Agricultural Systems Modeling and Simulation

EDITED BY

ROBERT M. PEART

R. BRUCE CURRY

The University of Florida
Gainesville, Florida

MARCEL DEKKER, INC. NEW YORK · BASEL · HONG KONG

ISBN: 0-8247-0041-4

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016
http://www.dekker.com

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

2

Simulation of Biological Processes

James W. Jones and Joep C. Luyten
University of Florida, Gainesville, Florida

I. INTRODUCTION

Biological systems are made up of interacting chemical and physical processes. Living systems are composed of many subsystems and components, each having its own unique characteristics and behavior while contributing to the overall form and function of an entire system. These systems are highly complex; many components interact simultaneously, and their interactions are highly nonlinear or chaotic in nature. These interactions and nonlinearities must be taken into account when attempts are made to understand or predict system behavior. Our understanding of these interactions is incomplete and often guided only by empirical evidence of overall system behavior instead of empirical data on processes that lead to overall system behavior. Because of these complexities, the classical mathematical methods used to study nonliving physical or chemical systems have been inadequate for living systems.

Simulation based on quantitative models of biological processes and their interactions can provide considerable insight into the behavior of living systems and into ways of managing these systems to achieve specific goals. During the last several decades, computer sim-

ulation has proved to be a powerful tool in basic and applied biological sciences. The incredible growth and acceptance of personal computers during this same time period has made it possible for simulation to become an integral part of biological research in many basic and applied fields. The purpose of this chapter is to present an introduction to the concepts and techniques used in the simulation of agricultural and biological systems with a few examples that demonstrate the approach.

Because of our own interest and experience in the simulation of agricultural production systems (particularly crop systems), we focus on the simulation of crop systems in this chapter and include examples related to plants, soil, and insects. Additional material on crop simulation can be found in a number of references, including Thornley and Johnson (1990), Leffelaar (1993), Penning de Vries et al. (1989), and Goudriaan and van Laar (1994). The concepts and methodology could have been presented equally well with examples on animal, disease, or other biological systems (e.g., see Curry and Feldman, 1987; Keen and Spain, 1990; France and Thornley, 1984; Odum, 1973, and Dent and Blackie, 1979).

II. TERMINOLOGY

A. System

A system is a collection of components and their interrelationships that have been grouped together for the purpose of studying some part of the real world. The selection of the components to include in a system depends on the objectives of the study and actually represents our simplified view of reality. Systems can be described as collections of mutually interacting objects that are affected by outside forces (Pritsker, 1995; Rabbinge et al., 1994). For example, typical crop models define the crop and soil root zone as components that interact in complex ways and are also affected by weather conditions and management practices. Other models may define a leaf or a cell as the system for which models are to be developed.

One of the complicating features of biological systems is that they are hierarchically organized and can be studied at a number of levels. Figure 1 shows hierarchical levels for crop systems. At different hierarchical levels, models and simulation analysis are being performed by scientists, engineers, and economists. Note the parallel between research and models at each of these levels. The fundamental goals and objectives of a model should guide one to determine which hierarchical level to use.

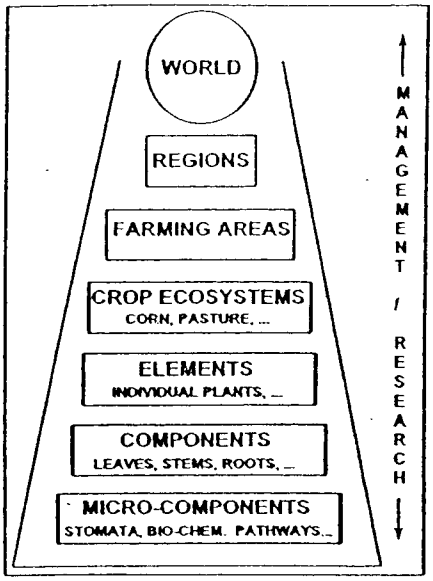


FIGURE 1 Hierarchy of models used in agriculture.

B. Environment and System Boundary

To consider the scope of a system, one must prescribe its boundaries and its contents. The environment of a system includes everything except the components of the system. A system boundary is an abstraction of the limits of the system components, separating them from the environment. The boundary may be physical; however, it is better to think of it in terms of cause and effect. The environment may affect the system in a number of ways, but the system does not affect the environment. For example, there may be a flow of mass and energy from the environment to the system that affects the system's behavior. The model of the system must take into account how these flows affect the changes in system components. In contrast, the environment is not affected by the system, and the environment does not have to be modeled. For example, in a crop model, the crop is affected by the temperature and radiation of the environment but does not itself influence these environmental conditions. Of course, crop processes do affect the microclimate and, in a very small way, other meteorological conditions, but these are usually assumed to be negligible for a number of purposes.

C. Model

A model is defined as a mathematical representation of a system, and modeling is the process of developing that representation. Conceptually, the process of developing a model of a system is different from and a prerequisite to computer simulation.

D. Computer Simulation

Computer simulation includes the processes necessary for operationalizing or solving a model to mimic real system behavior. Developing computer logic and flow diagrams, writing the computer code, and implementing the code on a computer to produce desired outputs are necessary tasks in the simulation process. In practice, modeling and computer simulation are usually interrelated. One should have in mind techniques for implementing the mathematical model as it is being conceptualized and developed. Some authors define simulation to include the modeling process (Pritsker, 1995). In this chapter, "simulation of biological systems" refers to the overall procedures outlined below, including mathematical model development and computer simulation to produce numerical model results for analysis.

E. Inputs and Outputs

Inputs to the system are factors in the environment that influence the behavior of the system but are not influenced by the system. Inputs are also referred to as exogenous variables, driving variables, or forcing functions. Inputs may vary with time, and numerical values for inputs are needed for all values of time for which the model is to be simulated by prescribed values, tables, or equations. Rainfall, temperature, and light are examples of inputs to crop systems. These environmental variables affect the soil water and crop growth dynamics in a field but are not themselves affected by the crop system. The choice of components and inputs for various models may differ depending on model objectives and the availability of data. For example, if one chooses to continuously monitor soil water conditions, then soil water could be an input to a crop model having no soil component. However, soil components are usually added because of the difficulty in monitoring soil water at any site where one wishes to apply the model. This is in contrast to the relative ease with which rainfall can be monitored and used as input to a crop model with a soil component.

Outputs from the system represent the characteristic behavior of the system that is of interest to the modeler. For example, outputs from a crop system model would include biomass of the crop and water content of the soil component.

F. Parameters and Constants

Parameters and constants are characteristics of the components of the model that are constant throughout simulated time. Usually, constants are quantities with reliable and accurate values that remain the same when experimental conditions are varied or when the model is applied to different genotypes or organisms (France and Thornley, 1984). Examples of constants are the molecular weight of glucose, the gravitational constant, and the number of seconds in a day. Parameters, on the other hand, are quantities whose values are less certain but are assumed to be the same throughout a simulation. For example, parameters may define the functional response of photosynthesis to light, the resistance to soil water flow, respiration loss for tissue synthesis, and the number of degree-days from crop emergence to flowering.

G. State Variables

State variables are quantities that describe the conditions of system components. These state variables change with time in dynamic models as system components interact with each other and with the environment. For example, soil water content and crop biomass are two state variables that change with time in most crop models. A model may have one state variable or it may have many to describe the various characteristics of a system that change with time. State variables are of critical importance because these are the dynamic characteristics of the crop or other system that are of interest to researchers.

H. Process Models

The interrelationships between components in a system, and therefore between state variables in the system, exist because of various processes. We sometimes use the term "process-oriented" to describe models that describe the flow and storage of mass, energy, or other substances. For example, the crop biomass state variable changes as a result of photosynthesis and respiration processes, and the soil water state variable changes as a result of rainfall, runoff, percolation,

and evapotranspiration processes. A crop model is the set of mathematical relationships that describe the changes in state variables as a result of the various processes that occur.

Continuous models are characterized by state variables that can change smoothly over small time intervals and are not restricted to integer values. *Discrete* models, on the other hand, are those in which the variables describing the system states take on integer values. Crop models are usually classified as continuous models. An example of a discrete biological model is one that predicts the birth and death of individual insects. The number of insects alive at any point in time is a whole number. Discrete models usually require information on the times required to complete activities, such as the time required for an insect egg to hatch. In contrast, continuous models require information on processes such as the flow rate of material or energy between components and between components and the environment.

Continuous models are usually represented by a set of differential or difference equations derived from the structure of the system and the interrelationships among components. Some systems can be modeled as either continuous or discrete, depending on the purpose of the model. For example, the fate (birth, death) of individual insects may be viewed as a series of discrete events. The population number at any time would be a count of the total number of individuals, each considered separately. Obviously, if the population is large, a lot of computer time is required to keep up with each insect. One may also view the insect population density (number of insects per unit area) as a continuous state variable for large numbers and when population processes (births, deaths) occur smoothly over time, and thus use differential equations to model their dynamics.

I. Verification

Verification involves the evaluation of the accuracy with which the computer code represents the mathematical model and the programmer's intentions. Verification also involves the careful checking of mathematical manipulations, units and their conversions, and programming logic and code to ensure that neither the mathematical model nor its translation into one or more computer programs has errors in it.

J. Calibration

Calibration consists of making adjustments to model parameters to give the best fit between simulated results and results obtained from

measurements on the real system. In other words, calibration involves adjusting certain model parameters by systematically comparing simulated results with observations of state variables. Model structure remains the same, and parameters are adjusted to more closely describe observed behavior. For example, suppose that partitioning of new crop growth to leaves varies with variety, all else being equal. An experiment could be conducted in which the total crop and leaf dry matter are measured and a leaf partitioning parameter estimated for each variety by fitting simulated results to match observed data. Calibration should be conducted only within the confines of a given data set. In many cases, calibration is the only practical way to estimate some parameter values that are used in biological models.

K. Validation

Validation is the process of comparing simulated results to real system data not previously used in any calibration or parameter estimation process. The purpose of validation is to determine if the model is sufficiently accurate for its application as defined by objectives of the simulation study. Simulated state variables are compared with measured values of state variables. Usually, in crop simulation studies only a few state variables out of many possibilities are measured, and thus a complete comparison is usually not possible. Validation involves subjective judgment. First, the areas for comparison must be selected. Then a measure of "accuracy" or "closeness of fit" must be established, such as the final crop yield. The choice of criteria and important state variables for comparison should be based on model objectives. Validation efforts are essential in the application of crop models.

III. SIMULATION APPROACH

Simulation involves the development of a model and its use in gaining insight about the system itself or its management. The simulation approach usually entails the use of a number of steps, which are summarized below. Following this summary, techniques for biological model development and implementation are presented along with an example of a simplified crop model.

A. Statement of Objectives

A clear statement of the reasons for undertaking the simulation study is essential. All remaining steps in the simulation study depend on

this initial step. The problem to be addressed and the information to be derived from the model should be explicitly identified. Although this is one of the most critical steps in a simulation project, it is often overlooked or is not clearly communicated. As a part of the statement of objectives, there should be a clear definition of the intended end product and the intended users of the models that will be developed within a project. Without such documentation, project goals may remain poorly defined and reduce the effectiveness of contributors to the project. This step is not as easy as it may seem. The objective of the simulation effort should be used to determine the level of detail needed in the model, the type of experiment that is needed, the type of data that should be collected, and every other step that follows.

Broadly speaking, there are two fundamental objectives for biological simulation models. First, scientists may wish to obtain a better understanding of the behavior of a system, the interactions that occur, and cause-and-effect relationships in a system. Thus, simulation of biological systems can help evaluate one's understanding by testing hypotheses about its behavior, using models and well-defined experiments (Boote et al., 1996; Sinclair, 1990; Monteith, 1990; Goudriaan and Monteith, 1990; Goudriaan, 1988). For example, a scientist may be interested in developing a quantitative model of crop canopy photosynthesis under varying temperature and carbon dioxide conditions to gain insight into possible crop responses to changing climatic conditions. A model would serve as the scientist's hypothesis concerning crop process and overall response, and experiments would be designed specifically for determining the adequacy or inadequacy of the hypothesis. The end result of such a study could be a model with improved capabilities for prediction, or it could lead to a new research when a model and current hypotheses are shown to be inadequate. This fundamental objective for modeling could be referred to as a scientific goal since the major expected result is an increase in knowledge.

The second fundamental objective is more problem-oriented or applied and could be referred to as an engineering goal. In this case, the major goal relates to better prediction of system behavior for use in improving control or management of a system. For example, a research team may wish to develop a model for use in a computerized irrigation system controller or to be used by water resource policy board members to help them decide how to allocate water during periods of drought. The end result of this type of project is a software product or combination hardware and software product designed for a specific application.

It can be argued that scientific models can also be applied and that there is really no distinction between these two goals. However, for problem-oriented or engineering objectives, it is highly important that the model predict the system behavior with an acceptable level of accuracy. The type of model is not important as long as it is reliable and performs as needed. In contrast, scientific models may not predict the system's behavior well at all, but the study itself could be very valuable. It could demonstrate where knowledge is inadequate and where additional research is needed. In this case, the form of the model is important; it should be structured to represent current scientific understanding of the processes and used to help design experiments and data collection methods for testing the understanding. Both scientific and engineering objectives are relevant. General arguments about which is the most appropriate approach serve no useful purpose. Applied models of biological systems evolve as understanding is improved; scientific modeling is a highly effective method for helping researchers improve their understanding of biological systems.

One can think of biological simulation efforts as being "product-oriented." The end product may be increased knowledge and an improved understanding of crop behavior, or the end product may be a tool that is designed for application to specific problems. This characterization of objectives sets the stage for all activities in a simulation project.

B. Definition of the System

The components to be included in the study and the system boundary should be identified. Inputs and outputs should be described. As additional information is gathered, it may later be necessary to redefine the system. Or, for example, lack of available data may force a change in the system definition.

C. Literature Review and Data Analysis

Prior to model development, a review of information availability is needed in order to evaluate the possibility of meeting objectives as stated. In this step, the essential features of the system are established and a conceptual model that will depend on the data available and the ability to quantify processes in the model is developed.

D. Model Development

In the model development step, diagrams may be used to represent the components in the system and to summarize their interrelation-

ships. The mathematical representation of the system should be developed, including specific functions and relationships to be used in the model. Experiments may be needed to develop functions and estimate parameters for the model. The model is translated into computer code for simulating the behavior of the real system. Computer flowcharts are usually helpful in translating the model to computer code and documenting the relationships between the model and the code. Techniques for model development are presented in Section IV.

E. Model Accuracy Evaluation

The question of whether a model is accurate must be followed by "for what objectives?" and "as compared to what?" In the final analysis, one must ascertain whether the cost of obtaining an additional unit of accuracy is greater than the benefits to be gained in terms of the study objectives. Accuracy is usually defined in terms of verification, calibration, and validation.

Dent and Blackie (1979) provide an excellent perspective on model evaluation:

> When we are confident that the behavior of the model is satisfactory, the formal validation process is over. Although a formal process of validation is always recommended, the process of gaining confidence in the model is generally a slowly emerging one over the period of model construction, through formal validation, to application of the model. During this time, assessment and modification will proceed, essentially two sets of judgments being made on the way:
>
> a. That the model is not different from the real existing system to a degree that will detract from the value of the model for the purposes for which it was designed;
> b. That if the model is accepted as being adequate then the decisions made with its assistance will not be measurably less correct than those made without the benefit of the model. Such an exact assessment is extremely difficult and cannot be other than subjective in nature. Subjective judgments in this regard are, however, not without value and this aspect of validation may often be the most relevant.
>
> It is because of the difficulties associated with these two sets of judgments that validation remains an elusive issue in the simulation procedure. Statistical tests are available to deter-

> mine whether the model behavior is different from real-system behavior to some stated level of significance. The precision of such tests should not overshadow the conceptual problems in applying them. There is no mechanical procedure that permits the conditions for acceptable validation by a single (or even a series of) statistical comparison(s) of the model-performance against some recorded or measured data from an existing system which the model represents. Any statistical tests that may be carried out and which are positive in favor of the model add to the model-builder's confidence that his creation is functioning well. The acceptable level of confidence is achieved normally by a series of assessments and subsequent modifications until such time as the model is to be applied in support of decision making.

F. Sensitivity Analysis

Sensitivity analysis involves exploring the behavior of the model for different values of parameters. This is done to determine how much a change in the value of a parameter influences the important outputs from the model. Techniques for performing sensitivity analysis are presented in Section VI.

G. Model Application

The model application may relate to management of a system. In this case, the model may be viewed as "complete," and its users may not have participated in its development. Documentation of the model and the computer code is essential. The original objective may be related to research and the use of systems modeling to assist in a more thorough understanding of the system. In this case, the use of sensitivity analysis in guiding research efforts may be the intended application.

Simulation steps should not be viewed as a strict sequential process. Rather, in most systems modeling and simulation studies, these steps are repeated in an interactive fashion as information is gained and progress is made.

IV. MODELING

Continuous system modeling is a process-oriented approach for describing the behavior of a system. Processes fall into three categories: transport (or flow), transformation, and storage (Smerage, 1977).

These processes are described by two classes of variables: *extensive* or through variables and *intensive* or across variables. Extensive variables are characterized by flow-through quantities such as mass, volume, electric charge, force, and heat flows. Intensive variables are measures of energy intensities or potentials across system components. Intensive variables represent the driving force for the extensive variables. Examples are pressure, temperature, voltage, and velocity.

Extensive and intensive variables are well-defined for many physical and chemical systems and have been widely used. In such systems, a component process is described by relationships between its extensive and intensive variables. For example, heat flow through a conductor is described by its heat conduction property, and the difference in temperature is the intensive variable. Water flow through soil is described by soil properties and the pressure gradient in the soil. The flow of charge (or current) through a resistor is described by the voltage drop across the resistor and its resistive or energy-dissipative property. Water flow through soil–plant–atmosphere systems is sometimes represented by resistances and water potential gradients. Extensive variables are measured at a point, and intensive variables are measured across an object of interest.

By proper identification of intensive and extensive variables and system components, a system diagram can be drawn that can be used to derive the mathematical model of the system. Because of this commonality of extensive and intensive variables among systems, similar methodologies can be used to model the results of analogous mathematical models from different systems. Koenig et al. (1967) and Martens and Allen (1969) describe in detail the use of these procedures for physical systems. Smerage (1977) extended these concepts to a broader class of agricultural, biological, and ecological systems.

One difficulty in modeling agricultural and biological systems using this methodology stems from the fact that we do not always know the intensive variable for a particular extensive or flow process, or there may be several mechanisms causing flow. In this case, the inclusion of the details of all intensive variables causing flow on a diagram may unduly complicate the model and detract from its purposes. In other cases, our objectives for modeling the system and the resultant scope and hierarchical detail may make it impractical to express flow processes mechanistically as related to intensive variables. For example, we know that the rate of flow of water through a plant is regulated by water potential differences between the leaf and soil and by resistances to flow along this path. However, if our

objective is to predict transpiration water use of a crop for application to irrigation management, it is neither necessary nor desirable to include these mechanisms in the model. Including these details would require model descriptions of water potentials of soil and plant and associated details of plant and soil parameters that might be useful for scientific purposes but not practical for these purposes.

A. Compartment Models

So-called compartment modeling is a useful tool in conceptualizing systems in which the primary emphasis is on the flow and storage of system variables. Compartment models have been widely used to schematically represent various agricultural, biological, and ecological systems (Dent and Blackie, 1979; Patten, 1973; Odum, 1973; Penning de Vries and van Laar, 1982; Keen and Spain, 1992; Leffelaar, 1993). This approach is particularly useful for conceptualizing continuous systems. A mathematical model in the form of a set of first-order differential equations describing the system structure can be obtained directly from the compartment model diagram. Although there are several variations in the symbols used to represent components in compartment models (Forrester, 1971; Patten, 1973), we will use the Forrester (1971) notation because of its widespread acceptance in agricultural and biological literature. Forrester (1961) originally developed the compartment diagrams to provide a pictorial representation of systems of equations for industrial dynamics models.

A *compartment* is defined to be a quantity or level of a state variable. Compartments represent the state variables in a system or transformations of state variables. Examples of these state variables are mass, volume, electric charge, and population. Many applications in agricultural and biological systems relate to mass flow among various compartments in the system. A conceptual model of this type consists of a network of compartments representing all relevant components of the system. The compartments are connected by lines that represent the flow rates of the quantities between compartments. Flow may also occur between the environment and compartments.

Figure 2 shows an abbreviated set of symbols from Forrester (1961). The level represents the state variable at a point in time and is shown by a rectangle. Sources and sinks represent the environment of the system in that flow can occur from a source in the environment into the system without affecting the environment. Likewise, flow

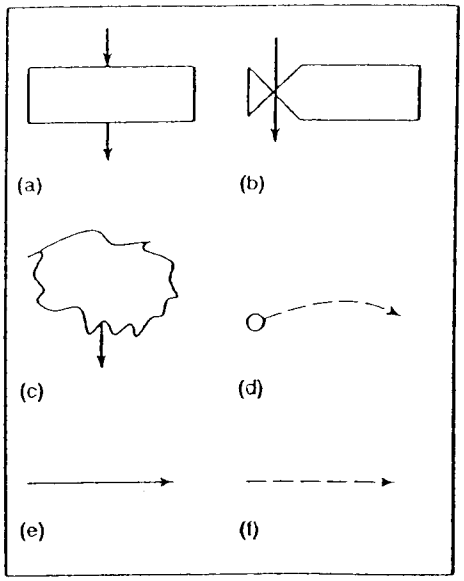


FIGURE 2 Compartment model symbols adapted from Forrester (1961). (a) Level; (b) Rate; (c) Source; (d) Auxiliary Variable; (e) Pathway for Material Flow; (f) Pathway for Information Flow.

from the system into a sink in the environment has no effect on the environment. The rate symbol represents flow from one compartment to another, from a source to a compartment, or from a compartment to a sink. Auxiliary variables are shown by circles and represent factors (inputs and parameters) that influence the rates. For example, temperature may affect the rate of biomass accumulation in a crop growth model and would be classified in this diagram as an auxiliary variable. Solid lines represent pathways for material flow whereas dashed lines represent information flow.

After the compartment model is constructed, one can write the mathematical model representing the system using structural constraints and component descriptions. The Appendix lists variables used in this chapter along with their definitions and units. From the diagram itself the structural constraint or continuity can be applied to each compartment level, or

$$\frac{dx_i}{dt} = \sum_j I_{i,j} - \sum_k O_{i,k} \tag{1}$$

where

x_i = level of ith variable
dx_i/dt = rate of change in the level of the ith variable
$I_{i,j}$ = rate of flow into level i from source j
$O_{i,k}$ = rate of flow out of level i to source k

By applying this condition to each compartment in the system, a set of first-order differential equations is derived.

B. Water Tank Example

For example, consider the simple water flow system in Fig. 3a. The volumes of water in each tank are the integrated extensive variables of interest in the system. Equations describing this system are

$$\frac{dV_1}{dt} = i(t) - f_{1,2}(t), \qquad \frac{dV_2}{dt} = f_{1,2}(t) - O(t) \tag{2}$$

Now we must describe the component flow processes. Since flow from each tank is controlled by an orifice, we know that flow rate is proportional to the square root of the height of the water surface above the orifice multiplied by 2 times the gravitational constant. Thus, we can further describe $f_{1,2}(t)$ and $O(t)$ as follows.

Component descriptions:

$$f_{1,2}(t) = C_1(2gH_1)^{1/2}, \qquad O(t) = C_2(2gH_2)^{1/2} \tag{3}$$

Since H_1 and H_2 can be expressed in terms of V_1 and V_2, we can write the complete model for this two-compartment system as (compare with Fig. 3b):

$$\frac{dV_1}{dt} = i(t) - C_1\left(\frac{2gV_1}{A_1}\right)^{1/2}, \qquad \frac{dV_2}{dt} = C_1\left(\frac{2gV_1}{A_1}\right)^{1/2} - C_2\left(\frac{2gV_2}{A_2}\right)^{1/2} \tag{4}$$

When developing the structural and component equations as in Eqs. (1)–(4), care should be taken to ensure the dimensional consistency of the relationships. The model developers must always make sure that all terms have the same dimensions and that the same measurement system is used for each parameter, input, variable, and process. Otherwise, the resulting model results will not be meaningful. Figure 3c shows an example of simulated results obtained by programming the model and simulating its behavior for 100 h, for the

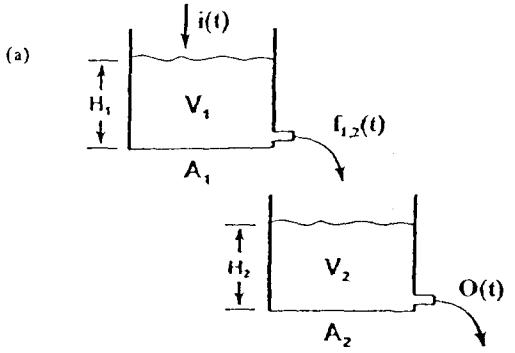


FIGURE 3 (a) Schematic of a two-tank water flow system; (b) the corresponding compartment model using Forrester (1961) notation; and (c) example results.

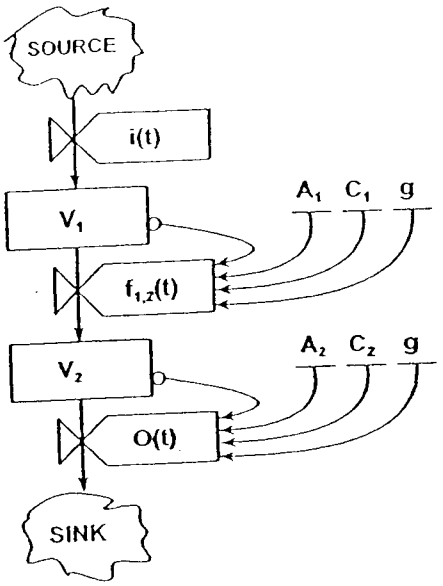
parameters shown in the figure. When the input to the first tank was set to 5 m^3/h for the first 5 h of simulation, then set to 0.0 for the remaining time, the volume of water in the first tank first increased to 16.9 m^3, then dropped to 0.0 after 19.1 h. Volume in the second tank lagged behind that in the first tank, peaking at only 7.1 m^3.

This example demonstrates the simplicity with which this approach can be used to obtain a first-order differential equation model of a system. Agricultural examples of compartment modeling are insect population dynamics and growth of crops. In some such systems, the mechanisms causing flow are not well enough understood or are so complex that research may be required to determine relationships between flow and other system characteristics and inputs.

In general, a model developed in this way will have state variables defined by levels $x_1, x_2, \ldots, x_n$. The mathematical model will be of the form

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, \ldots, x_n) + b_1(t) \\ &\vdots \\ \frac{dx_n}{dt} &= f_n(x_1, \ldots, x_n) + b_n(t) \end{aligned} \tag{5}$$

(b)



(c)

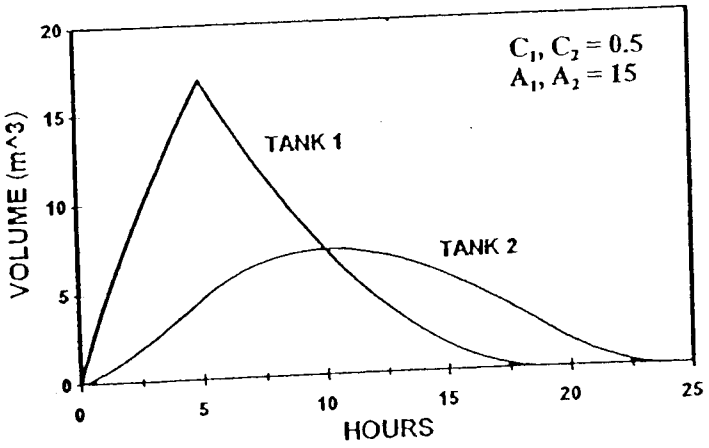


FIGURE 3 Continued

where $f_1(x_1, x_2, \ldots, x_n)$ represents the rates as affected by any of the variables $x_1, x_2, \ldots, x_n$, and $b_i(t)$ are constant or time-varying inputs to the ith storage compartment. For special cases when the relationships are linear, Eq. (5) can be abbreviated as

$$\frac{dx}{dt} = \mathbf{A}x + \mathbf{b} \tag{6}$$

where x and b are vectors and A is a matrix of coefficients.

Gold (1982) describes another useful diagrammatic approach called a signal-flow graph. In this approach, variables for the system, including state variables, flow variables, inputs, and parameters, are defined. The variables are connected by directional arrows in a diagram to denote functional dependency. For example, $A \rightarrow B \leftarrow C$ would indicate a functional dependence of A and C on B, i.e., $B = f(A, C)$. This approach is very useful during the early model formulation stages to obtain an overview of the needed mathematical relationships, such as $f(A, C)$ in the above example, and to identify important feedback loops that may occur in a system.

C. Distributed System Modeling

A distributed system is one in which the state variables and flow processes vary over the spatial region of interest. In other words, the state of the system may change through space as well as through time. The mathematical model that results from an analysis of a distributed system is a partial differential equation.

It is useful to break distributed systems into a set of connected, homogeneous elements for modeling purposes. For example, we may wish to study the flow of water and its state in a soil profile. At any point in time, soil water content and its flow may vary with depth. If we considered the entire profile to be a single compartment, it would not adequately represent the vertical variations in water content with time. Therefore, consider the soil profile to be divided into small, discrete layers, with each layer having a volume of water w_i stored in it. Each layer is connected to layers above and below by flow of water; $q_{i,j}$ is the volume flow rate of water from layer i to layer j. Figure 4 shows a schematic of the system and corresponding Forrester diagram representation. From this, we can write a series of first-order differential equations describing the structure of the system.

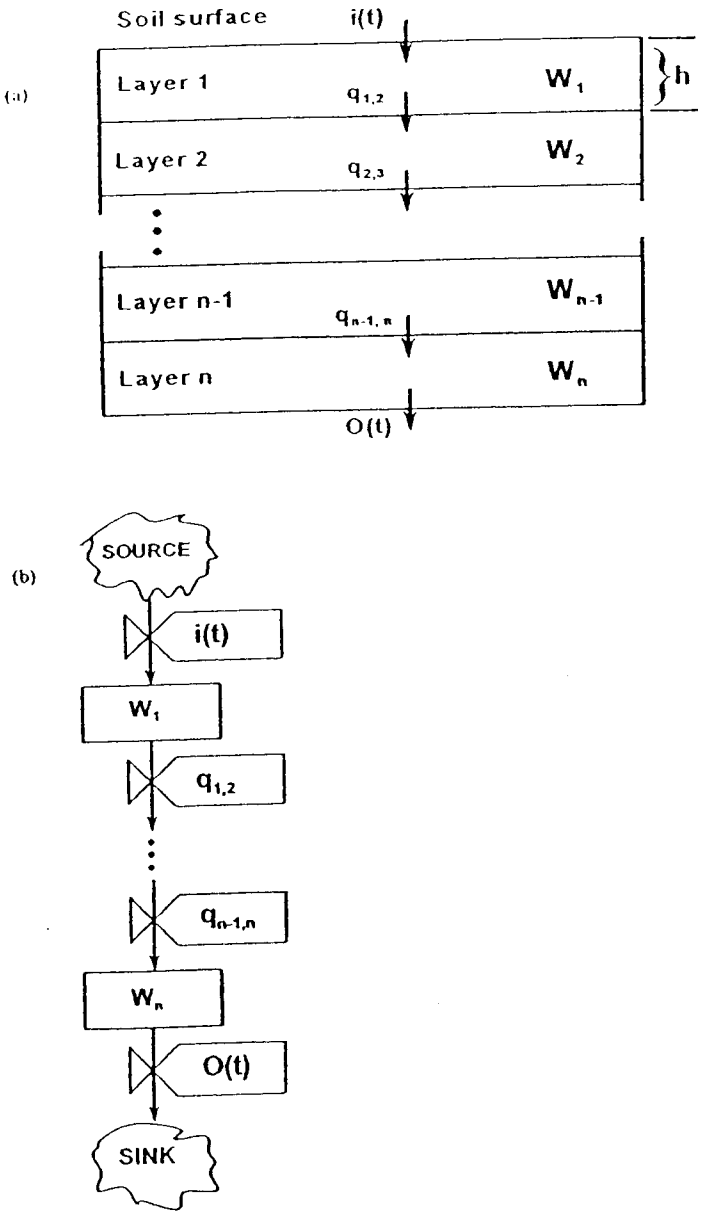


FIGURE 4 (a) Schematic of a distributed soil system (one-dimensional) for studying soil water processes and (b) the corresponding compartment model using Forrester (1961) notation.

$$\begin{aligned}
\frac{dw_1}{dt} &= i(t) - q_{1,2}(t) \\
\frac{dw_2}{dt} &= q_{1,2}(t) - q_{2,3}(t) \\
&\vdots \\
\frac{dw_n}{dt} &= q_{n-1,n}(t) - O(t)
\end{aligned} \tag{7}$$

Ignoring gravity effects, flow q can be represented by the diffusivity property of the soil, D, and the gradient of volumetric water content, θ (Hillel, 1971). Thus,

$$q_{i-1,i} = \frac{D_{i-1,i}(\theta_i - \theta_{i-1})A}{h_i} \tag{8}$$

where $\theta_i = w_i/V = w_i/Ah_i$ and h_i is the thickness of layer i.

This system of equations can be simulated if initial conditions $w_i(0)$ and boundary conditions $i(t)$, $O(t)$ are known for $i = 1, 2, \ldots, n$ and $t > 0$. Initial conditions refer to the values of all state variables at the start of the simulation, usually at time $t = 0$; they are required for simulating the behavior of a model. As an example, if a system has two state variables $x(t)$ and $y(t)$, then $x(0)$ and $y(0)$ must be specified in order to simulate system behavior for $t > 0$. Boundary conditions refer to values of variables at the boundaries of a system. Boundary conditions may refer to flows across the system boundary or to known values of intensive variables at the boundary. Values of boundary conditions must be specified for all time to be simulated.

Now we show how this set of equations is related to a partial differential equation. Let h_i and dt be approximated by Δx and Δt, respectively. Since $w_i = Ah_i\theta_i$, we can approximate dw_i/dt in difference form as

$$\frac{dw_i}{dt} \approx Ah_i \frac{d\theta_i}{dt} \approx A \frac{\Delta x(\theta_{i,t} - \theta_{i,t-\Delta t})}{\Delta t} \tag{9}$$

We can equate this expression with the right-hand side of Eq. (7) and expand it into

$$A \frac{\Delta x(\theta_{i,t} - \theta_{i,t-\Delta t})}{\Delta t} = A \left(\frac{-D_{i-1,i}(\theta_i - \theta_{i-1})}{\Delta x} - \frac{-D_{i,i+1}(\theta_{i+1} - \theta_i)}{\Delta x} \right) \tag{10}$$

Assuming D is variable, we can write

$$\frac{\theta_{i,t} - \theta_{i,t-\Delta t}}{\Delta t} = \frac{1}{\Delta x}\left(\frac{D_{i-1,i}(\theta_i - \theta_{i-1})}{\Delta x} - \frac{D_{i,i+1}(\theta_{i+1} - \theta_i)}{\Delta x}\right) \tag{11}$$

which can be written

$$\frac{\Delta\theta_i}{\Delta t} = \frac{\Delta(D_i\,\Delta\theta_i/\Delta x)}{\Delta x} \tag{12}$$

Taking the limit as $\Delta t \to 0$ and as $\Delta x \to 0$ results in

$$\frac{\partial\theta}{\partial t} = \frac{\partial}{\partial x} D \frac{\partial\theta}{\partial x} \tag{13}$$

where θ is a function of both x and t. Equation (13) is the partial differential equation for water flow in one dimension through a soil of unit area, the same equation presented by Gardner (1959). The first-order differential equation system [Eq. (7)] is an approximate "lumped" representation of the distributed (one-dimension) soil system more accurately represented by the partial differential equation, Eq. (13). When approximate solutions are adequate, the lumped representation provides a relatively quick and simple method for simulating the behavior of distributed systems.

The same procedure has been used to lump age structures in population dynamics models. Figure 5 shows the Forrester diagram for an insect model with the population divided into age classes to simulate age structure. The levels are population numbers for individuals in specific age classes. An age class may be defined as a range of ages, for example 0–10, 10–20, 20–30, and >30 represent four age classes of insects. Flow between age classes occurs as insects age. The partial differential equation describing this system is

$$\frac{\partial N(a, t)}{\partial t} + \frac{\partial N(a, t)}{\partial a} = -u(a, t)N(a, t) \tag{14}$$

where

$N(a, t)$ = population of insects age a at time t, dimensionless
$u(a, t)$ = net flux of insects age a at time t across the system boundary, s^{-1}

$u(a, t)$ represents migration and mortality factors.

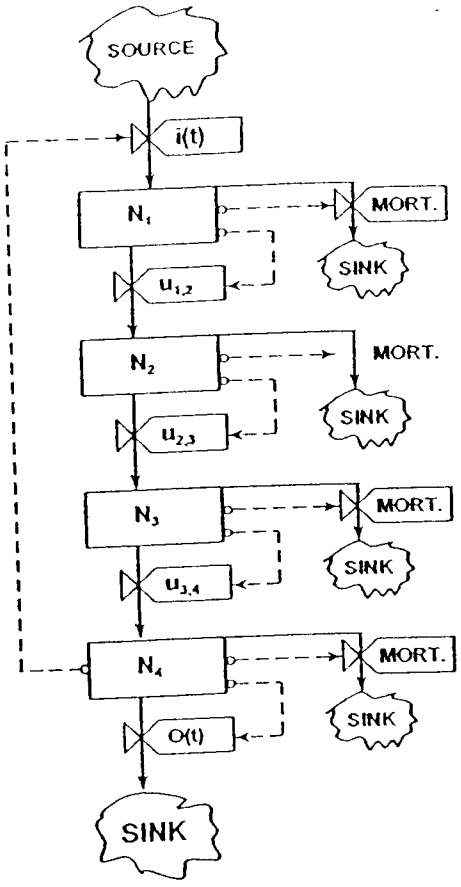


FIGURE 5 Compartment model of insect population, with insects distributed into four age classes.

V. COMPUTER SIMULATION

In continuous systems, computer simulation is performed to "solve" the set of system equations to result in a time series behavior of state variables of the system for a prescribed set of inputs. This involves one of several procedures for (1) advancing time in small steps and (2) calculating the state of the system for the new point in time. This can easily be shown for one time step using a simple model $dx/dt = ax + b$. Letting $dx \cong (x_{t+\Delta t} - x_t)$ and $dt \cong \Delta t$, we can write

$$\frac{dx}{dt} \cong \frac{x_{t+\Delta t} - x_t}{\Delta t} = ax_t + b \tag{15}$$

and

$$x_{t+\Delta t} = x_t + (ax_t + b)\Delta t \tag{16}$$

If x_t is known, we have all the information needed to calculate $x_{t+\Delta t}$ or x at time $t + \Delta t$. It is helpful to think of this in terms of a rate, expressed as

$$x_{t+\Delta t} = x_t + \text{Rate}_t\, \Delta t \tag{17}$$

where Rate_t is any expression for the rate of change of the state variable x at time t. In the example above, $\text{Rate}_t = ax_t + b$.

Equation (16) is a finite-difference expression of the differential equation. In general, Eq. (16) is also referred to as the rectangular or Euler method of integration of a first-order differential equation. The value of x is known at time t and possibly prior to time t. The model expresses the instantaneous rate of change of the variable x. Although we do not know the value of x at $t + \Delta t$, we can choose a small Δt and predict a value of x at time $t + \Delta t$ based on the current rate of change of x. This procedure is repeated as long as desired. For example, by continuing with x at $t + \Delta t$ and projecting forward by one more Δt we can estimate x for time $t + 2\Delta t$.

The result will be a time series behavior of x defined by the model. If the system has more than one first-order differential equation, the finite difference procedure can be applied to each. The rate expression for any one differential equation may be functionally dependent on any of the state variables as well as on inputs to the system. This method is dependent on the system model being described by first-order differential equations (or difference equations as described earlier). Second-order differential equations can be simulated by transformation to two first-order equations. For example, the second-order equation

$$\frac{d^2y}{dt^2} - a\frac{dy}{dt} + y = u(t) \tag{18}$$

can be transformed into two first-order equations by letting $x_1 = y$ and $x_2 = dy/dt$. It follows that $dx_2/dt = d^2y/dt^2$. Then the two equivalent first-order differential equations are

$$\frac{dx_1}{dt} = x_2 \qquad \text{and} \qquad \frac{dx_2}{dt} = ax_2 - x_1 + u(t) \tag{19}$$

Other numerical solution techniques are more accurate than the Euler method for the same size Δt. For example, the trapezoid integration method can be used to develop a "predictor-corrector" integration technique. In this case,

$$y_{t+\Delta t} = y_t + \frac{1}{2}(\text{Rate}_t + \text{Rate}^*_{t+\Delta t})\Delta t \tag{20}$$

is the formula for calculating $y_{t+\Delta t}$. Note that Rate^*_{t+t} is on the right-hand side of Eq. (20). We have to "predict" $\text{Rate}^*_{t+\Delta t}$ since we do not know y or Rate at time $t + \Delta t$. The Euler method is used to predict $y^*_{t+\Delta t}$ as a first approximation. Then that $y^*_{t+\Delta t}$ is used to compute $\text{Rate}^*_{t+\Delta t}$ using the first-order differential equation model. We then have all that is needed to compute $y_{t+\Delta t}$ above. The Euler and trapezoid methods are *explicit* methods; all terms on the right-hand side of the equation are known or can be computed from what is known.

In general, a Taylor's series expansion about y_t can be used to express $y^*_{t+\Delta t}$. This expression is

$$y_{t+\Delta t} = y_t + \Delta t\, y_t^{(1)} + \frac{(\Delta t)^2}{2!} y_t^{(2)} + \frac{(\Delta t)^3}{3!} y_t^{(3)} + \cdots + \frac{(\Delta t)^n}{n!} y_t^{(n)} \tag{21}$$

where $y_t^{(n)}$ represents the nth derivative of y with respect to t and $n!$ represents factorial numbers (i.e., $n! = 1 \times 2 \times 3 \times 4 \times 5 \times \cdots \times n$). If we approximate y_t by using the first three terms of Eq. (21), then

$$y_{t+\Delta t} = y_t + \Delta t\, y_t^{(1)} + \frac{(\Delta t)^2}{2!} y_t^{(2)} + \varepsilon_t \tag{22}$$

where ε_t is the total truncation error, or the error caused by not including all the terms in the Taylor expansion.

Since $y_t^{(2)}$ can be approximated by $(y^{(1)}_{t+\Delta t} - y_t^{(1)})/\Delta t$, Eq. (22) can be written

$$y_{t+\Delta t} = y_t + \Delta t\, y_t^{(1)} + \frac{(\Delta t)^2}{2} \frac{y^{(1)}_{t+\Delta t} - y_t^{(1)}}{\Delta t} + \varepsilon_t \tag{23}$$

Equation (23) is equivalent to the trapezoid predictor-corrector technique with the added error term. The Euler technique is equivalent to the Taylor series expansion with only the first derivative term included explicitly. Thus, the trapezoid method is a more accurate method for simulating first-order differential equations. However, the Euler method is used for most models because of its simplicity and ease of use.

Because that is a numerical approximation, it is subject to numerical errors. The local truncation error using the Euler method is $[(\Delta t)^2/2][d^2y(\xi)/dt^2]$, where ξ is some point in the interval $[t, t + \Delta t]$. The truncation error is the error with each integration step and can thus accumulate. Truncation error for the trapezoid predictor-corrector method is $[(\Delta t)^3/6][d^3y(\xi)/dt^3]$, which, for a small Δt, is less than that of Euler method (Ortega and Poole, 1981). Generally, the higher order methods are more stable than lower order methods.

VI. SENSITIVITY ANALYSIS

The purpose of sensitivity analysis is to study the behavior of the model. Sensitivity analysis can also be structured to determine important subsystems, relationships, and inputs. Sensitivity analysis should be designed with respect to the objectives of the study. A sensitivity analysis is the process by which parameters or inputs are evaluated with regard to their effects on simulated results. For example, rainfall would be an environmental input to a crop balance model for estimating crop irrigation requirements. Soil characteristics such as soil water-holding capacity or root zone depth are examples of model parameters. One may be interested in the sensitivity of estimated irrigation requirements to changes in rainfall, changes in soil characteristics, or both. A sensitivity analysis also provides a mechanism for testing the simulation in the extremes; that is, using the extreme values of parameters will rigorously test the model in terms of mathematical logic and stability.

Computer graphics are highly useful in sensitivity analysis. Graphical display of simulated results for a number of parameter values or inputs provides a visual image of model behavior over a range of parameter values. A mathematical approach to sensitivity analysis may also be a useful way to organize and present model behavior. This approach compares the change in one or more simulated outputs relative to the change in one or more parameters by approximating partial derivatives using numerical results. Absolute sensitivity, $\sigma(y|k)$, of some model output y to a variable k (a parameter of the model or input to the system) is defined by

$$\sigma(y|k) = \frac{\partial y}{\partial k} \approx \frac{y(k + \Delta k/2) - y(k - \Delta k/2)}{\Delta k} \tag{24}$$

Relative sensitivity, $\sigma_r(y|\kappa)$, is often used to provide a normalized measure for comparing the sensitivity of a model to several variables. Relative sensitivity is defined by

$$\sigma_r(y|k) = \frac{\partial y/y}{\partial k/k} = \sigma(y|k)\frac{k}{y} \tag{25}$$

For example, if $\sigma_r(y|k) = 2.0$ and k is changed by 30% ($\Delta k/k = 0.03$), we can expect y to change by 6%. Also, if $\sigma_r(y|k_1) = 0.5$ and $\sigma_r(y|k_2) = 1.5$, then y is more sensitive to a percentage change in k_2 than the same percentage change in k_1 (3 times as sensitive).

In simulation studies Δk is used to represent a small change in k so that changes in y, Δy, can be simulated. Then Eqs. (24 and (25) can be used to approximate $\sigma(y|k)$ and $\sigma_r(y|k)$ using the discrete changes in k and y, or $\partial y/\partial k \cong \Delta y/\Delta k$. By varying parameters through their expected ranges of extremes, one can use this approach to compare the sensitivity of model results.

Sensitivity analysis usually begins with the selection of model output results that are considered to be crucial to the study. A set of "base" conditions are then established; these usually comprise the set of the best estimates of each parameter and input. Base results are the simulation outputs obtained when base condition values are used. A range of values are then selected representing the extreme conditions associated with each parameter to be evaluated. Simulation runs are made using each value while holding all other base conditions constant. A comparison between changes in the base condition values and changes in the base results provides an indication of the relative importance of the variable. Results can be compared in graphs or tables or by computing the sensitivity variable values. Note that the values of σ and σ_r and graphical results depend on the "base" conditions selected.

Regression analysis or response surface techniques are also useful techniques in sensitivity analysis. Computer experiments are conducted by varying parameters over a range of interest followed by regressing simulated results against one or more parameters to determine if outputs are affected by changes in the parameter, and if so, to what extent.

VII. COMPUTER IMPLEMENTATION

There are many options for computer implementation of biological models today. Because there are many reasons for developing biological models, some consideration should be given to which option to choose for a particular model. Some biological models are developed by an individual or small team of scientists to test a hypothesis, and there is no attempt to make these models available to others except

through the scientific literature. Other models are developed by teams of researchers for use in a number of applications and for distribution to others who may want to use them for their own purposes. Biological models are also being used in formal university classrooms and in specialized short courses and training programs. Programs should be constructed using modules or objects designed for describing well-defined components of the system being modeled and for ease of documentation, revisions, replacement, and maintenance. Van Kraalingen (1995) describes a structure for programming in Fortran that provides these characteristics. In many applications, biological models need to be linked with other software, such as geographical information systems (e.g., Engel et al., 1996; Fraisse et al., 1994) and graphical user interface programs (e.g., Jacobson and Jones, 1996). Standardized formats for data storage and exchange are needed to facilitate model evaluation and to link models with such applications. Some progress has been made on standardizing data for crop models and their applications (Jones et al., 1994); this effort is continuing at an international scale (Ritchie, 1995). However, considerably more effort is needed by biological science communities on standard protocols for biological model module design and data storage and exchange.

There are many choices of programming languages and software to help those who are developing biological models. The choice of language is not a major issue today because of the ability to link together modules developed in different languages, particularly if the models use modular structure and adhere to well-defined data standards. However, each has its own advantages and disadvantages. Procedural languages such as Fortran, Basic, and Pascal are widely used for biological models. Most existing crop models, for example, are programmed in Fortran (Jones and Ritchie, 1990; Hoogenboom et al., 1994; Bouman et al., 1996). Procedural languages provide excellent capabilities for scientific computations, they can easily handle nonlinearities and other complexities common to many biological systems, and they can be used on many computer platforms with little or no modification. Because of limitations in these languages with respect to handling modern user interface programming, new applications of these models are being developed in which the model is treated as a module in an overall software package and the user interface and data manipulations are handled by one of the modern visual programming tools such as Visual Basic (Microsoft, 1995). Some biological models are now being developed using object-oriented programming languages such as SMALLTALK and C++

These new languages have potential for helping improve the modularity of biological models.

There are a number of specialized computer simulation languages that allow model developers with little or on computer programming skills to implement models. Continuous simulation languages solve systems of continuous differential equations. For example, the SLAM system (Pritsker, 1995) and FSE (van Kraalingen, 1995) packages provide modules that perform most of the required computer simulation tasks. Users only have to "program" their specific model equations and provide parameters and other inputs in order to obtain simulated behavior of a system. Probably the simplest and easiest to use simulation software allows users to build conceptual models on computer screens using icons, such as Forrester symbols presented earlier in this chapter. Two widely used packages are Stella (High Performance Systems, 1990) and VisSim (Visual Solutions, 1990), both available for use on personal computers. These packages guide users in linking together different icons to depict a system with its storage compartments and flow paths, then ask for needed parameters and other inputs to run the model. They will then simulate the system and graph or print results for users; no programming is needed. These packages have the distinct advantage of rapid evaluation of variations in model structure to test different hypotheses about the system. Each computer language has its advantages and disadvantages. The best choice for a particular model will depend on many factors such as model complexity, intended users of the model, how it will be used, the need to integrate it with other software and databases, and the need to maintain it.

VIII. DEVELOPING A CROP SYSTEM MODEL

Dynamic growth and yield models have been developed for a wide range of crops, including cotton, grain sorghum, wheat, corn, soybeans, alfalfa, and others (Jones and Ritchie, 1990). Researchers have used these models for a number of purposes, such as irrigation management (Boggess et al., 1981; Swaney et al., 1983), nutrient management (Keating et al., 1993; Singh et al. 1993), pest management (Pinnschmidt et al., 1990; Batchelor et al., 1993), land use planning (Beinroth et al., 1997), crop sequencing (Thornton et al., 1995), climate change assessment (Curry et al., 1995; Rosenzweig et al., 1995; Penning de Vries, 1993), and yield forecasting (Swaney et al., 1983). These applications have demonstrated conclusively the value of simulation applied to cropping systems. However, these models have limitations

because they do not include all factors that occur in reality and contain empiricism that may require calibration and testing for site-specific applications. Research directed to improving these shortcomings and limitations is needed to enhance the applicability of the crop models.

A useful characterization of crop models for application at the field scale was attributed to C. T. de Wit (Bouman et al., 1996; Penning de Vries and van Laar, 1982). Initially, four levels were defined, but these were later simplified to three levels (Lovenstein et al., 1993). First is the potential yield level; crop growth is dependent on weather factors, such as temperature, solar radiation, CO_2 concentration, and day length, and on genetic factors of the crop. This level of model includes basic crop growth processes such as photosynthesis, respiration, tissue growth, and development. Its main use is to gain an understanding of how these factors affect potential production of a crop, assuming that water and fertilizers are adequate and that no pest damage occurs. A second level of detail is characterized as attainable yield; at this level, the model has components that describe water and nitrogen fertilizer limitations to production. Attainable yield thus depends on water and nitrogen supplied by management as well as by rainfall and soil organic matter. Many existing crop models have the capability to describe crop growth and yield with these assumptions (Jones, 1993). The third level of detail is actual yield level, which attempts to explain all the factors that influence growth in a field, including pest damage and mismanagement. This characterization of levels of detail has been useful in providing both a practical pathway for model development and a conceptual framework for model applications. It has also served to determine data requirements for model development, taking into consideration the need for more comprehensive data sets as the level of model detail increases.

There are many choices of state variables for crop models. Simple models may have only one or two state variables, such as leaf area and dry weight, whereas others may have hundreds of state variables describing leaf area, leaf weight, number of leaves, stem weight, number of fruit, and weight of fruit for various age classes (Jones et al., 1991). Crop models with many state variables provide more details about crop behavior, including lags in fruit growth and size distributions, but require much more time to develop and test and more time to solve in a computer. For some problems, models with two to four state variables provide adequate predictions of crop responses. In this chapter, a simple crop model at the potential yield

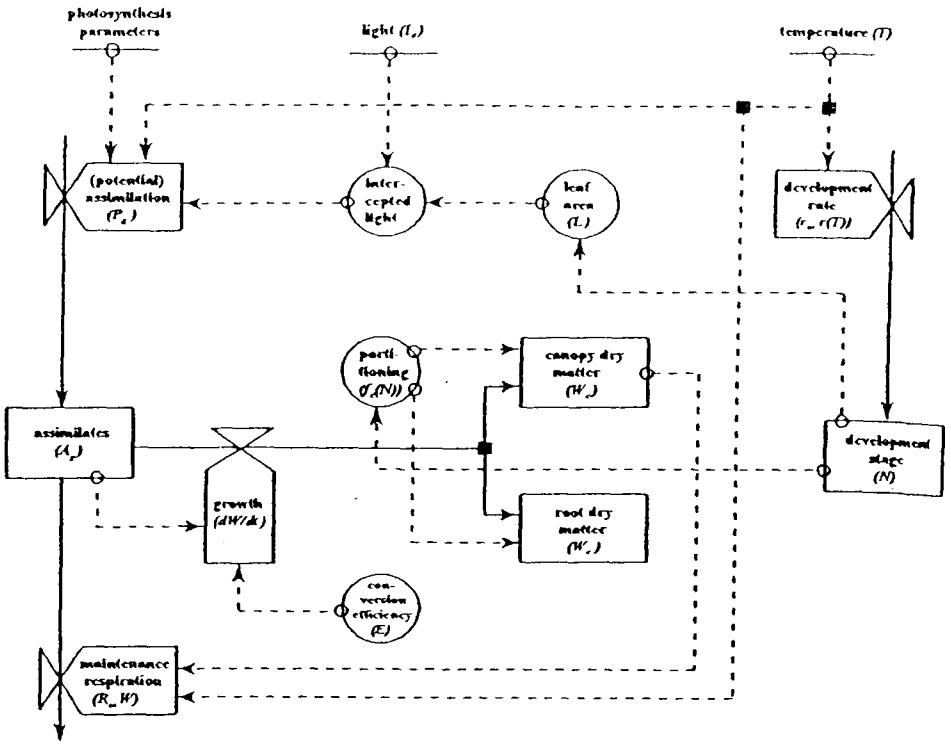


FIGURE 6 Forrester diagram for the crop model with three state variables at the potential production level of detail.

level of detail is presented along with its sensitivity to temperature and light.

Figure 6 shows a Forrester diagram of a simple crop model with three state variables: stage of development (N, number of vegetative nodes), canopy biomass (W_c, g/m^2), and root biomass (W_r, g/m^2). The diagram represents a potential yield level crop model as described by Lovenstein et al. (1993). The figure also shows a compartment for assimilates (A_p, g CH_2O/m). However, in the mathematical model, it is assumed that $dA_p/dt = 0.0$, and thus the assimilates do not accumulate but are transformed directly into plant tissue.

A. Crop Development

"Development" refers to the progression of the crop through its life stages, which may be measured by various vegetative and reproduc-

tive plant characteristics such as number of leaves on the plant or the appearance of the first flower or first fruit. Development rate is usually expressed as the inverse of time between events, such as the inverse of hours between successive leaves on the main stem or of days from plant emergence to flowering. These rates of development can thus be calculated from experiments in which stages or numbers of leaves are recorded. Development rates in plants are very sensitive to temperature and sometimes to day length but are usually insensitive to light, CO_2, and other factors such as water and nutrient stresses unless the stresses are severe. Crop development is usually insensitive to rate of dry matter production.

There have been many models for predicting crop development; most have related development rate to temperature and some, to day length. For example, Wolf et al. (1986) developed a model for scheduling field operations that predicted emergence, flowering, and harvest dates for field tomatoes within 3–5 days of actual dates. Kiker et al. (1991) described a system for predicting development stages for cucumbers, eggplant, and other crops growing in greenhouses. These models are mostly of the form $r = r(T, \eta)$, where r is the rate of development (h^{-1}), T is temperature, and η is day length. In models for predicting flowering data, r is the rate of progression toward flowering, and the state variable in these models is the development unit, which is 0.0 when the development starts and 1.0 at flowering.

For the three-state variable model, N is the state variable for leaf number for a vegetative plant or for node number for plants like tomatoes that alternately produce leaves and fruiting trusses on the main axis. Assuming that leaf appearance rate, dN/dt, depends only on temperature in this example, it can be computed as

$$\frac{dN}{dt} = r_m r(T) \tag{26}$$

where r_m is the maximum rate of leaf appearance (h^{-1}) and $r(T)$ is a function of temperature. Various functional forms have been used for $r(T)$. In the model presented here, a piecewise-linear form was used for $r(T)$, based on the tomato model (Jones et al., 1991), where $r(T)$ is 0 when T is below 8°C or above 50°C, and has the values of 0.55, 1.0, and 1.0 at temperatures of 12, 30, and 35°C, respectively. Between 30 and 35°C for this example, leaf appearance rate is maximum at r_m leaves per hour.

B. Dry Weight Growth

The rate of dry matter increase in crops is of major importance in determining yield of fruit and vegetative tissue, and it depends directly on photosynthesis. Photosynthesis converts carbon dioxide from the air into sucrose (CH_2O) in plant leaves. Some of the carbon in this CH_2O is combined with other elements (N, P, K, S, . . .) and retained in plant tissue, and some is used to provide energy for synthesizing tissue, releasing CO_2 in the process. Penning de Vries et al. (1989) describe a sound quantitative basis for computing the rates of tissue synthesis based on the photosynthetic rate and on the composition of tissue being synthesized. They computed the net product weights of protein, carbohydrate, lipid, lignin, organic acid, or mineral if 1 g of CH_2O was used to provide both the carbon in the final product and the energy for product synthesis. Thus, depending on the composition of plant tissue, a conversion efficiency, E, can be computed and used to convert photosynthesis rate into dry matter accumulation rate. For vegetative tissue, this conversion efficiency is in the range of 0.65–0.75 (g tissue)/(g CH_2O).

Respiration is the loss of CO_2 from plants in growth and maintenance processes. Growth respiration is accounted for in the conversion efficiency described above. Maintenance respiration is the loss of CO_2 due to the breakdown and resynthesis of existing tissue and depends on temperature. Gent and Enoch (1983) expressed maintenance respiration rate as

$$R_m = k_m \exp[0.0693(T - 25)] \tag{27}$$

where

R_m = maintenance respiration rate, (g CH_2O)/[(g tissue)·h]
T = temperature, °C
k_m = respiration rate at 25°C, (g CH_2O)/[(g tissue)·h]

The equation for crop dry weight growth rate is

$$\frac{dW}{dt} = E(P_g - R_m W) \tag{28}$$

where

dW/dt = rate of dry weight growth of crop, (g tissue)/(m^2·h)
W = total plant dry weight, g/m^2
R_m = maintenance respiration rate, (g CH_2O)/[(g tissue)·h]
E = conversion efficiency of CH_2O to plant tissue, (g tissue)/(g CH_2O)
P_g = canopy gross photosynthesis rate, (g CH_2O)/(m^2·h)

Total biomass growth rate is partitioned into canopy and root biomass with the function $f_c(N)$, or $dW_c/dt = (dW/dt)f_c(N)$, and $dW_r/dt = (dW/dt)[1.0 - f_c(N)]$. This partitioning of new growth between canopy and root varies with stage of development; however, in this example, we will assume that $f_c(N)$ is constant for all N.

An expression is needed to predict canopy gross photosynthesis rate, P_g, as affected by light, CO_2, temperature, and plant size. There are many models that describe leaf and canopy photosynthesis; Charles-Edwards (1981) gives the mathematics of photosynthesis processes in detail. The model of Acock et al. (1978) was found to adequately describe tomato canopy photosynthesis rates (Jones et al., 1991). This equation, modified to include temperature effects, is

$$P_g = D\frac{\tau C p(T)}{K}\ln\left[\frac{\alpha K I_0 + (1-m)\tau C}{\alpha K I_0 \exp(-KL) + (1-m)\tau C}\right] \tag{29}$$

where

D = coefficient to convert photosynthesis calculations from (μmol CO_2)/($m^2 \cdot s$) to (g CH_2O)/($m^2 \cdot h$)
τ = leaf conductance to CO_2, (μmol CO_2)/[(m^2 leaf)$\cdot$s]
C = CO_2 concentration of the air, (μmol CO_2)/(mol air)
$p(T)$ = photosynthesis reduction factor, dimensionless
α = leaf light utilization efficiency, (μmol CO_2)/(μmol photons)
K = canopy light extinction coefficient, dimensionless
I_0 = light flux density at the top of the canopy, (μmol photons)/[(m^2 ground)$\cdot$s]
m = light transmission coefficient of leaves, dimensionless
L = canopy leaf area index, (m^2 leaf)/(m^2 ground)

The function $p(T)$ expresses the effect of temperature on the maximum rate of photosynthesis for a single leaf, expressed as a quadratic equation with T:

$$p(T) = [1 - ((\phi_h - T)/(\phi_h - \phi_l))^2] \tag{30}$$

where ϕ_h is the temperature at which leaf photosynthesis is maximum and ϕ_l is the temperature below which leaf photosynthesis is zero. Hourly values of I_0 were computed using the method of Goudriaan (1986), assuming 12 h day lengths (i.e., $6 < t_h < 18$):

$$I_0 = I_m \sin\{2\pi[(t_h - 6)/24]\} \tag{31}$$

where I_m is the maximum light flux density (at noon) and t_h is the solar time in hours.

Equation (29) contains environmental variables (T, C, I_0), various parameters (τ, α, K, m, D), and the canopy leaf area, L, which depends on time. Leaf area ratio (square meters of leaf area per gram of plant) and specific leaf area, (m^2 leaf)/(g leaf), vary considerably with environmental conditions. Under cool temperatures, leaves appear more slowly [Eq. (26)], accumulate more dry weight during the extended growth period, but expand to about the same final area over a practical temperature range, as shown by Jones et al. (1991) for tomato. Under low light, leaves will tend to be thinner because of lower rates of photosynthesis per unit of development. The assumption that L is a function of N provides a model that mimics observed leaf area ratio responses to light, temperature, and CO_2. The expolinear equation (Goudriaan and Monteith, 1990) was used to fit leaf area vs. node number using data reported by Jones et al. (1991) for tomatoes grown at two CO_2 levels and three temperatures. The equation is

$$L = \rho(\delta/\beta)\ln\{1 + \exp[\beta(N - n_b)]\} \tag{32}$$

where

L = leaf area index, (m^2 leaf)/(m^2 ground)
ρ = plant density, number/m^2
N = leaf number

and δ, β, and n_b are empirical coefficients for the expolinear equation.

Therefore, the model predicts the rate of leaf appearance, which primarily depends on temperature, and calculates leaf area. The final three-state variable model can be expressed as

$$\frac{dN}{dt} = r_m r(T)$$
$$\frac{dW_c}{dt} = E(P_g - R_m W) f_c(N), \qquad \frac{dW_r}{dt} = E(P_g - R_m W)[1 - f_c(N)] \tag{33}$$

with Eqs. (27) and (29)–(32) required to compute R_m, P_g, $p(T)$, I_0, and L, respectively. These equations could be substituted into Eq. (33), but this is not necessary because computer simulation is used to solve the equations. For this study, a time step (Δt) of 1 h was used, and the model was simulated for 80 days. For each simulation, environmental inputs (C, T, and I_0) were held constant for each run, and a number of combinations were simulated to demonstrate model behavior. Table 1 shows the values for coefficients used in this model, most of which were taken from Jones et al. (1991).

Table 1 Values of Coefficients Used in the Crop Model Example

$\alpha = 0.056$	$\tau = 0.0664$	$r_m = 0.021$	$I_m = 1200$
$\beta = 0.38$	$\rho = 4.0$	$C = 350$	$K = 0.58$
$\delta = 0.074$	$k_m = 0.0006$	$D = 0.108$	
$\phi_h = 30.0$	$m = 0.10$	$E = 0.70$	
$\phi_l = 5.0$	$n_b = 13.3$	$f_c(N) = 0.85$	

There are limitations in the use of this model. It does not account for senescence or picking of plant material, nor does it partition dry matter into different plant components, such as fruit. The model could be extended to describe the growth of fruit and other organs, thereby creating more state variables, equations, and a need for more coefficients and relationships.

C. Model Behavior

Figures 7a and 7b show the effects of temperature, light, and CO_2 concentration on the rates of photosynthesis predicted by Eq. (29) and the effects of temperature on respiration for a canopy with a leaf area index (L) of 4.0. These results demonstrate the importance of weather on the rate of dry matter accumulation. Figure 7c shows the effect of temperature on leaf area expansion over time. Low temperature reduces the rate of node formation [Eq. (26)], which results in delays in leaf area development. When plants are young, rapid leaf area expansion is important to create a full canopy for capturing light for photosynthesis.

The three-state variable model, Eq. (33), was simulated for 80 days under assumed constant environmental conditions to demonstrate the integrated and cumulative effects of growth and development rates on yield. For each set of computations, all variables but one were held at reference values and one variable was changed for successive runs. Reference conditions were 30/20°C day/night temperatures, 12 h days, 1200 (μmol photons)/($m^2 \cdot s$) maximum light flux density, and 350 (μmol CO_2)/(mol air).

Figure 7d demonstrates a major effect of temperature on dry weight over the 80 day simulations. Temperature affects both development and photosynthesis, but its major effect is on development. For the 18/8°C temperature case, 59 days were required to reach node 10 in contrast at 23 days for the 30/20°C temperature case. Photosynthesis for the 18/8°C case was reduced by only about 30%. For this

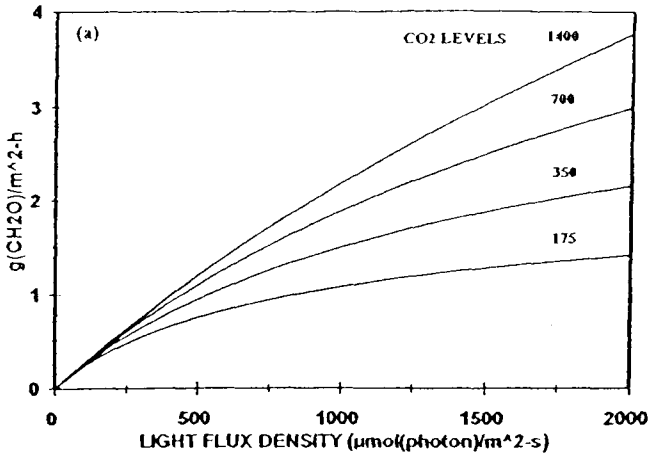


FIGURE 7 Simulated results from the three state variable crop model. (a) Photosynthesis vs. maximum light flux (I_m) at four CO_2 levels at temperature T = 30°C; (b) photosynthesis vs temperature at four maximum light flux levels for a canopy with leaf area index L = 4.0; (c) leaf area development vs time for four temperatures; (d) total dry weight vs. time for four day/night temperatures; and (e) total dry weight vs. time for four light levels at 30/20°C day/night temperatures.

coldest case, dry matter at 40 days was lower by a factor of 15. Note that growth was lower at 38/28°C than at 30/20°C day/night temperatures.

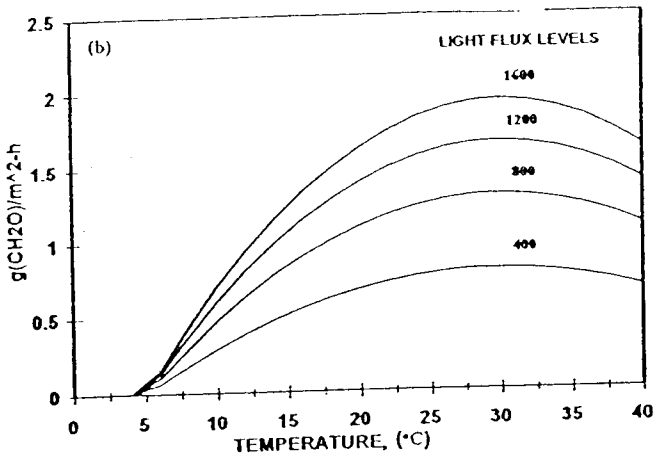
As light was varied from 400 to 1600 (μmol photons)/(m^2·s), dry weight yield increased by a factor of 2.5 at 80 days (Fig. 7e). Numbers of leaves and leaf area were not affected by light. Increasing CO_2 levels from an ambient level of 350 (μmol CO_2)/(mol air) to 700 resulted in a 28% increase in dry weight yield at 80 days (data not shown).

APPENDIX. VARIABLES USED IN THIS CHAPTER

Water Tank, Soil Layers, and Insect Population Examples

A_i area of water tank i, m^2
C_i constant, m^2 (water tank case)
D diffusivity of the soil, m/s

(*Appendix continues p. 57*)



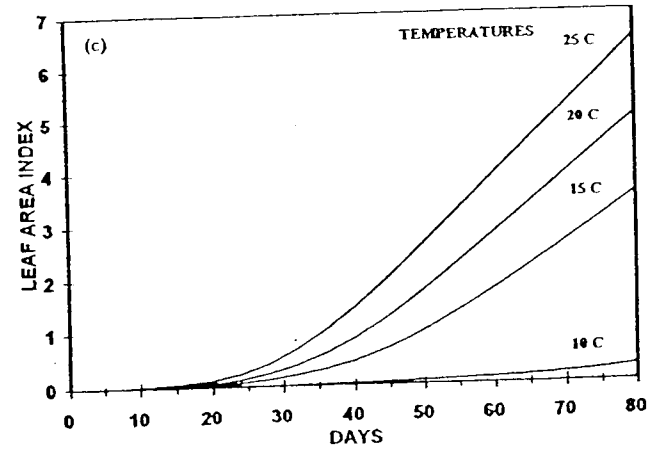
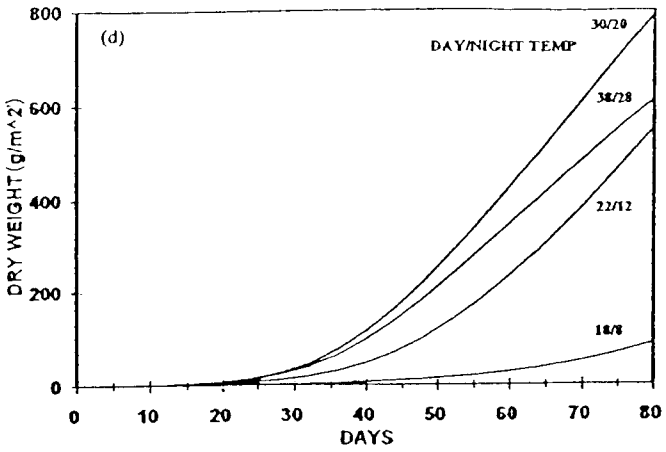


FIGURE 7 Continued



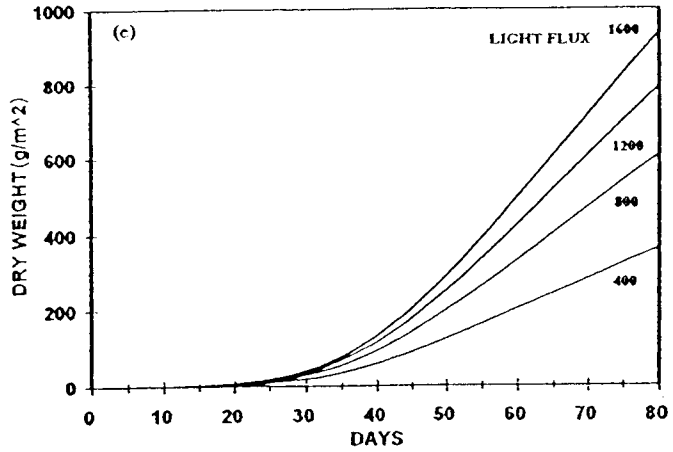


FIGURE 7 Continued

$f_{i,j}$ flow rate from tank i to tank j, m^3/s
g gravity constant, m/s^2
h_i thickness of soil layer i, m
H_i height of water in tank i, m
$I_{i,j}$ rate of flow into level i from source
$N(a, t)$ population of insects age a at time t, number
$O_{i,k}$ rate of flow out of level i to source
$q_{i,j}$ flow rate of water from layer i to layer j, m^3/s
$u(a, t)$ net flux of insects age a at time t across system boundary, s^{-1}
V_i volume in tank i, m^3
w_i volume of water stored in layer i, m^3
x_i level of ith state variable, various units
θ volumetric water content, m^3/m^3
$\sigma(y, k)$ absolute sensitivity of output y to input k, dimensionless
$\sigma_r(y, k)$ relative sensitivity of output y to input k, dimensionless

Crop Model Example

A_p assimilates, (g CH_2O)/m
C CO_2 concentration of the air, (μmol CO_2)/(mol air)
D photosynthesis conversion coefficient, (g CH_2O)/(μmol CO_2)
E conversion efficiency at CH_2O to plant tissue, (g tissue)/(g CH_2O)
$f_c(N)$ fraction of total crop growth partitioned to canopy, dimensionless
I_0 actual light flux density at top of canopy, (μmol photons)/($m^2 \cdot s$)
I_m maximum light flux density at top of canopy, (μmol photons)/($m^2 \cdot s$)
k_m respiration rate at 25°C, (g CH_2O)/[(g tissue)·h]
K canopy light extinction coefficient, dimensionless
L canopy leaf area index, (m^2 leaf)/(m^2 ground)
m light transmission coefficient of leaves, dimensionless
n_b empirical coefficient for expolinear equation, dimensionless
N leaf number or node number, dimensionless
$p(T)$ photosynthesis reduction factor, dependent on T, dimensionless
P_g canopy gross photosynthesis rate, (g CH_2O)/($m^2 \cdot h$)
r rate of development, h^{-1}
r_m maximum rate of leaf appearance, h^{-1}

R_m	maintenance respiration rate, (g CH_2O)/[(g tissue)·h]
T	temperature, °C
t_h	solar time, h
W	total plant dry matter weight, g/m^2
W_c	canopy dry matter weight, g/m^2
W_r	root dry matter weight, g/m^2
α	light utilization efficiency, (μmol CO_2)/(μmol photons)
β	empirical coefficient for expolinear equation, dimensionless
δ	empirical coefficient for expolinear equation, dimensionless
η	day length, h
ρ	plant density, m^{-2}
τ	leaf conductance to CO_2, (μmol CO_2)/[(m^2 leaf)·s]
ϕ_h	temperature at which leaf photosynthesis is maximal, °C
ϕ_l	temperature at which leaf photosynthesis is zero, °C

REFERENCES

Acock, B., D. A. Charles-Edwards, D. J. Fitter, D. W. Hand, L. J. Ludwig, J. Warren Wilson, and A. C. Withers. 1978. The contribution of leaves from different levels within a tomato crop to canopy net photosynthesis: An experimental examination of two canopy models. *J. Exp. Bot. 29*:815–827.

Batchelor, W. D., J. W. Jones, K. J. Boote, and H. Pinnschmidt. 1993. Extending the use of crop models to study pest damage. *Trans. ASAE 36*(2):551–558.

Beinroth, F. H., Jones, J. W., E. B. Knapp, P. Papajorgji, and J. C. Luyten. 1997. Application of DSSAT to the evaluation of land resources. In: *Understanding Options for Agricultural Production*. G. Y. Tsuji, G. Hoogenboom, and P. K. Thornton (Eds.). Dordrecht: Kluwer Academic. (in press).

Boggess, W. G., J. W. Jones, D. P. Swaney, and G. D. Lynne. 1981. Evaluating irrigation strategies in soybeans: A simulation approach. In: Proceedings of the ASAE irrigation Scheduling Conference, Dec. 14–15, 1981. St. Joseph, MI: ASAE, pp. 45–53.

Boote, K. J., J. W. Jones, and N. B. Pickering. 1996. Potential uses and limitations of crop models. *Agron. J. 88*(5):704–716.

Bouman, B. A. M., H. van Keulen, H. H. van Laar, and R. Rabbinge. 1996. The "School of de Wit" crop growth simulation models: A pedigree and historical overview. *Agric. Syst. 52*(2–3):171–198.

Charles-Edwards, D. A. 1981. *The Mathematics of Photosynthesis and Productivity*. New York: Academic Press.

Curry, G. L., and R. M. Feldman. 1987. *Mathematical Foundations of Population Dynamics*. TEES Monograph Series, No. 4. College Station, TX: Texas A&M Univ. Press.

Curry, R. B., J. W. Jones, K. J. Boote, R. M. Peart, L. H. Allen, Jr., and N. B. Pickering. 1995. Response of soybean to predicted climate change in the USA. In: *Climate Change and Agriculture: Analysis of Potential International*

Impacts. C. Rosenzweig, J. W. Jones, G. Y. Tsuji, and P. Hildebrand (Eds.). ASA Special Publication No. 59. Madison, WI: American Society of Agronomy, pp. 163–182.

Dent, J. B., and M. J. Blackie. 1979. *Systems Simulation in Agriculture*. London: Applied Science.

Engel, T., J. W. Jones, and G. Hoogenboom. 1996. AEGIS/WIN—A powerful tool for the visualization and comparison of crop simulation results. In: *Proceedings of the Sixth International Conference on Computers in Agriculture*. F. S. Zazueta (Ed.). Cancun, Mexico, June 1–14, 1996. ASAE Pub. 701P0396. St. Joseph, MI: ASAE, pp. 605–612.

Forrester, J. W. 1961. *Industrial Dynamics*. New York: Wiley.

Forrester, J. W. 1971. *Principles of Systems*. Cambridge, MA: Wright-Allen.

Fraisse, C. W., K. L.Campbell, J. W. Jones, W. G. Boggess, and B. Negahban. 1994. Integration of GIS and hydrologic models for nutrient management planning. Proceedings of the National Conference on Environmental Problem Solving with Geographical Information Systems. Cincinnati, OH. Sept. 21–23, 1994. EPA/625/R-95/004, pp. 283–291.

France, J., and J. H. M. Thornley. 1984. *Mathematical Models in Agriculture*. London: Butterworths.

Gardner, W. R. 1959. Solutions of the flow equation for the drying of soil and other porous media. *Soil Sci. Soc. Proc.* 23:183–187.

Gent, M. P. N., and H. Z. Enoch. 1983. Temperature dependence of vegetative growth and dark respiration: a mathematical model. *Plant Physiol. 71*: 562–567.

Gold, H. J. 1982. *Mathematical Modeling of Biological Systems: An Introductory Guidebook*. New York: Wiley.

Goudriaan, J. 1986. A simple and fast numerical method for the computation of daily totals of crop photosynthesis. *Agric. Forest Meteorol. 38*:251–255.

Goudriaan, J. 1988. The bare bones of leaf-angle distribution in radiation models for canopy photosynthesis and energy exchange. *Agric. Forest Meteorol. 43*:155–169.

Goudriaan, J., and H. H. van Laar. 1994. *Modelling Potential Crop Growth Processes: Textbook with Exercises*. Current Issues in Production Ecology, Vol. 2. Dordrecht: Kluwer.

Goudriaan, J., and J. L. Monteith. 1990. Mathematical function for crop growth based on light interception and leaf area expansion. *Ann. Bot. 66*: 695–701.

High Performance Systems. 1990. *Stella II User's Guide*. Hanover, NH: High Performance Systems.

Hillel, D. 1971. *Soil and Water Physical Principles and Processes*. New York: Academic Press.

Hoogenboom, G., J. W. Jones, P. W. Wilkens, W. D. Batchelor, W. T. Bowen, L. A. Hunt, N. B. Pickering, U. Singh, D. C. Godwin, B. Bear, K. J. Boote, J. T. Ritchie, and J. W. White. 1994. Crop models. In: *DSSAT v3*. Vol. 2. G. Y. Tsuji, G. Uehara, and S. Balas (Eds.). Honolulu: University of Hawaii, pp. 95–244.

Jacobson, B. M., and J. W. Jones, 1996. Designing a decision support system for soybean management. In: *Proceedings of the Sixth International Conference on Computers in Agriculture*. F. S. Zazueta (Ed.). Cancun, Mexico, June 1–14, 1996. ASAE Pub. 701P0396. St. Joseph, MI: ASAE, pp. 394–403.

Jones, J. W. 1993. Decision support systems for agricultural development. In: *Systems Approaches for Agricultural Development*. F. Penning de Vries, P. Teng, and K. Metselaar (Eds.). Dordrecht: Kluwer Academic, pp. 459–471.

Jones, J. W., and J. T. Ritchie. 1990. Crop growth models. In: *Management of Farm Irrigation Systems*. G. J. Hoffman, T. A. Howell, and K. H. Solomon (Eds.). St. Joseph, MI: Am Soc. Agric. Eng., pp. 61–89

Jones, J. W., E. Dayan, L. H. Allen, H. van Keulen, and H. Challa. 1991. A dynamic tomato growth and yield model (TOMGRO). *Trans. ASAE 34*: 663–672.

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. In: *DSSAT v3*. Vol. 2. G. Y. Tsuji, G. Uehara, and S. Balas (Eds.). Honolulu: University of Hawaii, pp. 1–94.

Keating, B. A., R. L. McCown, and B. M. Wafula. 1993. Adjustment of nitrogen inputs in response to a seasonal forecast in a region of high climatic risk. In: *Systems Approaches for Agricultural Development*. F. Penning de Vries, P. Teng, and K. Metselaar (Eds.). Dordrecht: Kluwer Academic, pp. 233–252.

Keen, R. E., and J. D. Spain. 1992. *Computer Simulation in Biology: A Basic Introduction*. New York, Wiley-Liss.

Kiker, G. A., J. W. Jones, and E. R. Muller. 1991. A model and database for predicting phenomenological stages. ASAE Paper No. 91-7035. St. Joseph, MI: ASAE.

Koenig, H. E., Y. Tokad, and D. R. Allen. 1967. *Analysis of Discrete Physical Systems*. New York: McGraw-Hill.

Leffelaar, P. A. (Ed.). 1993. *On Systems Analysis and Simulation of Ecological Processes, with Examples in CSMP*. Current Issues in Production Ecology. Vol. 1. Dordrecht: Kluwer Academic.

Lovenstein, H., E. A. Lantinga, R. Rabbinge, and H. van Keulen. 1993. *Principles of Theoretical Production Ecology*. Course book. Department of Theoretical Production Ecology. Wageningen, The Netherlands: Wageningen Agricultural University.

Martens, H. R., and D. R. Allen. 1969. *Introduction to Systems Theory*. Columbus, OH: Merrill.

Microsoft. 1995. *Microsoft Visual Basic, Programmer's Guide*. Redmond, WA: Microsoft Corporation.

Monteith, J. L. 1990. Conservative behaviour in response of crops to water and light. In: *Theoretical Production Ecology: Reflections and Prospects*. R. Rabbinge, J. Goudriaan, H. van Keulen, F. W. T. Penning de Vries and H. H. van Laar (Eds.). Simulation Monographs 34. Wageningen, The Netherlands: PUDOC, pp. 3–16.

Odum, H. T. 1973. An energy circuit language for ecological and social systems. In *Systems Analysis and Simulation in Ecology*. Vol. II. B. C. Patten (Ed.). New York: Academic Press, pp. 140–211.

Ortega, J. M., and W. G. Poole, Jr. 1981. *An Introduction to Numerical Methods for Differential Equations*. Marshfield, MA: Pitman.

Patten, B. C. 1973. *Systems Analysis and Simulation in Ecology*. Vol. II. New York: Academic Press.

Penning de Vries, F. W. T. 1993. Rice production and climate change. In: *Systems Approaches for Agricultural Development*. F. Penning de Vries, P. Teng, and K. Metselaar (Eds.). Dordrecht, The Netherlands: Kluwer Academic, pp. 175–189.

Penning de Vries, F. W. T., and H. H. van Laar (Eds.). 1982. *Simulation of Plant Growth and Crop Production*. Wageningen, The Netherlands: PUDOC.

Penning de Vries, F. W. T., D. M. Jansen, H. F. M. ten Berge, and A. Bakema. 1989. *Simulation of Ecophysiological Processes of Growth in Several Annual Crops*. Simulation Monographs 29. Wageningen, The Netherlands: PUDOC.

Pinnschmidt, H., P. S. Teng, and J. E. Yuen. 1990. Pest effects on crop growth and yield. In *Proceedings of the Workshop on Modeling Pest–Crop Interactions*. P. Teng and H. Yuen (Eds.). Honolulu: University of Hawaii.

Pritsker, A. A. B. 1995. *Introduction to Simulation and SLAM II*. 4th ed. New York: Wiley.

Rabbinge, R., P. A. Leffelaar, and H. C. van Latesteijn. 1994. The role of systems analysis as an instrument in policy making and resource management. In *Opportunities, Use, and Transfer of Systems Research Methods in Agriculture to Developing Countries*. P. R. Goldsworthy and F. W. T. Penning de Vries (Eds.). Dordrecht, The Netherlands: Kluwer Academic, pp. 67–79.

Ritchie, J. T. 1995. International Consortium for Agricultural Systems Applications (ICASA): Establishment and purpose. *Agric. Syst.* 49:329–335.

Rosenzweig, C., J. W. Jones, G. Y. Tsuji, and P. Hildebrand (Eds.). *Climate Change and Agriculture: Analysis of Potential International Impacts*. ASA Spec. Publ. No. 59. Madison, WI: American Society of Agronomy.

Sinclair, T. R. 1990. Nitrogen influence on the physiology of crop yield. In: *Theoretical Production Ecology: Reflections and Prospects*. R. Rabbinge, J. Goudriaan, H. van Keulen, F. W. T. Penning de Vries, and H. H. van Laar (Eds.). Simulation Monographs 34. Wageningen, The Netherlands: PUDOC, pp. 41–55.

Singh, U., P. K. Thornton, A. R. Saka, and J. B. Dent. 1993. Maize modeling in Malawi: A tool for soil fertility research and development. In: *Systems Approaches for Agricultural Development*. F. Penning de Vries, P. Teng, and K. Metselaar (Eds.). Dordrecht, The Netherlands: Kluwer Academic, pp. 253–273.

Smerage, G. H. 1977. On the extension of complementary variables modeling theory to biology and ecology. In: *Problem Analysis in Science and Engineering*. F. H. Branin, Jr. and K. Huseyin (Eds.). New York: Academic Press, pp. 427–462.

Swaney, D. P., J. W. Jones, W. G. Boggess, G. G. Wilkerson, and J. W. Mishoe, 1983. A crop simulation model for evaluation of within season irrigation decisions. *Trans. ASAE 26*(2):362–568.

Thornley, J. H. M., and I. R. Johnson. 1990. *Plant and Crop Modeling*. New York: Oxford Univ. Press.

Thornton, P. K., G. Hoogenboom, P. W. Wilkens, and W. T. Bowen. 1995. A computer program to analyze multi-season crop model outputs. *Agron. J. 87*:131–136.

van Kraalingen, D. W. G. 1995. *The FSE System for Crop Simulation, Version* 2.1. *Quantitative Approaches in Systems Analysis*, No. 1. July 1995. Wageningen, The Netherlands: DLO Research Institute for Agrobiology and Soil Fertility and C. T. de Wit Graduate School for Production Ecology.

Visual Solutions. 1990. *VisSim. User's Manual*. Westford, MA: Visual Solutions Inc.

Wolf, S., J. Rudich, A. Marani, and Y. Rekah. 1986. Predicting harvesting date of processing tomatoes by a simulation model. *J. Am. Soc. Hort. Sci. 111*(1) 11–16.