Published January, 1995

A Computer Program to Analyze Multiple-Season Crop Model Outputs

Philip K. Thornton,* Gerrit Hoogenboom, Paul W. Wilkens, and Walter T. Bowen

ABSTRACT

Management-oriented simulation models of the growth, development, and yield of annual crops are useful tools for screening management options on the computer. Until recently, a limitation of these models has been the inability to simulate more than one cropping season at a time. The capability to simulate long-term field experiments with such models now exists, in which the simulated soil water, N, organic C, and crop residue outputs from one model run become the input conditions for the next. Simulations of crop rotations can produce large quantities of data, especially if the simulation experiment involves replications across different years. Computer software was written to perform simple analyses of such simulation experiments. The major purpose of the software is to allow the user to investigate the stability and profitability of crop sequences. The program calculates summary statistics for model output variables; these are presented to the user in tabular and graphical forms. Net monetary returns or gross margins can also be calculated, and price and cost variability can be taken into account in the analysis. The program allows rapid, preliminary analysis of a particular crop sequence from replicated simulation experiments and can help the user to assess whether the sequence warrants further evaluation. The program can also be used to summarize the results from historical long-term field trials. The analyses performed constitute a first step in investigating the sustainability of a particular cropping sequence for a specified length of time.

Crop simulation models constitute one of the few tools that can be used to investigate the long-term economic feasibility of crop performance in a given environment. With appropriate soil and weather data, crop models can be used for studying sustainability issues at any number of sites. The management-oriented crop models associated with the International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT) project continue to be developed, in particular with respect to the capability to simulate crop rotations or sequences over a number of years. These include the CERES (Singh et al., 1988; 1993) and GRO (Hoogenboom et al., 1992) families of models. Simulation of a cropping sequence involves setting the initial conditions at the start of the run and then simulating crop growth during a number of successive growing seasons interspersed with fallow periods when there is no planned crop cover. Appropriate crop models are run in sequence, and the soil water, nutrient, and organic matter status at harvesting of one crop becomes the input status of the next model, be it another crop model or a simulated fallow period (Bowen et al., 1993). For a fallow period, no crop growth takes place, but water and nutrient dynamics in the soil are simulated.

The ability to accurately simulate crop rotations has a number of implications and uses. These include the screening or assessment of management options with time, given particular environmental conditions and resource constraints at the farm level. Such assessments may involve a number of replications across different sequence years to quantify the production risk associated with different weather patterns. Large amounts of model output data can be produced from such simulation experiments. The user is then faced with the task of analyzing these outputs to produce useful information for helping to make decisions.

This paper describes computer software that analyzes the results of simulation experiments involving annual crops run in sequences. Summary statistics for model outputs over time can be calculated and graphed. A simple economic analysis can also be performed, enabling the user to examine the economic feasibility of the particular management scenario and to help identify those that are most promising for a particular environment.

The software is intended for users of crop simulation models, especially those who would apply crop models to investigate specific production problems and sustainability issues. The software package can also be used to investigate and summarize the major results of historical long-term field trials. It could also be of use to model developers, allowing rapid appraisal of sensitivity analyses designed to identify possible problems in model specification, for example.

The software was designed as one module of IBSNAT's Decision Support System for Agrotechnology Transfer (DSSAT). DSSAT Version 3 can be used for storing information on field trials, extracting data from crop, soils, and weather databases in a format suitable for the simulation models, running the crop models over single seasons for purposes of validation or comparing management strategies, and performing simple analyses of the results of simulation runs (Jones, 1993). DSSAT-3 can be used to simulate and analyze crop sequences. Although the analysis program is an integral part of DSSAT-3, designed to analyze outputs from the IBSNAT crop models, it can also be run in a stand-alone mode. Thus the outputs from any other crop simulation model can be analyzed if the outputs have a compatible format. For example, the analysis program could be used to complement other long-term simulation packages such as the CENTURY model (Parton et al., 1992) and EPIC (Williams, 1990).

DEFINITION OF TERMS

To clarify terminology used in the discussion, we define the following:

1. A *sequence* refers to the growing of particular crops, one after another, for a specified length of time. The cropping system that is simulated may be a contin-

P.K. Thornton, P.W. Wilkens and W.T. Bowen, Res. & Dev. Div., Int. Fertilizer Dev. Ctr., P.O. Box 2040, Muscle Shoals, AL 35662; G. Hoogenboom, Dep. of Biological and Agric. Engineering, Georgia Station, Univ. of Georgia, Griffin, GA 30223-1797. Received 5 Mar. 1994. *Corresponding author (Email: ifdc.model@cgnet.com).

Published in Agron. J. 87:131-136 (1995).

Abbreviations: CGA, Color Graphics Adapter; CPF, cumulative probability function; DSSAT, Decision Support System for Agrotechnology Transfer; EGA, Enhanced Graphics Adapter; HPGL, Hewlett-Packard Graphics Language; IBSNAT, International Benchmark Sites Network for Agrotechnology Transfer; IFDC, International Fertilizer Development Center; K, Kilobyte; VGA, Video Graphics Array.

uous succession of the same crop, or it may consist of a rotation of different crops.

2. A *replicate* is obtained by simulating a cropping sequence using a distinct sequence of weather. Analysis of a number of replicates allows weather-induced variability in crop performance to be quantified and assessed.
3. *Variation* in a computer-based crop sequencing experiment arises between treatments, between replicates, and between years within a sequence. The software package described below examines variation between replicates and between years for a particular cropping sequence or treatment.

PROGRAM DESCRIPTION

The primary objective of the program is to analyze the results from a cropping sequence experiment that is run by using simulation models of annual crops. The experiment will be made up of a particular sequence, such as continuous maize (*Zea mays* L.) grown in a plot of land for 10 yr, and it may involve a number of replicates derived by using different weather sequences. The weather data that drive the crop models may be historical, or may be simulated by using a statistical weather generator. The program allows the user to assess the performance of the sequence in terms of absolute or relative yield levels, yield risk, N loss through leaching, or economic returns for the length of the sequence, for example. The software has both data analysis and graphics capabilities.

Input and Output Files

The software package requires and produces a number of files (Table 1). Three of these files are briefly described as follows.

The Model Summary Output File contains the results of the sequence experiment, stored as one model run per record (one line in the file). The output variables written to this file are listed in Table 2; the data format is described in detail by Jones et al. (1994).

The Price-Cost File contains time series of the prices and costs for each crop component for each year of the sequence (Table 3). Each price and cost may be fixed (deterministic), or it may be described by a normal distribution, in which case the user specifies a standard deviation in terms of a percentage of the mean. The time series may comprise real or nominal dollars; they may be historical or generated according to a trend specified by the user. If no price-cost file exists, then the software will generate a default file that can be edited as required.

Table 1. Input and output files associated with the multiseason crop model analysis program.

File	Description
Model summary output	Simulation results, stored as one model run per record (input file)
Summary output listing	List of model summary output files in the current directory (input file)
Price-cost file	Price and cost time series for analyzing the experiment (input file)
Graphics device file	Options for hard-copy graphics output in terms of printer, plotter, port, scaling, plotter language, destination (input file)
Variable abbreviations file	Descriptions and legends for all model output variables available for analysis (input file)
Analysis results file	Output file from the analysis program containing results and calculations (output file)
Plotter files	Optional file containing plot commands for sending to plotter, printer, or document (output file)

Table 2. Variables written to the summary output file of the multiseason crop model analysis program.

Variable	Description†
1	Simulation starting date
2	Planting date
3	Anthesis date
4	Physiological maturity date
5	Harvest date
6	Weight of planting material, kg DM ha^{-1}
7	Weight of plant tops at maturity, kg DM ha^{-1}
8	Yield at maturity, kg DM ha^{-1}
9	Yield at harvest, kg DM ha^{-1}
10	Harvested by-product (such as straw) dry weight, kg DM ha^{-1}
11	Harvest product individual weight, mg DM
12	Harvest product number at maturity, number m^{-2}
13	Harvest product number per unit at maturity, number $unit^{-1}$
14	Number of irrigation applications during the season
15	Total seasonal irrigation applied, mm
16	Total seasonal precipitation, mm
17	Total seasonal evapotranspiration, mm
18	Total seasonal surface runoff, mm
19	Total seasonal water drainage, mm
20	Extractable soil water in the profile at maturity, cm
21	Number of N applications during the season
22	Total inorganic N applied, kg ha^{-1}
23	Total N_2 fixed during the season, kg ha^{-1}
24	N uptake by the plant during the season, kg ha^{-1}
25	N leached from profile during the season, kg ha^{-1}
26	Soil inorganic N at maturity, kg ha^{-1}
27	N content of plant tops at maturity, kg ha^{-1}
28	Harvestable product N content at maturity, kg ha^{-1}
29	Total crop residue applied, kg ha^{-1}
30	Organic soil N at maturity, kg ha^{-1}
31	Organic soil C at maturity, t ha^{-1}
32	Number of P applications during the season
33	Total P applied, kg ha^{-1}
34	P in the plant tops at maturity, kg ha^{-1}
35	Soil P content at maturity, kg ha^{-1}

† DM, dry matter.

The Analysis Results File stores all calculations and analysis results of a session with the analysis program for subsequent manipulation by the user. The file can be edited as required outside the software package. For instance, a user may import the file into a spreadsheet program to produce graphs not available in the software package, or into a statistics package to perform statistical analyses of the simulated results.

Program Operation

The program can analyze summary output files with a maximum of 1200 separate simulation runs. This total number may be split between the number of replicates and the length of the sequence. For example, a maize-fallow sequence might be repeated over a number of years. With two model runs per calendar year (maize planted at the start of the growing season and a fallow model to cover the winter until the start of the growing season in the next calendar year), these 1200 runs could be made up of 10 replicates of 60 yr (2 model outputs per year $\times$ 60 $\times$ 10 = 1200 runs) or 20 replicates of 30 yr (2 $\times$ 20 $\times$ 30 = 1200 runs), for instance.

The main menu of the package lists a number of options (Fig. 1): the user may select a model-produced file for analysis, select options for changing the way in which hardcopies of graphs are produced, assess the crop sequence in terms of biophysical or economic outputs, or quit the program.

The user may select any one of the 35 model-generated output variables listed in Table 2. (These output variables are specific to the IBSNAT crop models, but other variables produced by other crop models may be substituted for those shown.) The program calculates the mean, standard deviation, minimum, and maximum values over all replicates and tabulates these by season. The user can then proceed to graph these results in five ways:

Table 3. Costs and prices used in economic analysis of the multiseason crop model analysis program.

Variable	Description	Units	Example value
1	Price of grain	\$ t^{-1}	130.00
2	Price of by-product†	\$ t^{-1}	0.00
3	Base production costs‡	\$ ha^{-1}	185.00
4	N fertilizer cost	\$ kg^{-1}	0.62
5	Cost per N fertilizer application	\$	12.00
6	Irrigation costs	\$ mm^{-1}	0.00
7	Cost per irrigation application	\$	0.00
8	Seed cost	\$ kg^{-1}	0.50
9	Organic amendments	\$ t^{-1}	0.00
10	P fertilizer cost	\$ kg^{-1}	0.00
11	Cost per P fertilizer application	\$	0.00

† Straw or stover, for example.
‡ Other costs of production (such as labor, fuel, etc.).

1. As a percentile plot, where the 0th, 25th, 50th, 75th, and 100th percentiles of the distribution of the output variable are plotted against time.
2. As a cumulative probability function (CPF) plot, where the distribution of the output variable for any year is ordered from smallest to largest value and plotted against equal increments of cumulative probability (a maximum of six CPFs is allowed per plot to avoid congestion).
3. As a mean-variance plot, where the mean of the output variable is plotted against its variance.
4. As a variance plot, where the variance of the output variable of interest is plotted against time.
5. As a coefficient of variation (CV) plot, where the CV of the variable, expressed as a percentage, is plotted against time.

Percentile and CPF plots are useful for examining the way in which an output distribution changes during the course of a cropping sequence. Mean-variance, variance, and CV plots can be used to examine changes with time in the stability of simulated output variables.

The user also has the option to perform a univariate regression of the simulated mean values of any output variable. The calculated regression can be plotted with the simulated mean values, and the residuals of the regression can also be plotted and examined. Examples of some of these graphs are shown in a subsequent section.

Economic Analysis

For an economic analysis, the user may select a price–cost file for carrying out the analysis. Prices and costs can be edited from within the program using a text editor. The user may choose to include price and cost variability in the economic analysis; future production costs and product prices will not be known with certainty. The analysis then proceeds by combining the simulated distributions of crop model outputs with distributions of costs and prices that are specified by the user in the price–cost file. The method used is discussed in Thornton and Hoogenboom (1994).

The output of economic analysis is a time series of distributions of gross margin or net return. This is tabulated and can be graphed, either in terms of the various crop components that make up the crop sequence or in terms of total returns. In some situations, the fallow period will have an economic cost. For this reason, the user may choose to include or exclude fallow periods in an economic analysis. The program also calculates the probability, with time, of failing to reach a critical level of

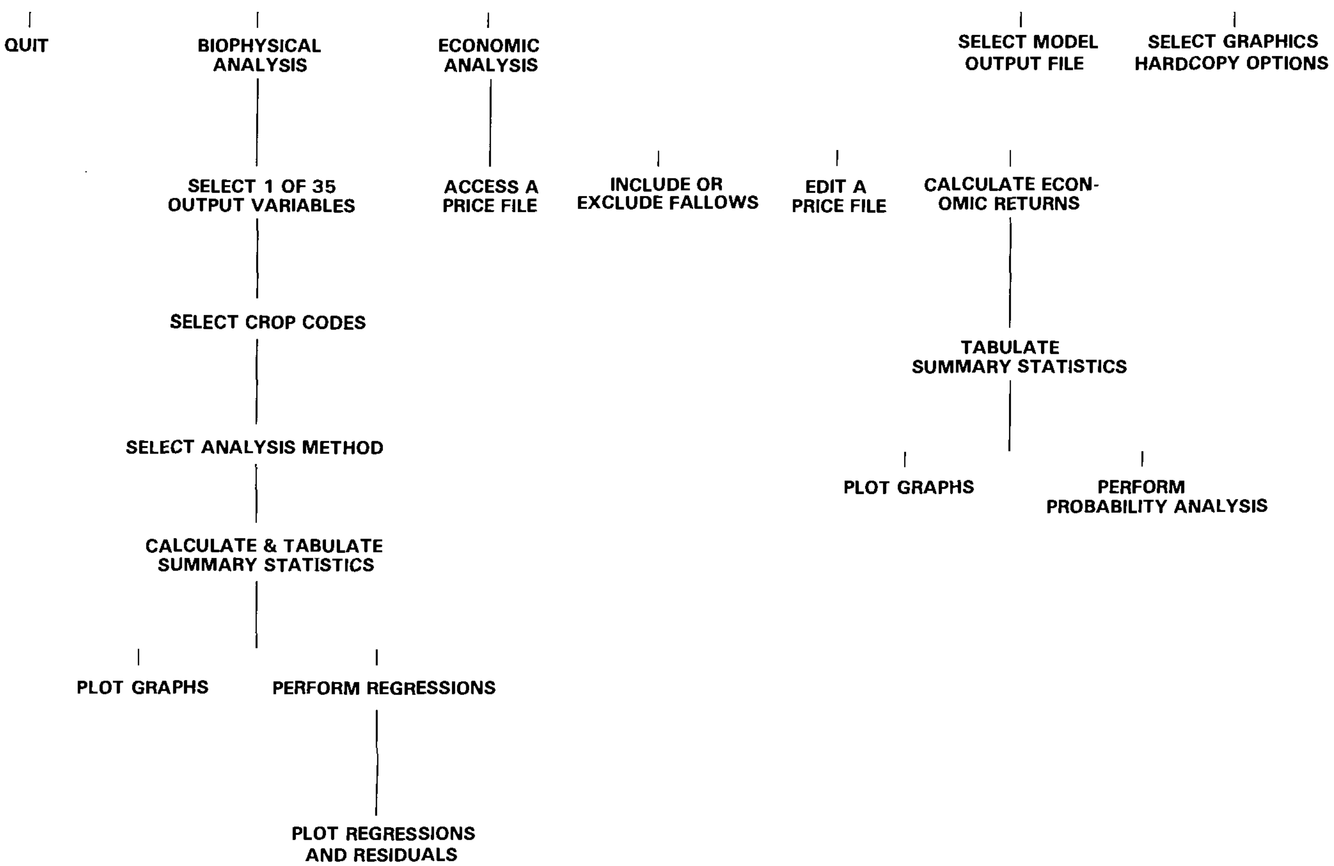

Fig. 1. Diagram of program options.

net return. Production systems that become less productive will become less profitable, and the probability of disastrous outcomes may well increase. The user may choose a critical level of returns, and the software will then calculate the probability of failing to reach this level each cropping season. This probability of failure is calculated using linear interpolation between successive data points on the CPF of annual economic returns. The probabilities may then be plotted as a time series.

AN EXAMPLE

To illustrate the use of this software package and the types of graphs that can be produced, a simulation experiment was performed with sequencing versions of CERES-Maize (Ritchie et al., 1989) and CROPGRO (Hoogenboom et al., 1994; Thornton et al., 1994) for a site in the wet-dry savannas of central Brazil (Bowen et al., 1993). Soil and weather data for the simulation site were provided by the Brazilian Corporation for Agricultural Research (EMBRAPA). The soil was a clayey oxidic isothermic Haplustox; historical weather data were from a site in Goiania (16°28′ S, 49°17′ W; altitude 823 m). The experiment, run as a 60-yr sequence, was replicated 10 times; simulated daily weather records for the site were obtained by using WGEN, a statistical weather generator (Richardson, 1985).

The experiment was made up of a repeating sequence of the maize model followed by the CROPGRO fallow model: in each year, maize was planted, followed by a fallow period that ran through to the beginning of the subsequent season. This rotation was repeated for 60 seasons. Crop management was the same in all seasons of the sequence; the variety used was Cargill 111, planted on 1 November each year, the beginning of the wet season. No N was applied, but all maize residue after harvesting the grain was returned to the soil.

The simulated yield distributions obtained from the 10 replicates of the sequence are summarized in Fig. 2 as percentile plots by sequence year, showing the 0th (lowest black square), 25th (lower of the joined bars), 50th (the diamond), 75th (upper of the joined bars), and 100th (topmost black square) percentile of the yield distribution for each season. A different way of presenting the yield dis-

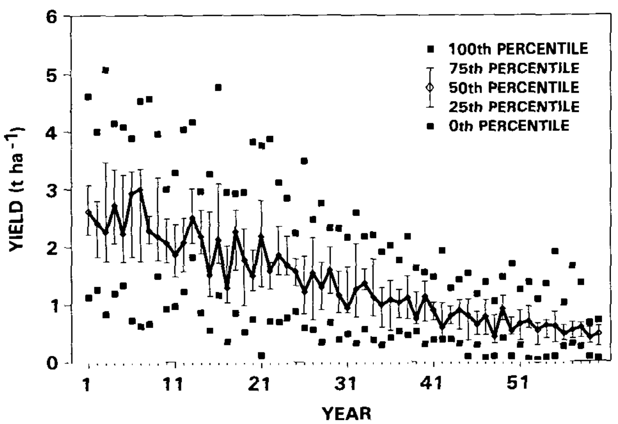

Fig. 2. Simulated yield distributions for continuous maize without N fertilizer for 60 yr at a site in central Brazil, replicated 10 times.

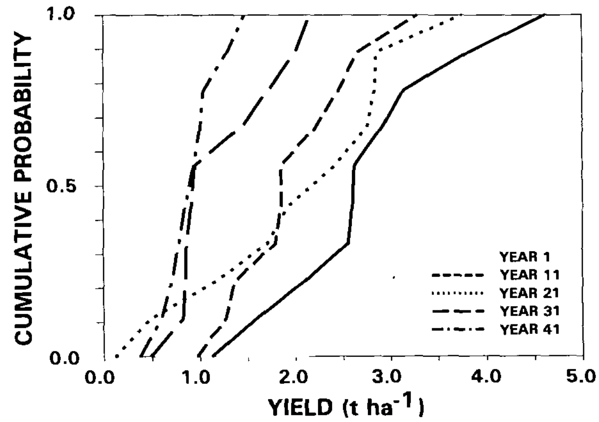

Fig. 3. Simulated yield cumulative probability functions for five selected seasons from a 60-yr sequence for a site in central Brazil, replicated 10 times.

tributions is shown in Fig. 3, where CPFs are plotted for Years 1, 11, 21, 31, and 41 of the sequence.

The software was used to fit a regression of the form

$$\ln(y) = \ln(c) + mx$$

(i.e., the exponential function $y = c \times e^{mx}$), where y is yield (kg ha^{-1}), x the year, c a constant, and m the gradient, through the means of these distributions (Fig. 4a). The decline in yield was highly significant ($\alpha < 0.01$); the value of r^2, calculated using the log-transformed data, for the regression of year on yield was 0.946. A plot of the regression residuals (Fig. 4b) suggests that the decline in maize yields with time was not exponential; yields for the first 10 yr of the sequence were comparatively stable at $\approx$2.5 t ha^{-1}, but then declined markedly.

Care is needed when interpreting the regressions produced by using the software package. A number of the assumptions of ordinary least-squares regression are routinely violated. For example, output variables are often not distributed normally; they are not independent, because outputs in one season will be dependent to some extent on outputs in previous years; and the variance of the output distributions may change during the sequence.

An economic analysis was carried out using the price-cost values shown in Table 3. The probability of the maize enterprise failing to generate a positive economic return per hectare in each season was calculated. These probabilities are plotted in Fig. 5. As might be expected, the probability of negative returns increased markedly with time, although the relative stability of maize returns during the first 10 yr of the sequence is apparent.

Sequences of Different Crops

The Brazilian experiment is a simple example of a crop sequence. More complex rotations can be analyzed, where there may be six or eight different sequence components or where more than one crop is grown during the growing season. Analyses can be carried out on any combination of the crop components (up to a maximum number of nine) that constitute the cropping sequence. If the user chooses more than one crop component of the sequence, the ap-

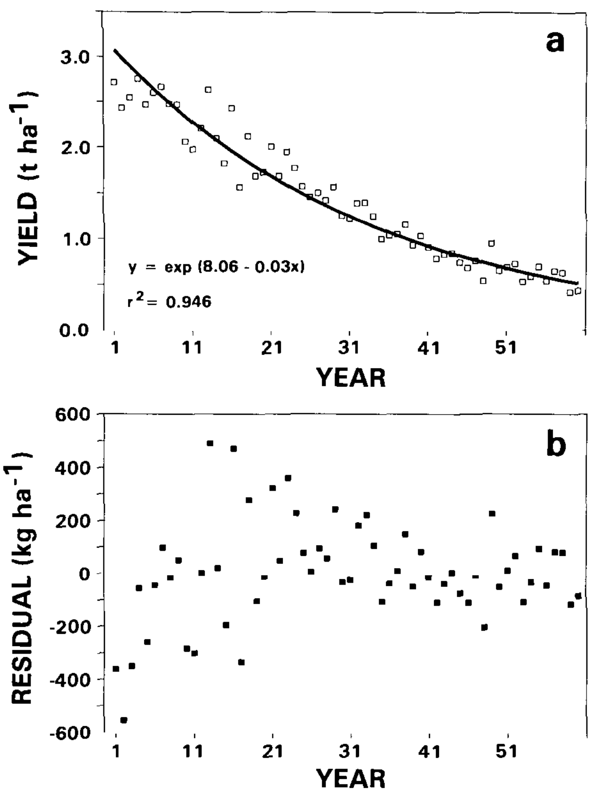

Fig. 4. Regression of sequence year number on simulated mean maize yield grown for 60 yr for a site in central Brazil: (a) regression line and mean yields; (b) regression residuals.

propriate model output data can be combined and analyzed in different ways. Assume, for example, that a replicated simulation sequence involves the growing of two successive crops within a single growing season each calendar year. There are three analysis options:

1. The output variables for each crop are simply added together to produce a single total output distribution by season or calendar year.
2. The output variables are averaged to provide a single output distribution by season or calendar yield.
3. A time series is produced. In this case, outputs are presented to the user as separate distributions, plotted or tabulated against the actual date of occurrence (thus yield distributions are plotted against date of harvesting, for example).

Each method is appropriate for a different type of output variable, but care is needed not to produce inappropriate analysis results. For example, averaging successive bean and maize yield distributions grown in the same calendar year will not normally be appropriate.

DISCUSSION

One of the major uses of crop sequence simulation experiments is to investigate productivity and profitability of a particular sequence for a number of years. The computer software described is intended to help crop model users and builders with this process, by offering users the

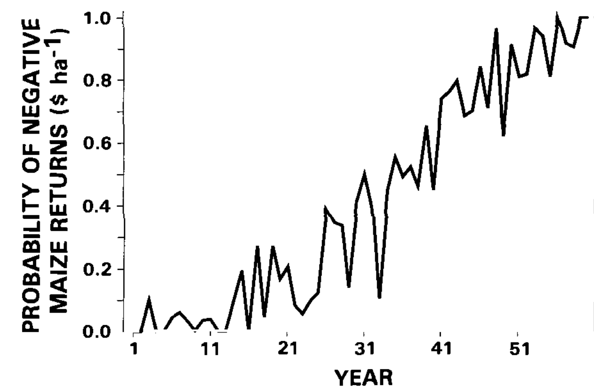

Fig. 5. The simulated probability of failing to obtain a positive economic return from the maize enterprise for 60 yr at a site in central Brazil, using the costs and prices shown in Table 3.

facility to investigate quickly the main results of multiple-season simulation experiments. Although it is difficult to know how much or how little to include in such a software package, the analyses included are those that we feel would be of value to users as a first step in analyzing the results of simulated sequence experiments.

Documentation

More detailed documentation is included with the program, together with information on input-output file structure. A text file containing this documentation is distributed with the program.

Hardware and Software Requirements

The analysis program is written in FORTRAN 77 with some Microsoft[1] extensions and compiled with the Microsoft FORTRAN compiler, version 5.1. The size of the executable file is 186K. The graphics component is written in Borland Pascal, and the size of the executable file is 176K. Approximately 550K of RAM is required to run the program. The program runs under DOS 3.1 or higher. To display the graphical results, the microcomputer must have a graphics adapter (IBM CGA, EGA, VGA, or equivalent) and a color or monochrome graphics monitor with matching screen resolution. The graphs produced on the monitor can be printed on any device that can emulate one the following printers: Hewlett-Packard LaserJet; Hewlett-Packard InkJet; Epson MX; Epson LQ; Epson FX; and Toshiba P series. The graphs can be plotted on any device that can handle Hewlett-Packard HPGL format or PostScript language format plots. Graphs can also be saved as a file on the hard disk in either Hewlett-Packard HPGL or PostScript language format for later printing, plotting, or incorporation into other graphics or word processing software. A mouse can be used with the package, but it is not required.

[1] Trade names and company names are included for the benefit of the reader and do not imply endorsement or preferential treatment of the product by the International Fertilizer Development Center or the University of Georgia.

Availability

The software, together with sample input files, is available from Dr. Gordon Tsuji, IBSNAT Project, University of Hawaii, 2500 Dole St., 22 Krauss Hall, Honolulu, HI 96822, either as part of IBSNAT's DSSAT-3 package or as a stand-alone program. Potential users are reminded that the quality of the analyses obtained using the program will depend on the integrity of the models used and the quality of their input data.

ACKNOWLEDGMENTS

The authors thank V.K. Chan and N.B. Pickering for their work on the graphics module, and J.W. Jones and G.Y. Tsuji for their input.

REFERENCES

Bowen, W.T., J.W. Jones, and P.K. Thornton. 1993. Crop simulation as a potential tool for evaluating sustainable land management. p. 15–21. *In* J.M. Kimble (ed.) Utilization of soil survey information for sustainable land use. Proc. Int. Soil Manage. Workshop, 8th. May 1993. USDA-SCS, Natl. Soil Surv. Ctr., Washington, DC.

Hoogenboom, G., J.W. Jones, and K.J. Boote. 1992. Modeling growth, development, and yield of grain legumes using SOYGRO, PNUTGRO, and BEANGRO: A review. Trans. ASAE 35:2043–2056.

Hoogenboom, G., J.W. Jones, P.W. Wilkens, W.D. Batchelor, W.T. Bowen, L.A. Hunt et al. 1994. Crop model user's guide. DSSAT Version 3.0. Vol. 2. Univ. of Hawaii Press, Honolulu (in press).

Jones, J.W. 1993. Decision support systems for agricultural development. p. 461–473. *In* F.W.T. Penning de Vries et al. (ed.) Systems approaches for agricultural development. Kluwer, Dordrecht, the Netherlands.

Jones, J.W., L.A. Hunt, G. Hoogenboom, D.C. Godwin, U. Singh, G.Y. Tsuji et al. 1994. General input and output file structures for crop simulation models. DSSAT Version 3.0. Vol. 1. Univ. of Hawaii Press, Honolulu.

Parton, W.J, B. McKeown, V. Kirchner, and D. Ojima. 1992. CENTURY users manual. Nat. Resource Ecology Lab., Colorado State Univ., Fort Collins, CO.

Richardson, C.W. 1985. Weather simulation for crop management models. Trans. ASAE 28:1602–1606.

Ritchie, J.T., U. Singh, D.C. Godwin, and L.A. Hunt. 1989. A user's guide to CERES-Maize. v. 2.10. IFDC, Muscle Shoals, AL.

Singh, U., J.T. Ritchie, and D.C. Godwin. 1993. A user's guide to CERES-Rice. v. 2.10. IFDC, Muscle Shoals, AL.

Singh, U., D.C. Godwin, J.T. Ritchie, G. Alagarswamy, S. Otter-Näcke, C.A. Jones, and J.R. Kiniry. 1988. Version 2 of the CERES models for wheat, maize, sorghum, barley and millet. p. 69. *In* Agronomy abstracts. ASA, Madison, WI.

Thornton, P.K., and G. Hoogenboom. 1994. A computer program to analyze single-season crop model outputs. Agron. J. 86:860–868.

Thornton, P.K., P.W. Wilkens, G. Hoogenboom, and J.W. Jones. 1994. Sequence analysis user's guide. DSSAT Version 3.0. Vol. 3. Univ. of Hawaii Press, Honolulu.

Williams, J.R. 1990. The erosion-productivity impact calculator (EPIC) model: A case history. Philos. Trans. R. Soc. Lond. B 329:421–428.