  Standard Soil Survey + Ritchie algorithms to obtain H2O limits
@ PUBLICATIONS
  None of direct application

!And specific information similar to that shown below is necessary
!for each soil.

*SOIL_PROFILE:KSAS200401
@ NAME
  Example soils file
@ COUNTRY     REGION           LOCATION
  USA         Kansas           Ashland
@ LAT         LONG        ELEV SLOPE
     37.1167    -90.4529    81   1.5
@ SLLB  CLAY  SILT  SLCF  SLOC SLPHW SLBDM
     5  7.92 16.31  3.25  1.89   8.5  1.42
    30  7.92 16.31  3.25  1.89   8.5  1.42
    60  6.35 10.61  1.52  0.88   8.5  1.67
    90  5.04 10.28  1.14  0.66   8.6  1.69
   190  5.72 12.08  1.72  0.99   8.6  1.53
```

```
!With the following added if available.
@ SL_SYSTEM   CLASSIFICATION
  NRCS(USDA)  Coarse-silty,mixed,calcareous,mesic typic udifluvent
@ SALB  SLU1  SLDR  SLRO  SLNF  SLPF  SLDN
  0.13    12  0.12    84   1.0   1.0  0.30
@ SLLB  SLLL SLDUL SLSAT SKSAT  SLNI SLPHB SLCEC SLADC SLRGF SCMPF
     5  0.12  0.32  0.48  12.8   2.0   8.5   255   0.0   1.0   0.0
    30  0.12  0.32  0.48  12.8   1.5   8.5   255   0.0   1.0   0.0
    60  0.09  0.22  0.38   6.1   1.0   8.5   255   0.0   0.4   0.6
    90  0.08  0.21  0.37   9.5   0.4   8.6   235   0.4   0.4   0.6
   190  0.09  0.24  0.43   6.4   -99   8.6   235   1.6   0.4   0.6
!And further data dealing with chemical characteristics that may be
!needed for special applications added if required and available.
@ SLLB  SLPX  SLPT  SLPO CACO3  SLAL  SLFE  SLMN  SLBS  SLPA  SLPB  SLKE  SLMG  SLNA  SLSU  SLEC  SLCA
     5   -99   -99   -99   -99   4.0   -99   -99   -99   4.0   4.0   5.0   -99   -99   -99   -99   3.0
    30   -99   -99   -99   -99   4.0   -99   -99   -99   4.0   4.0   5.0   -99   -99   -99   -99   3.0
    60   -99   -99   -99   -99   4.0   -99   -99   -99   4.0   4.0   5.0   -99   -99   -99   -99   3.0
    90   -99   -99   -99   -99   4.0   -99   -99   -99   4.0   4.0   5.0   -99   -99   -99   -99   3.0
   190   -99   -99   -99   -99   4.0   -99   -99   -99   4.0   4.0   5.0   -99   -99   -99   -99   3.0
```