Chapter 2
Experiments and Data for Model Evaluation and Application

Gerrit Hoogenboom, James W. Jones, Pierre C.S. Traore, and Kenneth J. Boote

Abstract Crop models and decision support systems can be very useful tools for scientists, extension educators, teachers, planners and policy makers to help with the evaluation of alternative management practices. Many of the current crop models respond to differences in local weather conditions, soil characteristics, crop management practices and genetics. However, computer-based tools require inputs in order to provide reliable results. Especially for those new to crop modeling, the data requirements are sometimes somewhat overwhelming. In this chapter we provide a clear and concise summary of the input data requirements for crop modeling. We differentiate between requirements for model evaluation, model application and model development and improvement. For model inputs we define daily weather data, soil surface and profile characteristics, and crop management. For model evaluation and improvement we define crop performance data as it relates to growth, development, yield and yield components, as well as additional observations. We expect that this chapter will make the use and application of crop models and decision support systems easier for beginning modelers as well as for the more advanced users.

Keywords Crop modeling • Simulation • Decision support systems • Minimum data set • DSSAT • Cropping System Model (DSSAT) • CROPGRO• CERES

G. Hoogenboom (✉)
AgWeatherNet, Washington State University, Prosser, Washington 99350, USA
e-mail: gerrit.hoogenboom@wsu.edu

J.W. Jones
Department of Agricultural and Biological Engineering, University of Florida, Gainesville, Florida 32611, USA

P.C.S. Traore
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Bamako, Mali

K.J. Boote
Department of Agronomy, University of Florida, Gainesville, Florida 32611, USA

J. Kihara et al. (eds.), *Improving Soil Fertility Recommendations in Africa using the Decision Support System for Agrotechnology Transfer (DSSAT)*,
DOI 10.1007/978-94-007-2960-5_2,

Introduction

With the increasing interest in the applications of crop modeling and decision support systems, there is a need to clearly define the type of experiments that are required for both crop model evaluation and application. Especially for those new to crop modeling it is unclear what types of experiments should be conducted and what information should be collected in these experiments. Over the years several publications have been written to document these requirements (IBSNAT 1988; Hunt and Boote 1998; Hunt et al. 2001). The most extensive ones can be found in the documentation that was developed for the Decision Support System for Agrotechnology Transfer (DSSAT) Version 3.5, especially Volume 4 (Hoogenboom et al. 1999). This information is still relevant and has been included as electronic documents in the documentation section of DSSAT Version 4.0 (Hoogenboom et al. 2004) and DSSAT v4.5 (Hoogenboom et al. 2010).

Volume 4.8 entitled "Field and Laboratory Methods for the Collection of the Minimum Data Set" by Ogoshi et al. (1999) is based on Technical Report 1 that was published by the International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) Project (IBSNAT 1988). It includes extensive documentation on data collection procedures for modeling. In volume 4.7 entitled "Data Requirements for Model Evaluation and Techniques for Sampling Crop Growth and Development" Boote (1999) provides detailed procedures on the actual sampling techniques for growth analysis and crop development. However, an easy to use summary is currently not available. The goal of this chapter is, therefore to provide a clear and concise summary for experimental data collection for model evaluation and application.

Overview

In order to run a crop model and to conduct a simulation, a set of data are required. Sometimes this is referred to as a "Minimum Data Set." The terminology Minimum Data Set was first introduced by the IBSNAT Project. Although the type and details required for model inputs might vary somewhat depending on the crop or agricultural model, in general we can differentiate between three broad levels or groups. Level 1 defines the data required for model applications, Level 2 defines the data required for general model evaluation, and Level 3 defines the data required for detailed model calibration and evaluation. Potentially this type of data can also be used for the development of a model for a crop for which currently no dynamic crop simulation model exists.

Level 1 includes daily weather data, soil surface characteristics and soil profile information, and crop management. Level 2 includes the environmental and management data from Level 1 and some type of observational data that are collected during the course of an experiment. At a minimum the two key phenological phases, i.e., flowering or anthesis and physiological or harvest maturity, and yield and yield components are needed for observational data. Level 3 would include the environmental, management and observational data described under Level 2 and additional observations related to growth and development, such as growth analysis, soil moisture content, and soil and plant nitrogen, phosphorus, potassium, and others, depending on the overall intended model application or evaluation.

Experiments and Modeling

It is important to understand that one rarely develops an experiment for modeling only, but that experiments should be conducted in such a manner that they also have a modeling component that can be used for either model evaluation or application or both. It is also important to keep in mind that some of the basic data that are required for any model application, especially those described under Level 1, should be a basic set of data that are collected for documentation of any experiment. For instance, for many experiments local weather and soil conditions have a major impact on the outcomes of an experiment and should be included as part of the overall analysis.

Location of Experiments

Normally data for model evaluation are obtained from experiments, although in some cases one might only have access to statistical yield and production data. Although this information can be used, one should understand the level of detail and the quality of this type of data and expected outcomes with respect to the accuracy of the evaluation of a model. In general experiments can be conducted under controlled management conditions, referred to as “on-station” and in farmers’ fields, referred to as “on-farm.” For Level 3 one normally would not use data from on-farm experiments, but the data can be useful for Level 2 model evaluation if one understands the limitations of the data, such as the lack of replications in most cases, variability of environmental conditions and uncertainty of the inputs. In some cases experiments can be conducted in growth chambers or in Soil-Plant-Atmosphere Research (SPAR) chambers where most environmental conditions can be controlled. However, for accurate model evaluation, on-station experiments with at least three or four replications are preferred.

General Purposes of Experiments

It is always important to keep the overall goal of the research in mind and design appropriate experiments accordingly, rather than concentrating on the model only. There is a wide range of applications with some of the key ones listed below.

– Technology evaluation, such as evaluation of new cultivars, inputs, including irrigation and fertilizers, and soil preparation, such as tillage and conservation agriculture.

– Characterization of yield limiting factors in order to focus on new technology development and evaluation.
– Understanding the interactions among management factors, such as water, nutrients, etc., and aiming at refining agricultural management technologies.
– Understanding the interactions of the environment, such as increases in temperature and CO_2.
– Understanding the interactions between genotype and environment (G x E).
– Long-term soil sustainability and soil health, including improvement of soil organic matter.
– Understanding environmental impact, such as nitrogen pollution due to different management practices.
– Potential application of agricultural crops for food, feed, fiber and fuel production.

General Purposes of Model Use

It is important to determine the overall purpose of the use of modeling and how it contributes to the overall research goal. In many cases adding a systems analysis and modeling component can strengthen the overall research approach. A partial list of model applications is listed below.

– Understand and interpret experimental results.
– Enhance quality of field research and the results that are derived from it.
– Diagnose yield gaps by looking at the differences between potential, attainable and actual yield from on-station and on-farm research, and to help develop technologies to test these under field conditions.
– Help publish results of field trials via systems and modeling analysis.
– Estimate impacts on production, water use, nitrogen use, and other inputs and determine various resource use efficiencies at scales from field to farm to watershed to region and higher.
– Estimate economic implications of different technologies.
– Estimate the impact of climate change and climate variability on crop production and develop adaptation scenarios.
– Plant breeding, Genotype * Environment interaction and the development of ideotypes.
– Enhance interdisciplinary research through interaction of soil scientists, agronomists, economists, engineers, GIS/remote sensing scientists, and others.

Level 1 Data

Level 1 data for model applications include daily weather data, soil characteristics and crop management. These data are an absolute requirement for any successful model evaluation and application. Well-documented experiments normally already include this information or researchers can provide you with the source of the data.

Weather Data Required (Daily)

1. Minimum and maximum temperature
2. Precipitation or rainfall
3. Total solar radiation or sunshine hours
4. Dewpoint temperature or relative humidity. If an automated weather station is used for weather data collection, this information is normally readily available
5. Average daily wind speed or daily wind run. Again, if an automated weather station is used for weather data collection, this information is normally readily available

Measured or recorded minimum and maximum temperature and rainfall have to be measured with a gauge, thermometer or environmental sensor, while in some cases other methods can be used to help estimate daily solar radiation based on satellites (White et al. 2011) or simple equations or solar radiation generators (Garcia y Garcia and Hoogenboom 2005).

It would be important to the overall value of an experimental data set to have an automatic weather station collect daily weather data in areas where experiments are to be conducted. In addition, rainfall is needed for all sites where experiments are performed, including on-farm as well as on-station due to its significant spatial and temporal variability. Otherwise, it will not be possible to interpret crop yield response or to understand the interactions of water, nutrients and management. Temperature and solar radiation could be measured in the areas where several experiments are conducted. Generally, temperatures and solar radiation may change little over distances of about 50 km, but this is not true near the coast or when elevation changes occur in the area. One could situate a weather station in the center of an area where experiments will be conducted either on-farm or on-station.

Soil Data

Soil information includes general site and soil surface information and soil profile characteristics.

1. General site information
   - Latitude, longitude and elevation
2. Soil surface information
   - Soil taxonomy (if available)
   - Soil slope

- Soil color
- Stones (%)

3. Soil profile data, for each soil horizon in which roots are likely to grow

- Soil texture, including % sand, silt, and clay and stones, especially for the surface layers
- Soil organic carbon
- Bulk density is desirable
- Lower Limit of plant extractable soil water (LL) or permanent wilting point and Drained Upper Limit (DUL) or field capacity. Field measurements are desirable, but there are various methodologies to help estimate these parameters

Initial Conditions

When running the model with the soil water, nitrogen, phosphorus, organic carbon and other soil components turned on, it is important to define the initial conditions of the soil profile at the start of the simulation. This might not necessarily be the planting date. If initial conditions are not defined, the model normally defaults to a 0 value for all soil nutrients, while soil moisture is set to the drained upper limit. We realize that for many scientists it is rather difficult to measure these initial conditions, or they do not have the resources. DSSAT, therefore, has tools to help you estimate values. Note that the surface and soil residues are the residues of the previous crop only. Any supplemental residues or manure are considered part of the inputs.

1. Previous field history

- For the simulation of the soil organic carbon balance a user has to know the previous crop history.

2. Initial soil profiles conditions

- Initial soil moisture versus depth
- Initial nutrients (NO_3, NH_4, P) versus depth
- Other soil chemical properties as needed for the experimental objectives

3. Surface residues at the start of simulation or at planting

- Crop type or manure type
- Total amount as dry weight
- %N and %C (and %P) contents
- Incorporation depth and % incorporation
- Note that if resides or manure are applied at planting they are considered to be part of crop management

Management Data

Crop management data seems like a relatively easy set of information to obtain. However, it is surprising to find that many experiments lack detailed information except if it is part of a treatment. For instance, if supplemental irrigation is applied but it is not part of the overall set of treatments, it is sometimes difficult to obtain the actual dates of the irrigation applications and the amount of water that was applied.

1. <u>Planting</u>
   - Date of planting
   - Plant spacing or density. This would be the official plant stand and not the seeding or sowing density.
   - Crop and cultivar name and its characteristics.
   - Planting material, e.g. seed, stick, etc.
   - Planting mode, e.g., row, hill, flat, ridge, etc.
2. <u>Input information</u>
   - Irrigation amount and the timing of the irrigation application
   - Fertilizer amount and type, timing of the fertilizer application, placement depth and application method
   - Amount of organic manure or residue, composition, time of the application, placement or incorporation depth and method of application
   - Amount and type of chemicals applied and for what purposes

Level 2 Data

Level 2 data includes the basic input data defined under level 2, including weather conditions, soil characteristics, crop management and initial conditions. In addition it includes some type of observational data in response to differences among treatments. This could range from different environments, including soil and weather conditions, different crops, different cultivars, varieties or hybrids, or different inputs.

Crop and Soil Response Measurements

1. <u>Treatments</u>. Although this seems obvious, it is important to understand the different treatment factors and their associated levels.

2. Yield and yield components

   - Grain yield (kg/ha)
   - Anthesis or flowering date and maturity date. Time of first seed or first grain would also be helpful
   - Number of main stem nodes
   - Above ground biomass, excluding grain
   - Plant density at harvest
   - Number of ears, pods, or other fruiting structures per unit area
   - Average weight per unit grain, seed, fruit or other harvested material
   - N and P concentrations of grain and other plant components

3. General observations

   - Weeds and weed management. It is important to document if the weeds affected the actual outcome of the experiment, such as yield and biomass
   - Pests and disease occurrence, including the date of the infection intensity, and actual damage
   - Damage due to extreme weather events, such as hail, rainstorms, wind gusts, etc.
   - General health of the crop

Level 3 Data

There are two types of data that provide crucial information related to growth response and water and nutrient use dynamics. These are detailed crop response measurements and detailed soil response measurements. However, they are expensive and time consuming, so it is impractical to collect these data in all experiments, especially if the experiments are conducted in farmers' fields. To obtain the data from these measurements in a few places would greatly increase the confidence in the use of models and the simulation of the responses in fields where they are not measured. In particular, water use by crops is highly critical to yield formation and is important in any technology aimed at increasing productivity and sustainability of soils in rainfed agriculture. The models predict root growth and water uptake, but credibility of those predictions will require good inputs as well as evaluation in soils and cropping systems in the region in which they are to be applied. Thus, the following measurements are suggested.

1. Growth analysis measurements. These measurements should be taken at least four to eight times during the growing season at regular spaced intervals, such as weekly or bi-weekly and one detailed sample at final harvest. Previous recommendations were to collect growth analysis samples at critical growth stages, but we have changed this approach and now prefer regularly spaced intervals for growth analysis sampling. Generally, these samples would be taken in on-station experiments only, and for some experiments only. This is mainly to make sure that you have good genetic coefficients and can simulate attainable yield or potential yield

- Above ground biomass
- Leaf, stem, seed, and ear (pod) mass
- Number of main stem nodes or leaves
- Leaf area index, if possible either destructively or in the field using a hand-held device
- N and P concentrations in plant parts
- Numbers of ears (pods) per square meter

2. Soil water content versus depth. There are various measurement techniques including gravimetric samples, time domain reflectometry (TDR), neutron probes and other electronic sensors. Samples or readings should be taken at least weekly or preferably more often. Moisture sampling after a prolonged wet period can be useful to determine the upper limits of water storage for the respective soil layers. Moisture sampling following a well-grown crop experiencing terminal drying is useful for determining the crop lower limits for water extraction.

3. Soil fertility versus depth

- Analysis for NO_3-N (ppm) for soil layers and sampling times are similar to what was discussed for soil water. However, these measurements can be taken less frequently due to the additional resources required and costs associated with sample collection and laboratory analysis.
- Soil Organic Carbon (%) for soil layers to rooting depth. At least one entire profile should be sampled for each soil type. This information is extremely critical, especially for eroded soils commonly found in the tropics. For other sites with the same soil type, a surface layer sample will suffice. In this case it is best to make the layer thickness applicable to the un-sampled profile. Total soil N could be useful to help determine the overall C:N ratio.
- Olsen-P, Total P and Org P for the top two soil layers should suffice.

Summary

This chapter provides a detailed review of the environmental and crop management data required for model simulation and the crop and soil response data for model evaluation. The DSSAT approach has been to emphasize model evaluation with locally available data to demonstrate that the model is able to simulate management and environmental responses and to determine local genetic coefficients. This is important, not only for a researcher, but also for those who will be using the outcomes of the simulation studies for policy and planning. Although the list of data might seem overwhelming for a novice modeler, most of the data presented here are in some form already available, especially for well-managed on-station experiments. Prior to designing and developing an experiment for model evaluation, make sure that this experiment has not already been conducted by a colleague at your university or a sister institution in the region. If you have to conduct an experiment, make sure to partner with others to compliment your skills and expertise. At the end the results will be much better and the outcomes of the modeling study much stronger. If questions still remain, make sure to use the resources that are available through international networks of model users, such as the DSSAT group (www.DSSAT.net).

References

Boote KJ (1999) Data required for model evaluation and techniques for sampling crop growth and development. DSSAT v3, vols 4–7. University of Hawaii, Honolulu, pp 201–216

Garcia y Garcia A, Hoogenboom G (2005) Evaluation of an improved daily solar radiation generator for the southeastern USA. Climate Res 29:91–102

Hoogenboom G, Wilkens PW, Tsuji GY (1999) DSSAT v3, vol 4. University of Hawaii, Honolulu, 286 pp. ISBN 1-886684-04-9

Hoogenboom G, Jones JW, Wilkens PW, Porter CH, Batchelor WD, Hunt LA, Boote KJ, Singh U, Uryasev O, Bowen WT, Gijsman AJ, du Toit A, White JW, Tsuji GY (2004) Decision support system for agrotechnology transfer version 4.0. [CD-ROM]. University of Hawaii, Honolulu

Hoogenboom G, Jones JW, Wilkens PW, Porter CH, Boote KJ, Hunt LA, Singh U, Lizaso JL, White JW, Uryasev O, Royce FS, Ogoshi R, Gijsman AJ, Tsuji GY (2010) Decision support system for agrotechnology transfer version 4.5. [CD-ROM]. University of Hawaii, Honolulu

Hunt LA, Boote KJ (1998) Data for model operation, calibration, and evaluation. In: Tsuji GY, Hoogenboom G, Thornton PK (eds) Understanding options for agricultural production. Systems approaches for sustainable agricultural development. Kluwer Academic, Dordrecht

Hunt LA, White JW, Hoogenboom G (2001) Agronomic data: advances in documentation and protocols for exchange and use. Agr Syst 70:477–492

IBSNAT (1988) Technical report 1, experimental design and data collection procedures for IBSNAT, 3rd edn revised. Department of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, University of Hawaii, Honolulu

Ogoshi RM, Cagauan BG Jr, Tsuji GY (1999) Field and laboratory methods for the collection of the minimum data set. DSSAT v3, vols 4–8, pp 217–286. University of Hawaii, Honolulu, Hawaii, 96822

White JW, Hoogenboom G, Wilkens PW, Stackhouse PW, Hoell JM (2011) Evaluation of satellite-based, modeled-derived daily solar radiation data for the continental United States. Agron J 103(4):1242–1251