Department of Agricultural Economics

STAFF PAPER SERIES

DIR 89-1
SP-4

JULY
1989

COMMENTS ON DEVELOPMENT AND USE OF BIOPHYSICAL SIMULATION MODELS

CARL R. DILLON
JAMES W. MJELDE
BRUCE A. MCCARL

Departmental Information Report

The Texas Agricultural Experiment Station
Neville P. Clarke, Director
Texas A&M University System
College Station, Texas 77843

DIR 89-1
SP-4

JULY
1989

COMMENTS ON DEVELOPMENT AND USE OF BIOPHYSICAL SIMULATION MODELS

CARL R. DILLON
JAMES W. MJELDE
BRUCE A. MCCARL

Department of Agricultural Economics
Texas A & M University
Texas Agricultural Experiment Station

Staff papers are available for distribution without formal review by the Department of Agricultural Economics.

The authors are, respectively: Research Associate, Assistant Professor, and Professor in the Department of Agricultural Economics, Texas A&M University.
Funded by Texas Agricultural Experiment Station projects 6507 and 3801.

AAEA paper presented at its annual meetings.
BATON ROUGE, LA JUL. 28 - AUG. 2, 1989

COMMENTS ON DEVELOPMENT AND USE OF BIOPHYSICAL SIMULATION MODELS

ABSTRACT

Issues concerning practical aspects of implementing and developing biophysical simulation models are presented. Several questions researchers should ask when considering using a simulation model are discussed. Unfortunately, there are no definitive answers to these questions. Standard programming, documentation and interaction between users and developers are recommendations made to developers.

KEY WORDS : Biophysical simulation models

COMMENTS ON DEVELOPMENT AND USE OF BIOPHYSICAL SIMULATION MODELS

Biophysical simulation models are becoming increasingly important tools in agricultural research. Applications of biophysical models include studies focusing on production practice decisions (Dillon, Mjelde, and McCarl), irrigation management (Boggess and Amerling; Bernardo et al.; Zavaleta, Lacewell, and Taylor; Mapp and Eidman, 1975, 1976), pest management (Reichelderfer and Bender; Nordh, Zavaleta, and Ruesink), and information systems development (Mjelde et al.). Previous authors have discussed the general principles of biophysical simulation model use (Musser and Tew; Boggess; Baier; Thornton; McKinion; Penning de Vries). The literature almost uniformly predicts an increasing role for biophysical simulation models to provide otherwise unavailable and very costly data. However, little attention has been devoted to the practical aspects of utilizing such models and little discussion has been addressed to the developers of biophysical simulation models.

Fundamentally, agricultural researchers use biophysical models for generating data. The acquisition of data for technical coefficients, resource endowments, parameter estimation, and hypothesis testing is not only a rudimentary requirement, it often constitutes a major and difficult portion of the study effort (Just; Gardener; Daniel). Biophysical simulation models (BPSMs) can aid researchers in overcoming certain types of data limitations. Namely, carefully controlled production data often may be generated using BPSMs. However, utilization of BPSMs to provide data requires a number of related considerations. Guidelines and recommendations on the practical aspects of utilizing such models are necessary.

This paper presents a discussion on the use of BPSMs in agricultural research, as well as implications for developers of BPSMs. The discussion is divided into three sections. First, a series of questions and related discussion pertaining to the use and implementation of BPSMs are presented. The second section presents modeling suggestions to the developers of BPSMs. Concluding remarks complete the paper.

QUESTIONS POTENTIAL USERS OF BPSMs SHOULD ASK

Six highly interrelated questions and associated discussion pertaining to the use of BPSMs are presented. There are no definitive answers to these questions, but that does not preclude a researcher from considering the issues raised if the use of BPSMs is being considered. All the questions need to be examined before BPSMs are used.

Should a biophysical simulator be used?

A researcher needs to decide if a BPSM should be utilized in the proposed study. In order to answer this, the type of data necessary to complete the proposed study, as well as the type of data the BPSM might be used for, has to be determined. Potential data sources of the data which the BPSM could generate need to be identified and explored. An advantage of using BPSM is that they can provide estimates of yield based on various management practices over a wide variety of weather conditions at the same location and using the same production technology. That is, the experimental setting can be tightly controlled. BPSMs should be used only when suitable data is not otherwise available.

Potential users of BPSMs must be aware that biophysical simulators are not a panacea to all data problems; they do not solve data problems but rather they create different data and validity problems. Using a BPSM causes additional data needs, those required by the BPSM. Further, the credibility of studies employing BPSMs is almost entirely tied to the credibility of the BPSMs used in the study. Finally, time constraints need to be considered.

Utilizing BPSMs is time consuming, even if previously developed BPSMs are used. Time and resources are required to understand the model, calibrate/validate the model, generate the necessary data, and complete the analysis.

The lack of data, which is often cited as the reason BPSMs are used, also affects the building of the models. Although the models are based on physical and agronomic principles, some functions within BPSMs are derived from experimental data. Consequently, these functions may not be estimated over a set of weather conditions, production practices, or geographical area relevant to the study at hand. Researchers must be aware that the model may be utilizing data for which it has not been estimated; there are out of sample error problems associated with BPSMs analogous to the problem in statistical studies. Also, BPSMs do not completely model all the factors influencing yield. Thus, interactions and constraints faced by a farmer may not be present in the generated data. Therefore, the decision to use a BPSM is not a trivial decision.

How will the biophysical simulator be used?

BPSMs may be used in several ways. The simplest procedure is to use the model in strictly a simulation setting. In a simulation setting, a model is used to simulate the consequences of various actions (e.g. management practices, weather conditions, soil types) on yield to gain an understanding of the problem at hand (Ahmed, van Bavel, and Hiller). BPSMs may also be used to generate data for further analysis. For example, BPSMs may be utilized to either estimate a yield management relationship (Mazzocco), to reflect management alternatives in an economic mathematical programming model (Dillon, Mjelde, and McCarl), or to describe distributions of profits (Cochran and

Mjelde). A third procedure incorporates the BPSM as a subcomponent within an overall economic model with interactive optimization of management practices (Harris and Mapp).

Should a previously developed BPSM be used or should a new model be developed?

Depending on the time and resources available, expertise of the researchers, objectives of the study, and existence and capabilities of relevant previously developed BPSMs, researchers should decide between using existing models or developing a new one. Development of new models involves considerable cost, time, programming expertise, agronomic expertise, and data. Because of these high costs, any researcher contemplating developing a BPSM should bear in mind that the model will most likely be used in different studies at some later time. These high costs and the increased availability of BPSMs are the major reasons the use of previously developed models is becoming more frequent and will continue to increase.

Which BPSM should be used?

Because most researchers lack the resources to develop a BPSM, they use previously developed models. Quite frequently, several BPSMs can be found which simulate growth for the same crop. Selection of which BPSM to use is crucial, because the credibility of the study is directly tied to the credibility of the model(s) used. Unfortunately, there is not an easy answer to the question of which model to use. Many issues must be weighed by the researcher in determining which model(s) to use, some of which are discussed here. An obvious issue is the availability of a particular BPSM.

Several other issues pertaining to the choice of previously developed model(s) need to be considered. First, is the model designed to generate the type of output data necessary to complete the study? In many cases utilizing a previously developed model results in a model being used in a manner different than the purpose for which originally designed. Second, does the model allow the researcher to alter the input variables of interest? If one is studying yield response to nitrogen for example, the explicit inclusion of that relationship should be incorporated into the model. Third, how available is the necessary input data over the time period to be simulated? For example, solar radiation data is not readily available for a long time period in most areas of the country. If the model under consideration requires solar radiation data, the model would not be very useful for generating a long time series of yield data. A model which internally calculates solar radiation may be better suited for the study.

Fourth, what are the computer requirements of the BPSM? If the model is to be transferred to the users' computer system, has this been done? In a study involving the authors, four BPSMs were obtained and transferred to our computer system. Because of compiler specific programming, several months were required to reprogram the models to operate on our computer system. Fifth, are there resources available to run the BPSM? These resources involve monetary and time resources to reformat and/or reprogram the model as necessary, validate/calibrate the model, and finally complete ultimate model simulations. Sixth, is the area for which the model was originally developed for identical to, comparable with, or quite different from the region to be studied? That is, are the assumptions used in developing the BPSM compatible with the study region? Seventh, has the model been published in a accepted journal? Although on the surface this appears trivial, in reality using a previously published model gives credibility to the new study. Eighth, is the BPSM accompanied with complete, understandable documentation? A model needs to be understood to be properly applied. Further, it is our contention that model developers which take the time to write understandable documentation most likely were also conscientious in the development of the model. Finally, we contend that a good working relationship between the model developers and the model users is important. Such a relationship will improve the use of the models, increase the validity of the BPSM generated data, and improve the BPSM development as the developers and users interact. This point is expanded on in the section directed toward model developers.

How can the BPSM be validated?

Once a model is selected, it must be validated/calibrated. Lack of production data is often the major reason for adopting the BPSM approach to generate data. This lack of data makes complete validation impossible. Because complete validation is not possible, many other procedures to access the validity of the BPSMs are used. Previous attempts at validating the BPSMs in question and their conclusions are often cited as the first means of establishing a model's validity. The applicability of these previous attempts to the particular setting at hand must be discussed. This discussion usually includes the model developers' validation/calibration of individual functions within the BPSM.

Another ad hoc validation procedure is to generate a series of yields under different conditions and see if the results conform to prior expectations. Usually this involves examining averages, as well as determining if yields resulting from different weather conditions are the same as expected agronomically for given management practice changes. For example, consider the prior expectation that an earlier planting date on average yields higher than a later planting date. Several different planting dates could be input into the BPSM and average yields over many types of weather conditions calculated to see if the generated average yields followed prior expectations Further, the individual years could be compared to see if a year with a late freeze had a lower yield for early planting dates. Although this procedure is very ad hoc, it is commonly used and can provide insights about the BPSM, especially if agronomic experts familiar with the study region are involved. Another commonly used procedure to establish a model's validity is to use what data is available and determine if the BPSM tracks this data "acceptably". Such a procedure provides only a partial validation; in most available data sets not all production practices are varied.

How should the BPSM be calibrated?

Calibration, which is highly interrelated to the incomplete validation procedures used with BPSMs, involves the tuning of parameters in such a way the model closely approximates an available data set. Documentation and a good working relationship with model developers is vital in calibrating a BPSM. Proper parameters to vary and acceptable ranges for the parameters can most easily be identified by the model developers. Several procedures such as the one developed by Talpaz, da Roza, and Hearn can aid in the tuning of parameters, but calibration will remain an art rather than be a science. Only a brief discussion of validation/calibration issues involved in the selection and use of BPSM has been presented. These issues remain some of the most important issues involved in using simulation models. Interested readers are referred to the rich body of literature concerning validation/calibration (McCarl).

RECOMMENDATIONS FOR MODEL DEVELOPERS

Because most developers of biophysical models are not primarily trained as computer programmers but rather as agricultural scientists with computer knowledge, several recommendations are made directed towards the developers. These recommendations should help to facilitate the working environment between user and developer. Developers and users of BPSMs are not disjoint;

BPSMs are usually developed for use in a specific study. As noted earlier, it is highly likely the model will be used again by either the developer or another researcher after the original study is completed because of the costs associated with developing a BPSM. In developing a model, this point should be kept in mind. The programming and documentation should not be done hastily, but rather the developer should use standard programming techniques with complete documentation both within the computer code and the user manual.

Developers should attempt to write the models in ANSI standard programming languages. This is important because model developers and potential model users may not share the same computer system. BPSMs should also be generated in modular (subroutine or procedure) form with independent modules for data input, simulation control, simulation execution, and response output. For developers of more than one BPSM, the development of common execution modules across simulators (as possible) for the major biophysical processes such as evapotranspiration, photosynthesis, and soil water balance would facilitate simulation comprehension, implementation, modification, and application. Further, the modules need to be tested so that repeated simulations can be done under the control of the simulation control module. Such a control module permits one to run multiple simulation under alterations of certain parameters (e.g., planting date), facilitating development of a wide variety of response data. This also means that developers should place all of the model's parameters under the user's control (but with recommended default values) rather than having them embedded within the code. The preceding formulation facilitates the performance of many repeated simulations. In addition, the models need to be tested using different compilers and computers to insure that repeated simulations can be performed accurately under the control of the main program. Currently, some BPSMs are written without this feature; much time and effort is expended to incorporate this feature into the models.

We agree with Musser and Tew that it is essential that a close working relationship be established between the model users and developers. Such interaction and feedback will improve the quality of the analysis performed with BPSMs, as well as the quality of models themselves. It is important that the model users and developers discuss the model data and identify which parameters should be used in calibrating model results. This allows the researchers to verify whether the model is applicable to the problem, thus helping to alleviate some of the problems discussed earlier. Also, while we recognize the difficulty of incorporating more detailed simulation, we feel users would be very interested in the inclusion of factors such as soil compaction, pests and diseases, soil nutrients, organic matter, harvesting conditions, salinity, grazing, previous crop planted, and irrigation regimes on yield along with the traditional management practices of planting date, hybrid, and population. We disagree with Musser and Tew that the results from such models cannot be interpreted and are not transferable to farm managers. We believe the resultant models, while complex, would still be manageable. However, developers should include recommended default settings for the parameters so that it does not put the responsibility solely on the user to choose appropriate parameter values.

More thorough validation of model results to currently incorporated input factors is as vital as expanding the number of input factors modeled. This will require designing experimental plot studies with the objective of validating/calibrating existing BPSMs, along with using currently available data. Finally, model developers and users would benefit from interactions concerning the desirability of possible program features.

CONCLUDING COMMENTS

Biophysical simulation models are useful research tools which possess both inherent advantages and disadvantages. Convenience, adaptability, expandability, dependability, function, and applicability are the chief benefits of the utilization of biophysical simulation models. Validity of the simulated results is the major disadvantage. The future of biophysical simulation modeling is even more promising as the total benefits of biophysical simulation models are as of yet unrealized. Further advancements in biophysical simulation enables conceptually unlimited applications of the models. To obtain the benefits of these advancements, feedback and interactions between model developers and users is necessary.

REFERENCES

Ahmed, J., C.H.M. van Bavel, and E.A. Hiller. "Optimization of Crop Irrigation Strategy under a Stochastic Weather Regime: A Simulation Study." Water Resources Res. 12:6(1976):1241-47.

Baier, W. "Crop-Weather Analysis Model: Review and Model Development." J. Appl. Meteorology. 12(1973):937-947.

Bernardo, D.J., N.K. Whittlesey, K.E. Saxton, and D.L. Bassett. "An Irrigation Model for Management of Limited Water Supplies." West. J. Agr. Econ. 12(1987):164-173.

Boggess, W.G. "Discussion: Use of Biophysical Simulation in Production Economics." So. J. Agr. Econ. 16(1984):87-89.

Boggess, W.G. and C.B. Amerling. "A Bioeconomic Simulation Analysis of Irrigation Investments." So. J. Agr. Econ. 15(1983):85-92.

Cochran, M.J. and J.W. Mjelde. "Estimating the Value of Information with Stochastic Dominance: An Application from Agricultural Crop Management." Paper presented at the Southeast Region of the Decision Sciences Institute meeting, Charleston, South Carolina, February 22-24, 1989.

Daniel, R. "The Dilemma of Agricultural Economists: Discussion." Amer. J. Agr. Econ. 65(1983):891-92.

Dillon, C.R., J.W. Mjelde, and B.A. McCarl. "Biophysical Simulation in Support of Crop Production Decisions: A Case Study in the Blacklands Region of Texas." Forthcoming in the So. J. Agr. Econ. 1989.

Gardner, B. "Fact and Fiction in the Public Data Budget Crunch." Amer. J. Agr. Econ. 65(1983):882-90.

Harris, T.R. and H.P. Mapp Jr. "A Control Theory Approach to Optimal Irrigation Scheduling in the Oklahoma Panhandle." So. J. Agr. Econ. 12(1980):165-72.

Just, R.E. "The Impact of Less Data on the Agricultural Economy and Society." Amer. J. Agr. Econ. 65(1983):872-81.

Mapp, H.P. Jr. and V.R. Eidman. "A Bioeconomic Simulation Analysis of Regulating Groundwater Irrigation." Amer. J. Agr. Econ. 58(1976):391-402.

Mapp, H.P. Jr. and V.R. Eidman. "Simulation of Soil Water-Crop Yield Systems: The Potential for Economic Analysis." So. J. Agr. Econ. 7(1975):47-53.

Mazzocco, M.A. "Valuing Climate Forecasts for Midwestern Grain Producers." Unpublished Ph.D. Dissertation. University of Illinois. 1988.

McCarl, B.A. "Model Validation: An Overview with Some Emphasis on Risk Models." Rev. of Mark. and Agr. Econ. 52(1984):153-173.

McKinion, J.M. "Dynamic Simulation: A Positive Feedback Mechanism for Experimental Research in the Biological Sciences." Agr. Systems 5(1980):239-250.

Mjelde, J.W., S.T. Sonka, B.L. Dixon and P.S. Lamb. "Valuing Forecast Characteristics in a Dynamic Agricultural Production System." Amer. J. Agr. Econ. 70(1988):674-684.

Musser, W.N. and B.V. Tew. "Use of Biophysical Simulation in Production Economics." So. J. Agr. Econ. 16(1984):77-86.

Nordh, M.B., L.R. Zavaleta, and W.G. Ruesink. "Estimating Multidimensional Economic Injury Levels with Simulation Models." Agr. Systems 26(1988):19-33.

Penning de Vries, F.W.T. "Evaluation of Simulation Models in Agriculture and Biology: Conclusions of a Workshop." Agr. Systems 2(1977):99-107.

Reicheiderfer, K.H. and F.E. Bender. "Application of a Simulative Approach to Evaluating Alternative Methods for the Control of Agricultural Pests." Amer. J. Agr. Econ. 61(1979):258-67.

Talpaz, H., G.D. da Roza, and A.B. Hearn. "Parameter Estimation and Calibration of Simulation Models as a Nonlinear Optimization Problem." Agr. Systems 23(1987):107-16.

Thornton, P.K. "The Use of Biological Simulation Models in Agricultural Management Information Systems." Agr. Systems 17(1985):51-57.

Zavaleta, L.R., R.D. Lacewell, and C.R. Taylor. "Open-Loop Stochastic Control of Grain Sorghum Irrigation Levels and Timing." Amer. J. Agr. Econ. 62(1980):785-92.