Articles

Comax: An Expert System for Cotton Crop Management

HAL LEMMON

Expert systems are computer programs that perform at the level of human experts. One expert system, Comax, has been developed that acts as an expert in cotton crop management. The system has a knowledge base consisting of a sophisticated cotton plant simulation computer program, a set of "if-then" rules, and a computer program called an inference engine. Comax determines the best strategy for irrigating, applying fertilizer, and applying defoliants and cotton boll openers. Sensors in the cotton fields automatically report weather conditions to the system, and Comax reevaluates its recommendations daily. Comax was tested on a large farm and demonstrated excellent results in reducing the unit costs of production.

TODAY THREE BALES OF SYNTHETIC FIBERS ARE MILLED FOR every bale of cotton. Further, the synthetic fiber industry has recently adopted a vigorous research program to produce fibers at still lower cost. For cotton to survive, research to lower production costs is imperative (*1*).

An expert system, Comax (COtton MAnagement eXpert), has been developed that advises cotton growers on crop management at the farm level. The expert system is integrated with a computer model, Gossym (from *Gossypium* and simulation), that simulates the growth of the cotton plant (*2*). This is the first integration of an expert system with a simulation model for daily use in farm management.

Gossym

Researchers began developing Gossym in 1973. The program was developed over 12 years with contributions from ten scientists at four institutions (*3*) in two countries. It simulates the growth and development of the entire cotton plant on an organ-by-organ basis: roots, stems, leaves, blooms, squares, and bolls. It also simulates soil processes such as the transfer of water and nutrients through the soil profile. For Gossym to accomplish this, it needs data from mechanical and chemical soil analyses of the farm field to which it is being applied. Such analyses can be performed by state-owned soil test laboratories, the Soil Conservation Service, or commercial laboratories. The specific data required are soil hydrologic properties, soil fertility, soil impedance (resistance to root growth), water release curves, and bulk density.

The model is driven by weather variables. It requires, on a daily basis, such data as the maximum and minimum temperatures, solar radiation, and rainfall. It was developed with SPAR (Soil-Plant-Atmosphere-Research) units, where cotton is grown under highly controlled conditions and the various rate processes can be determined, but it was extensively tested and validated against field data.

Gossym is capable of running on most computers, including microcomputers. A complete simulation, from emergence to harvest, can be done in 6 to 8 minutes on a VAX 750 computer, in 60 to 90 minutes on a microcomputer (an IBM PC, or equivalent, with a math coprocessor), and in 20 to 30 minutes on an advanced microcomputer (an IBM PC-AT, or equivalent, with a math coprocessor).

The development of microcomputers has expedited the movement of Gossym to the farm to assist in crop management. In 1984 a project to use Gossym on cotton farms was initiated by the USDA Agricultural Research Service in cooperation with the National Cotton Council, and microcomputers were provided for a 6000-acre farm in the Mississippi Delta (*4*) and a 1000-acre farm in the South Carolina Coastal Plain (*5*). In 1985 Comax was tested on the 6000-acre farm.

In the research laboratory, a multidisciplinary team of cotton experts provides Gossym with input and interprets its output. Comax was developed to provide the input and to perform the analyses when Gossym is used for practical, on-farm decision making. This is the first attempt I am aware of to integrate an expert system with a simulation model with the objective of optimizing crop production.

Comax

An expert system is a computer system with the capability of performing at the level of human experts in some particular domain. It is possible to build expert systems that perform at remarkable levels (*6*). While there are several methods for designing expert systems, rule-based systems have emerged as the popular architecture. Deriving their knowledge from relatively easily understood facts and rules, rule-based systems offer surprising power and versatility (*7*).

Comax is a rule-based expert system that operates Gossym the way a human expert would to determine three factors: irrigation schedules, nitrogen requirements, and the crop maturity date.

As shown in Fig. 1, Comax consists of a knowledge base, an inference engine, Gossym, a weather station, and data (for example, the seeding rate and soil parameters). The knowledge base is a set of rules and facts written in near-English. The inference engine examines the rules and facts to determine what is to be done. It prepares data files accordingly to hypothesize the weather and to hypothesize applications of water and nitrogen. Then it calls Gossym, which reads the data files prepared by the inference engine and simulates

The author is a computer scientist for the USDA Agricultural Research Service, Mississippi State, MS 39762.

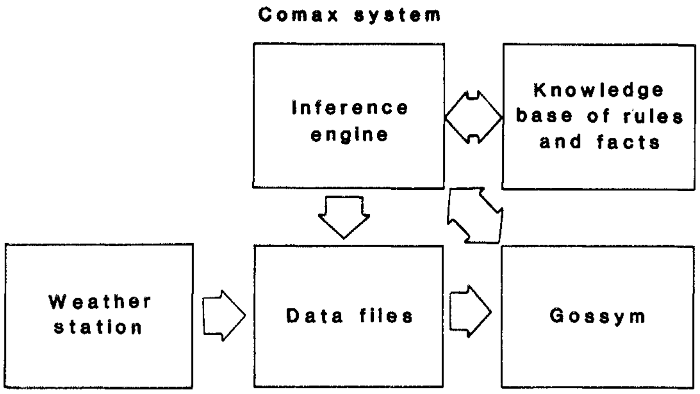

Fig. 1. The Comax components. The four components to the right reside in a microcomputer located at the grower's farm.

the growth of the cotton plant under the conditions specified in those files. Results from the simulation (such as the day the simulated crop goes into water stress) are saved as facts in the knowledge base.

The inference engine program and the Gossym program change little if at all. The knowledge base continuously changes as researchers and growers improve management strategies or observe the impact of different strategies.

Software, Hardware, and Data

The software components of Comax are the inference engine and Gossym. The inference engine is written in the LISP computer language, and Gossym is written in FORTRAN. The computer languages were selected on the basis of appropriateness for the task to be performed, LISP being appropriate for an expert system but inappropriate for simulation. The knowledge base, so far, has about 50 rules, the inference engine about 6000 lines of code, and Gossym about 3000 lines of code.

Comax was developed on a Symbolics 3670 computer and is down-loaded, unchanged, to the PC computers where it runs under Common LISP, offered by Gold Hill Computers. Gossym was developed on the VAX 750 computer and is also down-loaded, unchanged, to the PC computers and compiled using the FORTRAN 77 compiler offered by Ryan-McFarland.

The cotton grower who used Comax has a microcomputer (an IBM PC or equivalent) with a math coprocessor and a dot-matrix printer in his office. The system can automatically call the weather station daily by telephone but, if a phone line is not practical, the data may be entered into the computer manually. The microcomputer costs \$4000 to \$7000, depending on the configuration selected. The cost of the weather station is \$4000, which includes solar panels to provide power. Hardware for telephone connection is \$1200.

Comax Rules

Figure 2 shows some of the facts and one of the rules used in Comax. This rule, "find-water-stress-day," is one of the set of rules used to determine the optimum irrigation schedule. The rule is true if every term in the "if" part of the rule matches a term in the facts base. In this case, (run-number *?number*) of the rule matches the fact (run-number 1) if *?number* is assigned the value 1, and (hypothesized-weather *?weather*) matches the fact (hypothesized-weather hot-dry) if *?weather* is assigned the value hot-dry. Entries that begin with a question mark, such as *?number*, are treated as variables by the inference engine and are assigned values, as needed, to cause a match.

In the case shown in Fig. 2, the rule is true, and the inference engine will proceed with the actions in the "then" part of the rule. It first prints on the computer screen a message describing the action. Next, it runs the Gossym program using the hot-dry weather scenario. When Gossym is finished, the inference engine examines the results of the run and places new facts into the facts base. One of the new facts will be, for example, (w-stress-day 236), where 236 represents the day of the year the crop went into water stress.

The final action of the inference engine is to assert a new fact, (set-hypothesis-irrigation), into the facts base. The purpose of this new fact is to cause another rule, which is called "set-up-hypothesized-irrigation" and is not shown in the figure, to be true. That rule, a lengthy one, determines the day that irrigation should be applied. Conceptually, it does this by taking the water stress day, subtracting the application time given in the fact (irrigation application-time 4), determining the amount of water to be applied from the fact (irrigation amount 1), and asserting a new fact (hypothesized-irrigation 232 1). However, there are actually other considerations, such as how soon to harvest and how many days since the last irrigation, which this rule also considers.

Comax recomputes the optimum management scenario each day, prints a daily report that recommends crop management procedures and, if it is desired, summarizes the intermediate simulations to explain the basis for the recommendations. Comax can show the results of simulations either by tabular reports or by graphs on the dot-matrix printer.

Operating Comax on the Farm

Comax is designed to run continuously throughout the crop year on a dedicated microcomputer. Each day it computes the expected irrigation date, the expected date and amount of fertilization, and the expected date of crop maturity. These are computed daily because, as the hypothesized weather for each day is replaced by the actual weather for that day, the computed dates change.

Determining irrigation requirements. Comax begins each day by determining the expected irrigation date. It does this by running Gossym with a hypothesized weather scenario, noting the date the crop goes into water stress and subtracting the number of days it takes to apply the irrigation. Some irrigation systems, the center-pivot type, for example, take several days to apply water. Comax uses three different types of hypothesized weather scenarios: (i) normal weather, (ii) hot-dry weather, and (iii) cold-wet weather. The weather scenarios are specific to each farm. Comax first runs Gossym with the hypothesized hot-dry weather scenario. This establishes the earliest date that irrigation would be required. Comax then runs Gossym with the normal weather scenario to determine the most likely date that irrigation will be required. The results are presented in a report printed at the end of the daily Comax operation.

The report states, for example, that today is 1 July and irrigation will be required on 10 July if subsequent weather is hot and dry or on 17 July if subsequent weather is normal. The next day, 2 July, the hypothesized weather for 1 July is replaced with the actual weather for 1 July, and the irrigation requirement is redetermined. If 1 July was a cold and wet day, the new report may state that irrigation is required on 12 July if subsequent weather is hot and dry (instead of 10 July as reported the day before) or on 19 July if the subsequent weather is normal (instead of 17 July). Conversely, if 1 July is actually a hot and dry day, the irrigation date for hot-dry weather will still be 10 July, but the irrigation date for the normal weather hypothesis will be earlier, perhaps 15 July instead of 17 July.

Determining nitrogen requirements. With cotton, it is important not to overfertilize, not only because of the obvious economic waste but also because overfertilization can cause the plant to be in an undesirable state at time of harvest. To determine the nitrogen requirements, Comax first ensures that there is no water stress by calculating an additional series of irrigation dates. After each calculation Comax determines the day the simulated crop went into water stress and, on the basis of the assumption that the grower would irrigate to relieve that stress, it hypothesizes a date and amount of irrigation. It then runs Gossym again to determine the next date that the crop will be in water stress. This process is repeated until the end of the season is reached, and the result is an hypothesized irrigation schedule that should prevent the crop from ever being in water stress. This schedule is only for use in determining nitrogen requirements and is never followed. The actual irrigation schedule to be followed is determined as described in the previous section.

Comax is now ready to determine the minimum amount of nitrogen that can be safely applied. It does so by making a series of Gossym runs with the cold-wet weather scenario, to simulate the minimum plant growth and thus to estimate the minimum nitrogen requirement. Comax again makes a series of these Gossym runs and, after each run, the day the crop went into nitrogen stress is noted. Comax then enters into the calculation a predetermined amount of nitrogen, and runs Gossym again. If nitrogen stress occurs again, the amount of nitrogen hypothesized is increased. When too much nitrogen is applied, there will be an undesirable effect: after the bolls are mature, the plant will begin to grow vigorously. If such undesirable growth (shown in Fig. 3, row 4, third graph) occurs, Comax reduces the amount of nitrogen. This process is repeated until Comax has determined the amount of nitrogen just sufficient to relieve nitrogen stress. This value is printed in the Comax daily report and represents the minimum amount of nitrogen the grower should apply.

The process is repeated with the normal weather scenario. This tells the grower the most probable nitrogen requirement. Finally, the process is repeated a third time with the hot-dry weather scenario, and the result tells the grower the maximum nitrogen requirement. From these three figures and from his own assessment of the weather the grower decides the amount of nitrogen to apply.

The grower's safest strategy is to assume the cold-wet weather scenario will hold and apply the minimum amount of nitrogen. If the weather turns out to be better than this, the grower can apply additional amounts of nitrogen later in the season. The penalty for underestimating the nitrogen requirement is only the cost of applying the additional nitrogen. The penalty for overestimating the nitrogen requirement is the cost of the excess nitrogen plus, at harvest, the loss from its undesirable effects, which can be substantial.

There is an additional risk that nitrogen applied too early in the season can be lost because of leaching. Such a loss varies with soil conditions, rainfall, and irrigation. Gossym is capable of identifying the amount of nitrogen lost in this way.

Farms that do not have irrigation systems are handled in a different, simpler manner. Farms with trickle irrigation require a different set of rules, a problem which will be addressed this year.

Determining harvest date. Comax also informs the grower when the cotton is mature so he can apply defoliants and boll openers. This is particularly important in such locations as the Mississippi Delta, where early rains can physically damage the cotton, induce boll rot, and make the ground so muddy that the mechanical cotton pickers cannot operate. Near the end of each season the grower must decide either to wait until it is certain the cotton has reached its maximum yield or to proceed with the harvest before the rains begin. With Comax, the farmer knows weeks in advance when his crop will mature. This can only be an approximation because of uncertainty in the weather; but as each day passes, the hypothesized weather is replaced by the actual weather, and the projected maturity date becomes more reliable.

```
FACTS
  (run-number 1)
  (hypothesized-weather hot-dry)
  (irrigation amount 1)
  (irrigation application-time 4)
  ...
RULE find-water-stress-day
 IF
   (run-number ?number)
   (hypothesized-weather ?weather)
 THEN
   (printout "Finding water stress day")
   (run-gossym ?number ?weather)
   (assert (set-hypothesized-irrigation)
```

Fig. 2. Four of the facts and one of the rules used in Comax. The rules are discussed in the text.

Comax in operation. An example of the operation of Comax as it selects nitrogen and irrigation schedules is shown in Fig. 3. The graphs in each row are the results of a Gossym simulation run by Comax. In the first graph of each row, the circles represent nitrogen applications. The first three applications are actual, but the fourth application (on the first graph of rows 3, 4, and 5) is hypothesized by Comax. On this farm the grower has applied 55, 60, and 30 pounds of nitrogen per acre at the time of planting and at 33 and 63 days after the plants emerged, respectively. The line shows the nitrogen stress, computed as the ratio of the nitrogen used to the nitrogen needed by the plant for full growth of all organs. In the second graph of each row, the jagged line represents a measure of water stress in the plant, and the vertical bars indicate the amount of water applied or that is expected to be applied by either rain or irrigation. The third graph of each row shows the height of the plant, the number of squares (unpollinated flower buds), and the number of bolls. The number of squares increases with time and then decreases as some squares are shed (because of stress) and others turn to bolls. The fourth graph of each row shows the development of the predicted yield. The final yield, in bales per acre, is printed above the curve.

The first row of graphs were produced by Comax just after the third application of nitrogen. The second row of graphs is the last of a series of Gossym runs in which Comax has directed its attention to the water stress problem and hypothesized a heavier irrigation schedule with no additional nitrogen. The second graph of this row shows that increased irrigation resulted in reduced water stress and in intensified nitrogen stress. With increased water, the simulated plant has the capacity for increased growth, and therefore it needs even more nitrogen. Even though irrigation is increased, there is no increased yield.

In the third row, Comax has hypothesized an application of 30 pounds of nitrogen per acre. The nitrogen stress is reduced, and the yield is increased.

In the fourth row, Comax has hypothesized an additional 60 pounds of nitrogen per acre. The nitrogen stress is eliminated, and the yield has increased correspondingly. However, the third graph of this row shows that, after the bolls have all matured, the cotton plant has had a spurt of new growth and that it has started adding new squares that will never mature. At the point where the yield levels off, the crop should be harvested since no more cotton would be expected and delay would increase the risk of harvest losses due to inclement weather. To harvest cotton with modern equipment, it is necessary to apply a defoliant; however, this model plant would be so robust that the defoliant would not be as effective as it should be. The rules of Comax will cause this hypothesis to be rejected.

In the last row, Comax has selected 40 pounds of nitrogen per acre in conjunction with the indicated irrigation applications. This provides the maximum yield subject to the constraint of no secondary growth.

Constraints, such as irrigation capacity and the time required to irrigate, are provided for in the knowledge base. For example, on a field with pivot irrigation a typical constraint may be that 1 inch of water can be applied in 4 days. Constraints are considered on a farm-by-farm basis; as a consequence, the knowledge base varies somewhat from farm to farm.

Results from a Pilot Test

Comax was tested on the Mitchener farm (*4*) so that we could acquire experience in its practical operation under realistic conditions (*8*). In mid-July 1985 Comax predicted the need for nitrogen at the rate of 50 pounds per acre, as shown in the last row of Fig. 3. As a result, the grower, who had not planned to apply any additional nitrogen, applied 20 pounds per acre throughout the farm except on a 6-acre test plot where no nitrogen was applied on alternate eight-

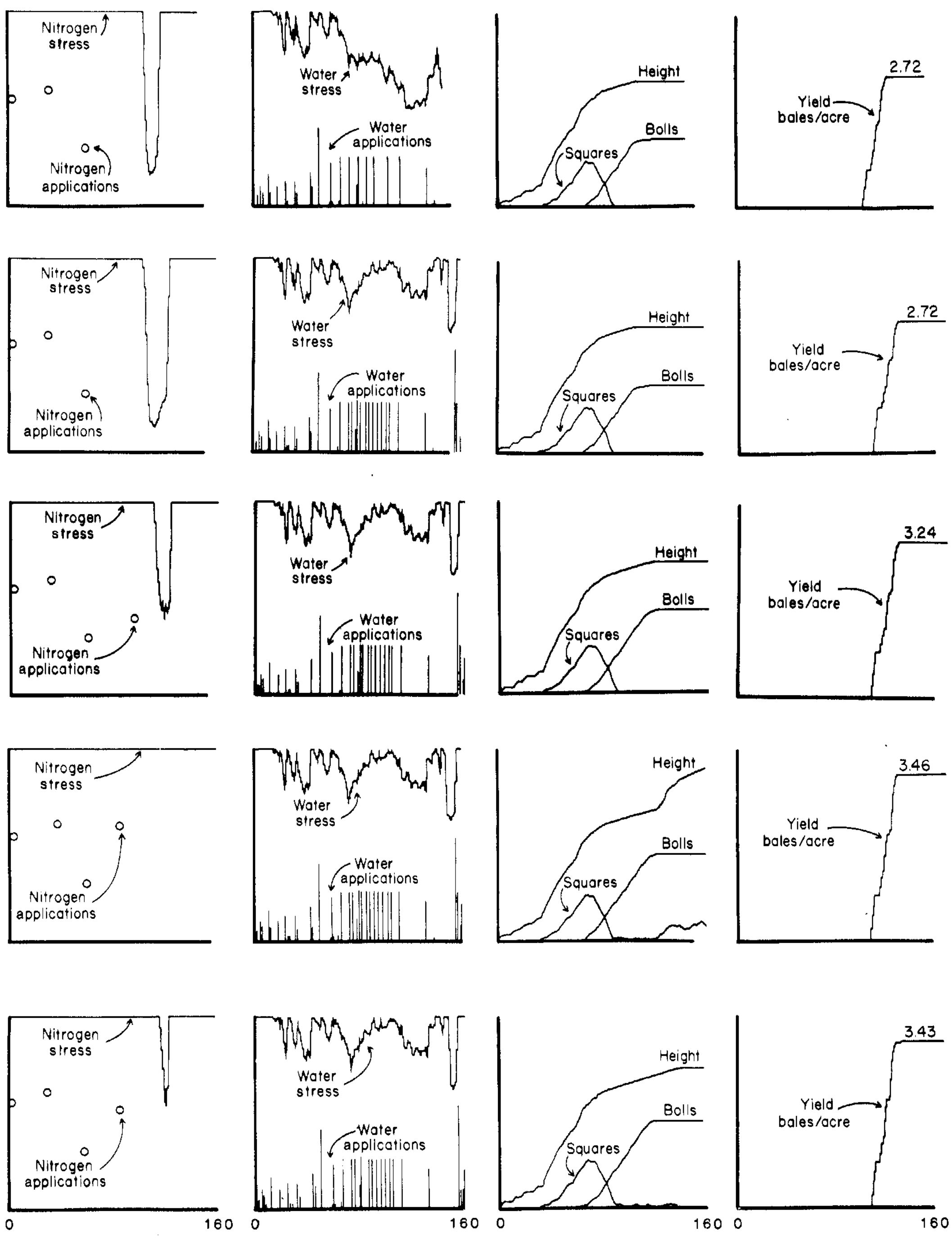

Fig. 3. Graphs produced by Comax from the results of Gossym simulations, showing the process whereby Comax reduces the water stress and then the nitrogen stress, as described in the text.

row strips. Comax predicted an additional 200 pounds of cotton lint on the cotton treated with nitrogen, with no delay in the date of maturity. At the end of the season, the test plots were picked, some by hand and some by machine. Although cotton is no longer picked by hand for commercial purposes, some rows of the test plot were so picked to obtain a precise figure to compare with the yield predicted by Comax. The hand-picked rows showed a net increase of 180 pounds per acre of cotton, and the machine-picked rows a net increase of 115 pounds per acre. The additional cotton (machine-picked) had an economic value of about \$71 per acre, the cost of the nitrogen was \$4 per acre, and the cost of application was \$5 per acre. Allowing for the cost of processing the additional cotton, there was a net gain of over \$60 per acre on this 6000-acre farm.

The grower believes, however, that it is the system's ability to pinpoint the day the crop is mature that is its most valuable feature. In the previous year (1984), the system predicted a maturity date of 1 September for the crop. Instead, the grower elected to use the widely accepted rule that a crop is not mature until 60% of the bolls are open and delayed harvesting until 21 September. Rain began on 6 October, and it was not possible to complete the harvest until November, which resulted in a loss of both yield and quality. The grower now believes that the maturity date of 1 September was correct and that, if the harvest had begun on that date, cotton production would have increased by approximately 4.3 million pounds and the quality would have been improved by an amount worth an additional \$0.11 per pound.

Future Outlook

During the coming crop year (1986), testing and development of Comax is continuing with 15 growers in five states and with a total cultivation of over 50,000 acres of cotton.

In the United States, there are 10 to 12 million acres (varying from year to year) of cotton on 30,000 farms. Approximately 1300 farms (4%) are of 1000 acres or more and account for 33% of the cotton, whereas 4000 farms are of 500 acres or more and account for 58% of the production (*9*). The former are obvious candidates for Comax; the latter are probable candidates.

REFERENCES AND NOTES

1. R. J. Kohel and C. F. Lewis, Eds., *Cotton* (American Society of Agronomy, Madison, WI, 1984), pp. xi–xii.
2. D. N. Baker, J. R. Lambert, J. M. McKinion, *S.C. Agric. Exp. Stn. Tech. Bull. 1089* (December 1983).
3. The scientists and their institutions are: D. N. Baker, J. M. McKinion, R. E. Fye, USDA Agricultural Research Service, Mississippi State, MS; F. D. Whisler, J. A. Landivar, D. R. Reddy, S. G. Kharche, A. Ben-Porath, Mississippi State University, Mississippi State, MS; J. R. Lambert, Clemson University, Clemson, SC; and A. Marani, Hebrew University of Jerusalem, Rehovot, Israel.
4. Mitchener farm, Sumner, MS.
5. McCoy farm, Sumter, SC.
6. R. O. Duda and E. H. Shortliffe, *Science* **220**, 261 (1983).
7. R. Davis, *ibid.* **231**, 957 (1986).
8. F. Mitchener, *Cotton Grower* 2, 42 (1986).
9. A. Jordan, National Cotton Council, Memphis, TN, personal communication.

10. I thank D. N. Baker, A. R. Grable, J. M. McKinion, and G. S. Hasegawa for their critical reading of the manuscript and for their many helpful suggestions.