PAPER NO. 82-4570

PARAMETER SENSITIVITY IN PLANT PROCESS MODELS

BY

F. D. Baker
Mathematical Statistician
USDA/SRS
Washington, D. C.

For presentation at the 1982 Winter Meeting
AMERICAN SOCIETY OF AGRICULTURAL ENGINEERS

Palmer House, Chicago, Illinois
December 14-17, 1982

SUMMARY:
Plant process simulation models contain many parameters, both single-valued and functional, which may influence final biomass significantly. Response surface techniques can identify sensitive parameters and indicate parameters which could be determined less accurately and with less costly field and laboratory experimentation.

1907-1982
75
Food, Fiber and the Future

American Society of Agricultural Engineers

St. Joseph, Michigan 49085

Parameter Sensitivity in Plant Process Models

Fred D. Baker (USDA/SRS)

INTRODUCTION

The application of various analytic techniques to plant process simulation (PPS) models was initiated by USDA as a tool by which agronomists, agricultural engineers, and other plant scientists could identify more clearly the structure of their plant process models and the effectiveness of those models in mimicing the actual growth of a specific kind of plant or plants. In addition, the sensitivities of various responses could be identified with respect to model inputs, parameters, and substructures.

Plant modelers have had two basic objectives -- to study the individual plant processes and their adaptation to various stresses and to model the actual plant biological yield. The intricacy of the modeled growth processes varies; many of the parameter values for differential equations which regulate growth and development and for functional expressions are determined by either experimentally-derived regression equations or calibration. In most cases a plant (or a small plot) is grown on a daily time step, where the biomass is generated as a function of heat units or calendar days. The plant(s) may become stressed as a reaction to initial conditions and daily input of environmental variables such as maximum temperature, minimum temperature, solar radiation and precipitation. Various models permit the user to apply fertilizers, pesticides, herbicides, and/or irrigation treatments. Other models consider such environmental factors as soil composition, insects and diseases.

If forecasts using a plant process model are desired, stochastic weather simulators (for example, Larsen and Pense (1981)) can be introduced to generate possible future weather conditions for given locations during the growing season. Actual weather data from a historical data base also could be used. Aggregation of sample field estimates (or forecasts) then could be utilized to obtain more universal values.

Problems in modeling arise whenever the model parameters are extremely sensitive to plant location or variety. Most plant process simulation models do not contain sufficient complexity to reduce this sensitivity. Since the aggregation of sample values occurs over an entire state, or over regions of the United States, these factors become very important. If any location and variety parameters are not sensitive, then the time and cost needed to ascertain these values can be reduced.

In the study of the structure of the model and of the interrelationships between variables and parameters, the model author can provide important technical support. The dependence of modeled yield on parameters and input variables can be identified for specific subsets of parameters and variables. Our concern has been the degree of interrelationships and the sensitiviity of the yield response to various factors.

The traditional approach to studying the effect of varying relevant parameters in a simulation model is the same as that used in any other interpolation problem: a multidimensional regula falsi technique. The use of fractional factorial designs to limit the number of simulation runs required, is clearly indicated when the number of factors becomes large; Montgomery (1979) presents a detailed discussion of the use of such techniques.

As discussed in previous papers (Baker and Bargmann (1981), 1982)), we have applied response surface techniques to plant process models. Many of the resultant surfaces have not been quadratic, but linear with large errors. The simple cubic response surface does not permit easy identification of any factor effect on the given response. With a scaled orthogonal central composite design, higher-order relationships (in the model) and the sensitivity of the yield response to these relationships (as presented in the given model), have been detected.

In this paper, we shall first indicate the assumptions under which we have applied response surface techniques. Then we shall discuss several ways that the sensitivity of parameters -- both single-valued and functional -- have been evaluated.

RESPONSE SURFACE ASSUMPTIONS

In addition to the usual assumptions that the experiments (simulated growths of a plant) can be performed and the response(s) measured at each of the design points, two special assumptions have been made. First, it is assumed that an available simulation program is the author's best description of the modeled plant processes. The model usually has been validated by the author for several locations and the forecasting (estimation) ability of the model should have been evaluated. Thus, the existing values of the factors (called the center point values) provide a good indication of the response and the response surface interecept should be close to the center point response. Second, in the calibration of the model parameters (selected as factors), a range of possible values (or a set of bounds) for each factor has been determined by the model author. By setting the factor levels to be + 1 and - 1 at these bounds, a design lattice is defined and provides a description of the variation of the "real" response of the model expected under variation of parameters.

RELATIVE SENSITIVITY

Although the physical units of the factors may be completely different and levels may even be categorical in nature, use of the model author's calibration limits in setting the factor levels standardizes the sensitivity of the response to all of the factors. Since the intercept of the response equation is the best estimate of the response at the center and "reality" lies inside the unit lattice, the contribution of each factor or factor combination at the boundary (lattice points), can serve as an index of sensitivity. The maximum contribution to the equation occurs when the factor level is "1" or "- 1" and equals the value of the coefficient. The ratio of this coefficient to the intercept has been defined to be the "relative sensitivity" of the response to variation of a factor or factor combination.

APPLICATIONS

Two models have been used during the development of the analytic techniques. A dynamic wheat growth and development simulator called TAMW ("Texas A&M Wheat Simulation Model") (Maas and Arkin, 1980a) provided an early forum for the practicality of the response surface technique. A early version of SOYGRO (a soybean growth simulation model developed at the University of Florida) (Wilkerson and others, 1981) was used to extend these techniques.

As Baier (1977) points out, a proper plant process simulation model should reflect differences in location and in variety. The current generation of plant process simulation models is usually calibrated to account for both of these factors with soil characteristics as a major contributing factor. Ideally a model could be corrected internally for location (with the same variety). Using the appropriate daily weather input, this model then would produce "similar" response surfaces. In these cases the generality of the response model would depend on, perhaps, the mean yield for each location with minor variation attributable to the specific factors.

TAMW was tested by its authors over ten different locations from Texas to South Dakota. Preliminary results (Maas and Arkin, 1980b) indicated that although TAMW simulated the date of floral initiation later than it was observed in the field, the simulated yield did not differ much from the actual yield for most fields. The larger differences occurred when unmodeled management practices or severe weather conditions were present. Daily weather data for the 1978-79 growing season and appropriate latitudes (for daylength) were available for each of these locations.

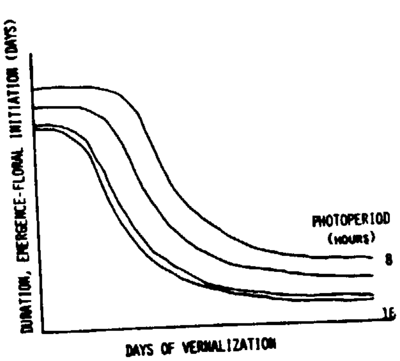
With the assistance of Maas, we selected three factors with yield dependence.

1. CVERN -- an input function which governs the effect of vernalization on the vegetative phase (see Figure 1).

2. BETA -- the limiting value (or the asymptote) for the number of heads (H) possible at head emergence, given the number of existing shoots (X) at terminal spikelet.
   $H = \beta * (1 - EXP(-\alpha * X/\beta))$

3. CGSET -- the fraction of wheat florets converted to grain

FIGURE 1: Picture of CVERN



DURATION OF THE PERIOD FROM EMERGENCE TO FLORAL INITIATION AS A FUNCTION OF PHOTOPERIOD AND VERNALIZATION.

The levels of CVERN and other precisions of measurements were identified with the assistance of Maas. The function (F) retained the same shape at each of the levels, but had different y-intercepts. In the initial testing, the precision of measurement was defined as a vertical shift of all four units. Similarly the precisions of measurement of BETA and CGSET determined the parameter values at the levels for these variables (See Table 1).

Table 1: Parameter Values for Factors from TAMW

Factor Level	-2	-1	0	1	2
CVERN	F - 8	F - 4	F	F + 4	F + 8
BETA	620	670	720	770	820
CGSET	.15	.20	.25	.30	.35

Response surfaces were generated for three data sets representing different environmental conditions. The field sites were located near Temple, Texas (a wet southern latitude growing season); Fort Pierre, South Dakota (a normal northern latitude growing season); and Brewster, Kansas (a severe early-season water stress growing season).

In this case the TAMS parameter values for all three factor variables are determined experimentally. If yields are not significantly affected by any or all of the factors, then less costly research could be possible in the determination of the corresponding parameter values. Also, by comparison of the response surfaces at different locations, the portability of the INTERNAL TAMW parameters can be evaluated. The results are shown in Table 2.

Table 2: Quadratic Response Surfaces for TAMW
(Coefficients Significant for $\alpha = .10$

FACTOR/LOCATION	TEMPLE	BREWSTER	FT . PIERRE
INTERCEPT	3.85	2.99	2.575
CVERN	- .13	-	- .10
BETA	.28	-	.15
CGSET	.77	.59	.52
CVERN*BETA	-	-	-
CVERB*CGSET	-	-	-
BETA*CGSET	-	-	-
CVERN**2	-	- .12	-
BETA**2	-	-	-
CGSET**2	-	-	-

The response surface for Brewster, Kansas, with $\alpha = .10$ is

YIELD = 2.99 + .59 * CGSET - .12 + CVERN**2

The relative sensitivity for CGSET is approximately .20 for each of the three locations $\left(\frac{.77}{3.85} \approx \frac{.59}{2.99} \approx \frac{.52}{2.575}\right)$. This large value indicates the importance of an accurate identification of the percentage of florets converted to grain. Smaller relative sensitivities for CVERN and BETA suggest that less precise experimentation may be sufficient to simulate wheat yield with TAMW. BETA is less important when early season water stress is present. Relative sensitivities for all factors are similar over location and the coefficients vary proportionally to the intercept (center-point value).

In a second application, the response surface techniques were used with early versions of SOYGRO. The behavior of daily photosynthesis in response to light had been modeled from experimental data (Ingram and others, 1981). Maximum available photosynthate is reduced by a series of effects including crop growth rate and temperature stress. These effects are modeled from experimental field data to give the fraction of available photosynthate which remains. The model authors had decided to model each of these effects with specific types of behavior, but were concerned about sensitivity to any changes in the corresponding functions.

For crop growth rate, the fraction of remaining photosynthate is dependent on leaf area index (LAI) and is assumed to be (1) equal to LAI for small values of LAI, (2) parabolically increasing to a maximum for larger values of LAI, and then (3) exponentially and asymptotically increasing to 1. The model equation is:

$$f_L = \begin{cases} L & 0 < L \le .1022 \\ .054 - .4778*L - .0623*L^2 & .1022 < L \le 3.835 \\ 1 - e^{-0.9144L} & L > 3.835 \end{cases}$$

If the model equation has this same structure, but with a different interval for the quadratic relationship, then the end-points -- p_1 and p_2 -- of this interval determine the coefficient of the quadratic equation.

$$f_L = \begin{cases} L & 0 < L \le p_1 \\ c + bL + aL^2 & p_1 < L \le p_2 \\ 1 - e^{qL} & L > p_2 \end{cases}$$

in fact,

$$a = \frac{-b}{2p_2} \text{ and } c = \frac{p_1^2\, b}{2p_2} - (b - 1)p_1$$

are known after calculation of

$$b = \frac{2p_2(1 - p_1 - e^{-qp_2})}{(p_2 - p_1)^2}$$

for any given value of q which determines the rate of asymptotic convergence. The points p_1 and p_2 have been chosen as the first two factors in the sensitivity analysis.

The last two factors are the endpoints in the function describing the temperature stress. The model equation is:

$$f_T = \begin{cases} 0.0 & T \leq 13 \\ \ln (T - 5)/8 & 13 < T \leq 27 \\ 1.0 & T > 27 \end{cases}$$

Assuming that the behavior is logarithmic between the endpoints, the two numeric constants can be expressed as a function of these endpoints.

$$f_T = \begin{cases} 0.0 & T \leq p_3 \\ \ln (T - a)/b & p_3 < T \leq p_4 \\ 1.0 & T > p_4 \end{cases}$$

Where

$$a = (p_3 e - p_4) /(e - 1)$$

and

$$b = (p_4 - p_3) /(e - 1)$$

Wilkerson suggested degrees of precision for the factors and we generated the quadratic response surfaces for each of the years 1978-1980 with weather data from Gainesville, Florida. The parameter values at each level are listed in Table 3 and the response surfaces are given in Table 4.

Table 3: Parameter Values for Photosynthesis Factors from SOYGRO

Factor/Level	- 1.141	- 1	0	1	1.141
LAILO (p_1)	.0794	.0822	.1022	.1222	.1250
LAIHI (p_2)	3.607	3.635	3.835	4.035	4.063
TPHLO (p_3)	11.86	12	13	14	14.14
TPHHI (p_4)	24.72	25	27	29	29.28

Table 4: Quadratic Response Surfaces for Photosynthesis Parameters
(Significant Coefficients for $\alpha = .10$)

FACTOR/YEAR	1978	1979	1980
INTERCEPT	104.46	369.89	217.87
LAILO	0.44	2.07	3.67
LAIHI	- 1.12	- 4.91	- 8.48
TPHLO	- 0.10	- 1.48	- 0.77
TPHHI	- 2.00	-14.48	- 5.66
LAILO*LAIHI	0.10	0.58	0.36
LAIHI*LAIHI	-	- 1.29	0.22
LAIHI*TPHHI	-	-	0.17
TPHLO*TPHHI	- 0.11	- 0.93	- 0.36
TPHHI*TPHHI	- 1.36	- 5.91	- 2.51

The relative sensitivities are all less than .05 and suggest that as long as the degree of precision is as defined in Table 4, and the behavior of the functions is consistent with the assumptions, then yield has minimal sensitivity to the choice of the points p_1, p_2, p_3, and p_4. The value for TPHHI (p_4) "contributes" most to the model. Larger values of p_4 require higher temperatures for the stress fraction to be 1.0 and thus decreases the amount of photosynthate available; consequently the biomass is smaller.

CONCLUSION

Response surface techniques in sensitivity analyses have helped to identify the relative importance of parameters in the plant process simulation of various crops. Coefficients of assumed functional relationships for the plant processes can be evaluated in terms of the response of yield to changes in the coefficients. Those parameters which have minimal effect on final yield could be established with less precise and less costly experimentation.

BIBLIOGRAPHY

(1) Baier, W., 1977: Crop-Weather Models and Their Use in Yield Assessments. WMO Bulletin 458, Technical Note 151.

(2) Baker, F. D. and R. E. Bargmann, 1981: Analysis of Simulation Models by Response Surface Techniques. Proceedings of the American Statistical Association, Section on Survey Research Methods, p 34-39.

(3) ______________________________, 1982: Orthogonal Central Composite Designs of the Third Order in the Evaluation of Sensitivity and Plant Growth Simulation Models. Presented at 1982 American Statistical Society Annual Meetings.

(4) Ingram, K. T., A. C. Herzog, K. J. Boote, J. W. Jones, and C. S. Barfield, 1981: Effects of Defoliating Pests on Soybean Canopy CO_2 Exchange and Reproductive Growth. Crop Science 21: 961-968.

(5) Larsen, G. A. and R. B. Pense, 1981: Stochastic Simulation of Daily Climate Data. USDA/SRS Staff Report No. AGES 810831.

(6) Maas, S. J. and G. F. Arkin, 1980a: TAMW: A Wheat Growth and Development Simulation Model. Research Center Program and Model Documentation No. 80-3, Texas Agricultural Experiment Station, College Station, Texas.

(7) ______________________________, 1980b: Initial Validation of a Winter Wheat Simulation Model. Technical paper No. 80-4010, presented at 1980 ASAE Summer Meeting, San Antonio, Texas.

(8) Montgomery, D. C., 1979: Methods for Factor Screening in Computer Simulation Experiments. Final Report, Office of Naval Research, Contract N0014-78-C-0312.

(9) Wilkerson, G. G., J. W. Jones, K. J. Boote, K. T. Ingram, and J. W. Mishoe, 1981: Modelling Soybean Growth for Crop Management. Presented at 1981 ASAE Summer Meeting, Orlando, Florida.