PRINCIPLES OF CROP MODELLING AND SIMULATION:
II. THE IMPLICATIONS OF THE OBJECTIVE IN
MODEL DEVELOPMENT

D. DOURADO-NETO\textsuperscript{1,6}; D.A. TERUEL\textsuperscript{1,7}; K. REICHAERTD\textsuperscript{2,5,6}; D.R. NIELSEN\textsuperscript{3}; J. A. FRIZZONE\textsuperscript{4,6}; O.O.S. BACCHI\textsuperscript{5,6}

\textsuperscript{1}Depto. de Agricultura-ESALQ/USP, C.P. 9, CEP: 13418-900 - Piracicaba, SP, Brasil.
\textsuperscript{2}Depto. de Física e Meteorologia-ESALQ/USP, C.P. 9, CEP: 13418-900 - Piracicaba, SP, Brasil.
\textsuperscript{3}Dept. of Land, Air and Water Resources, University of California, Davis, USA.
\textsuperscript{4}Depto. de Engenharia Rural-ESALQ/USP, C.P. 9, CEP: 13418-900 - Piracicaba, SP, Brasil.
\textsuperscript{5}Centro de Energia Nuclear na Agricultura/USP, C.P. 96, CEP: 13400-970 - Piracicaba, SP, Brasil.
\textsuperscript{6}Bolsista do CNPq
\textsuperscript{7}Bolsista da FAPESP

ABSTRACT: With the purpose of presenting to scientists the implications of the objective in model development and a basic vision of modeling, with its potential applications and limitations in agriculture, an integration of crop modeling professionals with agricultural professionals is suggested. Models mean modernization of the information, of the measurement process and of an efficient way to learn more about complex systems. They are one of the best mechanisms of transforming information in useful knowledge and of transferring this knowledge to others. One of the problems that impede a larger progress in modeling is the lack of communication between modelers and a frequent appearance of modelers without a global vision of reality.

Key Words: crop modeling, simulation

INTRODUCTION

No industry is built before a model is conceived where the controls of all its production processes are known. The Chemical Engineer needs a model for the chemical processes to plan the plant of a chemical industry. However, the Agricultural Engineer, whose function is to develop and to execute systems of environmental control for the largest and most important chemical industry of the world, which involves the transformation of luminous energy and minerals in essential and other fundamental raw materials for human existence, in most cases does not have an adapted model of the system with which it works.

The objective of this paper is to present to scientists (Agronomists mainly) directly involved with crop production, a basic vision of modeling, with its potential applications and limitations, in order that a larger integration is made possible with the professionals working with modeling of plant and agricultural processes. New problems for some can...
be old and soluble for others. Solutions can be transferred, if there is integration among professionals of the agronomic area.

THE IMPLICATIONS OF THE MODELING OBJECTIVE IN MODEL DEVELOPMENT

The level of details that the modeler includes in any mathematical description of plant growth depends on his purposes. The modeler intends to target the objectives with the generated model. Rarely the model in itself is an objective, therefore the modeler should have his objectives very well defined at the very beginning of his modeling exercise.

Let us consider a case where, for a global economic model, the modeler needed to foresee the world production of a crop in any year. It would probably be enough to use as input data in the model the annual average of world production of the crop in the 3 preceding years. In the case of a model with the objective of foreseeing the annual production of any country, probably a few more data would be necessary, like some estimates of the climatic conditions for that year and some knowledge of the agricultural politics of the country in question.

However, if the objective is to develop a growth model and crop development to help in the daily crop management in a specific place, it would be necessary to use a model sensitive to environmental fluctuations of short periods, with a number of variables and data of very large input. For this purpose, it would be necessary to develop a dynamic model of crop growth and development, that takes into account the effects of the management strategies and meteorological conditions on the crop performance, as it is in the case of the model SOYGRO (Jones et al., 1988), and to make local adjustments (Lasinski, 1993; Klosowski, 1994). This has to be done, of course, after (i) the objectives of the modeling were established, (ii) the type of needed model was determined, and (iii) the appropriate level of resolution at which model was defined.

LEVELS OF COMPLEXITY AND ACCURACY OF THE MODEL

Experience has shown that the model should be sufficiently simple to allow its manipulation and understanding, and sufficiently complex to allow the extrapolation of conclusions. It is interesting to mention the letter of Glyn Rimmington to Karin Wisiol in 1986 (Wisiol & Hesketh, 1987): “I believe that there will be a different model for each problem that the modeler is trying to solve. This vision is contrary to the notion that the modeler can not, without limits, increase the amount of information and complexity in a model. A case in subject is the structure of many commercial packages of aid to the crop management that break the problem of the handling in small parts - irrigation calendar, fertilizers and pesticides application, sowing date, and cultivar selection for example. In each case, only the necessary information to solve the immediate problem is included in each sub-model, maintaining the programs of the small and fast model”.

Another important thing is that a model should be built, in such a way that its segments can be removed and changed by better relationships among processes. For example, the dynamic procedure developed by Teruel (1995) to forecast the temporal variation of sugar-cane leaf area, when introduced in Barbieri’s (1993) model for estimating the potential productivity, substituting one of its segments, increased its accuracy.

The human mind is in general unable to store in memory more than seven facts of short duration (Simon, 1977). Therefore, the complexity of the model that the modeler builds, if he wants to understand its operation, is quite restricted. However, it is known that the accuracy of our simulation increases as the modeler increases the degree of complexity of the model. This solution is used in many growth and plant development models, that is the compartmentalization of the model is made in several hierarchical structures (Passioura, 1979; Penning de Vries, 1982; Thornley, 1980; de Wit, 1982), with each hierarchical structure being limited to seven state variables at the most. The number of hierarchical levels or sub-models contained in the general model will determine its degree of accuracy.

However, the modeler cannot hope to simply obtain better simulations and forecasts by increasing the number of hierarchical levels of a model and turning it more complex. The cost associated with the search for models is not negligible. The number of hierarchical levels that can be included in practice in a model, and the resolution that can be obtained, are subject to restrictions related to the characteristic times of response of the different levels (Ferrari, 1978; Stroosnijder, 1982).

The response time can be defined as the time that the system spends to return to a balance state after some disturbance. For example, the photosynthesis rate is dependent on the luminous...
intensity, but when it increases the incident intensity on a leaf increases, several minutes will pass until the liquid rate of photosynthesis reaches a new steady value.

When considering a biological community, the modeler can have response time of days, as in the case of verifying some effect of the water deficiency on the final yield, although at level of biochemical reactions the modeler has response time less than one second.

Summarizing, it can be said that the physical limitations of our computers, their operation speed (to take into account each biochemical reaction, the software would take hours or days to produce a result in terms of observed growth) and the level of organization of the system that the modeler wants to model are the largest restrictions for the detail level and accuracy of a viable model. This necessarily imposes limits on the accuracy of the forecast that the modeler can generate.

**TECHNIQUES OF MODELING**

Regression analysis is one of the most common statistical techniques used to adjust mathematical models or a curve to the experimental data. Being the simplest technique of modeling, it has been used since the beginning of the century. Its first use was in the estimate of agricultural yield by Sir Ronald Fisher, in 1924, to study the influence of rainfall on the wheat yield cultivated at the Experimental Station of Rothamsted, England (Pereira, 1987).

Different procedures of regression analysis can be used to adjust linear or nonlinear models. For many analytic purposes and for forecast, the regression analysis has provided a simple mathematical description of growth and plant development.

In agronomic sciences, four main aspects of the crop behavior can be modeled using regression methods: (i) production of total dry matter or of part of its cycle; (ii) numbers of several components of crop yield, as number of grains (or seeds) and number of beans, for example; (iii) crop phenological development; (iv) use of available resources for a crop and corresponding answers.

Adjusted models for regression have been used to foresee the effects on crop yield of management strategies, as application of fertilizers and variation of planting/sowing densities.

While at the beginning the use of regression analysis to adjust a linear function to a series of experimental data was a simple and direct exercise, the adjustment to nonlinear functions is a more difficult task, and should utilize an interactive method to minimize the sum of squares of the deviations (Dourado-Neto & Teruel, 1996).

Hunt (1978, 1981) revised the techniques and applications of regression type analysis. The resulting equations of a process of regression analysis are entirely empirical and their parameters do not have any physiologic meaning. For example, an equation of the following type can be obtained by this technique:

\[
Y = a + \sum_{i=1}^{n} b_i x_i + \sum_{j=1}^{m} c_j T_j + e
\]

where \( Y \) refers to the estimated crop yield; \( a, b \) and \( c \) to the regression coefficients; \( x_i \) to the climatic variables; \( n \) to the number of climatic variables; \( T_j \) to the technological variables; and \( m \) to the number of technological variables and \( e \) to the estimated error.

As climatic variable, monthly mean temperature of the crop cycle, rainfall and solar radiation can be used, as technological variables, amounts of fertilizers and soil type.

A model of this type, when developed from a long series of data, can be sufficiently exact to be used for a given condition. However, it only can be extrapolated with difficulty for different conditions from those it was developed.

It is interesting to point out that one of the oldest applications of regression analysis is still used today to foresee the phenological development of several crops, the concept of thermal sum or degree-day (Réaumur, 1735, mentioned by Wang, 1960).

Finally, the idea that should stay is that although not being able to extrapolate without adjustment for different conditions from those of its origin, a model obtained by regression analysis, when based on the understanding of processes of the development of a crop, is more worthy than a complex mechanistic model based solely on the knowledge of its answers (Rimmington & Charles-Edwards, 1987).

**EQUATIONS AND INDEXES OF GROWTH**

The concept of growth index comes from the idea of climatic index, where the relative development of some attributes of a plant, or crop or ecosystem is expressed as a function of one or more climatic variables. A lot of the agricultural zoning made in the past in Brazil was based on growth indices or climatic indices, as humidity index or index of aridity of the area (Ometto, 1981).
where GI is the growth index (the fraction of the maximum productivity that can be achieved at a given locality); LI the light index; TI the temperature index; and MI to the air relative humidity index.

Basically, LI, TI and MI represent, respectively, relationships among the observed solar radiation, the levels of registered temperatures and the rainfall/evapotranspiration conditions during the cycle of the crop, and those that would provide the maximum yield.

The model used by FAO (Doorembos & Kassam, 1979) for the calculation of the real yield is:

\[
GI = LI . TI . MI
\] (2)

MODELS OF DYNAMIC SIMULATION OF THE VEGETABLE GROWTH

In the development of models of dynamic simulation of the vegetable growth, the modeler assumed that the state of the plant at any moment in time can be quantitatively described by numeric values of a discreet number of defined state variables (Rimmington & Charles-Edwards, 1987).

These variables characterize the state of the plant growth system. They can be in particular chosen among properties of the components of a plant, as masses of the leaves or stem, leaf area, or number of leaves. The models assume that temporary variations in the values of these variables can be described mathematically. Simon (1961) describes the problem faced by the modeler elegantly: the modeler “lifted a problem giving the description of state of the solution. The task is to discover a sequence of processes that will produce the state-goal starting from an initial state. The conversion of the description of the processes to the description of the state facilitates us to recognize when the modeler had success.”

A dynamic simulator of the sugar-cane growth can be found in Pereira & Machado (1986). To facilitate the understanding, it belongs to a dynamic simulator (SIMCANA), and is briefly described:

\[
MSC_t = MSC_{t-1} + TCC_t \Delta t
\] (4)

where \( MSC_t \) refers to the crop dry mass at time \( t \) (g.m-2); \( MSC_{t-1} \) at time \( t-1 \) (g.m-2); \( TCC_t \) to the crop growth rate at time \( t \) (g.m-2.day-1); \( \Delta t \) to the considered time interval (days).

Mathematically, \( TCC_t \) is given by:

\[
TCC_t = EC(FC_t - M . MSC_t)
\] (5)

where \( EC \) is the conversion efficiency of photosynthate materials in mass, \( FC_t \) the crop gross photosynthesis rate at time \( t \) (g.m-2.day-1), M the maintenance coefficient of the crop physiological processes (day-1), and

\[
FC_t = F_{\text{MAX}} . LAI_t . FL . F
\] (6)

where \( F_{\text{MAX}} \) is the maximum photosynthesis rate for ideal conditions (g.m-2.day-1); \( LAI_t \) the leaf area index (m2.m-2) at that day, \( FL \) the ratio between day and night lengths; F the factor of adjustment of \( F_{\text{MAX}} \) for the environmental conditions, each term being explained by a new mathematical equation, that for a given time will give to the modeler an explanation for a new mathematical equation, until the modeler obtains in measurable variables, as air temperature and air relative humidity.

MODELS APPLIED TO CROPS

The development of applied crop models already started at least 65 years ago (Joyce & Kickert, 1987). In the last years, several review have been published describing the historical development of the modeling of crops.

Peart & Barret (1976) present a review of the pioneer work in modeling of crops, and they give detailed examples of the efforts of development of
computerized simulation seeking the handling of agricultural ecosystems.

Sakamoto & LeDuc (1981) present a panorama of the historical development of crop meteorological parameter models. Getz & Gutierrez (1982) describe a few historical papers on applied modeling the plant physiology and its impact on the ecological theory, biology of populations and on the development of new techniques of resource management and agroecosystem modeling.

**TYPES OF APPLIED MODELS**

**Empirical models:** In the case of applied models, an empirical model is typically a simplified mathematical model of the system, containing few variables. Most of the models of crop forecast as a function of climate, at national governmental level, are empirical. Although such models are useful analytic tools, they suffer for the lack of realism and generality (ability to be applied in conditions different from those for which they were built).

**Models of processes (dynamic simulators):** The processes are generally modeled trying to adjust the reply of the crop to the following factors: (i) meteorological, mainly light availability, heat and water; (ii) edaphic, mainly availability of nutrients and water, presence of toxic elements and physical characteristics of the soil; (iii) biological, mainly insects, diseases and competition with other plants.

**Models applied to soybean crop:** The models of dynamic simulation developed for the soybean crop, are those of largest international projection: (i) GLYCIM (Acock et al., 1985); (ii) SOYMOD (Meyer et al., 1979; 1981); (iii) SOYGRO (Jones et al., 1988). These models are described here with respect to their input parameters, processes used and output parameters.

**Glycim:** (i) output parameters (parameters simulated by the model): (i.1) dry mass of different parts of the plant, (i.2) number of beans and seeds, (i.3) leaf area, (i.4) phenological stages and plant height; (i.5) processes (biological concepts) used: flow of carbon, nitrogen and other mineral substances in the plant, and flow of water, heat, nitrate and oxygen in the soil; (ii) input parameters (factors affecting the growth): (i.i) solar radiation, (i.ii) minimum and maximum daily air temperatures, (ii.3) pluvial precipitation, (ii.4) wind speed, (ii.5) relative air humidity, (ii.6) carbon dioxide concentration of the air, (ii.7) attributes of the soil and density of plants.

**Soymod:** (i) output parameters: (i.1) yield of seeds, (i.2) mass of 100 seeds, (i.3) evapotranspiration, (i.4) plant height, (i.5) production of total dry matter; (i.6) processes used: changes of the dry matter of different organs of the plant given by the flow of carbohydrates, starch and proteins under internal control for the balance between carbon and nitrogen; (ii) input parameters: (ii.1) solar radiation, (ii.2) minimum and maximum air temperatures, (ii.3) rainfall, (ii.4) wind, (ii.5) soil type and (ii.6) sowing density.

**Soygro:** (i) output parameters: (i.1) yield of grain (or seed), (i.2) mass produced by unit of land area for: leaves, roots, stems and beans; (i.3) basic biological concepts: photosynthate production and partition for the plant depending on the phenological stage, accumulation and remobilization of proteins for the growth of the beans and seeds, and effects of defoliation and water stress; (ii) factors affecting the growth: (ii.1) photosynthetically active radiation, (ii.2) average hourly temperature, and (ii.3) soil water content in the root zone.

**The ideal applied model:** The advances in technology continually increase the possibilities of manipulating the atmosphere (Holt, 1985). These advances trigger new subjects for the researcher as well as for the administrator of resources. Reynolds & Acock (1985) verified that the existent models have been used to obtain answers on subjects for which they were not developed, like seeking answers to the cause of variations in atmospheric carbon dioxide concentrations for example.

With the expansion and larger accessibility of personal computers, some ideas were developed on an ideal model for use in management of resources, which are listed below (Joyce & Kickert, 1987): (i) to include the main contemporary biological concepts on the processes of plant growth which are sensitive management; (ii) to be capable to work at different scales, in time of any interval of time, using any space scale; (iii) to present possible mathematical suitable for portable or at least for microcomputers; (iv) to request a minimum number of input data and to execute the calculations in a minimum time; (v) to present an excellent representation of the plant behavior in the field, the model must be validated by experiments to be well documented, including a chart flow of the
embraced biological system, a logical chart flow of the computerized software and the programming code; (vii) to be at low price.

In a final analysis, the ideal model in this context should produce a result of forecast, faced to a specific management action. These approaches are important to define the measurements, and for more tools to become available and accessible to the professional working within management.

**FINAL CONSIDERATIONS**

The model means the modernization of the information, the measurement process, in that it stimulates the researchers to learn more regarding complex systems with which then work.

Day by day models will become the best mechanisms of transforming information in useful knowledge and of transferring this knowledge to the others.

One of the problems that impede larger progress in modeling is the dearth of modelers having a global vision of reality.

**REFERENCES**


DOURADO-NETO, D.; TERUEL, D.A. Raiz (v.1.00): Programa computacional para simulação de crescimento e estrutura de sistemas radiculares. /Apresentado no I Simpósio Nacional de Instrumentação Agropecuária, EMBRAPA-CNPDIA, 1996./


Recebido para publicação em 10/07/98
Aceito para publicação em 03/08/98