THE INSTITUTE OF STATISTICS

UNIVERSITY OF NORTH CAROLINA SYSTEM

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina

 Yu, Gold, &Linker

QSOY: AN EXPERT SYSTEM FOR PROTECT-

QSOY: AN EXPERT SYSTEM
FOR PROTECTING SOYBEAN AGAINST CORN EARWORM

Yanan Yu

Dept. of Mathematics

Alabama State University, Montgomery, Alabama

Harvey J. Gold

Biomathematics Graduate Program, Department of Statistics

N. C. State Univ., Raleigh, N. C.

H. Michael Linker

Crop Science Department

North Carolina State University, Raleigh, N.C.

Institute of Statistics Mimeo Series No. 2208

Biomathematics Series No. 35

Biomathematics Graduate Program

Dept. of Statistics

Campus Box 8203

Raleigh, N. C. 27695-8203

December, 1991

QSOY: AN EXPERT SYSTEM

FOR PROTECTING SOYBEAN AGAINST CORN EARWORM

Yanan Yu[1], Harvey J. Gold[2], H. Michael Linker[3]

[1]Dept. of Mathematics, Alabama State University, Montgomery, Alabama
[2]Biomathematics Graduate Program, Department of Statistics, N. C. State Univ., Raleigh, N. C.
[3]Crop Science Department, North Carolina State University, Raleigh, N.C.

ABSTRACT

An expert system, QSOY, was developed to determine to use insecticide against corn earworm (*Helicoverpa zea*) on soybean crops. Spray application thresholds are modeled based on general knowledge from agricultural extension publications, modified by knowledge obtained from interviews with extension entomology specialists and entomology researchers, and from judgments made by specialists on prototypical scenarios. The user is asked to supply information concerning the soybean field, the environmental conditions and the results of sampling for corn earworm larvae.

This report describes the process of knowledge acquisition used, the logic of the system and the way in which the system interacts with the user. The expert system development tool, ARITY/EXPERT is critically evaluated.

INTRODUCTION

This report describes the development and use of a rule-based expert system to assist extension agents and farmers in making decisions concerning use and timing of insecticides to protect soybean crops against corn earworm (Helicoverpa zea (Bodie)). Corn earworm is the most serious insect pest of soybean in North Carolina. It is a foliage feeder before bloom and a pod feeder after bloom. It can reduce the soybean yield and delay plant maturity.

Soybeans are grown on over 1 million acres each year in North Carolina. At least once in each season, a soybean grower be faced with a treatment decision to protect the crop against corn earworm. If a spray application is needed, a decision must be made as to which insecticide to apply and what rates to use.

A number of complicating factors affect these decisions, such as soil productivity, rainfall history and rainfall forecast, crop growth stage and vigorousness of the crop, and even the timing of nearby corn plantings. An incorrect decision not to spray may result in substantial crop loss; an incorrect decision to spray may result in wasted insecticide, reduction in beneficial insect, and a reduction of profits.

KNOWLEDGE ACQUISITION

A critical aspect of building an expert system is acquiring and distilling the required knowledge, and encoding it in a form that can be stored and utilized by one of the methods available to the software being used. A considerable literature is developing on the methodology of knowledge acquisition (Kidd, 1987). Other methodologies which are directly relevant are those concerned with problem structuring as part of the process of building system models (Gold, 1977), and as part of the process of decision analysis (Spetzler and Staël von Holstein, 1975).

Four separate steps were used in this work: review of published material related to the decision; unstructured discussion; structured discussion; and use of prototypical scenarios. In applying the last of these techniques, an interpolation model was used to reduce the number of scenarios to a tractable level. Although these steps are described as a sequence, it is important to note that the process iterates back to earlier steps as needed.

Step 1. Use of publications.

A first step is the use of publications, such as extension publications developed as guides for farmers or for extension agents (Fehr et al., 1980; Linker et al., 1987). These manuals give relatively simple rules for determining action thresholds under normal or "average" conditions. Other knowledge such as effectiveness of pesticides is obtained from experimental reports. This step serves to orient the "knowledge engineers" to the problem context, and to develop an initial conceptual structure to guide further work. This step included the writing of a prototype "novice" system, which helped to guide the choice of an expert system tool for later work.

Step 2. Unstructured discussions.

The next phase consisted of a series of intensive discussions between the "knowledge engineers" (H. J. G. and Y. Y.), the principal "domain expert" (H. M. L.) and other experts to expand the problem setting, obtain an initial understanding of the range of complications involved, and define specific problems for later discussion.

Step 3. Structured discussions.

The next series of discussions were each structured around a single or a small number of points. Discussion with other extension specialists helped to clarify questions related to crop development, insect population development, sampling accuracy, and conditions under which future infestations might be expected.

Disagreements on certain points were resolved through group discussion. Such a discussion process, which involves sharing of information and points of view is to be preferred, we feel, to more formal and impersonal methods, such as the well-known "Delphi" method.

Step 4. Prototype scenarios.

The rules developed to this point would have resulted in a system that might be expected to operate at the "advanced beginner" level, as defined by Dreyfus and Dreyfus (1986); i.e., being able to follow relatively straightforward rules. However, it was felt that a more complex set of rules was needed to address some of the complications that may affect the decision in the period before first bloom. A fourth phase of the knowledge acquisition process accordingly involved determining the judgments of a group of four extension specialists over a wide range of scenarios. This was done through a questionnaire describing the scenarios. Each specialist was asked to give a recommendation (spray or no action) for each scenario.

In developing the questionnaire, each scenario was initially characterized by stating the following variables:

a. Time before first bloom
b. Percent canopy coverage
c. Percent defoliation
d. Soil productivity (low, adequate, high)

e. Rainfall history (dry, normal, wet)

f. Rainfall forecast (dry, normal, wet)

However, covering the entire decision space with a grid on these six variables would have produced a completely unwieldy questionnaire. The initial questionnaire was therefore simplified by: first, noting that the characteristics listed above are all directed at determining the stress condition of the crop and the likelihood of recovery from stress; second, developing interpolation models to allow a more widely spaced grid; third, building into the system enough redundancy to offer a test of the interpolation models. The final questionnaire presented 72 scenarios, and could be completed in approximately 30 minutes. Format of the questionnaire is shown in Figure 1.

Results were used to determine corrections to the action thresholds previously arrived at. Disagreements between the specialists answering the questionnaires were relatively minor and usually consisted in one specialist making finer distinctions than another. Such disagreements were resolved in favor of the specialist making the finer distinction.

A complication not addressed by any of the rules, concerns the effect of the corn plantings in the surrounding agroecosystem on the timing of the peak corn earworm infestation. Under some conditions, while application of insecticide will appear to be the recommended action, the peak of the infestation has not yet occurred. In such a case, the insecticide may do more damage than good by killing beneficial insects. Accordingly, if the field is within two weeks of bloom, a recommendation to spray is accompanied by an explanation and a warning that it may be desirable to consult an extension agent or specialist.

MODELING THE THRESHOLDS

Before first bloom, thresholds are in terms of percent defoliation. After first bloom, they are in terms of the number of corn earworms of length 3/8 inch or larger.

Before First Bloom

In the period before first bloom, spray recommendations are based primarily on the percent defoliation due to corn earworm, but are modified on the basis of factors discussed in the previous section. According to the scouting manual (Linker et al., 1987), the nominal action threshold is 30%

defoliation up to 14 days before first bloom, and 15% from 14 days before bloom until bloom. This was modified to produce a smooth transition between the two threshold levels. Letting T_B be the action threshold (in percent defoliation), as a function of time before first bloom t, the form used was,

$$T_B(t) = \begin{cases} 15 & t < 10 \\ 30 - 15(18-t)/8, & 10 \leq t \leq 18 \\ 30 & t \geq 18 \end{cases}$$

As discussed above, the values determined from this function were modified on the basis of the stress history of the crop, and anticipated ability to recover from stress, as determined from percent canopy coverage, soil productivity, rain history, rain forecast and level of soil productivity. These values were interpolated between 10 and 18 days, in the manner discussed above for the nominal threshold.

A further modification was made on the basis of planting date. If the crop is planted after July 1, threshold is reduced by 5%.

After First Bloom

In the period after first bloom, recommendations are based on the number of corn earworms observed. The thresholds quoted in the scouting manual (Linker et al., 1987) depend upon the sampling device used, and take account of the relative efficiency of the devices (Braxton et al., 1988). The figures are based upon a nominal threshold of approximately 29,000 larvae per acre at least 3/8 inch long. Devices normally used are the rigid beat-cloth or sweep net for row widths up to 28 inches; the shake cloth for wider row widths.

If the user has sampled for corn earworm using a device other than those listed above, or under nonstandard conditions, the system is unable to apply a correction for sampling efficiency. In such a case, a threshold is computed in units of larvae per row-foot (based on the row spacing in the field), and the user is warned that no adjustment will be made for sampling efficiency.

OPERATION OF THE QSOY SYSTEM

The overall inference tree is shown in Figure 2. The system must first determine if the crop is before or after first bloom. If before first bloom, recommendations are based upon percent defoliation.

If defoliation is above 40% or below 5%, a recommendation can be made with no further inquiry. If it is between these extremes, the system seeks to determine the appropriate defoliation threshold, as described above. The following information will be requested (in some cases, if an estimate is not available, the system supplies a default, based on average conditions):

- estimated number of days till first bloom;
- percentage canopy coverage (default is based on "average" growth curve);
- adequacy of past rainfall (default is "adequate");
- rainfall forecast ("adequate" or "dry" – default is "adequate");
- soil productivity (default is "adequate");
- whether the crop has been planted late (after June 1 – default is "no").

When a default value is used, it is assigned a certainty of 0.5. Recommendations are displayed with a certainty factor to indicate if any default values have been used. In addition, the user may specify a certainty of less than 1.0 along with any answer.

If the response to any question is "why", the user will receive an explanation of why the information is needed, how it will be used, and the action taken by the system if the information is unavailable.

After first bloom, recommendations are based on the results of comparison of observed number of corn earworm larvae with a determined threshold. The following information is requested:

- row width;
- sampling device;
- number of larvae per sample, or number of larvae per row foot if a standard device was not used.

When crop and pest conditions lead to a recommendation to spray, the user is asked if rain is expected within the next 2 hours. If rain is certainly expected, the user is advised to wait; if there is a chance of rain, the user is advised to use more quick-acting insecticides. If the user is able to supply prices of the candidate insecticides, the system will recommend a specific insecticide on the basis of cost and effectiveness.

The compiled version of the system runs under DOS and can be run with an IBM compatible PC or PC/XT microcomputer.

SOME DIRECTIONS FOR SYSTEM EXPANSION

The system described in this report is currently being tested and expanded. The following lines of development are being pursued in collaboration with other scientists.

(1) Interfacing with a database which records and stores field monitoring information (Linker et al, 1986).

(2) Linkage with crop and insect simulation models (Wilkerson et al., 1983; Stinner et al., 1974), so as to take advantage of the extensive and detailed process level knowledge that has gone into development of those models.

(3) Linking with decision aids for other soybean pests, through a coordinating program, which will eventually act as a "diagnostician".

(4) Taking more detailed account of uncertainty. Decisions concerning pest control are made in the face of information that is incomplete and uncertain. The problem of how to treat uncertainties in an expert system environment has been an area of active interest within the past several years. Although some progress has been made in treating uncertainty in expert systems for medical diagnosis (see Spiegelhalter, 1986, for a review), sources of uncertainty in agricultural systems are quite different. They include uncertainties as to current pest levels, uncertainty in the models upon which future projections are made, uncertainty as to future weather conditions and future economic conditions. A useful approach in dealing with this problem is, we feel, through Bayesian probability theory and the paradigm of applied decision analysis (Gold, 1988; Gold et al, 1990).

(5) Incorporation of forecasts concerning sale price of the soybeans. In conjunction with yield-loss estimates, this will allow decisions to be based on economic optimization.

THE EXPERT SYSTEM DEVELOPMENT SHELL

The system was programmed in ARITY/EXPERT (Arity Corp., Concord, Mass.). This expert system shell was initially selected because of its versatility, in that it readily interfaced with the manufacturer's version of Prolog and with a variety of procedural languages.

Advantages we have found in working with ARITY/EXPERT are:

(1) The frame structure allows development of a logical hierarchical relationship between concepts, whereby concepts lower in the hierarchy inherit properties of the upper level concepts. The hierarchical relation between concepts in QSOY is shown in Figure 3.

(2) Ability to provide default values used by the system, when the user is not able to answer certain questions.

(3) Customizing abilities, including

(a) Ability to specify the combination and order of methods for obtaining needed information: application of rules; inquiry; or by default.

(b) Ability to customize wording of questions, of recommendations and of explanations, while using symbolic representations within the system itself.

(c) Ability to directly influence the order of goal satisfaction through the *precalc* command.

(4) A choice of methods for handling uncertainty, including the method of certainty factors introduced by Shortliffe and Buchanan (1975), a method based on fuzzy sets (Zadeh, 1982) and a method based on probability. Unfortunately, the method based on probability implicitly assumes that all individual pieces of information are independent.

(5) Ability to directly interface with PROLOG.

(6) Use of an interpretive mode for development, and a compiled mode for portability.

Principal disadvantages we found with the ARITY/EXPERT shell are:

(1) While the hierarchical structure is convenient for organizing concepts, it is, in implementation, inconveniently rigid. Syntactical rules imposed by the hierarchical structure are difficult to learn and easy to forget.

(2) Screen display for asking questions and making recommendations is austere. This, however, is subject to modification through interfacing with Prolog and use of a "screen design tool kit", which the company offers.

DISCUSSION

The system in its current form appears to function at the level of a reasonably experienced extension agent (although it obviously cannot make its own field observations), but must be tested to determine both accuracy and level of acceptability by farmers and extension agents. Although the system deals with both the period before bloom and the period after first bloom, it is the earlier period in which the problem is particularly complex.

Potentially, efficient expert systems which can be run on relatively inexpensive personal computers, can play a significant role in technology transfer, and would be of value to growers and inexperienced extension agents, Expert systems have been widely used in industrial production problems, business management, and medical diagnosis, and have recently begun to be applied to agricultural systems. Agricultural applications are complicated by the large number of uncertainties that surround the decision, and by the widely varying backgrounds of the decision makers.

The information required by QSOY is a combination of information that might come from field monitoring, weather forecasting, and information provided by the farmer, such as soil productivity, row width, planting date, etc. Ease of use of the system will be improved by linking to a data base which would contain much of this information (Linker et al, 1986), so that the user would be required to answer only a very few questions.

ACKNOWLEDGEMENTS

We wish to thank a number of scientists who contributed their expertise in the building of the knowledge base, through discussion and by answering the questionnaire. These are Drs. J. S. Bacheler, J. R. Bradley, E. J. Dunphy and J. W. Van Duyn. In addition, we thank Dr. R. E. Stinner and Dr. G. G. Wilkerson for critically evaluating the logical structure of the system at various points in its development.

This work was in part supported by grant no. SES 8819490 from the National Science Foundation.

REFERENCES

Braxton, W. M., J. W. Van Duyn, J. R. Bradley, Jr., E. P. Lampert, 1988. Sampling for corn earworm, (Lepidoptera: Noctuidae), in narrow-row soybean: comparison of three relative-density sampling devices to an absolute-density estimation technique, unpublished manuscript.

Dreyfus, H. L. and S. E. Dreyfus, 1986. Mind over machine, Free Press, New York, 231 pp.

Doluschitz, R. and W. E. Schmisseur, 1988. Expert systems: applications to agriculture and farm managment, Comput. Electron. Agric. 2:173-182.

Fehr, W. R. and C. E. Caviness 1980, Stages of Soybean Development, Special Report 80, Cooperative Extension Service, Iowa State Univ., Ames

Gold, H. J., 1977. Mathematical modeling of biological systems, Wiley, New York, 357 pp.

Gold, H. J., 1988. Decision analytic modeling for plant disease control, in Plant Disease Epidemiology, vol 2, eds. K. J. Leonard and W. E. Fry, Macmillan, New York, in press.

Gold, H. J., G. G. Wilkerson, Y. Yu, and R. E. Stinner. 1990. Decision Analysis as a Tool for Integrating Simulation with Expert Systems When Risk and Uncertainty Are Important. Computers and Electronics in Agric. 4:343-360.

Hink, R. F. and D. L. Woods 1987. How Humans Process Uncertain Knowledge: An Introduction for Knowledge Engineers, AI Magazine Fall, 41-53.

Hogarth, R., 1987. Judgement and choice, 2nd ed., Wiley, New York, pp. 311.

Kidd, A. L., ed., 1987. Knowledge acquisition for expert systems, Plenum Press, New York, 194 pp.

Jones, P., J. W. Jones, P. A. Everett, F. A. Johnson, and R. K. Sprenkle, 1986. Development of an expert system for making insect control decisions in soybeans., Int'l Conf on Computers in Agric. Extension Programs, Lake Buena Vista, Fla., Fla Cooperative Extension Service and Inst. of Food and Agric. Sci., Univ. of Fla., Gainesville.

Linker, H. M., J. R. Anderson, Jr., and K. L. Tyra. 1986. Crop Management Information System (CMIS): II Pest Data Collection, Field History and Information Delivery., in Proc. Intl. Conf. on Use of Computers in Agric. Ext. Programs, Orlando, Fla., eds. A. B. Bottcher and F. S. Zazueta, Fl. Coop Ext. Ser., p. 814-818

Linker, H. M., H. D. Coble, J. W. Van Duyn, E. J. Dunphy, J. S. Bacheler and D. P. Schmitt, 1987. Scouting Soybeans in North Carolina, N. C. Ag. Extension Service CS-IPM2v4, Raleigh, N. C.

McKinnion, J. M. and H. E. Lemmon, 1986. Expert systems for agriculture, Comput. Electron. Agric. 1:

Peart, R. M., F. S. Zazueta, P. Jones, J. W. Jones, and J. W. Mishoe, 1986. Expert systems take on three tough agrticultural tasks, Agric. Engineering May/June:8-10.

Shortliffe, E. H. and B. G. Buchanan, 1975. A model of inexact reasoning in medicine, Math. Biosci. 23:351-379.

Spiegelhalter, D. J. 1986. A Statistical View of Uncertainty in Expert Systems. in Artificial Intelligence & Statistics, ed. W. A. Gale, Addison-Wesley, Reading, Mass., p. 17-56.

Spetzler, C. S. and C. S. Stael von Holstein, 1975. Probability encoding in decision analysis, Management Sci. 22:340-358.

Stinner, R. E., R. L. Rabb and J. R. Bradley, Jr., 1974. Population dynamics of Heliothis zea (Boddie) and H. virescens (F.) in North Carolina: A simulation model, Environ. Entomol. 3:163-168.

Van Duyn, J. W. - to be added

Waterman, D. A., 1986. A guide to expert systems, Addison-Wesley, Reading, Mass., pp. 419.

Wilkerson, G. G., J. W. Jones, K. J. Boote, K. T. Ingram and J. W. Mishoe, 1983. Modeling soybean growth for crop management. Trans. ASAE 26:63-73.

Zadeh, L. A., 1982. Fuzzy probabilities and their role in decision analysis. IFAC Proc. on Theory and Application of Digital Control, p. 15-21.