INEEL/CON-01-00985
PREPRINT

From Prediction To Prescription: Intelligent Decision Support For Variable Rate Fertilization

Raymond K. Fink
Reed L. Hoskinson
J. Richard Hess

July 30, 2001

International Conference Of American Society Of Agricultural Engineers

Idaho National Engineering and Environmental Laboratory

From Prediction to Prescription: Intelligent Decision Support for Variable Rate Fertilization

Raymond K. Fink

Idaho National Engineering and Environmental Laboratory
P.O. Box 1625, Idaho Falls, ID 83415
rkf@inel.gov

Reed L. Hoskinson

Idaho National Engineering and Environmental Laboratory
P.O. Box 1625, Idaho Falls, ID 83415

J. Richard Hess

Idaho National Engineering and Environmental Laboratory
P.O. Box 1625, Idaho Falls, ID 83415

Written for presentation at the
2001 ASAE Annual International Meeting
Sponsored by ASAE
Sacramento Convention Center
Sacramento, California, USA
July 30-August 1, 2001

Abstract. *We describe the use of machine learning methods in the analysis of spatial soil fertility, soil physical characteristics, and yield data, with a particular objective of determining local (field- to farm-scale) crop response patterns. For effective prescriptive use, the output of these tools is augmented with economic data and operational constraints, and recast as a rule-based decision support tool to maximize economic return in variable rate fertilization systems. We describe some of the practical issues addressed in development of one such system, including data preparation, adaptation of regression tree output for use in a rule-based expert system, and incorporation of real-world limits on system recommendations. Results from various field trials of this system are summarized.*

Keywords. Computer techniques, decision making, fertilization, optimization, precision agriculture, site specific, statistical analysis

Introduction

Historically, most soil use and crop management decisions were based on the best available information, which was usually whole field or farm average yield. But over the past 10 years, producers have become overwhelmed with information (e.g., yield maps, digital soils maps, USGS digital elevation models, high precision elevation models, and remote sensed imagery). There is tremendous spatial variation in crop productivity, soil physical, and soil chemical properties. This variability is easy to recognize but very difficult to manage due to complex interactions.

This information, coupled with advances in GIS and GPS technology, yield monitors, and variable rate application technologies, can provide the basis for more effective spatial management of crop fertilization. In particular, methods from artificial intelligence and machine learning can provide a means for improved economic application of fertilizers.

Prediction Methods

We consider the case where the producer has site-specific data from previous growing seasons. For example, field nutrient information is obtained from standard soil sampling and analysis techniques, and crop yield data can be obtained using commercially available crop yield monitors and recorders. This historical site-specific data can be used to determine local (field- to farm-scale) crop response patterns.

A wide variety of statistical techniques can be used for analysis of this type of input/response data. In our experience, simple linear regression methods perform very poorly at capturing the complex nonlinear interactions inherent in crop response to spatial soil fertility and physical characteristics.

We describe the use of regression tree methods in the analysis of spatial soil fertility, soil physical characteristics, and yield data, with a particular objective of determining local (field- to farm-scale) crop response patterns.

Regression trees

Regression trees are a method for inductive learning in multivariate data sets. The INEEL Decision Support System for Agriculture (DSS4Ag) uses the CART regression tree algorithm (Breiman, et al. 1984). This is a tree-structured classification method based upon a recursive-partitioning algorithm. At each step of the analysis, every independent variable is examined in order to identify a splitting condition that maximizes the separation, or purity, of the two resulting subsets of the data. This step is repeated recursively for each subset until no more statistically useful splits can be identified. The result of the regression tree algorithm is a binary decision tree that can be used to predict a resulting crop yield for any combination of specified inputs. Each terminal node of the tree represents the outcome that is associated with a specific set of input conditions; these conditions are obtained by traversing the tree from the root node down through the branch nodes. Figure 1 shows a portion of one such tree.

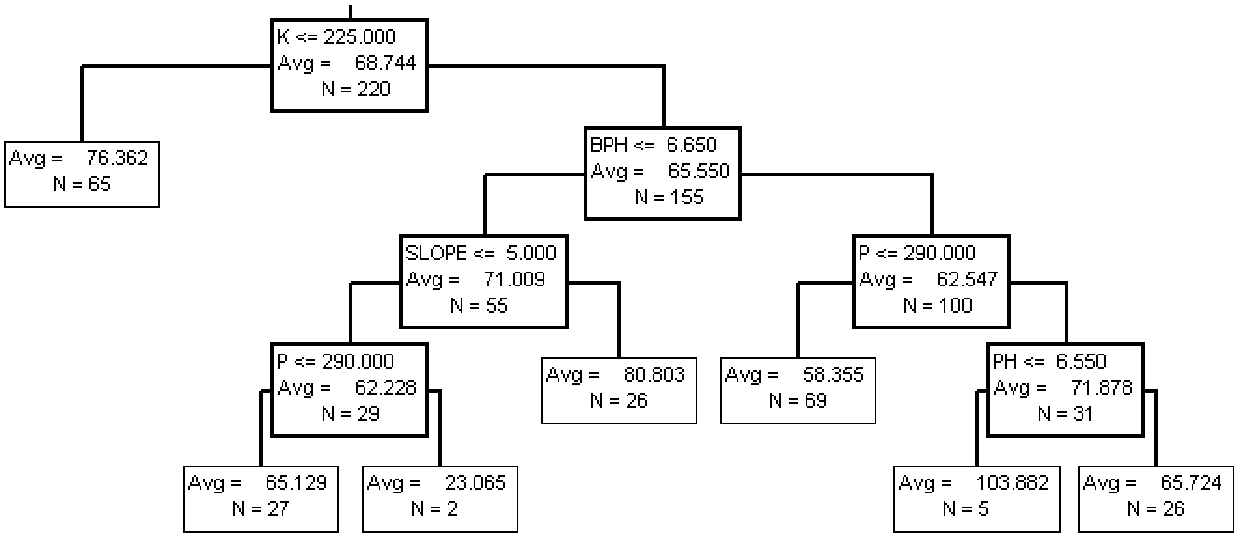

Figure 1. Sample output (partial) from CART algorithm

In practice we have used as many as 18 soil parameters (chemical and physical properties), slope and aspect, and selected bands from satellite imagery data. These input parameters and resulting yield data for a selected field are gridded using standard GIS tools; we have chosen grid sizes in the range of 10 to 30 meters square. These sizes are chosen to correspond to the path width of current variable-rate application equipment. This results in a few thousand input records for a typical field. The resulting trees typically have 100 to 200 terminal nodes.

The CART algorithm examines every unique value of each independent variable for every record in the applicable subset at a branch node. In order to reduce the computational overhead associated with this, we round off the independent variable values to reduce the number of distinct values. The rounding limits are chosen based upon measurement accuracy and estimated clinical significance for each parameter. The effect of this rounding step is to dramatically reduce the computation times required for a solution, without adversely affecting the quality of the solution.

Prescription Methods

The methods described above are predictive tools; for effective prescriptive use, the output of these tools is augmented with economic data and operational constraints, and recast as a rule-based decision support tool to maximize economic return in variable rate fertilization systems. The underlying objective of such a system is to identify locally significant crop response patterns and use those as a basis for variable rate application, rather than using typical uniform soil fertility goals as defined by regional recommendations. In this fashion, variable rate application is prescribed based on variable goals; the most appropriate goal is chosen economically on a very small scale.

One advantage of regression tree methods is that the transformation from prediction to prescription can be made in an explicit manner. In the INEEL DSS4Ag, the terminal nodes of the regression tree are used to define a set of mutually exclusive treatment alternatives. Some of the conditions for the terminal node refer to independent variables such as slope and aspect, or micronutrients that are beyond the grower's control. Other conditions refer to decision variables of interest, i.e., soil nutrients that can be changed by fertilizer application. The terminal node conditions are converted to a format suitable for a rule-based expert system engine. For example, the lower left terminal node shown in Figure 1 can be represented as a rule:

IF	K > 225 and
	buffered pH <= 6.65 and
	slope <= 5% and
	P <= 290
THEN	predicted yield = 65

For each location in the field, the current soil parameters are compared against the independent variables in each rule to determine which rules could potentially apply, after adjusting the dependent variables through addition of appropriate fertilizer amounts. The DSS4Ag chooses the fertilizer treatment that results in the greatest net economic gain, i.e., market value of the predicted crop yield less the cost of the added fertilizers. The rule processing is accomplished using CLIPS, a rule-based inference engine developed by NASA (Giarratano 1998).

Augmenting Statistical Models With Real-world Constraints

Real-world data sets contain implicit limits that are not easily detected by regression tree methods such as CART. For example, the range of potassium levels in a particular field might reflect a uniform application of fertilizer such that all locations in the field had adequate potassium. This is in contrast to test plot data that would contain some test zones with nutrient levels that are deliberately maintained at extremely low or high levels. Tree-based classification methods such as CART have no means of detecting these implicit limits. For example, a branch node condition might state "If K < 225 ppm, then yield =76, else yield = 66". Since there were no data reflecting what happens if the potassium level is extremely low there is no way for the classification algorithm to predict what would happen in that event.

In the DSS4Ag, this problem of implicit limits is addressed by augmenting the conditions for the terminal node with the observed limits in the prior data that falls within the terminal node. For example, given a terminal node that does not have a lower bound on potassium levels, the decision support system derives a minimum potassium level from observations in the historical data set.

Another issue that must be accommodated is the knowledge about the input data that is external to the data set itself. For example, sufficiently detailed data sets are relatively uncommon within current crop management practice; the only available historical data set may have been collected in an unusually low-yielding year due to adverse weather conditions. In such a case, one approach is to rely upon the producer’s experience and knowledge to estimate appropriate correction factors to normalize the data set to more typical yield response. A corollary issue here is that the growth patterns identified may be peculiar to the adverse weather conditions experienced; this highlights the value of multi-year data sets.

The producer’s crop management practices must also be taken into account. If the historical and current data sets reflect differing crop rotation patterns it may be necessary to adjust the resulting prescription accordingly. For example, wheat producers in Idaho will add additional nitrogen when replanting a field in wheat two seasons in a row. Although this does not affect the base prescription created by the DSS4Ag, an integrated decision support system must be able to include this information in its final output so as to reduce the number of post-processing steps that are needed.

Current fertilizer application technology also places constraints on the prescribed application. Upper and lower bounds on application rates due to machine limits must be accounted for, as well as upper limits on application rates to prevent unrealistically large amounts of fertilizer being requested due to a statistical anomaly in the data set. One approach to this is to simply "clip" the prescribed recipe to the appropriate limits; note that this clipping might result in yields considerably different than expected. The approach taken in the DSS4Ag is to eliminate choices that would result in limits violations, thus considering only those possibilities that can realistically be achieved.

In addition to application rate limits, in the future producers will increasingly have limits on total amounts of fertilizers applied to a given field. These global limits can be addressed within the DSS4Ag as a capital budgeting formulation. The discrete alternatives defined by the terminal nodes of the decision tree are then considered not only for each specific location in the field, but also are considered for their total effect against any overall limits.

To summarize, the INEEL Decision Support System for Agriculture (DSS4Ag) defines two key pieces of technology:

- A prototype expert system for site-specific management decision-making; and
- A partially automated decision support system architecture (Figure 2), that defines that interface between the decision-making tool, artificial intelligence information technologies and with the existing agriculture infrastructure and process for making management decisions

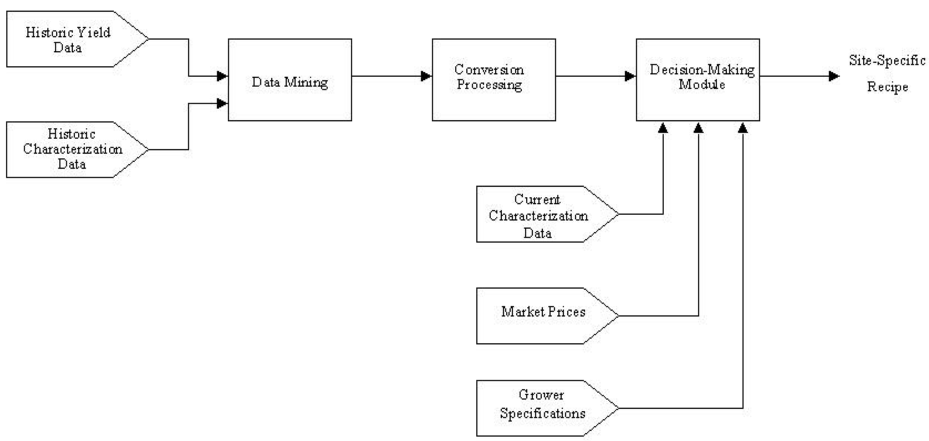

Figure 2. DSS4Ag Architecture

Discussion of DSS4Ag Field Experiments

Experimental field-scale testing of the INEEL DSS4Ag has been conducted at several locations over the past four years. The field tests have included wheat in Idaho (Hoskinson and Hess 1998), potatoes in Idaho (Hoskinson et al. 2000), corn in Kentucky (Shearer et al. 2000), and cotton in Mississippi (Thomasson et al. 2001). A variety of experimental designs have been used in these studies. The field-scale tests compared the performance of the DSS4Ag fertilization recipe against similar blocks or strips fertilized at a constant-rate, or using variable-rate fertilization to state agency recommendations. Results from yield monitoring at harvest were used to compare crop response across the differing fertilization treatments.

The results of each of the field tests are reported in the references, and will not be duplicated here. In most of the tests, the DSS4Ag recommendations were for less fertilizer than that of the constant-rate treatment or variable-rate constant-goal treatments. The general trend in the yield results is that differences in net yield have not been statistically significant. However, the DSS4Ag variable-rate recipes typically showed a slight advantage in net economic return; this increased return has generally been comparable the additional costs associated with the necessary soil testing and variable-rate application costs.

Conclusion

Data mining and artificial intelligence technologies have been used as the basis for a working prototype for site-specific management decision-making. The decision-making prototype, as part of a partially automated data/information handling architecture, identifies crop production success patterns within accurate precision agriculture data sets and uses these patterns to develop optimized site-specific management actions. The methodology relies upon spatial databases characterizing the crop production environment. This approach avoids constraints inherent with limited understanding of soil/plant/atmosphere mechanisms by basing decisions on documented site-specific patterns of crop and ecosystem variability.

Results from initial field-scale testing of the prototype have been encouraging but not statistically conclusive. Additional evolution and refinement of this methodology may provide a means for improving the economic effectiveness of site-specific fertilization practices.

Acknowledgements

Work supported through the INEEL Laboratory Directed Research & Development (LDRD) Program under DOE Idaho Operations Office Contract DE-AC07-99ID13727.

References

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, 1984. *Classification and Regression Trees.* Pacific Grove, CA: Wadsworth.

Giarratano, Joseph C. 1998. *CLIPS Users Guide*. http://www.ghgcorp.com/clips/download/documentation/usrguide.pdf

Hoskinson, R. L., and J. R. Hess. 1998. Using the Decision Support System for Agriculture (DSS4Ag) for Wheat Fertilization. In *Proceedings of the 4th International Conference on Precision Agriculture,* ed. P Roberts. St. Paul, MN: ASA, CSSA, and SSSA.

Hoskinson, R. L., J. R. Hess, and R. K. Fink. 1999. A Decision Support System for Optimum Use of Fertilizers. In *Precision Agriculture '99*, ed. J. V. Stafford. Sheffield, UK: Sheffield Academic Press Ltd.

Hoskinson, R. L., R. K. Fink, J. R. Hess, R.Oborn, T. Tindall, and L. Robinson. 2000. Using the Decision Support System for Agriculture (DSS4Ag) for Potato Fertilization. In *Proceedings of the 5th International Conference on Precision Agriculture,* ed. P Roberts. Madison, WI: ASA, CSSA, and SSSA.

Shearer, S. A., R. L. Hoskinson, R. K. Fink, T. Mueller, A. Thomasson, J. P. Fulton, and M. Ellis. 2000. Use of the Decision Support System for Agriculture (DSS4Ag) for Corn Fertilization in Kentucky. In *Proceedings of the 5th International Conference on Precision Agriculture,* ed. P Roberts. Madison, WI: ASA, CSSA, and SSSA.

Thomasson, J. A., S. A. Shearer, R. L. Hoskinson, J. R. Wooten, R. K. Fink, J. P. Fulton, S. F. Higgins, T. G. Mueller. 2001. Decision Support System for Agriculture (DSS4Ag) Applied to Fertilization of Corn and Cotton. ASAE Paper No. 01-1018. St. Joseph, Mich.: ASAE.