United States
Department
of Agriculture

Forest Service

Intermountain Research Station

Research Note
INT-390

December 1989

Evolution of a Research Prototype Expert System for Endemic Populations of Mountain Pine Beetle in Lodgepole Pine Forests

Dale L. Bartos
Kent B. Downing[1]

ABSTRACT

A knowledge acquisition program was written to aid in obtaining knowledge from the experts concerning endemic populations of mountain pine beetle in lodgepole pine forest. An application expert system is then automatically generated by the knowledge acquisition program that contains the codified base of expert knowledge. Data can then be entered into the expert system to generate predictions.

KEYWORDS: *Pinus contorta, Dendroctonus ponderosae,* knowledge acquisition

Expert systems (ES) are computer programs that process information in ways that mimic the deductive or inductive reasoning processes of a human expert (Negotia 1985). Expert systems are receiving greater attention in the area of natural resources as tools to organize existing knowledge for use by the land manager or research scientist. This paper details the work at a first-pass attempt in developing an ES for endemic populations of mountain pine beetle (MPB) (*Dendroctonus ponderosae* Hopkins [Coleoptera: Scolytidae]), in lodgepole pine (*Pinus contorta* Dougl. var. *latifolia* Engelm.) forests. This project was undertaken, in part, to add to our knowledge about how to develop, implement, evaluate, and modify expert systems in our area of expertise (Stock 1987).

Early examples of expert systems have been in such fields as mineral exploration where human experts are scarce and the cost of equipment lying idle is so high that the price of an ES can be recaptured rapidly; in medicine where extensive written records on the best known ways to diagnose diseases exist; and in finance and accounting because expert knowledge is relatively concrete and can be easily incorporated into a knowledge base (Negotia 1985).

"Fuzziness" due to the ill-defined nature of decision variables and processes is a part of problem solving in many fields, including natural resources. Where decisions are at best, semistructured human judgment is required, although computer processing and other analytical aids are valuable in discovering patterns as well as in spotting irregularities and other inconsistencies that are important to understanding processes at work. Expert systems enable us to work systematically with concrete knowledge as well as material containing a high degree of expert judgment. At any stage in the evolution of an ES, it might best be regarded as an adviser or source of "second opinion" that asks necessary questions and gives advice within a context of current but changing state-of-the-art knowledge and experience.

Difficulties encountered in the development and application of ES technology are varied and include: (1) selecting topics suitable for ES development; (2) maintaining an objective perspective necessary for recognizing ill-defined characteristics of subject matter and the patience to reformulate them during the knowledge acquisition/organization process; (3) capturing a base of knowledge adequate for generating valid predictions; and (4) helping end-users adapt selected aspects of their work to maximize the benefits of the system. Moreover, the system designer (knowledge engineer) must assure experts and end-users that the ES is not final although it may reach a stage suitable for use as a working model in the field—one that can be further modified as knowledge accumulates or as inconsistencies between model predictions and actual events are discovered.

[1]Dale L. Bartos is operations research analyst, Intermountain Research Station, Logan, UT. Kent B. Downing is associate professor, Department of Forest Resources, Utah State University, Logan.

PROJECT OVERVIEW

The goal of this work is to devise an information/ decision support system that provides expert knowledge to researchers and managers in a user friendly, easily revised format. Our ES will deal with endemic (initial stages prior to outbreak) populations of MPB in lodgepole pine forests in the Intermountain West. This work will complement the current research being conducted by Forest Service scientists with the mountain pine beetle project in Ogden, UT. We anticipate that the ES generated will approach that of a "research prototype" (Waterman 1986).

During the past 20 years the MPB project in Ogden has developed an extensive knowledge base about the epidemic phase of MPB infestation. These contributions have provided a good understanding of the dynamics of epidemic level populations in lodgepole pine (Amman and Cole 1983; Cole and Amman 1980; Cole and others 1985).

A hazard rating system to determine stand susceptibility to MPB in lodgepole pine forests has been developed. Also, researchers realized that direct control of outbreak level populations was a holding action at best and directed attention toward silvicultural means of suppressing MPB under outbreak conditions. These conclusions emphasized the need for a better understanding of the dynamics of low or endemic level MPB populations that should ultimately lead to development of preventive strategies.

Years of working with epidemic MPB populations have given rise to better understanding of endemic situations. Research dealing with low level population densities has shown the beetle is sometimes associated with diseased lodgepole pine trees, particularly those having armillaria root disease (Tkacz and Schmitz 1986) and commandra rust (Rasmussen 1987). The role of these associations and other leads developed from initial investigations of low level MPB populations will be tested in a study of the dispersion of low level MPB populations on the Medicine Bow National Forest in southeastern Wyoming. This field effort will provide a framework for development of the ES.

This paper details the steps followed in the development of several expert systems for field testing:

1. Elicit from experts (scientists and managers) a specific objective for the ES model (example: to predict new tree kills for individual lodgepole pine stands.)
2. Elicit knowledge and professional judgment necessary for formulating one or more predictive models.
3. Elicit predictions for each combination of conditions represented in each model.
4. Elicit an estimate of certainty associated with each prediction.
5. Convert the conceptual ES model(s) into an operational computerized ES program.
6. Generate field data sheets for use in collecting appropriate data for ES.
7. Verify ES—see how well the predicted mimics real life.

OBJECTIVES

This work addresses the following objectives:

1. To develop a model or models that, in the opinion of one or more experts, reflect conditions contributing to the initiation of outbreaks from within stands with endemic populations of MPB.
2. To assess the role of the expert system in facilitating the generation, clarification, refinement, and verification of scientific knowledge.

KNOWLEDGE ACQUISITION

We acquire specific programming objectives, facts, hypothesized relationships among facts, predictions, and estimates of certainty of prediction by extracting information from published literature and from experts in the field. Because much of the knowledge used for successful decision making in natural resources has evolved through professional experience, and because ES programming allows us to incorporate this kind of qualitative, heuristic reasoning, we view skilled practitioners and researchers as major sources of meaningful information for model building. We capture their thought processes and their knowledge of facts and how they interrelate those facts. This enables us to explain within the program how predictions or conclusions were generated. Capturing knowledge from experts is difficult and time consuming and something commonly referred to as the "knowledge acquisition bottleneck." The main problem arises from the inability of experts to describe their own reasoning processes—they are not "expert" at articulating what they know (Shapiro 1987).

STRUCTURING KNOWLEDGE

While knowledge acquisition has received little systematic research (Hoffman 1987), our approach to cope with this problem has included: (1) interviewing experts about conditions under which we might expect MPB buildup to occur as well as when we would expect buildup not to occur; (2) watching them go about their work in the field; and (3) feeding information back to them in different formats such as flowcharts, maps, three-dimensional graphics, and summary printouts of their responses.

Two kinds of knowledge processing programs are being used. The first, a knowledge acquisition program written for this project, enables us to prototype and revise information rapidly as experts enter, reorganize, and refine their thoughts in response to feedback. (In future studies we hope to analyze various knowledge acquisition techniques and compare them with procedures being used in this investigation.) This approach allows us the means to gather knowledge from the expert(s) and to develop a sound knowledge base to be used in the expert system.

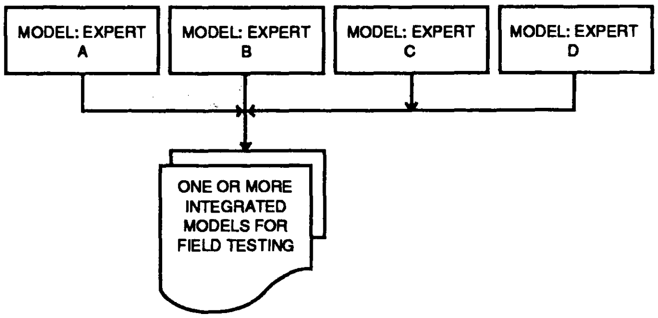

Figure 1—How models are integrated from different experts.

The second program, an applications expert system, is generated automatically by the first and contains the codified base of expert knowledge. Case data can be entered into the applications ES to generate predictions.

Because our efforts focus specifically on techniques for capturing scientific expertise in a form suitable for further scientific interchange, refinement, and field verification, our concern is with linking what is known with what is less certain as well as what is currently incapable of satisfactory explanation.

Moreover, we intend to use this process to capture the perspectives of several investigators and compare one model against another for similarities and differences in objectives, factors, factor categories, interrelationships, nature of predictions, and certainty levels. From this information one or more integrated models can be constructed easily in a form suitable for rigorous scientific verification (fig. 1).

SYSTEM DEVELOPMENT

The dynamics of MPB at endemic levels has received little attention, so our work is directed at capturing the experiences and judgments of scientists and pest management specialists who have significant experience in following the progress of MPB under epidemic conditions. Several applications expert systems have been built by working directly with three to four Forest Service researchers and university scientists in Utah and Colorado and two to three pest management specialists in Colorado, Montana, and Utah. Information was collected independently from each individual. The information from one was not shared with the others until all models had been captured, integrated to the extent possible, and sent as a complete set of protocols to the original participants and other specialists for review. This process was designed to elicit new perspectives as well as to provide a basis for "collective work" (Winograd and Flores 1987, p. 158) in support of focused scientific dialogue.

The interview-knowledge acquisition process used with the initial participants is currently being done as follows.

Interviewing, Observing, and Conceptual Diagramming

Each scientist and pest management specialist is briefed in general on what the model is being designed to explain or predict. The respondents are then asked to formulate a specific objective and list the kinds of information (factors) they would request before making a prediction about whether the MPB population (future tree kills or MPB population change) would increase, remain relatively static, or decline. Also elicited were appropriate response categories such as high, moderate, low, or isolated tree kills increasing to clumps of trees killed, clumps declining to scattered individuals, or attacking/emerging adult ratio trend.

Where factors themselves must be predicted, other variables are listed. For example, to estimate attack success last season, it may be necessary to make judgments about weather variables at flight time. Figure 2A represents the process of working backward from objective to factors and relationships during modeling sessions with the expert(s). Figure 2B illustrates a generalized flowchart view of the information captured. It is reviewed and refined repeatedly until it reflects the expert's views of processes at work. Other kinds of graphics (fig. 3) are used to trigger the expert's thought processes.

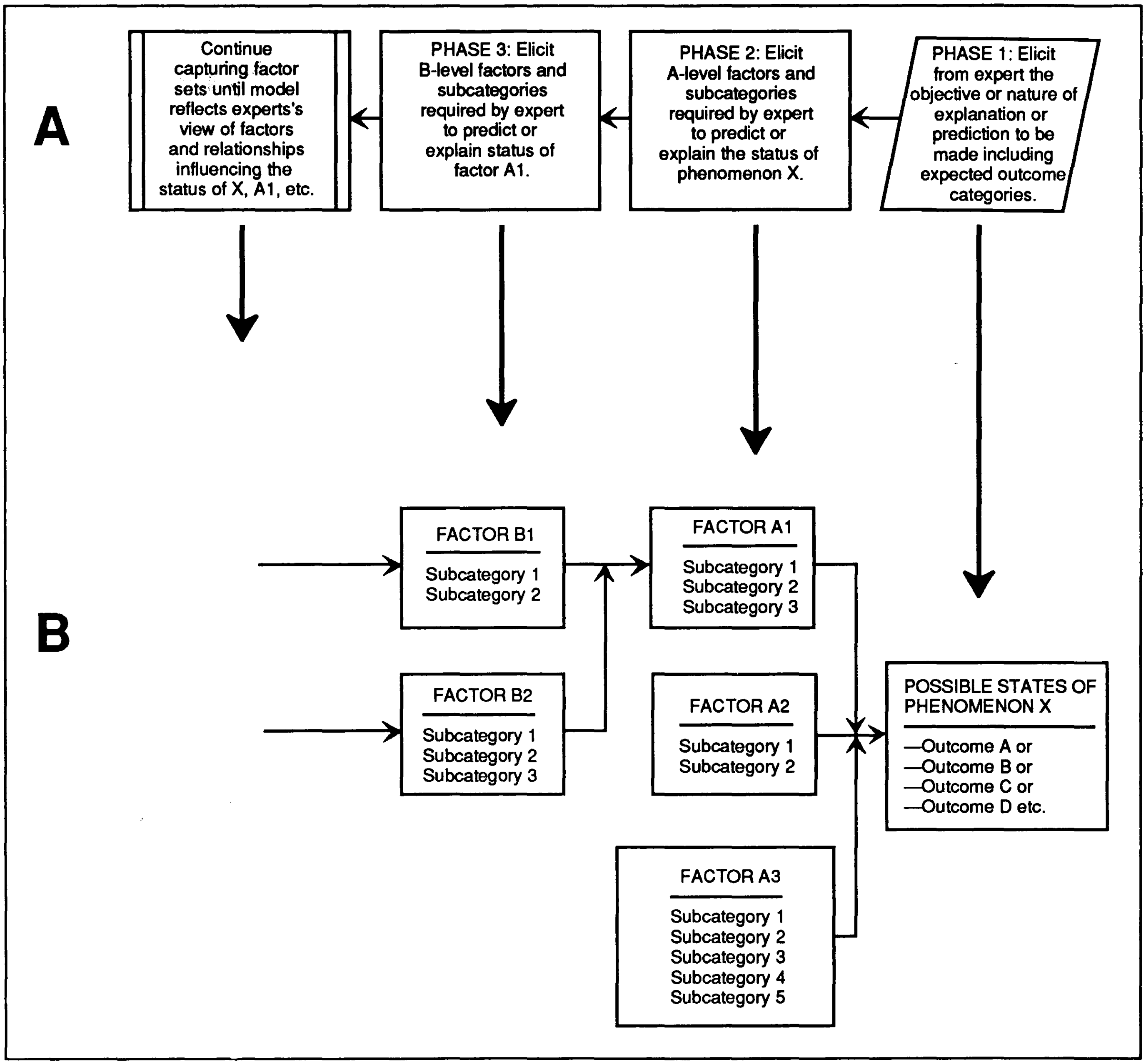

Figure 2—(A) Knowledge acquisition phases including factors and relationships. (B) Feedback to experts including the structure of the acquired knowledge base—factors and relationships.

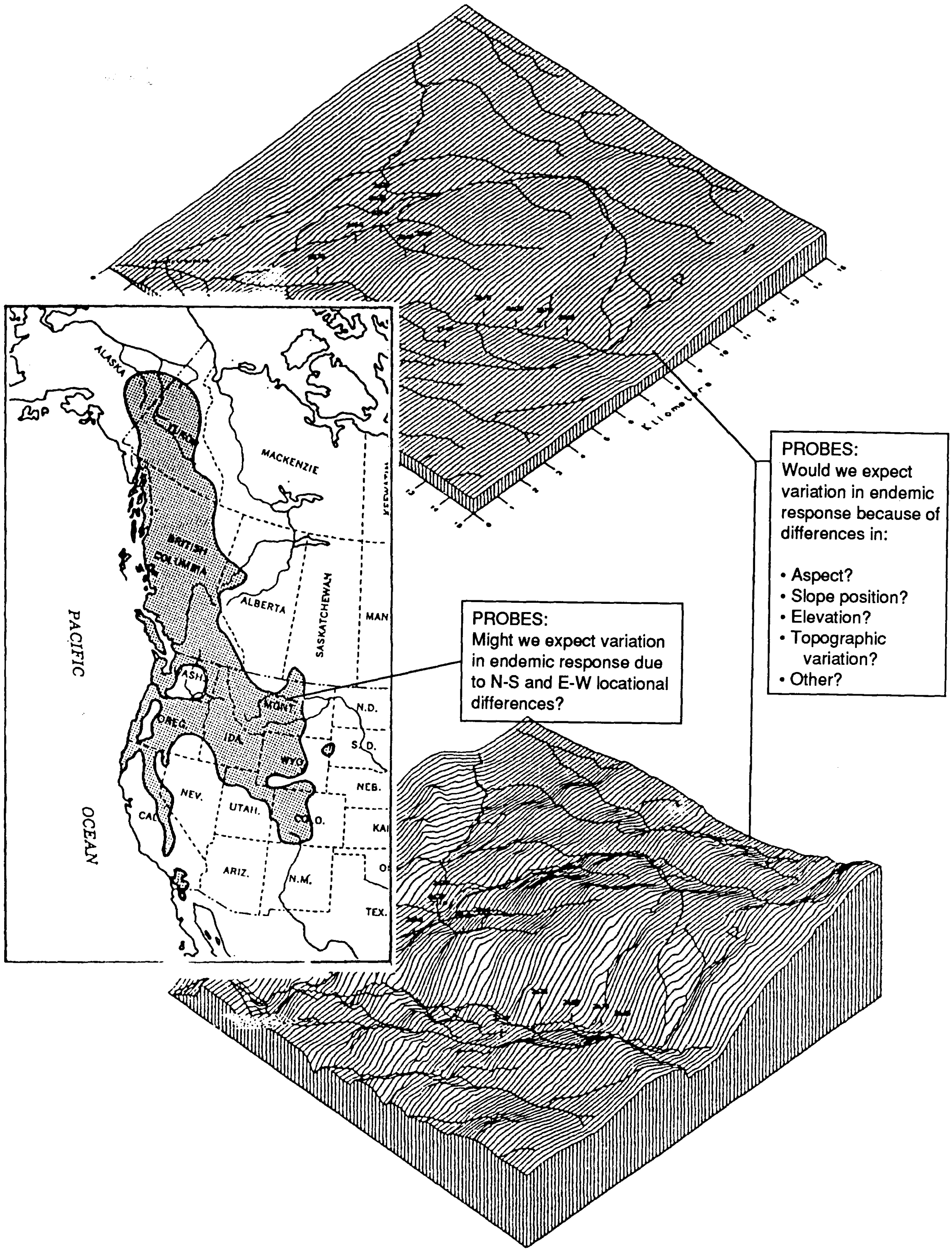

Figure 3—Graphics used to trigger expert's thought processes.

Generating Predictions for Factor Combinations

Once the diagram is finished, factors and subcategories are entered into the knowledge acquisition/processing program and a printout is generated listing all combinations of factors/subcategories given by the respondent (fig. 4).

The respondent then enters three types of information on the printout for each factor combination:

1. Predicted population change.
2. A numerical value that represents the degree of certainty the respondent has in the prediction.
3. One or more comments that qualify or otherwise explain the rationale for the prediction. Published reference citations that support or refute the prediction may be entered as well. Qualifying comments can be edited and added to the final applications program for recall at run time by users who want more information about the basis for questions asked the program.

Experience suggests it is easier for the respondent to enter responses directly on the printout rather than enter into the computer. Predictions, certainty factors, and comments can be quickly added to the program knowledge base later by the system designer.

Generating a Data Collection Form

A field data entry form is generated automatically for record entry. It mirrors the factors and factor categories reflected in the mental model as it has been designed by the respondent. As a result, provision is made to ensure that data required to verify the model(s) are collected.

```
=====================================================
OBJECTIVE: To predict/explain post-flight tree kill pattern.
-----------------------------------------------------
TOTAL RULES ENTERED: 4    RULE NUMBER: 1
PREDICTION: --> 1
Estimated certainty value [0-10 with 0 = none]: 2

          Case conditions
  ----------------------------
tree_kill_pattern_past_season      single trees widely scattered
over_wintering_brood_level         moderate
attack_success_at_flight_time      moderate

COMMENT 1: 3

COMMENT 2:

-----------------------------------------------------
=====================================================
OBJECTIVE: To predict/explain post-flight tree kill pattern.
-----------------------------------------------------
TOTAL RULES ENTERED: 4    RULE NUMBER: 2
PREDICTION: --> 1
Estimated certainty value [0-10 with 0 = none]: 2

          Case conditions
  ----------------------------
tree_kill_pattern_past_season      single trees widely scattered
over_wintering_brood_level         low
attack_success_at_flight_time      moderate

COMMENT 1: 3

COMMENT 2:

-----------------------------------------------------
```

Predictions, certainty values, and comments entered here.

Figure 4—Sample printout of knowledge acquisition/processing program that includes all combinations of factors/subcategories.

Generating Predictions From Field Data

Field data are processed through the applications ES, and expected outcomes or predictions are displayed. For research, the results can be matched against actual events—a key step in verification. For management, the predictions can be used to plan management actions.

RESULTS AND DISCUSSION

At this stage of the investigation we limit our description of results to two sets of general statements or propositions that seem to characterize the substantive and methodological findings. These propositions will be the focus of further investigation in both the substantive area of investigation (MPB dynamics) and knowledge acquisition methodology.

I. Behavior of Endemic MPB Populations in Lodgepole Pine

1. There are areas of agreement as well as disagreement among the researchers and specialists we have interviewed regarding priorities for investigation (research objectives), factors, interrelationships, predicted outcomes, and levels of certainty. Models being charted and programmed will provide a framework for focused discussion on priorities for further investigation as well as strategies for testing and validation.

2. A scientific model of population dynamics will differ from a management-oriented model. The scientific model that emerges will be a more complete description of how endemic populations change over time in response to short-term as well as long-term population, stand, and weather factors. The management model will emphasize long-term factors, particularly projected stand conditions, because of the need to plan remedial actions 3 to 6 years ahead of actual events. Short-term variation in weather, tree, or stand factors is likely to be less relevant in a management context.

II. Perceived Contribution of Expert Systems Development in the Scientific Process

1. There are competing models for endemic population dynamics, and this variation in perspective serves as a barrier to the interchange of ideas. Emphasizing one model and discrediting the possibilities offered by others complicates the research process. However, the process being used in this investigation promotes communication among researchers and yields a final model that is better because of the different perspectives contributed by each participant.

2. The possibility of rapid update, refinement, expression of certainty level, and guidance for future verification that is designed into the knowledge processing system makes professional speculation within a scientific context a tolerable activity. Scientists are willing to describe their views and to offer predictions even though they must extrapolate from related but not identical experiences or systematic investigations.

3. The potential for simultaneous testing of competing models is a positive prospect. Scientists are showing interest in how the models of their colleagues will compare. Where differences exist, alternative ES models can be constructed and tested simultaneously. It is not necessary to achieve consensus on one model initially. Rather, the objective is to formulate competing models so they can be subjected to valid testing.

4. Responses from experts to feedback during system development fall into several categories: (a) substance of the feedback is incorrect—the system developer misinterpreted the information; (b) feedback causes the respondent to think about additional factor combinations and suggests other avenues for consideration; (c) advice is needed from other specialists before a response can be made; (d) request for a prediction must be "cannot predict" because the combinations of factors extend beyond our current level of understanding. This modeling process helps identify topics that may be worthy of new research.

5. Information entered early in the process often appears faulty to experts when it is fed back in a different form. Rapid prototyping and feedback causes respondents to modify factors, interrelationships, and even the statement of objectives.

6. This modeling strategy facilitates interdisciplinary research. Individual scientists suggest that some of the uncertainty in their own models might be reduced by asking advice of specialists in other fields. For example, information from a silviculturist or pathologist may help refine a critical factor that an entomologist believes to be operating.

7. This modeling strategy provides a framework for integrated data collection. Both scientists and management specialists have observed that data collection for research and management decision making can be more closely coordinated where the model(s) in need of verification is clearly diagrammed.

8. Indirect feedback from respondents is of importance to the research coordinator and system designer. Examples include: (a) "this is too big a problem to handle easily—it should be stratified into smaller pieces"; and (b) "if I am asked to contribute so much time, I should have joint authorship on the resultant paper." Comments such as these lead to other, subset, propositions:

A. Knowledge acquisition is complex and time consuming. It places a time demand on the expert that is considerable. Automated knowledge acquisition procedures will reduce the impacts on participants.

B. Knowledge acquired through years of work and observation is valuable. The willingness of experts to share that knowledge must be publicly acknowledged and rewarded.

CONCLUSIONS

Our work is directed at (1) developing a model or models that reflect conditions contributing to changes in endemic populations of mountain pine beetle and (2) establishing a foundation for assessing the role that ES programming may play in furthering the generation, clarification, refinement, and verification of scientific knowledge.

We have briefly discussed two sets of propositions that have surfaced from our experience with knowledge acquisition and model building in a domain of uncertain and incomplete scientific knowledge and speculation.

We are currently at the stage where the models from different experts are being integrated. This task is being done independently by several investigators, including several that were not involved in the first phase of knowledge acquisition. The resulting integrated model(s) will be distributed to other MPB specialists throughout the West for review and comment.

It is now appropriate to explore opportunities for coordinating and modifying field inventory work already planned by research and management in order to collect additional data required for testing the model(s). The results will be the subject of future technical reports.

Research will also continue on knowledge acquisition. Procedures developed in this study are being refined, the programs are being revised, and we are applying them to other natural resource problems. Of particular interest is the development of ES packages that can be "localized" by management specialists to reflect variation in environmental, managerial, social, or economic condition peculiar to a specific area.

ACKNOWLEDGMENTS

The authors wish to acknowledge the willingness of the following individuals to participate in the knowledge acquisition process. Their time, effort, knowledge, and judgment were critical to the attainment of project objectives: G. Amman, D. Holland, M. Jenkins, K. Lister, K. Gibson, J. Long, L. Rasmussen, J. Schmid, R. Schmitz.

REFERENCES

Amman, G. D.; Cole, W. E. 1983. Mountain pine beetle dynamics in lodgepole pine forests. Part II: population dynamics. Gen. Tech. Rep. INT-145. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 59 p.

Cole, W. E.; Amman, G. D. 1980. Mountain pine beetle dynamics in lodgepole pine forests. Part I: course of an infestation. Gen. Tech. Rep. INT-89. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 56 p.

Cole, W. E.; Amman, G. D.; Jensen, C. E. 1985. Mountain pine beetle dynamics in lodgepole pine forests. Part III: sampling and modeling of mountain pine beetle populations. Gen. Tech. Rep. INT-188. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 46 p.

Hoffman, Robert R. 1987. The problem of extracting the knowledge of experts. AI Magazine. 8(2): 53-67.

Negotia, C. V. 1985. Expert systems and fuzzy systems. Menlo Park, CA: Benjamin/Cummings Publishing Company, Inc. 190 p.

Rasmussen, L. A. 1987. Mountain pine beetle selection of dwarf mistletoe and commandra blister rust infested lodgepole pine. Res. Note INT-367. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 3 p.

Shapiro, A. D. 1987. Structured induction in expert systems. New York: Addison-Wesley. 143 p.

Stock, M. 1987. AI and expert systems: an overview. AI Applications in Natural Resource Management. 1(1): 9-17.

Tkacz, B. M.; Schmitz, R. F. 1986. Association of an endemic mountain pine beetle population with lodgepole pine infested by armillaria root disease in Utah. Res. Note INT-353. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 7 p.

Waterman, D. A. 1986. A guide to expert systems. New York: Addison-Wesley. 419 p.

Winograd, T.; Flores, F. 1987. Understanding computers and cognition. New York: Addison-Wesley. 207 p.

Intermountain Research Station
324 25th Street
Ogden, UT 84401