Knowledge and decision support for variable rate application of materials in prescription farming

Jack Robert Ambuel
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, and the Bioresource and Agricultural Engineering Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/10757

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600
Knowledge and decision support for variable rate application of materials in prescription farming

by

Jack Robert Ambuel

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Agricultural and Biosystems Engineering
Major: Agricultural Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1995
# TABLE OF CONTENTS

**GENERAL INTRODUCTION**

| PART 1: DETERMINATION OF APPLICATION RATES: GENERAL STRATEGIES AND THE NATURE AND CAUSES OF YIELD VARIABILITY |
|---|---|
| CHAPTER 1: STRATEGIES FOR DETERMINING FERTILIZER APPLICATION RATES | 5 |
| CHAPTER 2: CAUSES OF YIELD VARIABILITY WITHIN FIELDS | 14 |
| CHAPTER 3: CONSISTENCY OF YIELDS AT SPECIFIC LOCATIONS WITHIN FIELDS | 26 |
| CHAPTER 4: CONCLUSIONS | 48 |

| PART 2: DETERMINATION OF APPLICATION RATES: FUZZY LOGIC EXPERT SYSTEMS |
|---|---|
| CHAPTER 5: INTRODUCTION | 51 |
| CHAPTER 6: CONTROL OF HYDROSTATIC TRANSMISSION OUTPUT SPEED: DEVELOPMENT AND COMPARISON OF PI AND HYBRID FUZZY-PI CONTROLLERS | 54 |
| CHAPTER 7: FUZZY LOGIC SOYBEAN PLANT EVALUATION PROGRAM | 83 |
| CHAPTER 8: A FUZZY LOGIC YIELD SIMULATOR FOR PRESCRIPTION FARMING | 101 |
| CHAPTER 9: CONCLUSIONS | 145 |

| PART 3: IMPLEMENTATION OF VARIABLE RATE APPLICATION OF MATERIALS IN PRESCRIPTION FARMING |
|---|---|
| CHAPTER 10: INTRODUCTION | 148 |
| CHAPTER 11: SATELLITE BASED POSITIONING SYSTEM FOR FARM EQUIPMENT | 149 |
| CHAPTER 12: VARIABLE RATE CONTROL | 161 |
| CHAPTER 13: YIELD MEASUREMENT | 164 |
| SUMMARY AND CONCLUSIONS | 168 |
ACKNOWLEDGEMENTS

I want to thank my advisor, Thomas S. Colvin, for giving me the freedom to pursue my own ideas as well as for always being available to talk and give excellent advice based on his considerable experience. I also want to thank my family and friends for their support.
GENERAL INTRODUCTION

Interest in prescription farming, also referred to as site specific farming or precision farming, has grown as the technology necessary for its implementation has become available (Peck, 1995). One major goal of prescription farming is to optimize application rates for seed, fertilizer, and other agricultural inputs as a function of location within a field. There are four main components of a variable rate application system:

1. A method of determining the location of farm equipment (e.g. combines, tractors) within a field.
2. A system for controlling the application rate of materials (seeds, fertilizers, pesticides).
3. The capability to measure the results using a real time yield monitor.
4. An algorithm or prescription for determining what rates of materials to apply at each location.

The technology for the first three components is being developed rapidly by agricultural equipment, position equipment, and information systems manufacturers. Location, control and monitoring technology exists today and is being used commercially (McGrath et al., 1990; Ahlrichs, 1993; Hammond, 1993; Macy, 1993; AgChem, 1994; Rockwell, 1994). The fourth component of variable rate application - determination of what rates to apply at each location - has not advanced as rapidly as the other three. Only recently has research into variable rate application of fertilizers, seeds and pesticides begun (Lowenber-DeBoer, 1995; Nielsen, 1995; Pierce et al., 1995; Robert, 1995; Sudduth, 1995; Pan et al., 1995). In this dissertation, all four aspects of variable rate application systems are discussed. However, the fourth component - methods of determining how much material to apply (with an emphasis on fertilizer application) - was the focus of this research. The main objective of this research was to develop a method of predicting yields as a function of position in the field that can be used to make fertilizer application rate decisions in a prescription farming system.
Dissertation Organization

The dissertation is organized into three parts. The first two parts are concerned with the question of how to determine optimum material application rates. The emphasis is on fertilizer application rates. Seeding, herbicide, and pesticide application rates are also considered. In the first part, general strategies for determining these rates are discussed. The success of most of these strategies ultimately depends on an understanding of the causes of yield variability within a field. Therefore, a review of the literature of research into in-field variability follows the discussion on strategies. Then, because application rate decisions can be much easier to make if areas of consistently high or low yield constitute a significant portion of a field, a review of yield studies is conducted to determine the significance of areas of consistently high and low yield on farms throughout the world and in Iowa.

In the second part of the dissertation, a specific method for determining fertilizer, seeding and/or weed and pest control application rates is presented. This method involves the use of a fuzzy logic expert system to predict yields as a function of position on a particular farm. First, a brief introduction to fuzzy logic and its applications is given. This is followed by a discussion of a common application of fuzzy logic in machine and process control. In particular, a fuzzy logic controller for a hydrostatic transmission is discussed. Then, a second common application of fuzzy logic to pattern recognition is described. Specifically, a fuzzy logic expert system used to evaluate soybean plant shape quality is presented. Finally, the application of fuzzy logic to a less common problem of biological modeling is discussed. In this section, the development of a fuzzy logic expert system model to predict three years of corn yields measured on small 3 row by 12 meter (40 foot) sections of a cooperator farm in Boone County, Iowa is described. The fuzzy logic controller and the fuzzy logic expert system yield model sections are reproduced from two published papers. The soybean plant shape evaluation section is reproduced from a paper submitted for publication.

In the third part of the dissertation, implementation of variable rate application in prescription farming is discussed. First a method of determining position using GPS receivers is described (reproduced
from a published paper). This is followed by short sections discussing some of the problems of controlling material application rates and monitoring the results (yield). Finally, summary and conclusions are presented.

References


PART I: DETERMINATION OF APPLICATION RATES:
GENERAL STRATEGIES
AND THE NATURE AND CAUSES OF YIELD VARIABILITY
CHAPTER 1: STRATEGIES FOR DETERMINING FERTILIZER APPLICATION RATES

There are a number of approaches being taken to determine what rates of fertilizer to apply at different locations within a field in a prescription farming system. Three of the more important methods are:

1. Adaptation of field scale Extension Service recommendations.
2. Real time on-the-go soil property sensing.
3. Development of prescriptions based on yield histories and weather data.

Each of these techniques are explained in more detail below.

Adaptation of Extension Service Recommendations

This method of determining application rates uses application rates developed for NPK for each soil type based on soil test levels. These rates have been developed over the years by State Extension Services (Iowa State University Cooperative Extension Service, 1988). The rates are based on replicated plot trials performed on the major soil associations in the state. The yields at different fertilizer levels are averaged for all tests on each soil association. Fertilizer recommendations are then developed from the results of these tests. There are two main methods used for making fertilizer rate recommendations. In one method a yield goal equation is used which relates the average yield to the number of pounds per acre of fertilizer applied and a soil factor that is a function of the soil association. The soil factor is determined from yield studies conducted on the soil association. No soil testing is required for this method.

The method for determining nutrient application rates relies only on soil test values and does not consider yield goals. The Extension Service sets a minimum level in the soil for the nutrient, above which the yield will not be limited by that nutrient and below which yield reductions will occur. This again is specific to a soil association. The farmer typically will have soil tests performed for the entire farm and obtain an average nutrient level for the farm. This will be subtracted from the recommended level for the main soil type or soil association on the farm and the net amount applied.
In Iowa, both methods are used to establish fertilizer recommendations (Iowa State University Cooperative Extension Service, 1988). For nitrogen, the yield goal approach is used. A nitrogen application level versus yield equation has been established for each of the major soil associations in the state:

\[ N = YG \times SF \]

where

- \( N \) = lbs/acre of Nitrogen
- \( YG \) = Yield goal in bushels/acre
- \( SF \) = Soil Factor in lbs N/bushel. The soil factor varies with soil association.

For example, for the Clarion-Nicollet-Webster soil association the soil factor is 1.22 lbs N/bu. If the yield goal is 150 bu/acre, then the farmer would apply 183 lbs/acre of Nitrogen to his fields. If the farmer had applied manure in the Spring or Fall or had grown a legume crop in one of the previous two years, then credits should be applied for that, and the application rate reduced. These credits can also be calculated using extension service formulas or tables.

For Potassium and Phosphorus, the second method is used. Recommended levels are published for the major soil associations in the state, independent of yield. The amount of P and K applied depends on the soil and subsoil test levels and the predominant soil association.

To use Extension Service recommendations in a prescription farming system, several modifications are normally made. For current farming practices, soil tests are performed and average levels for the entire field are obtained. Then one rate is applied for the entire field. In prescription farming systems, the objective is to vary the rate so that the amounts more closely match the different needs at different locations. Therefore, when performing soil tests, intensive sampling or grid sampling of nutrients is performed (Mann, 1993). The grid samples then may be statistically analyzed to give an estimate of fertility at all points in the field. Typical grid dimensions are from 50 to 100 meters. The grid samples are then combined with soil survey information and a combined nutrient level and soil type map is developed (Carr et al., 1991; Mausbach et al. 1993). Extension service recommendations are then used to calculate the amounts of nutrients to be applied in each square grid area (Anderson, 1995).
There are a number of problems with using this approach to determining fertilizer application rates. The main problem is that the Extension Service recommended rates were calculated based on average yields on a large number of test plots on each of the major soil associations (Iowa State University Cooperative Extension Service, 1988; Tisdale et al., 1985, pg. 50). This means that variations in weather and variability within each soil type are not reflected in the fertilizer recommendation levels. Because of this, their usefulness in making site specific fertilizer recommendations is limited. This is because the weather during the growing season, and local variations, inclusions and anomalies in the soil physical properties at each site, are often more important in determining yield for the year than local fertility levels and soil type (Lyon, 1932; Rennie, 1960; Hunsaker et al., 1987; Ferguson and Gorby, 1966; Colvin, 1993; Karlen, 1993; Ambuel et al., 1994).

While it is true that, on average, different soil types have different yield potentials (Odell and Smith, 1940; Shrader et al., 1957), in fact, it is very difficult to establish a strong correlation between yields and fertility or soil type at a specific location (Tisdale et al., 1985, Chapter 2; Karlen, 1993; Birrell et al., 1993). Consider, for example, the summary given by Birrell et al. (1993, page 7) of a two year study mapping yields and nutrient levels:

The use of soil nutrient maps to determine fertilizer recommendations is fairly straightforward. However, using soil and yield maps to determine the "cause and effect" relationship between the soil nutrient levels and crop yield is extremely challenging. Climatic conditions and the interaction of climate and landscape position can have a major effect on the yield response, masking any soil-yield relationships that may exist. Considering the major differences in yield patterns over the two years on the same field, the use of yield maps to determine yield goals for future fertilizer recommendations must be approached with suspicion unless reasonable historical yield data is available.

A second problem with the grid sampling approach is the cost. Intensive soil sampling is very expensive (Wollenhaupt and Wolkowski, 1994). In addition, in many cases, chemical analysis will not lead to the optimum fertilizer recommendation. Instead, additional information about the soil physical properties, drainage, topography, and depth to restricting layers must be collected. It is then necessary to combine that information with the nutrient information and use a crop model to estimate yield under various weather conditions for different fertilizer levels to obtain optimal fertilizer amounts. The final decision is then made.
based on long term weather forecasts and other considerations. The additional information required adds
significantly to the cost and the complexity.

Real-Time Soil Property Testing

A second method for determining fertilizer application rates is to perform real-time measurement of
one or more soil properties as the tractor is moving through the field and adjust rates based on the measured
values (Gaultney, 1989). One such device proposed for site specific farming is the soil organic matter
sensor (McGrath et al., 1990; Gaultney et al., 1988). The operation of such a system is simple: fertilizer rates
are varied by some algorithm relating fertilizer amount to organic matter percentage.

The main problem with this technique is that it is experimental. While some studies have reported a
relationship between yield and soil organic matter (Mulla, 1993), as indicated previously and as will be shown
below, this is far from universally applicable. Furthermore, very little if any testing has been done to verify the
algorithms used to relate spreading rate and organic matter or other soil properties. Very few in-field yield
test results have been reported. A second major problem is that these sensors are new and still under
development. Viable real time sensors exist to measure only a few of the soil properties affecting yield.
However, this approach may become much more useful in the future as the technology advances. Research
into new methods of characterizing soil properties, such as conductivity meters (Jaynes, 1993) and other non-
contact electromagnetic and acoustic sensing, may result in substantially reduced costs if significant
correlation between the sensor output and underlying soil characteristics and yield potential can be made.

Prescriptions Based on Yield Histories and Weather Data

Development of an expert system for determining fertilizer application rates based on yield and
weather data is a third method proposed for use in prescription farming systems (Colvin et al., 1991; Colvin,
1992; Kachanoski et al., 1992, pg. viii). In this technique, yield as a function of position is measured using
an on-the-go yield monitor. The yield data is collected over a period of years and stored. In addition to yield
data, daily weather data (temperature and rainfall) is collected during the growing season. After 4 to 5 years of data are collected for each crop grown, the data are combined and analyzed. Fertilizer application rates are then established based on the results of that analysis, or further testing could be performed by applying varying amounts of fertilizer in adjacent strips at different locations in the field.

In its simplest implementation, the yield and weather data are used to identify areas of consistently high, and low yields, and then variable rates of application are established for these areas. On areas of consistently medium yield or on areas with no pattern, standard fertilizer application rates would be used based on Extension Service recommendations. On areas of consistently high or low yield, the standard rates would be modified to more optimal levels.

These areas of consistently high or low yield may be either independent or dependent on the weather. For example, areas of very low fertility, excessive drainage, or very poor drainage may have low yields every year. Similarly, a well drained area with high organic matter content and high water holding capacity may have high yields every year. Then there may be other areas where yield is weather dependent. For example, some areas may do well in wet years and poor in dry years or vice versa.

Determination of the optimal levels of fertilization for the areas of high or low yield would depend on the causes of the high or low yields. For high yielding areas, higher rates of seeding, fertilization, and pesticide application could be used to increase the yields even further. For lower yielding areas, fertilization rates could be increased if the low yields were due to low fertility. If, on the other hand, low yields were due to moisture problems (either excessive or insufficient), then seeding, fertilization, and pesticide application rates could be reduced. In some cases of particularly poor soil conditions, low yielding areas may be taken completely out of production. The rates selected would depend on the causes of the high or low yield and would differ from farm to farm.

Analysis of the causes of the low and high yielding areas can vary in complexity. The simplest approach would be to observe the areas and make educated guesses, using soil survey data, of the causes for
the low or high yields. A more sophisticated approach would be to combine visual observation with selective testing of the physical and chemical properties of the soil at those locations.

The effect of variable rates can be observed by examining subsequent yield and weather data and comparison with the results from prior years. Site specific farming also offers the opportunity to perform more controlled tests, by varying the rates of material application in adjacent rows in each of the areas of high or low yield and monitoring the results. These tests, however, would cause some inconvenience in harvesting. The test areas would have to be harvested separately, because typical combines in use on commercial farms harvest from 4 to 8 rows at a time. If tests were performed that correspond to the width of the combine, then resolution would be lost, although in some cases this may not matter.

There are a number of potential problems with the yield history method. One important problem is the yield history approach may require a long period of time to complete. At this time it is not known how many years of yield data would be required to establish the yield potential at each location in a field (Colvin, 1994). For locations that show consistent patterns of relative yield, independent of weather conditions, relatively few years would be required. However, if the relative yield potential at a point is weather dependent, the number of years required would depend on the weather. For example, if at one location the relative yield potential was strongly determined by the amount of moisture, yield data would have to be collected until at least one year of severe drought and one year of excessively high moisture was obtained. If too many years of yield data are required, this reduces the appeal of the yield history approach. For example, suppose five or more years of yield data must be collected for each crop grown in order to have confidence that the patterns observed are real. For a corn-soybean rotation, this would require ten years of data before the prescription farming program could be implemented. In addition to the time factor, a second problem with the method is that the results are only applicable to the farm on which it is used. Results from one farm cannot be applied to any other farm.
Comparison of the Three Methods

Each of the three strategies for determining application rates in prescription farming has drawbacks. The grid sampling technique is expensive and not optimal. The on-the-go soil property sensing method is experimental and untested. The yield history method is time consuming. In spite of the problems, all three methods have merit and are worth testing as possible techniques for determining application rates. It is likely that in the future, prescription farming systems will combine elements of all three techniques.

Although all three methods are worthy of further research, given the current state of development of variable rate application of materials in prescription farming, the yield history method is probably the most practical and profitable approach. Until years of additional plot studies are performed to determine optimal site specific application rates on a wide variety of soil conditions, the validity of the grid sampling technique will not be known. Similarly, the technology for on-the-go sensing is much too preliminary to be of use in commercial or research applications of site specific farming. The yield history approach, by initially focusing only on those areas of consistently high or low yield in a field (and not on the entire field), is much more likely to have results that are clear and unambiguous. The causes of consistently low and high yields will in general be easier to identify and material application rates easier to determine. Furthermore, the effects of the modified application rates will be more easily seen. And it is possible that the yield history approach can be implemented with little or no soil testing.

Testing of the yield history approach can be most easily performed if it can be demonstrated that areas of consistently high and low yield exist in fields and if the causes of that yield variability can be identified. Therefore, the remainder of this section is concerned with the causes of yield variability and its consistency. It is shown that the main causes of yield variability are diverse and are different in different locations. Several studies are also described that show areas of consistently high and low yields. But the number of these studies are insufficient to determine if areas of consistently high and low yield constitute a significant part of most fields. The answer to that question will have to await the results of future on-the-go yield monitoring of commercial and research farms.
References


CHAPTER 2: CAUSES OF YIELD VARIABILITY WITHIN FIELDS

In order for a site specific farming system to be successful, yield variability within fields must be identified and some method developed to determine how much material to apply, based on that variability. Ideally, a relatively simple algorithm or model would be developed that would relate inputs to yield at each location in the field and could be used to determine those application rates. Unfortunately, no such model exists. Yield depends on a complex mixture of soil properties, topography, weather, and cultivar characteristics. There is a natural tendency to seek a simple solution to any problem and a number of people involved in agriculture have attempted to develop easy to apply productivity indices, crop suitability ratings, etc. Under current farming practices, where one farm is treated uniformly, the use of such indices, which are derived statistically from trials on many different plots, is appropriate. Under prescription farming practices, results cannot be averaged and so the current sets of indices may not be applicable.

In the next section, a number of studies of yield are reviewed and some of the complexities of yield are discussed. In particular, yield dependence on crop type and soil characteristics is examined.

Comparison of Studies on Yield

From examinations of studies that attempt to identify the major parameters controlling yield, it becomes clear that such parameters depend on crop, location, and weather. It is probably impossible to develop a small set of key parameters that can be applied universally. In this section, examples of this complexity will be illustrated by discussing some selected studies of yield.

One important factor that determines whether the yield of a crop will be high or low in any particular location is the crop itself. In a study of soft winter wheat in the Palouse region of southeastern Washington, Cih a (1984) studied the effect of slope position. Cih a studied yields on the shoulder (just below the hilltop), on the side slopes and on the toe slopes (just above the bottomlands). The soil type on the shoulders was a Palouse silt loam (fine-silty, mixed, mesic pachic ultic haploxerolls with 3 to 7% slopes); on the side slopes was Naff silt loam (fine-silty, mixed, mesic ultic argixerolls with 15 to 45% slopes); and on the toe slopes was
a Snow silt loam (fine-silty, mixed, mesic cumulic haploxerolls with 7 to 15% slopes). Ciha found that in a normal year, yields were highest on shoulders. The author attributed this to less moisture runoff and deeper soil on the top of the hill as compared with the side slopes. The side slopes had the lowest yields. Yields increased somewhat on the toe slope where soil depth began to increase again due to accumulation and where moisture due to seepage increased. Bottomlands and hilltops were not studied. In an abnormal year with cold weather resulting in winter kill, yields on the hilltops are the lowest. In contrast to this study, Mahler (1979) studied the effect of topography on the yield of dry peas in the Palouse region of Washington. In this study, the results were opposite. The author studied yields on the bottomlands (xeric argiubolls), south side slopes (ultic argixerolls), and ridgetops (pachic haploxerolls). Yields of peas, which were spring sown, were highest in the depressions and lowest on the hill tops and south hill slopes. The yields on the bottomlands were four times higher than on the ridge tops and three times higher than on the south side slopes. The author attributed this to greater fertility, greater soil depth, and more water in the depressions.

In a study conducted in the loessial plain of West Tennessee, Fribourg (1989) compared the yields of a number of forage crops on the following soil types: Calloway (fine-silty, mixed, thermic Glossaquic Fragiudalfs), Grenada (fine-silty, mixed, thermic Glossic Fragiudalfs), Henry (coarse-silty, mixed, thermic typic Fragiqualfs), Lexington (fine-silty, mixed, thermic typic Fragiudalfs), and Memphis (fine-silty, mixed, thermic typic Hapludalfs). The most important differences between soil types were natural drainage and water holding capacity. The yields of the forages were compared with the yields of tall fescue on the same soils. Yields ranged from 40% to 160% of the tall fescue. The authors found that some forage crops showed little variation in yield from soil to soil, while others showed large variations.

Daniels, et al. (1989) conducted a study on the productivity of eroded soils in the North Carolina Piedmont area. Soil types studied were Cecil, Georgeville, Cullen and Vance (all typic Hapludults). The main conclusion of the study was that comparison of eroded soils in one landscape position with non-eroded soils in another landscape position was not valid, because of the different yield potentials at the different positions. However, the study also resulted in information of value to prescription farming. The authors found that small
grain yields on head and foot slopes were often low, due to the fact that small grains in the region were winter
grown and those landscape positions were often wet in the winter. In contrast, the study revealed that slightly
eroded head and foot slopes were most productive for corn and soybeans because they receive runoff and
sediment from adjacent areas and therefore may have extra moisture supplied during dry spells.

From 1964 to 1968, Lee and Spillane (1970) measured the yield and quality of spring wheat. The
yields were measured on the Clonroche, Screen, Rathangan, and Macamore soil types in Ireland. Clonroche
and Screen soils are well drained and excessively well drained respectively, with coarse texture and low water
holding capacity. The Rathangan soils are poorly drained with high silt content, weak structure and somewhat
clayey subsoil. The Macamore soils are clay loams with a heavy clay subsoil responsible for their poor
drainage. The study showed differences in yield on all four soil types. For the first three years of the study the
Atle variety of wheat was used. The differences between high and low yielding soil types was 27%. The last
two years a new variety, the Quern, was used. Yield differences between soil types diminished. The
differences between maximum and minimum yields dropped to 12.5%. In addition, with the variety change,
overall yields increased significantly, and the soil type with the highest yields changed. This was complicated
by the fact that the weather was drier during the last two years than in the first three.

It is seen from the above experiments, that the crop grown is an important determinant of the yield at
different locations. The responses of different crops at the same location are often so different that different
site specific strategies would have to be developed for different crops. For example, in the Palouse, seeding,
fertilizer, and pesticide application rates may be increased in the depressional areas and decreased on the
hilltops when peas are grown. The opposite approach may be taken for winter wheat. Furthermore, there are
not only differences between crops, but between varieties of the same crop.

Another important factor that determines whether the yield of a crop will be high or low in any
particular location is the topography or landscape position. However, as with type of crop, the effects of
landscape position vary from country to country, state to state and even on the same farm. In an experiment
done in California in 1977 (Whitman et al., 1985), barley yields were found to vary by a factor of four at
different topographic locations. The experiment was conducted on a Schorn-Balcom soil complex (Entic Chromoxeret and Typic Xerochert). The highest yields were obtained on the hilltops and south and west facing slopes. The lowest yields occurred in the low-lying areas. Hill slopes varied from 0 to 40%. During the growing season, rainfall was 27 inches, approximately twice the normal amount. In another study in the Palouse region of Washington discussed above (Ciha, 1984), soft winter wheat yields were highest near the hilltops, lowest on the sideslopes and in between on the toe slopes during a year with normal winter temperatures. During years with winter kill, the hilltop yields were the lowest.

In Saskatchewan, Canada, studies were conducted on three different sites in 1956 and 1957 (Rennie and Clayton, 1960). At each site there was a different soil association. Spring wheat was seeded at each site following summer fallow the previous year. Both unfertilized and fertilized plots were sown and yields measured. At the Oxbow soil association site, maximum yields were on the intermediate well drained slopes. Then in order of decreasing yields: lower slope (moderately well drained); upper slope (excessively well drained) and depression (poorly drained). The yields on the intermediate slope were substantially above yields at the other locations. At the Weyburn soil association site, maximum yields were on the uneroded-eroded well drained upper slopes and the lower slopes (both poorly drained and moderately well drained). The intermediate yields occurred on the uneroded-eroded intermediate slopes, eroded upper slopes and eroded summits. The lowest yields occurred on the upland depressions and eroded intermediate slopes. At the Kindersley complex location, yields were all approximately the same on the lower, middle and upper slopes. Yields on the depressional areas were substantially lower without fertilization, but were brought up significantly with fertilization. The results on these three associations were all obtained for the 1956 growing season, when rainfall was abundant. In 1957, the Kindersley association was tested again. In 1957 there was little precipitation from April to August. The maximum yields then occurred on the poorly drained depressional areas and were substantially lower on the medium to well drained lower and upper slopes.

In this Saskatchewan study, it is evident that the effect of topography on yields is also a function of soil type and weather. It also indicates the difficulty of making general fertilizer recommendations. The study
showed great differences in the effects of fertilization from location to location. Other results of interest to site specific farming was the failure of fertilization to completely correct for the loss of fertility due to erosion.

In another study done in Saskatchewan, Canada (Spratt and McIver, 1972), wheat was grown in southeastern Saskatchewan from 1965 to 1968. The plots were located on transects from top to bottom of hills. The maximum slopes on the hills were from 8 to 12%. All were located on one soil association - the Oxbow-Ryerson. Wheat was sown up and down the hills. Four adjacent fertilizer treatments were applied (none, P, N, NP). The authors found that with or without fertilizer, the yields were maximum at the depressions (bottoms of the hills) and decreased monotonically toward the summits. The authors attribute the yield pattern to increasing moisture availability from top to bottom of the hill. Fertilization raised the yields on the upper slopes and summit, but not to the level of the depressions. Fertilization had little effect at and near the bottoms of the hills.

The results of this study are particularly interesting, because they are quite different from the results obtained by Rennie (1960), described above. In the Spratt study yields increased from top to bottom of the hill. In the Rennie study yields increased from bottom to near the top and began to decline at the top.

The causes of these results cannot be known from the information given in the articles. The Rennie study was conducted for one year, in which there was abundant moisture. The Spratt study was conducted over a period of four years and the results were averaged. Therefore, the differences may be weather related, or they may be caused by local variations in the same soil association.

In another study of yield, topography and soil type, tests on sunflower and barley yields were done in North Dakota (Malo and Worcester, 1975). A transect was run from the bottom of a hill to the top of a hill. Sunflowers were grown on nine plots along the transect in 1972 and barley in eight plots in 1973. Yields for both sunflower and barley were highest on the back slope and decreased on the foot and toe slopes. There was a local minimum on the summit. The authors attribute the lower yields at the low positions to excessive moisture and salt content. At the summit, erosion, excessive drainage and poor water holding capacity were cited as causes of lower yields.
Another study of yield, topography and soil type was performed in North Dakota for five years from 1985 to 1989 on spring wheat (Halvorson and Doll, 1991). Tests were performed, at two different locations, on four different soil series, each corresponding to a specific topographic location: the Zahl series (fine-loamy, mixed, Entic Haploborolls) on hilltops and shoulders; the Williams series (fine-loamy, mixed, Typic Argiborolls) on sideslopes and hilltops; the Bowbells series (fine-loamy, mixed, Pachic Haploborolls) on foot slopes and toe slopes; and the Tonka series (fine, montmorillonitic, frigid Argiaquic Argialbolls) in small undrained depressional areas. At each location there were two fields. From four to six sites were used for each soil type. The sites were divided between the two different fields at each location. In both locations, the mean grain yields increased with decreasing elevation. This was consistent over the duration of the study. The authors attribute the yield increases to increasing available water. These results match the results obtained by Spratt in Saskatchewan, but not those of Malo in North Dakota, described above.

In Manitoba, Canada, a study on response of wheat to phosphorus was performed on two soil associations over a period of 4 years from 1959 to 1962 (Ferguson and Gorby, 1966). Soil types in each association were related to topography. Soils in the Miniota association were developed in coarse textured alluvium. Soils in the Waksada association were developed in medium textured glacial drift. Test sites were located on two farms, one for each soil association. On each farm, three sites on which wheat had been planted were selected each year. Each site had been summer fallowed the previous year. Six rates of fertilization were used at each site (including no fertilization). On the Miniota association, yield increased monotonically with decreasing elevation in two years (1960 and 1961). In 1959 the yields on the side slope were highest and the depression second highest. Only the side slope in 1959 responded to fertilizer. On the Waksada association, yields were highest on the summits and sideslopes (being approximately equal) and lowest on the depressions in all years and at all levels of fertilization. Response to fertilization was mixed. In some years there was no response and in others a significant response. The location of the greatest response varied from year to year.
Weather information was not given in the report for all of the years so it is difficult to analyze the results. However, some hypotheses can be developed. The Miniota soils were coarser than the Waksada soils, thus were more susceptible to drought stress. In addition, the water table was closer to the surface of the Miniota depressions than of the Waksada depressions. These two differences may account for yield increases with decreasing elevation on the Miniota soil but not the Waksada soil. The depressions in the Miniota association were relatively more wet than the summits. In the Waksada association, there may have been little difference in moisture and other factors determined yield. Whatever the cause, the results show that yield response at different landscape positions can vary drastically from soil type to soil type.

The results of these studies of yield and topography are summarized in Table 1. It is easy to see from these results, that topographic position by itself cannot serve as an indicator of productivity. This is true even within the same area as seen from the studies done in North Dakota and Saskatchewan. In addition to type of crop and topographic location, there are other important causes of yield variability within a field. Examples include depth to a restricting layer, amount of erosion, various hydrological properties and variability within soil types.

In Kentucky, corn yields were measured on test plots with varying depths to fragipan over a four year period (Pagoulatos, 1989; Frye et al., 1983). The study was conducted on a Zanesville silt loam (fine-silty, mixed, mesic Typic Fragiaudalf). The series was developed in loess parent material overlaying residuum from an acid sandstone. A strongly developed fragipan horizon existed at the loess-residuum interface. Fragipan depths ranged from .3 to .6 meters. In the yield studies conducted on this soil, varied results were observed.

In the one wet year when precipitation was evenly distributed, yields decreased with fragipan depth. Fragipans closer to the surface likely held the same amount of water in a smaller volume than those of greater depth. Volumetric moisture was greater and less root growth was required. More biomass was consequently distributed in the grain.
<table>
<thead>
<tr>
<th>LOCATION</th>
<th>CROP(S)</th>
<th>Relative Yield</th>
<th>Relative Yield</th>
<th>Relative Yield</th>
<th>Relative Yield</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>Barley</td>
<td>High</td>
<td>(1)</td>
<td>(1)</td>
<td>Low (Due to Excessive Moisture)</td>
<td>(1) Yields high on south and west slopes.</td>
<td>Whitman, et al., 1985</td>
</tr>
<tr>
<td>Washington: Palouse Region</td>
<td>Soft Winter Wheat</td>
<td>-</td>
<td>Normal Weather High; Winter Kill Low</td>
<td>Low</td>
<td>Medium</td>
<td>-</td>
<td>Ciha, 1984</td>
</tr>
<tr>
<td>Saskatchewan: Oxbow Soil Association</td>
<td>Spring Wheat</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>Rennie and Clayton, 1960</td>
</tr>
<tr>
<td>Saskatchewan: Kindersley Soil Association</td>
<td>Spring Wheat</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Rennie and Clayton, 1960</td>
</tr>
<tr>
<td>Saskatchewan: Kindersley Soil Association</td>
<td>Spring Wheat</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Rennie and Clayton, 1960</td>
</tr>
<tr>
<td>North Dakota</td>
<td>Sunflowers and Barley</td>
<td>Medium Low (due to erosion)</td>
<td>Medium</td>
<td>High</td>
<td>Medium Low</td>
<td>Low from salt and excessive moisture</td>
<td>Malo and Worcester, 1975</td>
</tr>
<tr>
<td>North Dakota</td>
<td>Spring Wheat</td>
<td>Low (due to low available water)</td>
<td>Medium Low</td>
<td>Medium</td>
<td>Medium High</td>
<td>High (due to high available water)</td>
<td>Halvorson and Doll, 1991</td>
</tr>
<tr>
<td>Manitoba: Minolta Soil Association</td>
<td>Spring Wheat</td>
<td>Low</td>
<td>Medium Low</td>
<td>Medium</td>
<td>Medium High</td>
<td>High (due to high available water)</td>
<td>Ferguson and Gorby, 1966</td>
</tr>
<tr>
<td>Manitoba: Waskada Soil Association</td>
<td>Spring Wheat</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Medium Low</td>
<td>Consistent for all 3 years of study</td>
<td>Ferguson and Gorby, 1966</td>
</tr>
</tbody>
</table>
In the one year with below normal precipitation, the yields increased with fragipan depth to the medium to maximum depth range. The yields decreased slightly on the soils with maximum fragipan depths. This pattern likely resulted from insufficient precipitation received during the growing season for maximum yields. In soils with shallower depth to fragipan, there was not enough stored moisture due to smaller volume to compensate for the small amount of precipitation. The optimum depth occurred midway between the middle and maximum depths. Yield reductions occurred due to insufficient water in soils with shallower depths to fragipan. For greater fragipan depths yield reductions occurred due to reduced volumetric moisture content requiring more root growth. In the two normal years, the shallow soils had lower yields. For soils with depths to fragipan from the middle to maximum distance, yields were at or near the maximum. This indicates that normal precipitation is insufficient to establish maximum corn yields.

The explanations proposed for the observed yields in the above experiment are speculative, because not enough data are available. However, the results do indicate that in some locations consistent yields from year to year can be observed if the years are grouped by similar weather patterns. The results are also another indication of the wide variety of the major factors that determine yield, and in comparison with the other studies reviewed above, shows how these factors are quite different from location to location.

Many studies measured the effects of erosion on yield. Most have considered the yield effects on an average basis (Pierce et al., 1983). However, some more recent studies have been on a site specific basis (Kachanoski et al., 1992). Of course, erosion is correlated with topography.

In Israel a study of soil surface curvature effects on soil moisture and yield was performed on a 70 square meter field planted in winter wheat (Sinai et al., 1981). The land was unfertilized and not irrigated. Average annual rainfall was under 200 mm per year. On the selected field, the average slope was 12%. Results showed large differences in yield between ridge areas and trough areas, with trough areas having higher yields. The authors attribute the difference to water flowing laterally in the soil under the troughs and being depleted from under the ridges.
A number of other studies have demonstrated productivity variation within the same soil type in the same area (Murray et al., 1939; Graveel et al., 1989).

Implications for Prescription Farming

In conclusion, the causes of yield variability are complex. The above discussion has shown that the yield obtained at any one location depends on both the crop variety and crop hybrid planted, the soil characteristics, the topography, and the weather. It would be very difficult to develop a simple index or algorithm relating yield to inputs for all crops. Furthermore, if one index were developed for each crop, it would be difficult to make that index generally applicable throughout the country. And as was shown in the studies of yield variation within soil types, it would also be difficult to create a regional or local index useful to site specific farming.

Current agricultural models are also of insufficient accuracy and require too many measurements for use in site specific farming. For example, a study was performed using the yield component of DRAINMOD to predict yields on two research farms in Iowa (Kanwar et al., 1994). Predicted average yields on the two sites were compared to average measured yields over a five year period on each site. Statistically, 54% of the variations in measured yields were predicted by the model. But the model requires extensive measurements (especially hydrological and climatological). And the model only predicts average yields for the entire field instead of predicting yield for smaller sections of the field. This may change in the future as modeling advances are made and less expensive methods of measuring soil characteristics are perfected. However, presently, an empirical approach to site specific farming decisions may be most cost effective and the best approach during the initial development of this new farming technique. As discussed above, an empirical approach using yield histories is one possible method. As noted previously, application of this method becomes much easier if significant areas of consistently high, medium or low yield exist in the field.
References


CHAPTER 3: CONSISTENCY OF YIELDS AT SPECIFIC LOCATIONS WITHIN FIELDS

Studies in which yield is measured in small segments of a field over a period of years are not common. Although a great amount of research has been done over the last 70 years addressing fertilizer, cultivation, herbicide, crop yield, hydrological, erosion, and water quality effects on yields, seldom are yields reported from small individual sections of a field over a period of years. Instead, a typical procedure is to statistically analyze the results from many plots and report the results of that analysis. Occasionally results from multiple small plots are reported (so-called raw data). Of these an even smaller number will report yields from the same small plots over a number of years. In this section, some of these experiments will be discussed. The section is divided into two parts: studies conducted outside of Iowa and studies conducted inside Iowa.

Yield Studies Outside of Iowa

A series of studies conducted in England from 1975 through 1978 reported patterns in crop growth at four different locations (Evans, 1990; Evans and Catt, 1987). At Maxey, Cambridgeshire, England from 1975 to 1978, barley was grown three years and winter wheat one year. Before the experiment was conducted, the site had shown variable crop growth patterns in aerial photographs. These photographs were used to select four transects on the site for measuring crop yield. The transects all showed consistently low, medium or high yields over the four year period. The cause of the observed yield pattern was topsoil depth to gravel. The deeper the topsoil, the greater the yield. These results are summarized in Figure 1, which shows the average annual yields for each transect.

The second study site was West Wratting, Cambridgeshire, from 1976 to 1978. At this location, winter wheat was grown one year and spring barley for two years. At this site there are alternating wide (5-7m) and narrow (4m) stripes of soil over chalk. The depth to chalk under the narrower stripes (65-90 cm) is deeper than under the wider stripes (45 cm). Sampling points were established along both a broad and a
FIGURE 1: YIELDS ON TRANSECTS AT MAXEY CAMBRIDGESHIRE SITE:
BARLEY IN 1975, 1977, 1978; WHEAT IN 1976 (FROM DATA IN
TABLE 1 OF EVANS AND CATT, 1987)
narrow stripe. In years when drought stress occurred (1976 and 1978), yields on the broad stripes were significantly larger than those on the narrow stripes. In 1977 when there was no drought stress, yields on the narrow stripe were slightly larger than on the broad stripe. The authors hypothesize that in years with drought stress, plant roots on the wide stripes can reach the chalk layer and extract water from the chalk which has a higher water holding capacity than the soil. In years with no drought stress, there is little difference between yields on the two stripes. The yields on the two stripes are summarized in Figure 2.

A third experiment was performed by Evans and Catt in Toseland, Cambridgeshire from 1976 to 1978. At this location, the crop patterns were due to a series of parallel valleys and ridges. The valley soil was darker than the ridge soil, which was mixed with more chalk fragments. One transect was established along a valley and two transects along adjacent ridges. Winter wheat was grown the first two years and winter barley the third year. In all three years, the yield in the valley was significantly higher than the yield on the ridges (see Figure 3). The authors hypothesize that the lighter ridge soil reflected more solar radiation and was thus cooler than the valley soil. The authors measured soil temperatures for both ridge and valley and found that on sunny days with dry soil, the darker color of the valley floor caused it to become warmer for several hours during the day. The authors believe that this resulted in earlier germination, earlier tillering and higher yields.

Other experiments performed by the authors revealed other causes of consistently high and low yields in all years or in years under specific climatic conditions. Clay soils adjacent to loamy soils showed consistently lower yields when plowing and or planting was done soil that was too wet. Shallow soils with many stones yield less than soils with fewer stones. Shallow soils over hard or soft rock or gravels and sands exhibit lower yields than on adjacent deeper soils. In drier parts of the country these differences appear with regularity.
**FIGURE 2:** YIELDS AT WEST WRATTING SITE: BARLEY IN 1977, 1978; WHEAT IN 1976  
(FROM DATA IN TABLE 2 OF EVANS AND CATT, 1987)

**FIGURE 3:** YIELDS AT TOSELAND SITE: WHEAT IN 1976, 1977; BARLEY IN 1978  
(FROM DATA IN TABLE 3 OF EVANS AND CATT, 1987)
Evans (1990) concluded that crop patterns occur regularly on 20% of the land in England. Because 50% of the land is arable, 40% of the arable land exhibits crop patterns. In dry years that increases to 50%. Such a large percentage of the land exhibiting crop patterns, indicates that site specific farming in England should be relatively easy to implement and has a high probability of success.

In southeast Nebraska a study of the relationship of landscape position to yield was performed from 1985 to 1986 (Jones et al., 1989). Five locations were studied, each on a different soil series. Yields were measured at six different topographic positions at each location: upper interfluve (UI), lower interfluve (LI), shoulder (S), upper linear (UL), lower linear (LL), and foot (F). On two of these locations the same crop (sorghum) was grown in each year. The relative yields were fairly consistent at both locations over the two years (see Figures 4 and 5). Although the yield differences were not as dramatic as those shown in the English study, and the study only ran for 2 years, there was some consistency at both locations.

In Canada a 5 year program examining the relationships between land form, soil properties, soil loss, tillage management, and crop yield was conducted (Kachanoski et al., 1992). A number of sites were selected at specific topographic locations (crest, shoulder, backslope, and footslope). Measurement of soil erosion and extensive soil physical and chemical testing was performed at each site. Each site consisted of a 6 by 6 meter square section. Three tillage methods were tested and compared: moldboard, minimal, and no till. Yields were recorded for each year. Since the sites were located on commercial farms, management practices were different on each farm. The following were some of the results obtained:

1. The main cause of erosion was tillage translocation not runoff water.

2. The effect of erosion was greatest on sandy soil (>70% sand) where yields were reduced by 37%.

3. Converging land forms had higher yields.

4. Diverging shoulder and backslope positions had significantly lower corn and soybean yields than other positions. The authors attribute this entirely to tillage translocation induced erosion.

5. Reasonably consistent patterns of yield data were found from year to year at a number of the sites.
FIGURE 4: RELATIVE SORGHUM YIELDS ON CRETE SOIL SERIES IN NEBRASKA (FROM DATA IN TABLE 2 OF JONES ET AL., 1989)
FIGURE 5: RELATIVE SORGHUM YIELDS ON HASTINGS SOIL SERIES IN NEBRASKA  
(FROM DATA IN TABLE 2 OF JONES ET AL., 1989)
Concerning the last result, some fields showed strong yield patterns. In Figures 6 and 7, yields at nine locations on both a spring disced field and a no-till field near Strathroy are shown for four years. Both fields showed points of consistently high, low and average yields. On other fields, the yield patterns were not so pronounced. In Figures 8 and 9, on a fall soil save tillage field near Pottruff, some locations in the field showed consistent levels of yield, but the consistency was not as pronounced as on the Strathroy fields and the differences between high and low yields were not as great. In Figure 8, four years of corn and one year of soybean yield data are displayed. In Figure 9, only the four years of corn data are shown. The consistency in the corn yield data is slightly greater than the combined corn and soybean yield data. The authors also suggest a strategy for making site specific decisions: "to measure and map yield response directly from year to year and identify sensitive areas from changes in yield from year to year especially during climatic stress conditions."

_Yield Studies Inside Iowa_

There are few published studies of yield in Iowa that provide raw data from which one can make an assessment of the consistency of yields from year to year on farms. However, there are a number of studies that reveal some of the major causes of yield variability and yield reduction in Iowa. In the North-central region of Iowa, known as the Des Moines lobe, drainage problems have been shown to be a major cause of yield reduction (Kanwar et al., 1983; Kanwar et al., 1984). Another important cause of yield reduction in all of Iowa has been found to be reduction in topsoil thickness caused by accelerated erosion (Kazemi, et.al, 1990) (Englestad, et.al, 1961).

In the last 6 years several studies have been conducted at Iowa State where detailed yield histories from farm fields over a period of years have been developed (Colvin, 1995). In one study conducted on a commercial farm in Boone county, six years of yield data have been collected from four 16 ha (40 acre) fields. On each field eight transects have been established. On each transect, yields have been measured for 12 meter (40 foot) distances using a three row combine. This results in twenty eight yield measurements for each transect.
FIGURE 6: STRATHROY SITE RELATIVE YIELDS ON SPRING DISC TILLAGE PLOTS (FROM DATA IN APPENDIX V.I. PAGE 10 OF KACHANOSKI ET AL., 1992)

FIGURE 7: STRATHROY SITE RELATIVE YIELDS ON NO TILL PLOTS (FROM DATA IN APPENDIX V.I. PAGE 10 OF KACHANOSKI ET AL., 1992)
FIGURE 8: POTTRUFF SITE RELATIVE YIELDS ON FALL SOIL SAVE TILLAGE PLOTS.
SOIL SAVE TILLAGE = MINIMUM TILLAGE USING MODIFIED CHISEL PLOW.
(FROM DATA IN APPENDIX V.1 PAGE 8 OF KACHANOSKI ET AL., 1992)

FIGURE 9: POTTRUFF SITE RELATIVE YIELDS ON FALL SOIL SAVE TILLAGE PLOTS
(CORN YEARS ONLY)
On two of the fields, the crop rotations (corn, soybeans, oats, meadow) have been such that not more than two years of yield data has been established for each crop. On another of the fields, the topography and soil types are relatively uniform. There has not been much variability in yield. However, on the fourth field, the topography and soil types are more varied. The main soil association is Clarion-Nicollet-Webster. Clarion and Nicollet soil types are Typic Haplaquolls and the Nicollet soil type is an Aquic Hapludolls. On the fourth field there are significant inclusions of Okoboji (Cumulic-Haplaquoll), Harps (Cumulic Calciaquoll), Zenor (coarse-loamy, Typic Hapludoll), and Storden (Calcicous Typic Udorthent) soil series. Yield variability has been pronounced. Furthermore, there have been areas of consistently high and consistently low yields over the period of the study.

The results for this field for corn years on transects 5 and 8 are shown in Figure 10. In this figure sample numbers correspond to the adjacent 12 meter (40 foot) lengths for which yield was measured on each transect. The yields are shown as relative values computed by dividing the measured yield at the sample number (plot) in one year by the average yield for the entire transect in that year.

It can be seen from these graphs that there are a number of locations where yields are consistently low or consistently high. For sample numbers 16 and 17 on transect 8, for example, yields for both corn and soybeans were below average in all three years. This position in the field is on a relatively steep hill (6-8% slope) that is severely eroded. The soil types at these locations include Storden and Zenor soil series. There are other positions that are consistently above average in yield, and there are several positions that have yields that appear to be weather dependent. At sample number 7, for example, there is a small drainageway. In 1993 which was a particularly wet year, yields at sample point 7 were depressed.

There were also locations on transect 5 that displayed consistently high or low yields depending on the weather. In particular, for sample points 15 through 24, yields were low in years when there were any periods of excess moisture. Even in 1989, which was very dry, several points in this section had below average yields. These positions correspond to a poorly drained section of the field that is frequently subject to excess moisture problems. The predominant soil types at these locations are Harps, Canisteo and Okoboji.
**FIGURE 10:** BAKER FARM SOUTH FIELD RELATIVE YIELDS: YIELD SAMPLE NUMBERS = YIELD FROM 12 METER BY 3 ROW PLOTS. SUCCESSIVE SAMPLE POINTS ARE ADJACENT.
In a separate study done in adjacent Story County on a commercial farm (Black farm), a field has demonstrated a consistent pattern of yield. Starting in 1992, a number of lines at the same locations were harvested each year. The lines were all approximately 294 meters (970 feet) in length. Each line or transect in corn years was three crop rows wide and each transect in the year soybeans were grown was five crop rows wide. Harvesting on each line was done in approximately 20 meter (66 foot) intervals. Thus each harvest section corresponded to three or five rows of 20 meters in length.

The relative yields of each of these sections on each harvest line, from 1992 through 1994, are shown in Figures 11 - 18. The yields for each year are relative to the yield average for all transects harvested in that year. Corn was planted in 1992 and 1994 and soybeans in 1993. The harvest line numbers began at the southern end of the field (harvest line number 1). The higher the harvest line number, the farther north it was from the southern boundary of the field. The harvest section numbers began at the eastern boundary of the field (harvest section number 1) and ended at the western edge (harvest section 14 or 15). The harvest line lengths were not all equal, but the eastern and southern boundaries were fixed. The relative yield was calculated by dividing the actual yield as measured in a yield section by the average yield for all sections for that year. Yields were measured for 8 harvest lines for the corn years (1992 and 1994) and for 4 harvest lines for soybeans.

Because there were three rows in the corn transects and five rows in the soybean transects, the transects do not match. Therefore, when graphing the yields for corn and soybeans, soybean yields were only included with those graphs of corn transects where the corn transects were completely contained within the soybean transect. Thus, for example, harvest line 1 of the corn is completely contained within the first harvest line of soybeans. Similarly, corn harvest line 3 is completely contained within the second harvest line of soybeans, corn line 5 in soybean line 3 and corn line 6 in soybean line 4. All other corn harvest lines overlap two soybean harvest lines. The soybean yields were therefore included only on those corn harvest lines which were 100% contained within the soybean lines. This is not a perfect comparison, since the soybean yields include soybeans from adjacent corn transects.
FIGURE 11: BLACK EAST RELATIVE YIELDS: HARVEST LINE 1
FIGURE 12: BLACK EAST RELATIVE YIELDS: HARVEST LINE 2
FIGURE 13: BLACK EAST RELATIVE YIELDS: HARVEST LINE 3
FIGURE 14: BLACK EAST RELATIVE YIELDS: HARVEST LINE 4
FIGURE 15: BLACK EAST RELATIVE YIELDS: HARVEST LINE 5
FIGURE 16: BLACK EAST RELATIVE YIELDS: HARVEST LINE 6
FIGURE 17: BLACK EAST RELATIVE YIELDS: HARVEST LINE 7
FIGURE 18: BLACK EAST RELATIVE YIELDS: HARVEST LINE 8
As can be seen in the graphs of relative yield, nearly all harvest lines exhibited the same yield pattern. Yields were highest on the east end (with a local minimum at harvest section 2 in each line) and decreased going toward the west edge of the field. This pattern was very consistent.

Although in the last two years the number of studies designed to accumulate detailed yield histories has increased, the number of existing yield histories in Iowa is small. As a consequence, it is not clear whether crop patterns, or areas of consistent yield are common throughout the state, and if they are, how significant they are. Many additional studies of this type are needed before the nature of yield variability on Iowa farms can be determined.

References


CHAPTER 4: CONCLUSIONS

Causes of yield variability are numerous and complex. At any one location, however, there will be a smaller number of critical factors that have the greatest effect on yield. As agricultural studies around the world have shown, these critical factors tend to be specific to location and crop. Because of this complexity it would be difficult to develop a general crop yield model. Furthermore, even local crop yield models that are deterministic, require extensive measurements and have not demonstrated adequate precision.

Strategies for determination of material application rates include: modification of current extension service procedures to develop prescriptions for smaller areas of the farm; application of emerging sensor technology in a real-time feedback system where material rates were varied depending on the output of the sensors; and use of expert systems to establish rates, based on yield histories and analysis of the causes of yield variation. The extension service recommendations have two drawbacks. One is that the recommendations are based on averages of a large number of field trials on a specific soil type. There is large variation of productivity within a single soil type and so recommendations based on averages will in many cases not be optimal. The second drawback is that extensive and expensive testing is required.

The problem with real time sensors is that the technology is emerging and there is little experimentally verified correlation between sensor output and inherent soil productivity. Finally the problem with a yield history-expert system approach is that it requires a long time period to implement. There are also some additional potential problems with this method. As described above, one approach to the yield history-expert system strategy is to concentrate on areas of a field that show consistently high or consistently low yields, and to vary the standard application rates only on those areas. For this approach to be profitable, a significant proportion of the fields must consist of such areas. As seen in the review of the literature above, such areas do exist, but their extent is unknown.

At this stage in the development of variable rate application of materials in prescription farming, the greatest need is for more information. How much variability in yield is there? How consistent is that variability from year to year? That information can be supplied if farmers and researchers begin to use yield
monitors to establish yield histories for their farms. As those yield histories are acquired, the best method or combination of methods for determining material application rates should become clearer. Or, if yield histories reveal little variability and consistency in yield, then implementation of variable rate application of materials in prescription farming may have to await additional advances in agronomic science and technology.
PART 2: DETERMINATION OF APPLICATION RATES:
FUZZY LOGIC EXPERT SYSTEMS
CHAPTER 5: INTRODUCTION

With the rapid pace of development of higher speed and more powerful computers, mathematical methods have been used to analyze more and more complicated physical problems. The accuracy of these results, however, has been somewhat disappointing. Consequently, a number of alternative non-mechanistic approaches have been developed in recent years. These approaches include expert systems, neural networks, fuzzy logic, chaos theory and genetic algorithms. These methods were developed as alternatives to solving problems for which standard numerical analysis and statistical techniques did not yield satisfactory results or were too difficult to apply. For example, consider the comments of Lofti Zadeh (Kandel and Langholz, 1994, pp xvii-xviii), the inventor of fuzzy logic:

When I wrote my first paper on fuzzy sets in 1965, my expectation was that most of the applications of the theory would be in those fields in which the conventional mathematical techniques are of limited effectiveness. This was, and still is, the case in biological and social sciences, linguistics, psychology, economics and, more generally, in the soft sciences. In such fields, the variables are hard to quantify and the dependencies are too ill defined to admit precise characterization in terms of difference or differential equations.

Zadeh goes on to explain the relationships among these new methods that are emerging:

Today...we are observing a paradigm shift from traditional, hard computing to what may be called soft computing (SC). As its name suggests, soft computing is concerned with modes of computation which are approximate rather than exact. At this juncture, the principal components of soft computing are fuzzy logic (FL), neural network theory (NN), and probabilistic reasoning (PR), with the latter subsuming belief networks (BN), genetic algorithms (GA) and the theory of chaotic systems (CT). There is substantial overlap between FL, NN, and PR but in the main, FL, NN and PR are complementary rather than competitive. For this reason, there are many situations in which FL, NN and PR may be used to advantage in combination rather than exclusively.

Zadeh concludes by explaining the main aspects of each of the three modes of computing:

Within soft computing, FL is concerned in the main with imprecision and approximate reasoning; NN with learning and curve-fitting; and PR with uncertainty and propagation of belief. In the final analysis, the role model for soft computing is the human mind.

The number of papers describing applications of these techniques to agricultural problems in the last few years has increased dramatically. Most of these articles have described application of neural networks or fuzzy logic. These applications span all areas of agricultural engineering and agronomy. In the area of
hydrology, Bardossy and Disse (1993) describe two fuzzy rule-based models for infiltration that require fewer input parameters, use input parameters that are relatively easy to measure, and execute faster than the Green-Ampt model and Richards equation which are currently two of the more popular models in use.

These techniques have also been applied to crop scouting problems. Burks (1994) developed a neural network system to classify plant canopy status using textural features as the input to the model. Zhang (1994) applied neural networks to the identification of weeds in a Kansas wheat field.

A third area of application is in the grain handling and food processing industry. Romaniuk et al. (1993) present a multi-layer neural network classifier used for identification (or separation) of three barley seed varieties. Seong (1994) discussed a combined neural network - fuzzy controller for the baking process.


In this part of the dissertation, the application of fuzzy logic to three increasingly complex problems is presented. In the first application, a traditional application of fuzzy logic to the control of a physical process is described. A fuzzy logic controller for the output speed of a hydrostatic transmission was developed and tested and compared with a more conventional PI controller. In the second application, a fuzzy logic expert system for the evaluation of soybean plant shape is described. In the final application, a fuzzy logic expert system for the prediction of corn yields in one field is described. The corn yield model has potential application in site specific farming as a decision support tool for selecting variable fertilizer and seed application rates.
References


CHAPTER 6: CONTROL OF HYDROSTATIC TRANSMISSION OUTPUT SPEED:
DEVELOPMENT AND COMPARISON OF PI AND HYBRID FUZZY-PI CONTROLLERS

A paper published in the Transactions of the ASAE

J. Ambuel, L. Steenhoek, R. Smith, T. Colvin

ABSTRACT

A program designed to control a hydraulic motor using either PI or hybrid Fuzzy Logic - PI control was written. The motor under control was part of a hydrostatic transmission. The real time closed loop control software was written to run on an IBM PC AT or equivalent. Additional software was written to generate the fuzzy rule set to be used with the hybrid Fuzzy-PI controller.

Test results indicate that hybrid fuzzy logic - PI control of the motor speed was more effective than pure PI control for set point changes. Load changes were not tested. KEYWORDS: Controls, Hydraulics, Fuzzy Logic.

---

1Reprinted with permission from Transactions of the ASAE. 1993, 36(4): 1057-1064. Copyright © 1993, Transactions of the ASAE.

2J. Ambuel, L. Steenhoek, and R. Smith, Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, Iowa; T. Colvin, USDA-ARS National Soil Tilth Lab, 2150 Pammel Drive, Ames, Iowa 50011

3Author for correspondence.
INTRODUCTION

In the last ten years fuzzy logic based control research and applications have undergone explosive growth. Invented by Zadeh [1965], application of fuzzy logic to control was initially slow in developing. Significant early work was conducted in England by Tong[1977], by Mamdani and Assilian [1975] and by Procyk and Mamdani [1979]; in Denmark by Holmblad and Ostergaard [1981]; and in Japan and other locations. However it was the results of the work done in the Japanese fifth generation project on computing that gave applications of fuzzy logic control the biggest boost.

The results of the Japanese and other research efforts have led to the development of a large number of consumer products and industrial processes employing fuzzy logic to control some aspect of their operation (Williams, 1991; Mamdani, 1984; Yamakawa, 1989; OMRON).

Based on the success of the application of fuzzy logic control to a wide variety of control problems and its continued rapid growth, the authors became interested in the suitability of fuzzy logic control for hydraulic systems. In particular, there was an interest in comparing the quality of a fuzzy logic-based controller to a conventional PI (proportional-integral) controller.

Several theoretical studies of a similar nature have been published. Daley and Gill [1989] compared the performance of a fuzzy self organizing controller (SOC) with a PD (proportional-derivative) controller used for the attitude control of a flexible satellite with significant dynamic coupling. A state variable model for the satellite was derived and the control results simulated. The authors found that the fuzzy controller performance was equal to but did not exceed the performance of the PD controller. A significant advantage to the fuzzy controller was that knowledge of the satellite transfer function was not necessary to establish control. In a recent paper, Buckley and Ying [1990] have shown that a linear fuzzy controller can be constructed that gives precisely the same control as a PID controller.

The main purpose of the research described in this paper was to determine the feasibility of using a hybrid fuzzy-PI controller to regulate the speed of a hydraulic motor. The method used to determine feasibility was to compare the difference in quality between the hybrid fuzzy-PI controller and a standard PI controller.
The reason fuzzy logic control should be considered for applications such as the regulation of hydrostatic transmissions is that it offers several advantages over more conventional techniques. The two most important of these are that fuzzy logic controllers require only a qualitative understanding of how the targeted device or system operates, and that fuzzy controllers have an inherent ability to control systems with non-linearities.

To establish fuzzy control it is not necessary to mathematically model the system to derive a system transfer function. Instead, fuzzy logic control is based entirely on empirical knowledge of the system and system response under various operating conditions. This knowledge may be obtained from operators that have controlled the system manually, or it may be developed through trial and error in systems that are not subject to manual control. The empirical basis of fuzzy controllers leads to their inherent ability to govern systems with non-linearities. As long as an empirical basis for control can be established, rules can be developed and fuzzy control implemented.

DESCRIPTION OF SYSTEM

A simplified block diagram of the hydraulic system and control electronics used to compare the performance of Fuzzy-PI control with PI control is shown in Figure 1. The system consisted of an adjustable fixed displacement hydraulic pump driving a fixed displacement hydraulic motor. The shaft of the motor was connected to a flywheel used to simulate an inertial load and to a hydraulic pump and variable relief valve arrangement used to simulate an external torque load. The motor shaft also drove a combined speed and torque meter which provided electrical signals proportional to motor speed and motor torque. The pump used to drive the motor was driven by a 3 phase ac induction motor. The swash-plate angle on the pump was controlled by a torque motor-actuated, two-stage valve. The torque motor drove a double-nozzle flapper valve which in turn drove a double spool valve. The spool valve controlled the pump swash-plate angle. Protection in the main hydraulic circuit was provided by two relief valves which limited differential pressure to 17.25 MPa (2500 psi).
The operation of the hydraulic system was not complicated. The pump supplied a constant flow proportional to the swashplate angle. The magnitude of this flow determined the motor speed, which increased as pump flow increased. The dynamics of the system were determined by the hose compliance, the pump and motor leakage, the flywheel moment of inertia and the load damping.

In the lower half of the block diagram the control electronics are displayed. Inputs to the control system were the motor speed signal and the motor shaft torque signal. These signals from the Himmelstein torque and speed meter were connected to a Himmelstein Model 66032 signal conditioner and display unit*. The conditioned motor speed and torque signals were connected to two input channels on an Analog Devices RTI 820 data acquisition system.

The output signal from the data acquisition board was connected to a torque motor drive circuit. The output of the drive circuit was connected to the input of the Sauer Sundstrand MCV104A electronic displacement control unit. The intermediate torque motor driver circuit was designed to boost the current output of the data acquisition board to levels required by the torque motor contained in the MCV104A. The electronic displacement control unit was connected to the pump and set the angle on the pump swashplate. Therefore, by varying the output signal level, the control program set the pump swash-plate angle which in turn set pump flow and motor speed for a given load.

In summary, motor speed and torque signals were monitored by the control program. The program then set or changed the output signal to a level that eliminated any difference between the measured motor speed and programmed speed.

*Trade names are included for the benefit of the reader and do not imply endorsement or preferential treatment of the products by Iowa State University or the USDA.
DESCRIPTION OF CONTROL SOFTWARE

Operator Interface

The control software written in C, was an interactive program that performed the following six functions: file operations; display of fuzzy control rules; manual control and monitoring of the data acquisition board; configuration of the controller for PI or Fuzzy-PI control; activation and deactivation of control; and tracing of control variables while control was active. These six functions were accessed through the operating system display screen.

Real Time Control Flow

After automatic control was started, the control software performed motor speed regulation using the selected control algorithm (PI or Fuzzy-PI). Motor speed and torque values were read and pump output value was calculated every 50 ms.

PI Control Algorithm

A standard algorithm was employed to implement proportional-integral control. The PI control flow chart is shown in Figure 2. The proportional term was set equal to the product of the proportional gain and the speed error (the difference between the motor speed setpoint and measured motor speed). The integral term computation was a two step process. First the integral term increment was obtained from the product of the integral gain and the speed error. This was then added algebraically to the last integral sum (= the sum of all prior integral increments) to give the current integral sum. The pump output was then set equal to the sum of the proportional term and the integral term (= current integral sum).

The PI control algorithm also employed bumpless transfer when switching from manual to automatic control. Bumpless transfer involves setting the current integral term equal to the current manual output level before activating automatic control. Thus if automatic control is initiated and there is no difference between the measured speed and the speed setpoint, the output will not change and the speed will not change.
Fuzzy-PI Control

When Fuzzy-PI control was activated the controller performed fuzzy logic control whenever the motor speed error exceeded a user defined minimum value. When the speed error dropped below this minimum value, the controller reverted to PI control to drive the error to zero. Of course a pure fuzzy logic controller could be implemented by setting the minimum speed error to 0. This was not done for the tests described in this report. Instead, comparisons were made between hybrid Fuzzy Logic - PI control and standard PI control. As will be seen in the discussion on results, this did not obscure the benefits of fuzzy logic control since the performance of the hybrid controller was significantly better than that of the pure PI controller.

A fuzzy logic controller is a rule based controller. The objective of fuzzy logic control is identical to all other techniques of control - to keep the variable under control at some prescribed level (or setpoint). For the process considered in this article the objective was to maintain the motor speed at the desired value. The implementation of fuzzy logic control, however, is quite different from other techniques. A fuzzy logic controller operates on a set of rules that relate combinations of input variables at various levels to output control actions. These rules are generally based on an expert knowledge of system operation. The first step, therefore, in the development of a fuzzy logic controller is to establish a set of rules.

The final rule set employed for the hydraulic motor speed controller is shown in Figure 3a. The rule set was based on two variables: motor speed error and load torque. In the matrix representation of the rule set, motor speed error levels are assigned to the seven columns of the matrix, beginning with Very Small (VS) speed error in column one and progressing up to Very Large (VL) speed error in column 7. Similarly, load torque levels are assigned to the seven rows of the matrix beginning with Very Small torque in row 1 and ending with Very Large torque in row 7.

Each combination of motor speed error magnitude and load torque magnitude has an associated output change magnitude. For example if the motor speed error magnitude is very small and the load torque magnitude is very small, then the output rule states that the output change should be very small. This output
change is indicated in the upper left corner square of the matrix which is at the intersection of the speed error VS column and the load torque VS row. Similarly, if the motor speed error is medium large (ML) and the load torque is medium small (MS), then the output rule for this combination states that the output change should be medium large. This output change is indicated at the intersection of the speed error ML column and the load torque MS row.

The output sign is equal to the sign of the speed error. Therefore, if the motor speed setpoint were greater than the measured motor speed, the speed error would be positive. In this case the sign of the output change would be positive and the output change would be added to the current output level. If the speed error sign were negative, the output change would be subtracted from the current output.

Since there are seven motor speed error magnitude ranges and seven load torque ranges, there are a total of $7 \times 7 = 49$ output rules. In the rule set presented in Figure 3a, not all of the output change magnitudes are specified. There are eight rules where a " - " is located in the output square. In those cases, the output is to remain unchanged.

The rule set as described above is no different from any other rule set used in expert systems or rule based controllers, if the input variable values are assigned to one of the seven ranges (VS, S, MS, M, ML, L, VL) in a conventional manner. Conventional in this case means that each range is unique and does not overlap with any other range. For example, for the motor speed error, let the following ranges be defined:

- **VS:** Speed Error < 100 RPM
- **S:** 100 RPM <= Speed Error < 200 RPM
- **MS:** 200 RPM <= Speed Error < 300 RPM
- **M:** 300 RPM <= Speed Error < 400 RPM
- **ML:** 400 RPM <= Speed Error < 500 RPM
- **L:** 500 RPM <= Speed Error < 600 RPM
- **VL:** Speed Error > 600 RPM
For ranges defined in this way, any one value of speed error will have 100% membership in one of the ranges and 0% membership in all of the rest. For example, if the speed error is 575 RPM, it will have 100% membership in the Large range and 0% membership in all other ranges.

In contrast to a conventional rule-based controller, the ranges in a fuzzy controller overlap. In addition, variables can have any membership value between 0 and 100% in any range. This is illustrated in Figure 3b, which shows the range functions used for both torque and speed error in the hydrostatic transmission controller. For both variables, the middle five ranges are represented by overlapping triangle functions and the two end ranges by half trapezoids. The Large range for speed error, for example, is a triangle function with apex at 600 RPM and end points at 500 and 700 RPM. Similarly, the Small range for torque is a triangle function that begins at 2.3 N-m (20 in-lb) and ends at 11.3 N-m (100 in-lb) with a peak at 6.8 N-m (60 in-lb).

Each range function above overlaps the adjacent range functions. For example, the medium range for speed error is represented by a triangle function with apex at 400 RPM and endpoints at 300 and 500 RPM. The end point at 300 RPM coincides with the apex of the adjacent medium-small range. The endpoint at 500 RPM coincides with the apex of the adjacent medium-large range. The very small and very large ranges only have one adjacent range each and thus only overlap with one other range.

Membership of a variable within a particular range depends on the position of the variable within the range. This is shown in Figure 3b where the membership values for a speed error of 575 RPM and a torque of 7.3 N-m (65 in-lb) are shown. Membership of a variable within any of the fuzzy controller ranges can be determined either analytically or graphically. For the purposes of illustration, the graphical technique is best and will be described here. The first step in determining membership of a specific value is to construct a perpendicular line on the membership function graph through that value. In Figure 3b this is done for 575 RPM and 7.3 N-m (65 in-lb). For all ranges where the perpendicular line does not intersect the range membership function, the variable value has 0% membership. For any range where the perpendicular line intersects with the range membership function, the membership in the range will be non zero. The degree of
membership (i.e. its magnitude) will be given by the ratio of the height of the point of intersection to the height (apex) of the range triangle function. The height of the point of intersection can be determined by similar triangles or can be determined analytically.

For the membership functions shown in Figure 3b, any variable will only intersect two membership functions and so will have non zero membership in only two ranges. For example, the speed error of 575 RPM intersects the medium large and large range functions. In Figure 3b, the height of each range triangle function has been set equal to 4096. Therefore, by similar triangles, the height of the point of intersection of the 575 RPM line with the large range membership function is:

\[
H = \frac{(575 -500)}{(600 -500)} \times 4096 = 3072
\]

Since 3072 is 75% of the height of the triangle, the degree of membership of 575 RPM in the large speed error range is 75%. Similarly, 575 RPM has 25% membership in the medium large speed error range and 7.3 N-m (65 in-lb) has 87.5% and 12.5% membership in the small and medium small torque ranges.

The use of fuzzy membership functions creates a problem when deciding what rule is satisfied for a particular combination of variables. For conventional membership functions, in which a variable has either 100% membership or 0% membership, only one rule will be invoked. If the speed error is 575 RPM and the torque is 7.3 N-m (65 in-lb), the speed error is large and the torque is small when conventional non-overlapping ranges are used. Only one output is active: the output at the intersection of the large speed error column and small torque row in Figure 3a. Under these conditions then the output would be changed by a very large (VL) amount.

When fuzzy membership functions are used for the rule set shown in Figure 3a, there will be 24 rules where at least one of the inputs to the rules is non zero and there will be four rules where both inputs to the rule will be non-zero (the shaded cells in Figure 3a). In effect, instead of the conditions for one rule being
completely fulfilled, and the conditions for all other rules being completely unfulfilled, there are a number of rules in which the conditions are partially fulfilled.

The technique used in fuzzy controllers to solve this problem and generate a final output from a collection of partially fulfilled rules is to quantify the degree of fulfillment of each rule and then use the degrees of fulfillment of each rule as weighting factors to calculate a final output. In fuzzy rule based controllers, the degree of fulfillment of any rule is defined as the minimum value of the degrees of membership of all input values to the rule. For example, a 575 RPM speed error has 75% membership in the large speed error range and 7.3 N-m (65 in-lb) torque has 87.5% membership in the small torque range. The degree of fulfillment for the rule that is invoked when speed error is large and torque is small is 75%. It should be noted that for the membership functions shown in Figure 3b, at most only 4 rules at any one time will have non zero degrees of fulfillment. This is because for each variable there are only two ranges in which the variable has non zero degrees of membership and each rule has two inputs giving a total of 4 rules with non-zero degrees of fulfillment.

Once the degrees of fulfillment of all rules are established, they are used as weighting factors to determine the final output. This procedure is illustrated in Figure 3c for a speed error of 575 RPM and torque of 7.3 N-m (65 in-lb). In that figure the rule output levels of VS, S, MS, M, ML, L and VL were assigned (arbitrarily) numerical values of 1 through 7 respectively. The magnitude corresponding to each rule output was then weighted with the rule degree of fulfillment to establish a final output of 6.5 which is halfway between the VL and L output levels. This technique of generating a final output is called defuzzification and is identical mathematically to the procedure for finding the center of mass. In this case the degree of fulfillment of the rule is analogous to the weight or mass of the object and the output magnitude of the rule is analogous to the position of the object with respect to the origin.

In summary, the fuzzy controller performs three main tasks when generating a new output to control a process. The controller first measures current values of all inputs and determines their degrees of membership in each of the defined ranges of values. Then the controller uses the degrees of membership to
determine the degrees of fulfillment of each rule in the rule set. Finally, the controller calculates a weighted average of all rules with non zero degrees of fulfillment, which is used as the final output. This procedure is summarized in Figure 4.

One other comment should be made on the fuzzy logic control algorithm. The algorithm generates changes to current output levels instead of generating absolute output levels for each rule. The fuzzy control algorithm is therefore a form of integral control, in which increments are being algebraically added to the current output. However, the control is not necessarily linear, as is the case with standard PI control, where the integral increment is equal to the product of the integral gain and the error. For a fuzzy controller the increments may be non linear, depending on how the rules are established. For example, in the rule set employed in this experiment, the output increments are not decreased in a linear manner as the speed error decreases. Nor are they decreased linearly as the torque increases. In fact there is an abrupt step change to no change in output when the torque enters the very large range. This provides added flexibility when attempting to control systems with non-linear responses.

TEST STRATEGY

The first control algorithm to be tested was pure PI control. Standard tuning techniques such as the Zeigler & Nichols method were not used for two reasons. One was that many of these methods place the system in an unstable state while making measurements. Due to the age of some of the equipment in the test stand, it was desirable to minimize the amount of stress on the system. The second reason that these methods were not used was the existence of significant non-linearities in the system due to the startup deadband and the relief valves. Therefore, the PI settings were obtained empirically.

Initially, low values of proportional and integral gain were selected. The response of the system under PI control was tested by applying step changes in the motor speed setpoint, starting at 0 RPM. The endpoints were normally confined to the range between 400 and 1000 RPM. At low values of proportional
and integral gain the times to reach the new setpoint were long. The gains were gradually increased to reduce the response time, until overshoot became too large and the system took too long to reach steady state.

Based on the results of the manual open loop control and automatic closed loop PI control tests performed on the system, an initial set of fuzzy control rules were developed. The rules were developed to account for the limiting effect of the relief valve on system response to large step changes in output to the actuator. The rule set was designed to boost the output substantially for large errors and low torques. However, since system pressure is proportional to load torque, as torque increases the pressure will approach the relief valve trip point. Therefore the fuzzy control rules were set up so that output changes were reduced as the torque increased to prevent the valve from opening. When the load torque entered the very large (VL) range, output changes were set to zero. The Very Large range of load torque was set by examining trace data collected during PI control. The trace data, which consisted of motor speed, motor speed error, load torque, integral sum and controller output values, revealed at what levels of load torque the relief valve fired. Initially the Very Large range was set somewhat below that level.

As will be shown in the next section, the control rules were modified several times until an optimum response was obtained. This response was significantly better than the best response obtainable from the pure PI controller.

**RESULTS**

Set point change tests were performed by starting with the motor speed set to zero, entering a non zero setpoint and then switching to automatic control with the trace function enabled. With the trace function enabled, the controller captured control data for the first 500 output updates (corresponding to the first 25 seconds of control with a 50 ms update period). Included in the trace data was motor speed, motor speed error, load torque, integral sum, fuzzy output increment and pump output. This trace data was saved to files, converted to Lotus compatible format and used to generate response curves for the various control modes and control settings.
Graphs for the response of the PI controller to setpoint changes are shown in Figures 5 and 6 which show the response of the controller as it attempts to drive the motor speed from 0 to 800 RPM (which corresponds to 164 A to D converter counts as noted on the figures). In Figure 5 the response with a relatively large integral gain is shown. The rise time is adequate, however, the overshoot is large and the time to reach steady state long. For this response the integral gain is probably at its maximum desirable setting. In Figure 6 the response with lower values of integral and proportional gain are shown. The rise time is relatively slow, but the output does not oscillate around the setpoint. The response at these settings is close to the boundary between critically damped and underdamped control.

Graphs for the response of the Fuzzy-PI controller to setpoint changes are shown in Figures 7 and 8. These figures show the response of the controller as it attempts to drive the motor speed from 0 to 800 RPM. In Figure 7 the response using the first rule set developed is shown (see Figure 11). While this response was better than either of those for the pure PI controller, there was room for improvement. The two main problems were that the start of the Very Large (VL) torque range was set too far below the point where the relief valve opened, and at low to medium values of torque the output increments were set too small. The first problem resulted in the "kink" in the fast rising portion of the response curve. The second problem resulted in a slower than necessary rise time.

In Figure 8 the response using an improved rule set (Figure 3) is shown. The response is almost optimum. There is little overshoot and no oscillation around the set point. This response was made possible by correcting the problems described above. The Very Large (VL) range starting point was set to a value (= 29.4 N-m = 260 in-lb) that was close to the value of torque at which the relief valve was observed to open. In addition, the output increments in the rule set for all torques below the Very Large range were increased. The net effect was that the controller drove the motor hard until just before the relief valve opened. At the threshold of the relief valve the fuzzy rule set shut down, maintaining the output at a value just below the firing point of the valve. This minimized both the rise time and the overshoot.
When the speed error entered the Very Small range (< 100 RPM), fuzzy control was suspended and PI control took over and drove the error to zero. The integral gain was set low to prevent oscillations from occurring around the setpoint.

The testing at 800 RPM was completed in January of 1992. A year later additional tests were run to obtain a record of the responses of the PI and Fuzzy controllers at different setpoints. The integral and proportional gains for the PI controller were maintained at those values used to obtain an overdamped response at 800 RPM. The Fuzzy rule set was the same as that used to obtain the "optimum" response in the original system.

During the time between the tests, the hydrostatic transmission was used for classroom laboratory exercises and special projects, and was moved to a different location within the building. When the second set of tests were run, it was discovered that the response of the system had changed. In particular, the response of the system to a step change in input had slowed. The cause for this change was not known. However, as systems age their operating characteristics frequently change. Therefore this situation provided a nice opportunity to compare the sensitivities of the two control techniques to changes in the system transfer function.

The response of the fuzzy controller at two different setpoints under these conditions is shown in Figure 9. The rise time at 800 RPM was increased compared to the rise time of the original system. However, the shape and overall quality of the response was unaffected. This is in contrast to the response of the PI controller at 1000 RPM. The PI settings which produced an overdamped response at 800 RPM in the original system, produce an underdamped response in the aged system.

CONCLUSIONS

In the tests described in this paper, speed control of a hydraulic motor using PI and hybrid Fuzzy Logic - PI controllers were compared. It was found that the performance of the Fuzzy-PI controller was significantly better than the performance of the PI controller for set point changes. Use of the Fuzzy-PI
controller resulted in faster rise times and less deadtime with minimal overshoot and a stable steady state output. It was also found that the fuzzy-PI controller was less sensitive to changes in system response than the PI controller.

It can be argued that similar results to those obtained using fuzzy logic control could be obtained by enhancing the PI control with more sophisticated techniques such as feedforward or adaptive control. That is undoubtedly true. However, the purpose of this investigation was to determine the viability of fuzzy logic control and not to determine the best technique for controlling the motor speed of the motor under test. In that respect, fuzzy logic control showed promise as a control technique.

Note on Control Software

The control and fuzzy rule programming software is available to the public without charge. The software can be obtained by sending a written request to:

Jack Ambuel  
National Soil Tilth Lab  
2150 Pammel Drive  
Ames, Iowa 50011

Intended use (commercial, academic or personal) should be indicated. Included with the request must be a self addressed, stamped diskette mailer, with pre-formatted diskette (3 1/2 - LD or HD; or 5 1/4 HD). Only one copy will be mailed per request. If any of these conditions are not met, the software will not be mailed.
REFERENCES


FIGURE 1: SYSTEM BLOCK DIAGRAM
Compute Integral Term Using Integral Gain $K_I$

$I = K_I \cdot E_k + \sum[K_i \cdot E_j] \quad 1 \leq j \leq k-1$

Measure Motor Speed and Compute Speed Error at Time $T = T_k$
$E_k = \text{Speed Error} = \text{Setpoint} - \text{Speed}$

Compute Proportional Term Using Proportional Gain $K_P$
$P = K_P \cdot E_k$

Add Integral Term and Proportional Term to Determine Total Output
$O = P + I = K_P \cdot E_k + \sum[K_i \cdot E_j] \quad 1 \leq j \leq k$

**FIGURE 2: PI CONTROL FLOW CHART**
<table>
<thead>
<tr>
<th>Torque</th>
<th>VS</th>
<th>S</th>
<th>MS</th>
<th>M</th>
<th>ML</th>
<th>L</th>
<th>VL</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS</td>
<td>VS</td>
<td>MS</td>
<td>M</td>
<td>ML</td>
<td>L</td>
<td>VL</td>
<td>VL</td>
</tr>
<tr>
<td>S</td>
<td>VS</td>
<td>MS</td>
<td>M</td>
<td>ML</td>
<td>L</td>
<td>VL</td>
<td>VL</td>
</tr>
<tr>
<td>MS</td>
<td>VS</td>
<td>MS</td>
<td>M</td>
<td>M</td>
<td>ML</td>
<td>L</td>
<td>VL</td>
</tr>
<tr>
<td>M</td>
<td>VS</td>
<td>MS</td>
<td>MS</td>
<td>M</td>
<td>M</td>
<td>ML</td>
<td>L</td>
</tr>
<tr>
<td>ML</td>
<td>VS</td>
<td>MS</td>
<td>MS</td>
<td>MS</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>L</td>
<td>-</td>
<td>S</td>
<td>MS</td>
<td>MS</td>
<td>MS</td>
<td>MS</td>
<td>MS</td>
</tr>
<tr>
<td>VL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**FIGURE 3a: FINAL FUZZY RULE SET**
FIGURE 3b: SPEED ERROR AND TORQUE MEMBERSHIP FUNCTIONS
FIGURE 3c: DEFUZZIFICATION OF OUTPUT USING CENTER OF GRAVITY TECHNIQUE

\[ \text{OUTPUT} = \frac{[1024 \times 6 + 3072 \times 7 + 512 \times 6 + 512 \times 5]}{[1024 + 3072 + 512 + 512]} = 6.5 \]
FUZZY CONTROL FLOWCHART

Measure Load Torque and Motor Speed
Compute Speed Error = Setpoint - Speed

Determine Degrees of Membership of Torque and Speed Error in Each Fuzzy Subrange:

\[ DOM_x = \frac{DOM_{\text{max}}}{C_x - C_{x-1}} \cdot [V - C_{x-1}] \quad \text{for } V < C_x \]

\[ DOM_x = DOM_{\text{max}} - \frac{DOM_{\text{max}}}{C_x - C_{x-1}} \cdot [V - C_{x-1}] \quad \text{for } V \geq C_x \]

where \( C_x \) = Magnitude at center point of kth subrange
V = Current value of variable

Determine Degrees of Fulfillment for Each Rule:

\[ DOF_{jk} = \min(DOM_j(\tau), DOM_j(\epsilon)) \]

where \( DOF_{jk} \) = Degree of Fulfillment for jkth Rule
\( DOM_j(\tau) \) = Degree of Membership of torque in subrange \( j \)
\( DOM_j(\epsilon) \) = Degree of Membership of speed error in subrange \( j \)

Set Addition to Output Equal to Center of Gravity of the Degrees of Fulfillment for all Rules:

\[ OUTPUT = COG = \frac{\sum \{DOF_{jk} \cdot IO_{jk}\}}{\sum \{DOF_{jk}\}} \]

where \( IO_{jk} \) = Ideal Output for jkth Rule
\( DOF_{jk} = 100\% \)

FIGURE 4: FUZZY CONTROL FLOW CHART
FIGURE 6: OVERDAMPED PI CONTROL

Motor Speed (ADC Input Counts)

PI Control: Kp=0.1; Ki=0.05; SP=164 (-800 RPM)

Time (Seconds)
FUZZY-PI Control: $K_p=0.1; K_i=0.05; SP=164$

FIGURE 7: FUZZY-PI CONTROL WITH INITIAL RULE SET
FUZZY-PI Control: \( K_p = 0.1; \ K_i = 0.05; \ SP = 164 \ (\approx 800 \text{ RPM}) \)

**FIGURE 8: FUZZY-PI CONTROL WITH OPTIMIZED RULE SET**
FUZZY-PI Control for Two Setpoints

Kp = 0.1; Ki = 0.05; SP1=164(=800RPM); SP2=226(=1103RPM)

FIGURE 9: FUZZY-PI CONTROL AFTER CHANGE IN SYSTEM RESPONSE
PI Control:  $K_p=0.10; K_i=0.05; SP=205 \ (=1000 \text{ RPM})$

**FIGURE 10: PI CONTROL AFTER CHANGE IN SYSTEM RESPONSE**
<table>
<thead>
<tr>
<th>Torque</th>
<th>VS</th>
<th>S</th>
<th>MS</th>
<th>M</th>
<th>ML</th>
<th>L</th>
<th>VL</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td>S</td>
<td>MS</td>
<td>M</td>
<td>ML</td>
<td>L</td>
</tr>
<tr>
<td>S</td>
<td>VS</td>
<td>VS</td>
<td>S</td>
<td>MS</td>
<td>M</td>
<td>ML</td>
<td>L</td>
</tr>
<tr>
<td>MS</td>
<td>VS</td>
<td>VS</td>
<td>S</td>
<td>S</td>
<td>MS</td>
<td>M</td>
<td>ML</td>
</tr>
<tr>
<td>M</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td>S</td>
<td>S</td>
<td>MS</td>
<td>M</td>
</tr>
<tr>
<td>ML</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>L</td>
<td>-</td>
<td>-</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
</tr>
<tr>
<td>VL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

FIGURE 11: INITIAL FUZZY RULE SET
CHAPTER 7: FUZZY LOGIC SOYBEAN PLANT EVALUATION PROGRAM

A paper to be submitted to Breeding Science

J.R. Ambuel and S. Ninomiya

ABSTRACT

Soybean plant shape evaluation is an important part of the soybean plant breeding process in Japan. To be selected, a new variety must have what the soybean breeder considers to be a good plant shape. This selection process is currently performed by visual inspection by the soybean plant breeder. This paper describes a method to evaluate soybean plant shape quality automatically.

A program to evaluate soybean shape quality was written. The program developed was an expert system using fuzzy logic rule sets to evaluate soybean shape quality. The program was written using a spreadsheet. The program operated on variables extracted from digitized images of each soybean plant. The program placed the shape of each soybean plant into one of three categories: good (3), fair (2), and poor (1). These were the same categories used by plant breeders to select soybean varieties. Only those rated as good were selected by the soybean breeders. The goal was to develop a program that would give the same ratings as those given by the soybean plant breeders.

The shape quality evaluation program results were approximately the same as those obtained using statistical discriminant analysis. The results indicate that with further refinements to the ruleset, it may be possible to obtain very close agreement between the evaluation of the program and that of the breeders.

S. Ninomiya, National Institute of AgroEnvironmental Science, Tsukuba, Ibaraki 305, Japan.
INTRODUCTION

In soybean breeding, one important factor used in plant selection is plant shape. Plant shape is believed to be related to lodging resistance, light interception, yield and suitability for machine harvesting (Ninomiya and Shigemori, 1991).

Currently, soybean plant shape quality is evaluated visually by expert soybean breeders. Recently, attempts have been made to automate the process using image analysis and statistical discriminant analysis (Ninomiya 1991). The purpose of that approach was to develop quantitative measures of soybean plant shape based on the heuristic rules developed by breeders, and to provide an automatic method of evaluating soybean plant shape to assist the breeders in the selection process. Subsequent to the application of discriminant analysis, the image data was processed using two other techniques: a neural network based selection program and an expert system based on fuzzy logic (Ninomiya et al., 1993). In this paper, the fuzzy logic expert system plant evaluation technique is described.

There were several reasons why it was decided to investigate the use of a fuzzy logic expert system for evaluation of soybean plant shape. One reason was that fuzzy logic evaluation is based on a heuristic instead of mathematical approach and is therefore inherently easier to understand. The rules are described linguistically in common language. This makes working with the evaluator much easier for researchers when tuning the system and for plant breeders when applying the system. In contrast, the techniques based on discriminant analysis and neural networks do not have the direct connection with the breeder's rules for evaluating plant shape.

A second reason for using fuzzy logic was that because it is an expert system, there are no constraints on the type of processes it is used to describe or simulate. This is in contrast to discriminant analysis where there are a number of constraints. For example, when discriminant analysis is used on the 18 shape variables to characterize soybean plant shape, the shape variables are assumed to have a normal distribution. Because the fuzzy expert system is not subject to constraints, it has the potential to provide a better match with the breeders' evaluations.
This paper describes the development of the expert system and compares the results of the program when used to evaluate the shape of 875 soybean plants with the evaluations of a team of soybean plant breeders and with the results of the evaluation based on discriminant analysis.

**FUZZY LOGIC EXPERT SYSTEMS**

Fuzzy logic, invented in 1965 (Zadeh), has in the last 10 to 15 years been applied in a large number of consumer goods and in industrial and transportation control (Williams 1991; Kosko, 1993). Originally envisioned as a technique for analyzing complex biological systems that are not subject to mathematical analysis, fuzzy logic has found its greatest use to date in process and device control. However, recently more attempts have been made to apply the techniques of fuzzy logic to problems of the life sciences.

The soybean plant shape evaluation program described in this paper was a modified version of a fuzzy logic expert system yield model (Ambuel 1993). For the details of that program, implemented in a spreadsheet, the reader is referred to the paper discussing the yield model (Ambuel). A brief review of the operation of this model is given below.

The fuzzy logic soybean plant shape evaluator was developed using a spreadsheet. Contained in the spreadsheet were the fuzzy logic rulesets, the soybean plant shape variable membership functions, the soybean plant shape variable raw data, and the programs (macros) used to link the raw data with the rule sets and extract the results. The program also calculated the percentage of matches between program and plant breeder evaluations.

The program used simple overlapping triangle functions for the membership functions just like those used in the fuzzy logic yield model. The program used two variables per rule set and organized the rule sets hierarchically, with the output of the top level rule set being the plant shape quality. Again, this structure was identical to that employed in the fuzzy logic yield model.

The program was tested on the images from the 875 soybean plants of the above mentioned study that used discriminant analysis to evaluate plant shape (Ninomiya 1991).
FUZZY LOGIC PLANT SHAPE EVALUATION TOOL DEVELOPMENT

The development of the plant shape evaluation tool was an iterative process. The basic steps in the development were:

1. Selection of the image variables to be used in the rules.
2. Selection of the number of fuzzy ranges to be used for each rule.
3. Assignment of the fuzzy ranges for each of the image variables used in the program.
4. Establishing how each of the selected plant shape image variables contributes to plant shape quality. This information was extracted from the experience of plant breeders.
5. Use of the plant breeders knowledge to develop the fuzzy rules.
6. Application of the complete program to evaluate soybean plant shapes.
7. Comparison of program plant shape evaluation with breeder evaluation. Iteration of either part of or all of the above six steps based on the results. Modifications include addition of new plant shape image variables and modification of the rule sets.

These steps are described in detail in the following paragraphs.

Selection of Image Variables: The variables used in the initial fuzzy logic plant shape evaluator were selected from the variables employed in the study of plant shape classification based on discriminant analysis (Ninomiya 1991). In that study, 18 plant shape indicators were extracted using image analysis techniques from the digitized images of each of the 875 soybean plants. The plant shape quality was then obtained by performing discriminant analysis on the 18 variables for each plant.

For the initial implementation of the fuzzy logic program, it was decided to use a reduced set of variables in order to simplify the development process. Therefore, only the variables that were most important in determining shape quality (as defined by expert breeders) were used.

The most important characteristics for good soybean plant shape quality are that the plant be symmetrical in shape (not skewed), that the stem be straight, that there be a moderate amount of leaves (not
too many and not too few) and that the plant be relatively slim (height greater than width). From these considerations the following four shape indicators were selected for development of the first fuzzy rule sets:

1. The degree of occupancy shape indicator:

\[ D = \frac{\text{AREA}}{\text{WDT} \times \text{HGT}} \]

where \( \text{AREA} = \) the surface area covered by plant leaves in the two dimensional image of the plant.

\( \text{WDT} = \) the maximum width of the plant

\( \text{HGT} = \) the height of the plant

The degree of occupancy is a fairly good indicator of plant leaf area.

2. The normalized width of the plant:

\[ \text{SWDT} = 1000 \times \frac{\text{WDT}}{\text{HGT}} \]

This is an indication of the plant slimness. Note that the height of all plants was fixed at 1000 mm so that the normalization factor was 1000/\( \text{HGT} \) where \( \text{HGT} = \) actual height.

3. The skew of the plant with respect to the horizontal:

\[
\text{XSK} = \sum_{i=1}^{N} \left[ \frac{\text{SWDT}}{N} \cdot \left( i \cdot \frac{1}{2} \right) - X_m \right]^2 \cdot \frac{f_x(i)}{XSD^2}
\]

where

\[
X_m = \sum_{i=1}^{N} \left[ \frac{\text{SWDT}}{N} \cdot \left( i \cdot \frac{1}{2} \right) \cdot f_x(i) \right]
\]

\[ XSD = \sqrt{\sum_{i=1}^{N} \left[ \frac{\text{SWDT}}{N} \cdot \left( i \cdot \frac{1}{2} \right) - X_m \right]^2 \cdot f_x(i)} \]

\( f_x(i) = \) the normalized height of the plant at the \( i \)th position, also known as the horizontal frequency distribution.

\( N = \) the number of sections that the image is partitioned into in the \( x \) direction.
This is the third moment of the horizontal distribution and is an indication of horizontal symmetry.

4. The lag in the horizontal direction:

\[ XD1 = AXS^* - XM \]

where \( AXS^* \) = normalized location of the plant main axis. The main axis is the linearized plant stem.

\( XM \) = the normalized location of the mean of the plant horizontal distribution.

This indicator is a measure of the degree of bending of the plant.

Some of these parameters are illustrated in Figure 1. For a complete description of the variables, the reader is referred to Ninomiya (1991).

Rule Descriptions: Having selected the variables for use in the shape evaluation program, the next step was to develop a general statement of the rules for what constitutes a good plant shape. As seen above, this step was actually a part of the variable selection procedure. The main requirements for good plant shape on which the variable selection was based were:

1. A plant is symmetrical.
2. The stem is straight not bent.
3. There are a medium amount of leaves.
4. Height is larger than its width.

It is these four heuristic rules combined with the ranges of the selected variables that were used to develop the fuzzy rules described below.

Determination of Ranges: The third step in the program development was the assignment of fuzzy ranges to each of the selected variables. For the program developed for this application, the structure used was identical to that used in the fuzzy logic expert system yield model (Ambuel 1993). In that program, as described above, simple overlapping triangle functions were used to represent membership functions for each variable. Rule sets were limited to two variables each, and rule sets were combined hierarchically to generate
the final output (in this case plant shape quality). The number of fuzzy ranges for each variable was set equal to seven in the following order:

- \( \text{VS} = \text{Very Small} \)
- \( \text{S} = \text{Small} \)
- \( \text{MS} = \text{Medium Small} \)
- \( \text{M} = \text{Medium} \)
- \( \text{ML} = \text{Medium Large} \)
- \( \text{L} = \text{Large} \)
- \( \text{VL} = \text{Very Large} \)

The procedure used to assign specific values to each of the fuzzy range centerpoints (peaks of the range triangle membership functions) was identical for each variable. First, the data for the 875 plants was split into three groups of 297, 298, and 280 plants. Then the first group was selected and a sort was performed on each variable. From the sort, the range (max and min) of the variable and its mean were determined. The mean value of the variable was then assigned to the Medium range centerpoint. Values close to the maximum and minimum values of the variable were then assigned to the Very Large and Very Small range centerpoints respectively. The values of the remaining range centerpoints were set so that the separation between ranges was a constant.

The results of the range assignments for the four variables are summarized in Table 1. In addition, the maximum, minimum and mean values for the variables are listed. These values were derived from the 297 plants in the first data set.

Development of fuzzy rule sets: The final step in the program development was the transformation of the general rules for good soybean plant shape into corresponding fuzzy rules. As noted above, because the program was developed in a two dimensional spreadsheet, the number of variables per rule set was limited to two. Since four variables were used in the initial implementation, the number of rule sets required was three.
Two on the first level to process the four variables and one on the second level to combine the outputs of the two first level rulesets and generate the final output (plant shape rating).

The development of the initial ruleset was a somewhat uncertain process. The strategy was to use the general rules obtained from expert soybean plant breeders as a guide. The plan was to make an initial estimate of the fuzzy rule sets and then refine the estimate after analyzing the results. From the breeders' rules on soybean shape the following guidelines for use in the fuzzy rule sets were obtained:

1. The degree of occupancy (D) should be Medium. Not too large and not too small
2. The slimness (SWDT) should be Medium Small or lower.
3. The skewness (XSK) should be Small or Very Small.
4. The lag (XD1) should be low, Small or Very Small.

Using these guidelines, the rule sets shown in Figure 2 were developed. The first rule set combined the leaf area (D) and the slimness (SWDT) variables. The output of this rule set was named light interception (LI), because the slimness and leaf area affect the amount of light intercepted. It could also have been called the yield potential output because at the optimum combination of leaf area and slimness, the yield will be a maximum.

The second rule set combined the horizontal skew (XSK) and the horizontal lag (XD1) variables. The output of this rule set was called the width symmetry (WS) although perhaps a better name would be width shape quality. It is a function of the horizontal skew and the straightness of the stem. Among other things, a plant with good width symmetry or quality will be resistant to lodging.

The third rule set on the second level uses the outputs of the first two rulesets as its input variables. The output of this rule set for the initial shape evaluation program was the plant shape quality (PQ). Since seven levels were used for both inputs to and outputs from the fuzzy rule sets, the output of the plant shape quality rule set had to be transformed into the three levels assigned by the breeders: good, fair and poor. This was done with the following mapping:
GOOD(3): PQ >= 500
FAIR(2): 300 <= PQ < 500
POOR(1): PQ < 300

The transformed output was then compared with the breeder evaluations and with the evaluations performed using discriminant analysis.

RESULTS

The initial shape evaluation program (version 1) was used to evaluate the shape of the plants in the first data set (297 plants). For the purpose of comparison, the statistical discriminant analysis program was used to evaluate the plant shape for the first 297 plants using just the same variables used in the fuzzy logic shape evaluator. The results are summarized in table 2. In that table, the match ratio is defined as the number of shape evaluations that matched (Good, Fair or Poor) between breeder and program divided by the total number of plants. The match percentage is the match ratio multiplied by 100. The discriminant analysis evaluation was slightly better than the fuzzy logic shape evaluation program in terms of the number of matches, and both were better than the random match percentage of 33%.

The next step taken was to modify the fuzzy shape program and add another variable identified as important to shape by the breeders. That variable was what is referred to as heavy head, where too much of the plant material is in the upper 1/2 of the plant. When this was done, there was little change in the results (for either fuzzy or discriminant analysis). Therefore, attention was returned to the original program and attempts were made to improve the results of the fuzzy shape evaluation program by modifying the rule sets.

The approach taken to modifying the rule sets was to examine the variables for those cases where there were mismatches and to attempt to determine what changes in the rule sets would result in a match. This was a brute force approach, and after four iterations, a distinct improvement was obtained. The rule set used (fifth rule set) is shown in Figure 3 and the results are shown in table 3. In this case the program was run on all 3 data sets, and the discriminant analysis was also run separately on the same three data sets. The
results using the modified rule set with the original four variables are close to the results obtained using statistical discriminant analysis.

DISCUSSION

The results of the application of the first fuzzy logic shape evaluation program were encouraging. The results were comparable to those obtained using discriminant analysis on the same variables. However, further improvement is necessary before the program can be used as a tool to assist soybean plant breeders in plant selection.

In order to obtain further improvement in the program results, it was decided that a detailed investigation of plants where program and breeder ratings were mismatched should be performed. This analysis was performed on a selected number of plants from the 875 plants that were evaluated - the first 60 plants from the first data set.

The analysis was performed by looking at the values of each of the four variables used in the evaluation, looking at video tape pictures of the plants, looking at silhouettes of the plants and the branches, and then checking the fuzzy logic plant quality evaluation. After conducting these examinations it was possible to determine the reason why there was disagreement between plant breeder and program evaluations. After this process was completed, soybean plant breeders at an experimental research farm were consulted. The breeders were asked why a particular rating was given to each plant in the first 60 plants where the breeder rating did not match the program rating. The results of the analysis and the breeder consultation were combined and 8 reasons were identified for the 23 mismatches in the first 60 plants of the first data set. These reasons are listed in Table 4.

From these 8 categories of reasons for mismatch, a strategy for improvement of the rule set and program can be developed. The most significant contribution to mismatch in the first 60 plants was breeder error. Upon further inspection by the soybean breeders, six of the mismatched ratings were changed to matched ratings. This raised the match percentage for the first 60 plants from 61.7% to 71.7%. For these
plants, the program operated as it should have and corrected the breeders ratings. No changes to the program are necessary in these cases.

For two of the eight categories of mismatch, it may be possible to eliminate the mismatch by adding variables to the rule set. These two categories are: too much vertical skew; and plant leaf distribution is not tight. Vertical skew is a variable already calculated by the image analysis program and can be easily added to the program. Tightness of leaf distribution is not directly related to any one of the 18 shape indicators and so may be a little more difficult to accommodate.

For three of the eight categories of mismatch, it may be possible to eliminate the mismatch by modifying the rules. These three categories are: the stem is bent but all other characteristics are good; the plant is too slim; and the plant has too few leaves. For the three plants in the unknown category, the action required cannot be determined. Further study is required.

The final category is mismatch caused by a data problem. The stem is bent but the corresponding variable (lag) does not reflect this. This is generally caused by stems which are bent which curve back upon their main axis. Because a linear approximation is done to estimate the location of the stem axis, the net bending can be too low if the stem curves back upon itself. It may be necessary to solve this problem by adding another variable to measure stem curvature.

The next step in the development of the fuzzy logic soybean plant shape evaluator will be to attempt to eliminate the problems identified by the analysis described above. The goal is to develop a program that can be used by soybean plant breeders to evaluate plant shape. However, addition of more variables will result in increased difficulty in selecting and adjusting the fuzzy rules. It has been found in a number of other cases (Ambuel et al., 1994) that as the complexity of the system increases, the number of variables increases and the development of the rule sets becomes more complicated and time consuming. Some automatic method of rule set adjustment or some alternative method, such as neural networks, in which automatic adjustment is inherent, is required.
CONCLUSIONS

The fuzzy logic evaluation of soybean plant shape is about as good as that obtained using discriminant analysis. It is a useful check and can cause breeders to re-evaluate their classifications. In all likelihood the fuzzy logic system can be improved.

BIBLIOGRAPHY


TABLE 1: SHAPE INDICATOR VARIABLES AND FUZZY RANGES FOR INITIAL SHAPE EVALUATOR PROGRAM

<table>
<thead>
<tr>
<th>VALUE</th>
<th>D</th>
<th>SWDT</th>
<th>XSK</th>
<th>XD1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>0.54362</td>
<td>1091</td>
<td>0.826</td>
<td>322.0</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.14684</td>
<td>332</td>
<td>0.00033</td>
<td>0.690</td>
</tr>
<tr>
<td>Mean</td>
<td>0.3452</td>
<td>712</td>
<td>0.41313</td>
<td>161</td>
</tr>
<tr>
<td>VS</td>
<td>0.20</td>
<td>400</td>
<td>0.070</td>
<td>10</td>
</tr>
<tr>
<td>S</td>
<td>0.25</td>
<td>500</td>
<td>0.185</td>
<td>60</td>
</tr>
<tr>
<td>MS</td>
<td>0.30</td>
<td>600</td>
<td>0.300</td>
<td>110</td>
</tr>
<tr>
<td>M</td>
<td>0.35</td>
<td>700</td>
<td>0.415</td>
<td>160</td>
</tr>
<tr>
<td>ML</td>
<td>0.40</td>
<td>800</td>
<td>0.530</td>
<td>210</td>
</tr>
<tr>
<td>L</td>
<td>0.45</td>
<td>900</td>
<td>0.645</td>
<td>260</td>
</tr>
<tr>
<td>VL</td>
<td>0.50</td>
<td>1000</td>
<td>0.700</td>
<td>310</td>
</tr>
</tbody>
</table>

D = Degree of Occupancy. As D increases, leaf area increases.
SWDT = Normalized Width. As SWDT increases, slimness increases.
XSK = Horizontal Skew. As XSK increases, horizontal symmetry decreases.
XD1 = Horizontal Lag. As XD1 increases, stem bending increases.
### TABLE 2: RESULTS OF INITIAL FUZZY LOGIC SHAPE EVALUATION PROGRAM (VER. 1) USING FIRST RULE SET

<table>
<thead>
<tr>
<th>EVALUATION METHOD</th>
<th>PERCENTAGE MATCHED WITH BREEDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy Logic</td>
<td>51%</td>
</tr>
<tr>
<td>Discriminant Analysis</td>
<td>57.6%</td>
</tr>
</tbody>
</table>

### TABLE 3: RESULTS OF INITIAL FUZZY LOGIC SHAPE EVALUATION PROGRAM (VER. 1) USING FIFTH RULE SET

<table>
<thead>
<tr>
<th>DATA SET</th>
<th>EVALUATION METHOD</th>
<th>PERCENTAGE MATCHED WITH BREEDERS</th>
<th>PERCENTAGE OF GOOD PLANTS FOUND *</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fuzzy Logic</td>
<td>56%</td>
<td>83.75%</td>
</tr>
<tr>
<td></td>
<td>Discriminant Analysis</td>
<td>60.6%</td>
<td>77.5%</td>
</tr>
<tr>
<td>2</td>
<td>Fuzzy Logic</td>
<td>57%</td>
<td>70.3%</td>
</tr>
<tr>
<td></td>
<td>Discriminant Analysis</td>
<td>60.1%</td>
<td>73.5%</td>
</tr>
<tr>
<td>3</td>
<td>Fuzzy Logic</td>
<td>50%</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td>Discriminant Analysis</td>
<td>50%</td>
<td>75%</td>
</tr>
</tbody>
</table>

* This is the percentage of plants classified as good by the breeders that were also classified as good by the program.
### TABLE 4: REASONS FOR MISMATCHES BETWEEN FUZZY LOGIC SHAPE EVALUATION PROGRAM RATINGS AND BREEDER RATINGS IN FIRST 60 PLANTS

<table>
<thead>
<tr>
<th>Reason for Mismatch</th>
<th>Number of Plants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical Skew Too High: too few leaves in upper or lower half of plant</td>
<td>3</td>
</tr>
<tr>
<td>Breeder Error. Not a problem with rule set.</td>
<td>6</td>
</tr>
<tr>
<td>Too much space in some sections of plant. Plant leaf pattern is not tight</td>
<td>3</td>
</tr>
<tr>
<td>Stem is bent (high lag) but all other characteristics are good. Rating should be higher</td>
<td>2</td>
</tr>
<tr>
<td>Data problem: stem is bent but lag is low. Rating should be higher. Not a problem with rule set</td>
<td>4</td>
</tr>
<tr>
<td>Plant is too slim</td>
<td>1</td>
</tr>
<tr>
<td>Plant has too few leaves</td>
<td>1</td>
</tr>
<tr>
<td>Unknown problem with rule set.</td>
<td>3</td>
</tr>
</tbody>
</table>
FIGURE 1: DIGITIZED REPRESENTATION OF SOYBEAN PLANT SHAPE
### Plant Quality Ruleset

<table>
<thead>
<tr>
<th></th>
<th>LIVS</th>
<th>LIS</th>
<th>LIMS</th>
<th>LIM</th>
<th>LIML</th>
<th>LIL</th>
<th>LIVL</th>
</tr>
</thead>
<tbody>
<tr>
<td>WSVS</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>WSS</td>
<td>100</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>300</td>
</tr>
<tr>
<td>WSMS</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>WSM</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>400</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>WSML</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>500</td>
<td>500</td>
<td>600</td>
</tr>
<tr>
<td>WSL</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>500</td>
<td>500</td>
<td>600</td>
<td>700</td>
</tr>
<tr>
<td>WSVL</td>
<td>100</td>
<td>300</td>
<td>300</td>
<td>500</td>
<td>600</td>
<td>700</td>
<td>700</td>
</tr>
</tbody>
</table>

### Light Interception Ruleset

<table>
<thead>
<tr>
<th></th>
<th>LAVS</th>
<th>LAS</th>
<th>LAMS</th>
<th>LAM</th>
<th>LAML</th>
<th>LAL</th>
<th>LAVL</th>
</tr>
</thead>
<tbody>
<tr>
<td>XLMVS</td>
<td>100</td>
<td>400</td>
<td>600</td>
<td>700</td>
<td>600</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>XLMS</td>
<td>100</td>
<td>400</td>
<td>600</td>
<td>700</td>
<td>600</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>XLMMS</td>
<td>100</td>
<td>300</td>
<td>500</td>
<td>600</td>
<td>500</td>
<td>300</td>
<td>100</td>
</tr>
<tr>
<td>XLMM</td>
<td>100</td>
<td>300</td>
<td>400</td>
<td>500</td>
<td>400</td>
<td>300</td>
<td>100</td>
</tr>
<tr>
<td>XLMMML</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>300</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>XLMVL</td>
<td>100</td>
<td>200</td>
<td>200</td>
<td>300</td>
<td>200</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>XLMVVL</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

### Width Symmetry Ruleset

<table>
<thead>
<tr>
<th></th>
<th>XSKVS</th>
<th>XSKS</th>
<th>XSKMS</th>
<th>XSKM</th>
<th>XSKML</th>
<th>XSKL</th>
<th>XSKVL</th>
</tr>
</thead>
<tbody>
<tr>
<td>XD1VS</td>
<td>700</td>
<td>700</td>
<td>600</td>
<td>400</td>
<td>300</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>XD1S</td>
<td>700</td>
<td>700</td>
<td>500</td>
<td>400</td>
<td>300</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>XD1MS</td>
<td>600</td>
<td>500</td>
<td>400</td>
<td>400</td>
<td>300</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>XD1M</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>300</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>XD1ML</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>XD1L</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XD1VL</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

**FIGURE 2: INITIAL RULESET FOR PROGRAM VERSION 1**
### Plant Quality Ruleset

<table>
<thead>
<tr>
<th></th>
<th>LIVS</th>
<th>LIS</th>
<th>LIMS</th>
<th>LIM</th>
<th>LIML</th>
<th>LIL</th>
<th>LIVL</th>
</tr>
</thead>
<tbody>
<tr>
<td>WSVS</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>WSS</td>
<td>100</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>300</td>
</tr>
<tr>
<td>WSMS</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>WSM</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>400</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>WSML</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>500</td>
<td>500</td>
<td>600</td>
</tr>
<tr>
<td>WSL</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>500</td>
<td>500</td>
<td>600</td>
<td>700</td>
</tr>
<tr>
<td>WSVL</td>
<td>100</td>
<td>300</td>
<td>300</td>
<td>500</td>
<td>600</td>
<td>700</td>
<td>700</td>
</tr>
</tbody>
</table>

### Light Interception Ruleset

<table>
<thead>
<tr>
<th></th>
<th>LAVS</th>
<th>LAS</th>
<th>LAMS</th>
<th>LAM</th>
<th>LAML</th>
<th>LAL</th>
<th>LAVL</th>
</tr>
</thead>
<tbody>
<tr>
<td>XLMVS</td>
<td>100</td>
<td>200</td>
<td>500</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>500</td>
</tr>
<tr>
<td>XLMS</td>
<td>100</td>
<td>200</td>
<td>500</td>
<td>700</td>
<td>700</td>
<td>700</td>
<td>500</td>
</tr>
<tr>
<td>XLMMS</td>
<td>100</td>
<td>200</td>
<td>500</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>400</td>
</tr>
<tr>
<td>XLM</td>
<td>100</td>
<td>200</td>
<td>400</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>300</td>
</tr>
<tr>
<td>XLMML</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>200</td>
</tr>
<tr>
<td>XLML</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>100</td>
</tr>
<tr>
<td>XLMVL</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

### Width Symmetry Ruleset

<table>
<thead>
<tr>
<th></th>
<th>XSKVS</th>
<th>XSKS</th>
<th>XSKMS</th>
<th>XSKM</th>
<th>XSKML</th>
<th>XSKL</th>
<th>XSKVL</th>
</tr>
</thead>
<tbody>
<tr>
<td>XD1VS</td>
<td>700</td>
<td>700</td>
<td>550</td>
<td>300</td>
<td>200</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XD1S</td>
<td>700</td>
<td>700</td>
<td>500</td>
<td>300</td>
<td>200</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XD1MS</td>
<td>600</td>
<td>550</td>
<td>450</td>
<td>200</td>
<td>200</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XD1M</td>
<td>500</td>
<td>450</td>
<td>400</td>
<td>200</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XD1ML</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XD1L</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XD1VL</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

**FIGURE 3: FIFTH RULESET FOR PROGRAM VERSION 1**
CHAPTER 8: A FUZZY LOGIC YIELD SIMULATOR FOR PRESCRIPTION FARMING

A paper published in the Transactions of the ASAE

J. R. Ambuel\textsuperscript{1,2}, T. S. Colvin, and D. L. Karlen\textsuperscript{3}

ABSTRACT

Interest in prescription farming has grown as the technology necessary for its implementation has become available. The central concept of prescription farming is that materials (chemicals, fertilizers, seeds) are optimally applied as a function of position within the field. Therefore, profits are maximized and potential adverse environmental effects are minimized. Our objective was to describe how fuzzy logic could be used to develop a crop yield simulator for assessing spatial variability with sufficient accuracy for optimizing application rates. The method is based on predictive yield models developed using field-scale research techniques. Two conceptual, expert system models were developed using fuzzy logic rules. In one model, chemical and physical characteristics of the soil were measured and combined with local meteorological data as input parameters. In the other model, soil properties were estimated rather than measured. The fuzzy logic rule sets were implemented using a spreadsheet. Rule sets were developed to simulate yields for two 16 ha fields in central Iowa. Predicted yields were then compared with measured yields for those fields. Our results indicate that on a relative basis, predicted yields generally agreed with measured yields.

\textsuperscript{1}Reprinted with permission from Transactions of the ASAE. 1994, 37(6): 1999-2009. Copyright © 1994, Transactions of the ASAE.

\textsuperscript{2}Author for correspondence.

\textsuperscript{3}J. R. Ambuel, Dept. Ag. & Biosys. Eng., Iowa State Univ., Ames, IA 50011; T. S. Colvin and D. L. Karlen, USDA-Agricultural Research Service (ARS), National Soil Tilth Laboratory, 2150 Pammel Dr., Ames, IA 50011.
INTRODUCTION

Prescription or site specific farming research is in its infancy, but there is significant interest and activity among farmers and researchers in the development of these practices (Bae, 1987; Elliot, 1987; Schueller, 1988; Brown and Saufferer, 1991; Carr et al., 1991; Kachanoski et al., 1992; Aspinall and Kachanoski, 1993). The goal of prescription farming is to optimize application rates for seed, fertilizer, and other agricultural inputs as a function of location within a field. This contrasts with current farming practices, where uniform rates are generally applied to entire fields. To effectively implement prescription farming, the variability associated with small areas in each field must be determined so that crop yields can be predicted and the potential for response to various inputs determined.

Simulation Modeling

The amount of spatial and temporal data being collected by farmers, agricultural chemical dealers, and researchers working on prescription farming is increasing, but since data can’t be collected from every field, simulation modeling will presumably become more important as the technology evolves. There are many algorithms and programs available for modeling crop growth and predicting yields. These models can be classified according to a number of different criteria including: (1) Time dependence — static vs. dynamic models; (2) Scope — comprehensive vs. limited; (3) Basis — direct process, statistically based, indirect process, or knowledge based.

In static models, average total yield or crop growth is predicted for an entire growing season. Cumulative values of inputs required for growth are used and perhaps adjusted for events that reduce yield such as pest or disease infestations, periods of drought, or excessive wetness. Static models predict a final value for yield or dry matter production. Diminishing return algorithms, such as the Mitscherlich equation or regression equations based on factorial experiments, are examples of static
models that predict yield as a function of one or more inputs (Tisdale, 1985). Dynamic models are time dependent and usually predict crop growth on a daily basis.

Scope provides a second criteria for classifying simulation models. In a limited scope model, effects of one or two inputs on crop yield are considered. In a comprehensive model, all significant factors affecting yield are included. Both types can be either static or dynamic. Limited scope models include the diminishing return nutrient models. There are a large number of models that can predict yield effects of input deficiencies or excesses (Hardjoamidjojo et al., 1982) or response to specific management practices (Seymour et al., 1992). Limited scope models are frequently used as components in comprehensive growth and yield models for several crops.

A third criterion for classifying crop models is the basis or method used to compute crop growth. In the direct process based models, yield is generally a function of the input of energy (radiant and thermal), materials (water and nutrients) and genetic potential (crop variety). Deductions are made for disease, insect, and weed pressures. Computations are based on a basic knowledge of plant physiology. Empirical equations based on material and energy flows and balances are generally used to simulate the accumulation of dry matter or grain. Statistically based models use regression equations or multivariate analysis to predict crop growth or yield. The equations are empirical and not founded on underlying physiological processes. Models based on indirect processes, such as neural networks (Hirafuji, 1991; Uhrig et al., 1992) or inductive learning techniques (Tarbell et al., 1991) have recently been developed to simulate plant growth.

Another recent development in crop growth modeling is the use of knowledge based or qualitative models. Simulation using artificial intelligence techniques (Puccia and Levins, 1985; Round, 1989) and fuzzy logic expert systems (Schmoldt, 1991; Bardossy and Disse, 1993) are examples of knowledge based modeling.
A number of crop growth model programs are listed in Table 1. Most of these are dynamic, comprehensive, direct process based models.

**Modeling for Prescription Farming**

Yield modeling for prescription farming systems has some unique requirements that are not satisfied by many currently available simulation models. To predict yield as a function of position, the model must estimate yields for many small pieces of land. For example, a 32 ha (80 acre) field with 3 by 15 meter (10 by 50 foot) sections would require approximately 7000 simulation runs to obtain yield estimates for all possible cells. What is needed is a procedure that will process the data for all cells in one run.

The model must be easily adapted to local conditions. Yield adjustments for sections of the field with drainage, seepage, erosion, or textural variations must be easily made. For economic reasons, the model should not be dependent on intensive soil testing. If testing is required, the frequency should be low. Finally, adapting the model should not require a complete, analytical understanding of the underlying biological, chemical, and physical processes that affect plant growth, development, and yield response. For these reasons, an expert system approach was selected as the basis for this research. The approach was based on fuzzy logic, because of its simplicity for implementation and demonstrated power in other applications (Kosko and Isaka, 1993).

**Concepts of Fuzzy Logic**

Fuzzy logic was developed by Zadeh (1965) as an alternative method for analyzing complex biological and social science systems. However, the greatest applications of fuzzy logic to date have been in the control of industrial processes and the operation of consumer products (Mamdani et al., 1984; Yamakawa 1989; Lee, 1990a, 1990b; OMRON, 1991; Williams, 1991; Ambuel et al., 1993; Kosko and Isaka, 1993). Fuzzy logic expert systems have their foundation in fuzzy set theory.
Fuzzy sets are best described by comparison with conventional sets (sometimes referred to as "crisp" sets). Conventional and fuzzy sets for different ranges of adult male human heights are illustrated in Fig 1. Seven sets are defined: Very Short (VS), Short (S), Medium Short (MS), Medium (M), Medium Tall (MT), Tall (T), and Very Tall (VT). The sets are shown as conventional (crisp) sets on top and as fuzzy sets in the bottom half of the figure.

In conventional sets, any one value has 100% membership in one set and 0% membership in all other sets. For example, a height of 175.25 cm (5'9") falls within the 172 cm (5'8") to 182 cm (6') range in the conventional set representation. Referring to the upper half of Fig. 1, this range corresponds to the set of Medium heights. Therefore, in conventional set representation, the height of 175.25 cm has 100% membership in the Medium height set and 0% membership in all other sets. Graphically, set membership functions are represented by adjoining, non-overlapping rectangles (Fig. 1). Any value within the rectangle is a member of the set and any point outside the rectangle is not a member.

In contrast to conventional sets, fuzzy sets overlap and membership can have any value between 0% and 100% in any set. This is illustrated in the bottom half of Fig. 1. For the middle five sets, membership functions are defined by overlapping triangular functions. The two end sets are defined by half trapezoids. The fuzzy set membership function for Medium heights is a triangle with apex at 177 cm (5'10") and end points at 167 cm (5'6") and 187 cm (6'2"). Similarly, the set of Medium Short heights is a triangle function that begins at 157 cm (5'2") and ends at 177 cm with a peak at 167 cm. It should be noted that selection of membership functions and the amount of overlap is somewhat arbitrary. Triangle functions were used for the purpose of illustration and to develop the actual yield models. However, membership functions in general can take on any form and are not limited to triangular shapes.
The degree of membership within a particular set depends on the position of the variable within the range defining the set. This position can be used to determine the membership either analytically or geometrically. Using the geometric approach shown in the lower half of Fig. 1, membership values for a height of 175.25 cm are determined as follows. First, a perpendicular line is constructed through the selected value on the membership function graph. For all sets where the line does not intersect the appropriate function, the value has 0% degree of membership. For all sets where the line intersects, membership will be in the non-zero range. The degree of membership will be determined by the ratio of the value at the point of intersection to the value at its peak. The magnitude of the intersection point can be determined analytically or geometrically (using similar triangles).

For the triangle membership functions (Fig. 1), any value will intersect only two membership functions, and therefore, will have non-zero membership in only two sets. The height of 175.25 cm has non-zero membership only in the Medium Short and the Medium sets. In Fig. 1, the magnitude of each triangle at its peak has been set equal to 1. Thus, by similar triangles, the intersection point for the 175.25 cm line with the Medium height membership function is shown in Eq. 1:

\[ M = \frac{(175.25 - 167)}{(177 - 167)} = .825 \quad \text{Eq. [1]} \]

The degree of membership for 175.25 cm in the Medium set is therefore 82.5%, and using the same procedure, 17.5% in the Medium Short set.

In summary, when conventional (crisp) sets are used to characterize the level or magnitude of any variable (height, weight, age, yield, etc.), membership functions for each set are non-overlapping. Any one value has 100% membership in only one set and 0% membership in all other...
sets. In contrast, when fuzzy sets are used, membership functions are overlapping and any one value can have membership in more than one set.

**Fuzzy Logic Yield Models**

An expert system model, based on fuzzy logic, is a rule based procedure that is identical to all other rule based modeling programs. The objective is to establish a functional relationship between an output variable (result) and one or more input variables. These functional relationships are expressed linguistically, in contrast to mathematical models which use equations to establish relationships between inputs and outputs. However, in an expert system model, transformation rules will exist for converting numerical input values to linguistic descriptions of input levels and for converting linguistic descriptions of output levels to numerical values.

Functional relationships in an expert system are established by a set of rules that are derived from expert knowledge on how the system operates. The first step toward development of a fuzzy logic modeling program is to design the model architecture. This requires selecting the input variables and establishing the relationships between input variables, intermediate outputs, and final output.

Model architecture and selection of input variables are important because those factors will determine the cost of applying the model. Models requiring a large number of measurements will generally be more expensive, so the design should minimize the number of measurements required, but maintain sufficient accuracy to meet the desired application.

The number of measurements required in an expert system model can vary greatly depending on model structure. All yield models, both qualitative and quantitative, are based directly or indirectly on the underlying physiological processes of plant growth. For a given amount of solar energy input, with specific climatic conditions (average and daily temperatures), maximum yield will occur at some specific combination of available water and nutrients. The number of measurements
required for modeling or simulation will depend on the detail to which the processes are modeled and whether the model is static or dynamic. For example, simulating the water flow component within a detailed, dynamic crop yield model requires a large number of measurements. These include (1) the initial soil water content, hydraulic conductivity at many points, and distribution of water and plant roots within the root zone; (2) the water holding capacity of the soil; (3) surface drainage characteristics; (4) the amount, rate and timing of all rainfall during the growing season; (5) solar radiation, air temperature, humidity, and wind velocity as a function of time; (6) weed and crop populations; (7) rooting efficiency, and other variables. The amount of water flow would depend on the crop growth, therefore, growth estimates would be required as a form of feedback input to the water flow component. In a complete yield simulation model, water flow, nutrient flow, and energy flow components would be combined with stress factors (disease, weeds) to determine crop growth. Crop growth would then provide feedback input for the water and nutrient flow components.

A large number of measurements are required for dynamic process modeling, and many of the measurements must be made at each time step for which growth and development are calculated. The number of measurements required can be significantly reduced by using a static instead of a dynamic model. In a static model, yearly totals (in the case of rainfall), average values (for temperature) and starting values (for nutrients) are used as input variables, and yield is estimated based on those parameters. The number of measurements required is reduced, but accuracy is also impaired. Further reductions in the number of required measurements can be obtained by modeling the various components with less detail. For example, instead of measuring the inorganic and organic concentrations of the major limiting nutrients in the soil, an estimate of fertility could be used. This value could be based on a single measurement of organic matter at each location or simply by making estimates based on long-term management practices.

The objective for this project was to examine the potential for using fuzzy logic to develop efficient simulation models, with the minimum possible number of measurements, that could predict
METHODS AND MATERIALS

The field data used to develop fuzzy-logic rules for our simulation models were obtained from two 16 ha (40 ac), farmer-managed fields in central Iowa. Production practices included a corn and soybean rotation and the use of conventional tillage and nutrient management practices since 1957. For the years 1989 through 1993, crop yield data were available for each 12.2 m (40 foot) section along eight transects within each field. Transects were approximately 400 m (1320 ft) long and spaced about 50 m (160 ft) apart throughout each field.

Meteorological information was available from a nearby weather station for all five years, but in general, the amount of information available for use in either a dynamic or static yield model was insufficient for either field. Soil nutrient data was available from a portion of the north field for one of the five years. In the south field, a detailed soil survey that provided a more precise characterization of soil series within the field, than was available from the county soil survey, had been prepared by Steinwand (1992). Changes in elevation in the north field were generally gradual, and soil characteristics appeared to be relatively uniform. In the south field, changes in elevation were more abrupt, with well defined hills and depressions. Soil characteristics appeared to be quite variable, a fact confirmed by the detailed soil survey.

The South field exhibited significant variation in topography and soil texture, making it a good candidate for a heuristic yield model. Soil properties at each location, where yields had been measured, were estimated visually for development of the fuzzy-logic model. The North field was relatively uniform in topography and soil texture, making it more difficult to develop a less detailed model by using estimated values for inputs instead of measured values. Therefore, a detailed model
with incomplete data was developed for this field. The model was used to verify the operation of fuzzy logic processing and to determine if additional data collection might be justified.

Yield Model for the North Field

A block diagram showing the rule sets used to develop a fuzzy-logic model for the North field is shown in Fig. 2. Each block represents a physiological function which can affect crop yields, and constitutes one rule set that consists of two inputs and one output. Outputs from lower level rule sets are used as inputs to higher level rule sets. For example, soil water and fertility function as inputs to the materials rule set. Output from the materials rule set, coupled with the energy variable, provide input to the rule set for growth potential. The yield rule set, at the top of the hierarchy, uses outputs from the growth potential and growth inhibition sets as inputs.

The rule sets were implemented in a hierarchical manner using a spreadsheet. This approach is the best for fuzzy rule sets composed of two inputs and one output, since a spreadsheet is basically a two dimensional array or matrix. A hierarchical arrangement was necessary because more than two factors can affect crop yield.

The intermediate fertility rule set #1 is shown as a 7x7 matrix in Fig. 3. Soil pH levels are assigned to the seven rows of the matrix, beginning with Very Low (VL) in row one and continuing to Very High (VH) in row seven. Similarly, soil nitrate (N) levels are assigned to the seven columns of the matrix, beginning with Very Low in column one and ending with Very High in column seven. The combination of soil pH and N provides a partial indication of soil fertility status. For example, a Medium Low (ML) soil pH and a Medium High (MH) soil N concentration would result in a High (H) soil fertility rating, relative to these two parameters. This soil fertility level is determined by the matrix cell where the ML soil pH row and the MH soil N column intersect. Since there are 7 soil pH and 7 soil N levels, there are 49 rules in soil fertility set #1 (one for each of the 49 cells in the 7x7
rule matrix). The output of this set is then combined with the output of fertility rule set #2 (based on %C and Bray extractable P) to establish the total soil fertility rating.

In standard expert systems, values are assigned to the seven ranges (Very Low through Very High) in a conventional manner. That means that each range is unique and does not overlap with other ranges. Each range defines a conventional (crisp) set. Values will therefore have membership in only one set and only one rule will be satisfied.

For yield modeling with fuzzy sets instead of conventional sets, values will have membership in more than one set and more than one rule will be satisfied. This is illustrated in Figure 4 where the 7 levels of pH and N are shown. These sets are similar to those described in the example for adult male heights. Overlapping triangle functions were used to define the fuzzy set ranges so that any value of pH or N will have non-zero membership in exactly 2 of the 7 fuzzy sets.

The use of fuzzy membership functions with yield models creates a problem when trying to decide which rule is satisfied for a particular combination of soil pH and N. For conventional membership functions, only one rule will be satisfied for any combination of pH and N, because each input value has 100% membership in only one set and 0% membership in all of the rest. The rule satisfied will be found at the intersection of the two sets in the rule matrix. However, when the fuzzy membership functions are used (Fig. 4), each value of soil pH and N will have partial membership (0% < degree of membership <100%) in two fuzzy sets.

For example, in Figure 4, the intersection of a soil pH value of 5.75 and a soil N level of 36.7 ppm is shown. The soil pH value intersects two fuzzy sets: the Medium Low set with a membership value of 75% and the Low set with a value of 25%. Similarly, the soil N value intersects the Medium High and Medium sets with membership values of 81.4 and 18.6%, respectively. A total of four soil fertility rules (Fig. 3) will be partially satisfied because each variable has non-zero membership in two sets. The cell corresponding to each of these four rules is highlighted in Fig. 3. Linguistically, the four rules are: (1) if soil pH is Low and soil N is Medium then fertility is Medium;
(2) if soil pH is Low and soil N is Medium High then fertility is Medium; (3) if soil pH is Medium Low and soil N is Medium then fertility is Medium High; and (4) if soil pH is Medium Low and soil N Medium High then fertility is High.

Each rule is only partially satisfied, therefore, the final output will be determined by some combination of outputs for each individual rule. In fuzzy rule based simulation or control, the method to determine the final output from a number of partially satisfied rules is a two step process. First, the degree of fulfillment is established for each rule. When the degree of fulfillment for each rule is determined, the outputs of all partially satisfied rules are combined in a weighted average using the degree of fulfillment as the weight.

The degree of fulfillment for any rule is set equal to the minimum value of the degrees of membership of all input values to the rule. For example, in soil fertility rule set #1, a pH value of 5.75 has a degree of membership of 25% in the Low pH set, while a soil N value of 36.7 ppm has a degree of membership of 18.6% in the Medium set. Therefore, the degree of fulfillment of the corresponding rule is 18.6%. Once the degrees of fulfillment for all partially satisfied rules are determined, the values are used as weighting factors in the calculation of the final output. This procedure is analogous to determining the center of mass of a group of particles. In Fig 5, output levels of VL, L, ML, M, MH, H, and VH were arbitrarily assigned numerical values of 1 through 7 respectively. These output levels are positioned along the x-axis. The y-axis represents the degree of fulfillment for each rule. The degree of fulfillment for each rule is plotted at each output position.

The center of gravity equation used to calculate the overall fertility level, based on the four partially satisfied rules, is given at the bottom of Fig. 5. Therefore, the fertility rating for a soil with pH 5.75 and nitrate N concentration of 36.7 ppm (Fig. 4.) would be 5.23. This corresponds to a fertility level between Medium High and High. The remaining rule sets for the soil fertility component are shown in Figs. 6 and 7. Operation of these rule sets are similar to those operating on pH and N.
Yield Model for the South Field

The yield model for the south field was functionally and architecturally the same as that developed for the north field. It was a hierarchical collection of fuzzy logic rule sets operating in a spreadsheet. The cell layout and programming were identical for each rule set, but agronomically, the yield model for the south field was substantially different from that for the north. The model developed for the south field relied mainly on estimated instead of measured values for soil characteristics. These soil characteristics were combined with measured weather data to form the complete model (Fig. 8).

In the model for the south field, starting from the top of Figure 8, yield was predicted by combining a reduction factor with a growth factor in the overall rule set. The yield reduction factor was derived based on the severity of ponding and drought in the yield reduction rule set. Drought severity was a function of the drought potential and drought index. The drought potential was a location specific variable, that was initially estimated by visually examining the soil texture and slope at each yield section. The drought index was a function of the reduction in rainfall compared to normal levels during the growing season. For rainfall at or above normal, the drought index was at a minimum. As rainfall amounts dropped below average, the drought index was increased. Similarly, ponding severity was a function of the ponding potential and the ponding index. The ponding potential, like drought potential, was location specific and was initially estimated visually. The ponding index was a function of both the yearly and the monthly above normal rainfall during the growing season. The ponding index was at a minimum for rainfall at or below average and increased as the amount of rainfall increased above average. Monthly averages were considered, as well as yearly averages, so that ponding events due to heavy rainfall in one or two months would be detected in years where the rainfall was at or near normal for the year.
Yield potential was a function of energy, water, and total fertility inputs. The energy-water factor was a function of the moisture deviation and the temperature deviation. The moisture deviation was derived from the difference between total rainfall during the growing season and the normal values, while temperature deviation was derived from the difference between average and measured seasonal temperatures. Temperature in this case was used as an estimate of energy delivered. The total fertility was a function of the inherent fertility (or soil organic matter) and the applied fertilizer. The applied fertilizer level was considered constant. The inherent fertility was location specific, and was initially estimated for each yield segment.

In summary, there were five inputs to the fuzzy-logic model: ponding potential, drought potential, organic matter, rainfall, and temperature. The first three were location specific and were initially estimated by observation. The last two were constant for the entire field and were measured values from a local weather station. Operation of the rule sets was identical to those in the model for the north field. This structure was similar to the yield model structure used in the crop growth submodel of the Erosion-Productivity Impact Calculator (EPIC) program. In that model, yield is proportional to the difference of the total above ground biomass and the total root weight. The total biomass is equal to the total potential biomass multiplied by a stress reduction factor. The total potential biomass is a function of energy input (radiation and heat) and a crop conversion factor. Four stress factors are used: nutrient stress, water stress, aeration stress and temperature stress. In the fuzzy logic yield model for the south field, most of the components of the EPIC crop growth model are present. Drought severity corresponds with water stress and ponding severity corresponds with aeration stress. Temperature stress and nutrient stress, however, are not separate entities in the fuzzy logic model. Instead they are incorporated in the fertility and energy-moisture rule sets that determine yield potential. Also, in the fuzzy logic model, radiation is not considered. It is assumed that temperature and radiation are not independent.
RESULTS AND DISCUSSION

Application of the North Field Yield Model

Before the model could be applied, it was necessary to map the fuzzy set ranges to the actual input variable and output yield ranges. The output yield was based on the Iowa average of 8.2 metric tons/ha (130 bu/ac), which was used to set the center of the Medium (M) output range. The remainder of the output ranges were set so that most of the measured yields fell within the extremes (VL and VH). This was accomplished by setting equal increment ranges of 1.26 metric tons/ha (20 bu/ac). The resulting map showing fuzzy ranges and corn yields was as follows: VL = 4.4 t/ha (70 bu/ac); L = 5.7 t/ha (90 bu/ac); ML = 6.9 t/ha (110 bu/ac); M = 8.2 t/ha (130 bu/ac); MH = 9.4 t/ha (150 bu/ac); H = 10.7 t/ha (170 bu/ac); and VH = 12.0 t/ha (190 bu/ac). For input variables, there is no preferred procedure for matching fuzzy set ranges to actual measured values. There is wide latitude in how the values are mapped, because it is the organization of the rule set that is the main determinant of accuracy for yield prediction. In general, it is best to attempt to match the mean measured values to the center of the fuzzy input ranges and to set end points of the fuzzy input ranges close to the minimum and maximum measured values. This will result in maximum sensitivity of the rule set throughout the range of measured values.

Mapping of Bray extractable P and nitrate N to their respective fuzzy ranges was done by setting the Medium (M) fuzzy range centerpoint equal to the mean measured values and the Very Low (VL) and Very High (VH) range centerpoints close to the minimum and maximum measured values. For %C, mapping was skewed somewhat by setting the Medium range centerpoint somewhat lower than the mean measured value. Finally, for soil pH, mapping was done so that the ranges corresponded with what was generally perceived to be Very Low, Medium, and Very High values. There was no compelling reason for making an exception from the general rule of thumb for pH and
The results would probably have been equally satisfactory if the general rule had been used for all four inputs. Mapping of input values to fuzzy range centerpoints is summarized in Table 2.

Having matched the measured values to fuzzy ranges for the four input variables, the next step was to develop the rule sets described above. Again, there are different procedures for developing the rule sets. The procedure used was an iterative one. Initial rule sets were developed based on a general understanding of crop response to the various input levels. The model was run and the predicted yields compared with the measured yields. In this case, the purposes of the testing were satisfied with the initial rule sets and no further refinement was necessary.

The general characteristics of the rule sets were as follows: The intermediate fertility rule set with soil pH and nitrate N as inputs and fertility level as the output showed an increasing level as N was increased. The fertility levels were sharply decreased at very low and very high pH values. For the other intermediate fertility rule set, fertility levels were increased for both increasing Bray P and %C. Sensitivity to C was set much greater than the sensitivity to P in order to account for the physical benefits of increased organic matter such as increased water availability (Hudson, 1994). Output for the overall fertility rule set was made nearly linear and proportional to the two intermediate fertility levels. After the rule sets were in place, the model was run and yields were obtained as a function of position.

Predicted and measured yields for 2 of the 5 transects in the north field are shown in Fig. 9a and 9b. The best match occurred on transect 1 (Fig. 9a). The worst match occurred on transect 5 (Fig. 9b). Agreement between measured and predicted values on the other transects, although better than transect 5, was not very close. The model developed for the north field was incomplete and used only the fertility data. Therefore, this response was not surprising. However, the results did verify the operation of the fuzzy logic yield model using a spreadsheet. The program was then modified to develop a more complete model for the south field.
Application of the South Field Yield Model

The south field simulation model used seven rule sets corresponding to the seven functional blocks shown in Fig. 8. Each rule set was identical in structure to those used for the north field. Each set had 2 inputs and 7 fuzzy levels per input. The conversions between physical values and fuzzy levels for yield and for the measured input values are shown in Table 3. Also shown are the formulas for calculating the drought and ponding indices and the temperature and moisture deviations.

The model was developed and adjusted for corn on transects 5 and 8 in the south field for the three years that corn was grown (1989, 1991, 1993). The rule sets, conversion factors, and estimated values for ponding potential, drought potential, and organic matter were all adjusted so that predicted and measured yields were as close as possible. The result was a common rule set and common conversion factors for both transects and all three years of corn. As expected, the estimated soil characteristic values changed from location to location because soil characteristics were location specific. The adjustment process may seem arbitrary, but with one exception, it is essentially no different from the processes used to develop empirical equations (curve fitting) or neural networks when they are used in modeling. In those situations, the model is adjusted so that the output of the model matches, as closely as possible, the actual measured value being simulated. The one difference between development of a fuzzy-logic model and the more conventional models is that some of the inputs were also adjusted (ponding potential, drought potential, and organic matter). This occurred because the inputs were estimated, not measured. This is legitimate if at some point in time the model becomes stable and can be used to predict yield with no further adjustments.

The model was then tested on transect 7. First the location specific parameters for transect 7 were optimized for the corn yields measured in 1991. Using those parameters, the model was then run for 1989 and 1993. Next the parameters were adjusted using results of 1989 and 1993 and the model was then rerun for all three years. The rules developed for transects 5 and 8 were not changed.
Results for the south field yield model are shown in Figs. 10 and 11. In Fig. 10, the test results for transect 7 are shown using the first set of parameters for the 1989 and 1991 corn crops. In Fig. 11 the test results using the adjusted parameters are shown. Both plots of measured and predicted yield as a function of distance and scatter plots of yield versus measured yield are presented. Each value of yield corresponds to one of the 12 m (40 ft) harvest sections. In Table 4, the average yield and standard error of estimate is given for all three transects in all three years. The predicted and measured yields are generally much closer for this model than the one used on the north field.

On transect number 7 there are two locations at approximately 61 m (200 ft) and 183 m (600 ft) from the eastern border of the field where yields were consistently low. These correspond to two eroded sections on hillslope positions. There are similar locations on transect 8. On transect number 5, there is a large section of very poorly drained silty clay loam soil (Okoboji) from 200 m (650 ft) to 290 m (950 ft) west of the eastern boundary. This area had below normal yields for all three years, and was completely unproductive in the two years with ponding problems (1991 and 1993). In 1991, the spring was excessively wet, while in 1993, most of the growing season was extremely wet. The model identified these areas quite well.

Predicted yields for the most part followed the observed yields with a fair degree of accuracy. Therefore, relative yields predicted by the model followed the relative measured yields well, with locations having relatively high, medium, or low predicted yields also having high, medium, or low measured yields. However, during the adjustment process when parameters for transect 7 were modified to obtain better results, it became apparent that either the rule sets were incomplete or were constructed improperly. Therefore, to obtain greater accuracy, the rule sets would have to be changed.

This type of model may be useful for situations where moderate accuracy is satisfactory. The model could be used in a prescription farming system sometime after the system was established. The following additional steps would have to be taken before applying the model to the field on which it was developed: (1) make further adjustments to the model to improve the accuracy for those years in which yield data is available; (2) take one or two more years of yield data and make any additional adjustments necessary; and (3) begin
variable rate application of fertilizer and use results to make adjustments so that the model makes accurate predictions of yield as a function of applied fertilizer. After completion of these steps, the model could then be used to assist in future variable rate application decisions. It could be run in a "what if" mode to determine yields under various weather conditions and fertilization rates.

There are a number of problems with this model. A major one is that it is time consuming and difficult to adjust manually. This is partly because the model is run in a spreadsheet, so it takes a substantial amount of time to obtain the results after changes are made. The other cause of this problem is that deciding what changes to make is also time consuming. A second problem is associated with the use of estimated soil properties. If the characteristics change over time, perhaps because of site specific farming decisions, then the model would have to be adjusted periodically. Deciding when to make those changes and making them would not be convenient. It would be more satisfactory if the estimated characteristics were replaced with measured values. If the measured values required re-measurement only every 5 to 10 years, then the added cost of the measurements might be more than offset by the cost associated with adjusting the model.

These problems will be addressed in the next phase of this project. A different model is currently being developed for the south field. In this model the estimated soil characteristics will be replaced using measured parameters from the intensive soil survey. In addition, the model being developed will be automatically tuned. This will be accomplished either by replacing the fuzzy logic expert system completely with a neural network, or by using a neural network to automatically adjust the rules for the fuzzy logic model.

**SUMMARY AND CONCLUSIONS**

The objective of this report was to discuss our preliminary experiences with the use of fuzzy logic for developing a crop yield simulation model that may be useful for site-specific or prescription farming. The approach can be implemented using spreadsheet computer technology, but it will require substantial modification before it will be fully operational. The technique, however, does have merit and appears to warrant further research.
REFERENCES


FIGURE 1: CONVENTIONAL AND FUZZY SETS FOR ADULT MALE HEIGHTS
FIGURE 2: YIELD MODEL BLOCK DIAGRAM FOR NORTH FIELD
**SOIL NITRATE NITROGEN**

<table>
<thead>
<tr>
<th></th>
<th>VL</th>
<th>L</th>
<th>ML</th>
<th>M</th>
<th>MH</th>
<th>H</th>
<th>VH</th>
</tr>
</thead>
<tbody>
<tr>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>ML</td>
<td>M</td>
</tr>
<tr>
<td>L</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>M</td>
<td>MH</td>
<td>H</td>
</tr>
<tr>
<td>ML</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td>M</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>VH</td>
<td>VH</td>
</tr>
<tr>
<td>MH</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td>H</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
</tr>
<tr>
<td>VH</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
</tr>
</tbody>
</table>

**FIGURE 3: SOIL FERTILITY INTERMEDIATE LEVEL RULE SET 1 FOR NORTH FIELD YIELD MODEL**
FIGURE 4: MEMBERSHIP FUNCTIONS FOR SOIL pH AND NITRATE NITROGEN
DEGREE OF
FULFILLMENT

VL L ML M MH H VH

OUTPUT

\[ F = \frac{(18.6)(4) + (25)(4) + (18.6)(5) + (75)(6)}{18.6 + 25 + 18.6 + 75} = 5.23 \]

FIGURE 5: FERTILITY LEVEL OBTAINED USING CENTER OF GRAVITY TECHNIQUE
## CARBON

<table>
<thead>
<tr>
<th></th>
<th>VL</th>
<th>L</th>
<th>ML</th>
<th>M</th>
<th>MH</th>
<th>H</th>
<th>VH</th>
</tr>
</thead>
<tbody>
<tr>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
</tr>
<tr>
<td>L</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td>ML</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td>M</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td>MH</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td>H</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td>VH</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>VH</td>
<td>VH</td>
</tr>
</tbody>
</table>

### PHOSPHORUS

**FIGURE 6: SOIL FERTILITY INTERMEDIATE LEVEL RULE SET 2 FOR NORTH FIELD YIELD MODEL**
**INTERMEDIATE FERTILITY OUTPUT 2**

<table>
<thead>
<tr>
<th></th>
<th>VL</th>
<th>L</th>
<th>ML</th>
<th>M</th>
<th>MH</th>
<th>H</th>
<th>VH</th>
</tr>
</thead>
<tbody>
<tr>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>VL</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>L</td>
<td>VL</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>ML</td>
<td>ML</td>
<td>ML</td>
</tr>
<tr>
<td>ML</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>ML</td>
<td>ML</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>M</td>
<td>VL</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>MH</td>
</tr>
<tr>
<td>MH</td>
<td>VL</td>
<td>ML</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>MH</td>
<td>H</td>
</tr>
<tr>
<td>H</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>VH</td>
<td>L</td>
<td>ML</td>
<td>M</td>
<td>MH</td>
<td>H</td>
<td>H</td>
<td>VH</td>
</tr>
</tbody>
</table>

**FIGURE 7: TOTAL SOIL FERTILITY LEVEL RULE SET FOR NORTH FIELD YIELD MODEL**
FIGURE 8: YIELD MODEL BLOCK DIAGRAM FOR SOUTH FIELD
FIGURE 9a: PREDICTED AND MEASURED YIELDS FOR FIRST TRANSECT ON NORTH FIELD
FIGURE 9 b: PREDICTED AND MEASURED YIELDS FOR FIFTH TRANSECT ON NORTH FIELD
FIGURE 10 a: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING FIRST PARAMETER SET: 1989 CORN
FIGURE 10a: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING FIRST PARAMETER SET: 1989 CORN
FIGURE 10b: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING FIRST PARAMETER SET: 1991 CORN
FIGURE 10 b: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING FIRST PARAMETER SET: 1991 CORN
FIGURE 11a: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING SECOND PARAMETER SET: 1989 CORN
FIGURE 11 a: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING SECOND PARAMETER SET: 1989 CORN
FIGURE 11 b: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING SECOND PARAMETER SET: 1991 CORN
FIGURE 11 b: TEST RESULTS FOR SEVENTH TRANSECT ON SOUTH FIELD USING SECOND PARAMETER SET: 1991 CORN
TABLE 1: PROCESS BASED DYNAMIC CROP GROWTH MODELS

<table>
<thead>
<tr>
<th>NAME</th>
<th>CROPS</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPIC</td>
<td>Multiple</td>
<td>Williams,J.R. 1991.</td>
</tr>
<tr>
<td>CORNF</td>
<td>Corn</td>
<td>Arkin and Strapper. 1980</td>
</tr>
<tr>
<td>CORNMOD</td>
<td>Corn</td>
<td>Baker and Horrocks. 1976</td>
</tr>
<tr>
<td>SIMAIZ</td>
<td>Corn</td>
<td>Duncan, W.G. 1975.</td>
</tr>
<tr>
<td>SORGF</td>
<td>Sorghum</td>
<td>Maas and Arkin. 1978.</td>
</tr>
<tr>
<td>PUTU</td>
<td>Wheat</td>
<td>Singels A. 1992</td>
</tr>
<tr>
<td>SOYGRO</td>
<td>Soybeans</td>
<td>Hoogenboom,G., J.W.Jones, and K.J.Boote. 1990</td>
</tr>
<tr>
<td>SOYMOD</td>
<td>Soybeans</td>
<td>Jones. 1990.</td>
</tr>
<tr>
<td>GLYCIM</td>
<td>Soybeans</td>
<td>Jones. 1990.</td>
</tr>
<tr>
<td>REALSOY</td>
<td>Soybeans</td>
<td>Jones. 1990.</td>
</tr>
</tbody>
</table>
TABLE 2: MAPPING OF INPUT VALUES TO FUZZY RANGE CENTER POINTS

<table>
<thead>
<tr>
<th>INPUT</th>
<th>UNITS</th>
<th>VL</th>
<th>L</th>
<th>ML</th>
<th>M</th>
<th>MH</th>
<th>H</th>
<th>VH</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td></td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>PPM</td>
<td>10</td>
<td>17</td>
<td>24</td>
<td>31</td>
<td>38</td>
<td>45</td>
<td>52</td>
</tr>
<tr>
<td>Phosph.</td>
<td>PPM</td>
<td>9</td>
<td>17</td>
<td>25</td>
<td>33</td>
<td>41</td>
<td>49</td>
<td>57</td>
</tr>
<tr>
<td>Carbon</td>
<td>%</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
<td>2.5</td>
<td>3</td>
<td>3.5</td>
<td>4</td>
</tr>
</tbody>
</table>
### TABLE 3: PARAMETER CALCULATION AND CONVERSION FORMULAS

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>FORMULA</th>
<th>NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly Ponding Index = MPI</td>
<td>Sum of the monthly rainfall for months when rainfall exceeds monthly threshold</td>
<td>Monthly thresholds: April, May = 5.1 cm; June, July, Aug, Sept = 7.6 cm</td>
</tr>
<tr>
<td>Yearly Ponding Index = YPI</td>
<td>Sum of the monthly deviations in rainfall from average from April through September plus 20% of the deviation from October to March</td>
<td>Monthly deviation = Total rainfall during the month - average rainfall for the month</td>
</tr>
</tbody>
</table>
| Ponding Index = PI            | YPI if YPI > 0  
MPI if YPI < 0                                                           |                                                                    |
| Drought Index = DI            | Sum of the monthly deviations in rainfall from average from April through September plus 20% of the deviation from October to March | If DI > 0 then DI set = 0                                           |
| Moisture Deviation = MD       | Sum of the monthly deviations in rainfall from average from April through September plus 20% of the deviation from October to March |                                                                    |
| Temperature Deviation = TD    | Sum of the monthly deviations in temperature from average from April through September | Monthly deviation = average temperature for the month - historical average for the month |
| Fuzzy Ponding Index = FPI     | FPI = 29.5* PI + 100 for PI in cm  
FPI = 75*PI + 100 for PI in inches |                                                                    |
| Fuzzy Drought Index = DI      | FDI = -27.6* DI + 100 for DI in cm  
FDI = -70*DI + 100 for DI in inches |                                                                    |
| Fuzzy Moisture Deviation = FMD | FMD = 5.9*MD + 400 for MD in cm  
FMD = 15*MD + 400 for MD in inches |                                                                    |
| Fuzzy Temperature Deviation = FTD | FTD = 22.5*TD + 400 for TD in degrees C  
FTD = 12.5*TD + 400 for TD in degrees F |                                                                    |
| Fuzzy Yield = FYI             | FYI = 38.2*Y + 100 for Y in t/ha  
FYI = 2.4*Y + 100 for Y in bu/ac |                                                                    |
## TABLE 4: SUMMARY OF RESULTS OF YIELD PREDICTION ON SOUTH FIELD

<table>
<thead>
<tr>
<th>TRANSECT</th>
<th>Average Yield (t/ha)</th>
<th>Standard Error of Estimate (t/ha)</th>
<th>1989</th>
<th>1991</th>
<th>1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>8.9</td>
<td>1.14</td>
<td>7.75</td>
<td>1.48</td>
<td>3.0</td>
</tr>
<tr>
<td>8</td>
<td>9.16</td>
<td>1.44</td>
<td>10.72</td>
<td>1.13</td>
<td>5.44</td>
</tr>
<tr>
<td>7 (P1)</td>
<td>8.92</td>
<td>2.35</td>
<td>8.08</td>
<td>1.09</td>
<td>5.55</td>
</tr>
<tr>
<td>7 (P2)</td>
<td>8.92</td>
<td>1.34</td>
<td>8.08</td>
<td>1.43</td>
<td>5.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 9: CONCLUSIONS

The emerging field of soft-computing shows promise in modeling and control of complex agricultural and biological processes. In this section the application of a branch of soft-computing, fuzzy logic expert systems, to three problems has been described: control of the motor speed on a hydrostatic transmission; classification of soybean plant shape quality; and prediction of corn yields. It was seen that as the complexity of the problem increased, the training and optimization of the fuzzy rules became more difficult.

In order for application of fuzzy logic to complex problems to become practical, the training (adjustment) of the rule sets must be automated. This can be accomplished by using neural networks to adjust the rule sets (Kosko, 1993; Kandel and Langholz, 1994). Alternatively, a neural network can be used in place of fuzzy logic. For the simpler hydrostatic transmission control problem, this is not necessary. But for the soybean plant shape evaluator, and the corn yield model, neural net training would be required.

Another consideration in the further development of the corn yield model is whether more measured physical parameters should be added to the model or not. In the model presented in this paper, the only measured values were weather data. Properties of the soil were included in the model but were not actually measured. Instead they were estimated and then adjusted using one or two years of yield data for training. The advantage of this approach is low cost, because weather data was published and no additional measurements were required. Using a neural network in place of the fuzzy logic model may require even fewer variables. It would be an interesting experiment to see if a suitable neural network yield model could be developed to estimate yields using only past yield data, weather data, and fertilization rates. It may be that most of the chemical, physical and biological characteristics of the soil are fully revealed in several years of relative yield history. If this were the case, then the data required for variable rate application could be obtained very inexpensively, requiring only on the go yield monitoring and perhaps on site weather data collection.
References


PART 3: IMPLEMENTATION OF VARIABLE RATE APPLICATION OF MATERIALS IN PRESCRIPTION FARMING
CHAPTER 10: INTRODUCTION

As noted previously, there are four major subsystems required for the implementation of variable rate application of materials in prescription farming: a method for determining material application rates as a function of location in the field; a system for determining absolute location of the farm equipment in the field; a system or systems for varying the application rates of materials to the field as a function of position; and a system to monitor the results (measure yield) so that application rates can be optimized. The first two parts of this dissertation were devoted to the first of these subsystems - the determination of application rates. In this part, the last three subsystems will be discussed.

As discussed in the introduction to this dissertation, the technology for the last three subsystems is being developed rapidly by agricultural equipment, position equipment, and information systems manufacturers. This is paralleled by academic research into positioning, control and monitoring technology. In this part, a brief review of possible implementations of these three subsystems will be given. The second chapter discusses positioning systems with a focus on those based on the Global Positioning System (GPS). The next chapter contains a brief discussion of control system requirements for the various materials. In the fourth chapter, yield monitoring research and development is reviewed.
A satellite-based positioning system is being developed to determine the location of farm equipment while working in the field. A satellite receiver mounted in a moving tractor or combine calculates position coordinates based on signals received from a selection of satellites in the Department of Defense Global Positioning System (GPS). The position coordinates are retrieved from the receiver by a computer and combined in a common data base with additional information collected by the computer such as yield, soil fertility, soil moisture, temperature, implement draft, and fuel consumption. There are a number of possible applications for this system. One is the automation of position information and data collection for research purposes. Comparison of results between years can be easily made. Another application is the optimized and automated application of manure, fertilizers and chemicals. Profits can be increased and environmental damage minimized.


Author for correspondence.

J. R. Ambuel, Dept. Ag. & Biosys. Eng., Iowa State Univ., Ames, IA 50011; K. Jeyapalan, Dept. of Civil Engineering, Iowa State University, Ames, IA 50011; T. S. Colvin, USDA-Agricultural Research Service (ARS), National Soil Tilth Laboratory, 2150 Pammel Dr., Ames, IA 50011.
INTRODUCTION

This paper describes the development of a satellite positioning system used to determine location of farm equipment while working in a field. The paper is organized into four sections. In the first section, the purpose and applications of the system are described. This is followed by a review of current techniques of position determination and recent research into positioning. The next section describes the operation of the system being developed at the National Soil Tilth Laboratory and Iowa State University. The final section discusses the current status of the system and future development.

PURPOSE OF POSITIONING SYSTEM

Position information for farm equipment is useful for a wide variety of applications both in production and research. On commercial farms, operators are searching for ways to reduce monetary costs and environmental damage by minimizing the amounts of fertilizers, herbicides, and pesticides applied. A number of systems have been or are being developed that combine soil characteristics, yields, fertility, weed pressure and pest problems in a common data base with all data referenced to position in the field. These data are then used to minimize the amount of chemicals applied to the crops and soil, consistent with economic production. Obviously, accurate and timely position information is essential to proper operation of these systems.

Position information systems are also important in University and on-farm research. Research trials typically run from 3 to 10 years or longer in duration. Accurate and repeatable positioning information is a requirement for the meaningful comparison of results from year to year. An example of on-farm research being conducted by the National Soil Tilth Laboratory is shown in Figure 1. That figure shows the map of soil types on two 16.2 hectare (40 acre) tracts. The two tracts are on adjacent farms in Boone County, Iowa. The management of one farm follows conventional tillage, crop rotation and chemical treatment practices. The management of the other farm uses alternative tillage practices, applies herbicides and pesticides selectively and at low rates and substitutes manure, sewage sludge and legumes for commercial fertilizers. The purpose
### Key to Map

<table>
<thead>
<tr>
<th>MAP UNIT</th>
<th>SOIL TYPE</th>
<th>CLASSIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Okoboji silt clay loam, 0 to 1% slope</td>
<td>Cumulic Haplaquoll</td>
</tr>
<tr>
<td>55</td>
<td>Nicollet loam, 1 to 3% slope</td>
<td>Aquic Hapludoll</td>
</tr>
<tr>
<td>107</td>
<td>Webster silt clay loam, 0 to 2% slope</td>
<td>Typic Haplaquoll</td>
</tr>
<tr>
<td>138B</td>
<td>Clarion loam, 2 to 5% slope</td>
<td>Typic Haplaquoll</td>
</tr>
<tr>
<td>138C</td>
<td>Clarion loam, 5 to 9% slope</td>
<td>Typic Haplaquoll</td>
</tr>
<tr>
<td>138C2</td>
<td>Clarion loam, 5 to 9% slope, moderately eroded</td>
<td>Typic Haplaquoll</td>
</tr>
<tr>
<td>138D2</td>
<td>Clarion loam, 9 to 14% slope</td>
<td>Typic Haplaquoll</td>
</tr>
<tr>
<td>507</td>
<td>Canisteo silt clay loam, 0 to 2% slope</td>
<td>Typic Haplaquoll</td>
</tr>
</tbody>
</table>

**FIGURE 1.** RESEARCH SITE 32.37 HECTARES (80 ACRES) FOR COMPARING CONVENTIONAL (LEFT) AND ALTERNATE (RIGHT) FARMING SYSTEMS IN BOONE COUNTY, IOWA.
of the Tilth Laboratory research is to compare the effects of the two management systems on the agroecosystem of each area. The project is multidisciplinary. Comparison of the effects of the two management systems on rainfall infiltration, water gradients, soil physical characteristics, soil chemistry and nutrient levels, soil biology (focusing on earthworm populations), crop yields and microclimates are being made. All of these investigations are tied to soil type and location.

An example of data being collected is shown in Table 1. Corn yields as a function of soil type are shown for 1989. These yields were determined by harvesting small plots along transects of the research tracts. The yields for all plots were grouped by soil type and averaged to determine the mean yields. The position of the harvester was determined manually (dead reckoning) and the yield measurement was performed as a batch as opposed to a continuous process. Thus the entire 12.2 x 2.1 meter (40 x 7 foot) plot was harvested and then the weight and moisture measured with the combine stopped.

This manual process of determining position and yield has two major problems: it is subject to error and it is inefficient. Automated position and yield monitoring can eliminate both problems. Use of GPS to determine position in real time and a computer to continuously monitor the yield and integrate that information with the position information into a common database is one of the goals of this project. The database will then be used as an input into a Geographic Information System (GIS).

TECHNIQUES OF POSITION DETERMINATION - CURRENT AND EMERGING

There are five main categories of position sensing that have been used for farm vehicles: mechanical systems, leader-cable, ultrasonic, ranging, and navigational.

Mechanical sensing systems include those that employ contact sensors mounted on the tractor to detect crop rows, or sensors that follow furrows. Neither of these techniques are suitable for automatic position sensing for the purpose of data collection or selective control of the application of fertilizers and pesticides.
## TABLE 1. CONVENTIONAL VS. ALTERNATIVE FARM SYSTEM CORN YIELDS FOR 1989.

<table>
<thead>
<tr>
<th>MAP UNIT</th>
<th>Conventional Yield in Mg/ha (Bu/Ac)</th>
<th>Alternative Yield in Mg/ha (Bu/Ac)</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>13.3 (152.5)</td>
<td>13.1 (150.4)</td>
</tr>
<tr>
<td>138B</td>
<td>13.0 (148.7)</td>
<td>12.2 (139.8)</td>
</tr>
<tr>
<td>507</td>
<td>13.3 (153.0)</td>
<td>12.2 (139.8)</td>
</tr>
<tr>
<td>107</td>
<td>12.2 (140.4)</td>
<td>13.3 (152.7)</td>
</tr>
<tr>
<td>138C2</td>
<td>11.3 (129.5)</td>
<td>11.8 (125.2)</td>
</tr>
<tr>
<td>138D2</td>
<td>---</td>
<td>9.9 (114.1)</td>
</tr>
</tbody>
</table>
In leader cable systems, wire is buried under the ground and energized with audio frequency current. The magnetic field is detected by search coils mounted on the farm vehicle. The system is not suitable because it does not permit easy changing of field row patterns, and is extremely expensive to install.

Ultrasonic systems are similar to mechanical contact sensing with the contact feelers being replaced with ultrasonic emitters which detect presence of crop rows, or furrows by sensing reflected sound energy. Like mechanical sensing, this technique is not suitable for the applications under consideration.

Ranging systems are based on surveying techniques where some form of electromagnetic energy (RF, infrared, light) is generated on the moving vehicle. This energy returns to the moving vehicle after striking two or more reflectors at known locations. The tractor position is then determined by triangulation. Palmer (1989) has employed this technique successfully using RF energy with fixed transmitters and a mobile receiver. The disadvantages of this system are limited range, and the need to maintain the reflectors or transmitters under adverse environmental conditions. Gordon and Holmes (1988) used laser energy and achieved six meter accuracy, but only over a range of 300 meters. Heil et al. (1986) developed a microwave positioning system for agricultural vehicles and reported good results at distances of up to 578 meters. Interest in navigational systems, particularly those tied to the Global Positioning System, has increased in recent years. Larsen et al. (1988) discussed a system similar to the one described in this paper.

**POSITION INFORMATION SYSTEM -DESCRIPTION OF OPERATION**

The position information system being developed at the National Soil Tilth Laboratory and Iowa State University uses a satellite receiver to collect signals from the Department of Defense Global Positioning System (GPS) satellites. The receiver, which is mounted on the farm equipment, uses these signals to calculate the position of the equipment in the field. That position information is then collected by a computer and may be combined with other information of interest in a common database for future analysis. Or it may be combined with position dependent information collected at some earlier time and used to perform real time control of equipment operations. As described previously, this real time control can include optimizing
application of manure, fertilizer and chemicals, so that excess application of material is avoided and environmental damage is prevented. Both the data collection and real time control functions are done automatically, without the intervention of the operator. A detailed description of operation will now be given.

**Operation of the Global Positioning System**

The Global Positioning System consists of a number of satellites in fixed, known orbits around the earth. Each satellite transmits two carrier frequencies: the L1 carrier at 1,575.42 MHz and the L2 carrier at 1,227.6 MHz. The carriers are modulated with two types of code and a navigation message. The two codes used to modulate the carriers are the P code (precision code) and the C/A code (course/acquisition code). Either code can be used to determine position, with the P code providing 10 to 20 meter accuracy and the C/A code providing 20 to 30 meter accuracy (Georgiadou and Doucet, 1990). Only the C/A code is available for unrestricted civilian use.

A receiver based on earth that acquires signals broadcast by 3 or more satellites can use the signals to compute its latitude and longitude. The receivers may be operated in one of two modes: stand alone or differential. In the stand alone mode, the receiver operates exactly as described above, receiving signals from satellites and calculating position using those signals. The receiver calculates the position in the following manner. As mentioned previously, the satellites modulate the L1 and L2 carriers with the P code, C/A code and navigation information. The navigation information includes the orbital position of the satellite. Therefore, by demodulating the carriers, the receiver can obtain the position of the satellite. The receiver can also measure the time required for each satellite signal acquired to travel from the satellite to the receiver. The receiver accomplishes this by generating a code identical to the satellite code (P code for military receivers and C/A code for commercial receivers). The receiver then code locks this replica with the received code by shifting the start time of the replica until maximum correlation is obtained. Since the receiver knows the nominal starting time, Ts, for the received code (which is repeated at regular predetermined intervals) and it knows the time shift, Tr, required to obtain code lock, it knows the time for the signal to travel from satellite to
the receiver, which is just the difference between the nominal start time for the satellite signal and the start time for the receiver replica. Multiplying this transit time by the speed of light gives the nominal distance (or pseudo range) between the satellite and the receiver:

\[ P = (T_r - T_s)c \]

This distance can also be expressed as the vector distance between satellite and receiver using earth based coordinates:

\[ P = \sqrt{(U_s - U_r)^2 + (V_s - V_r)^2 + (W_s - W_r)^2} \]

This equation contains three unknowns, the position coordinates of the receiver \( (U_r, V_r, W_r) \). If signals from three satellites are acquired, then these unknowns can be determined:

\[ P_1 = \sqrt{(U_{s1} - U_r)^2 + (V_{s1} - V_r)^2 + (W_{s1} - W_r)^2} \]

\[ P_2 = \sqrt{(U_{s2} - U_r)^2 + (V_{s2} - V_r)^2 + (W_{s2} - W_r)^2} \]

\[ P_3 = \sqrt{(U_{s3} - U_r)^2 + (V_{s3} - V_r)^2 + (W_{s3} - W_r)^2} \]

If signals from four satellites are acquired, then a term can be added to correct for the receiver clock error giving the following equations (Leick, 1990, pp. 205-206):

\[ P_1 = \sqrt{(U_{s1} - U_r)^2 + (V_{s1} - V_r)^2 + (W_{s1} - W_r)^2} + dT_r \times c \]

\[ P_2 = \sqrt{(U_{s2} - U_r)^2 + (V_{s2} - V_r)^2 + (W_{s2} - W_r)^2} + dT_r \times c \]
\[
P_3 = \sqrt{\left((U_{S3} - U_{tr})^2 + (V_{S3} - V_{tr})^2 + (W_{S3} - W_{tr})^2\right)} + d_{Tr} \times c
\]

\[
P_4 = \sqrt{\left((U_{S4} - U_{tr})^2 + (V_{S4} - V_{tr})^2 + (W_{S4} - W_{tr})^2\right)} + d_{Tr} \times c
\]

There are a number of errors associated with the stand alone mode of operation. These include errors in the satellite atomic clocks, geometric resolution errors, and errors associated with propagation through the atmosphere. All of these errors can be eliminated by operating the system in the differential mode. In differential mode the receiver, in addition to monitoring satellite signals, will receive error information from a remote base station located at some known position. The base station will also be monitoring satellite signals. In addition the base station will have preprogrammed into its memory the precise position at which it is located. The base station will compare that position with the position computed using the satellite signals. The difference between known and calculated locations will then be transmitted to the receiver mounted on the equipment. The receiver will adjust its calculated position using that difference. This entire process is accomplished in real time.

**Operation of the Farm Equipment Position Information System**

The farm equipment position information system being developed at the National Soil Tilth Laboratory is shown in Figure 2. The system as shown consists of a GPS mobile receiver, radion, RFmodems, a GPS base station receiver, a datalogger, and a computer. The purpose of this particular system is to collect equipment performance information as a function of equipment position.

The data collection is accomplished as follows. A variety of sensors monitoring tractor performance are connected to the datalogger. These sensors include fuel input and return sensors, a radar ground speed sensor, an axle speed sensor, and an equipment drawbar draft sensor. The datalogger monitors the sensors and stores their current values in memory.
At the same time that data is being collected, the equipment position and velocity are being determined by the GPS satellite receiver mounted on the farm equipment. The receiver is operated in the differential mode and is therefore linked via radio to a base station positioned at some known location. The base station also receives signals from the GPS satellites and calculates position based on those signals. The position is then compared with the known position of the base station and the difference (or error) is calculated. The error is then sent to the mobile station which uses the error to determine its actual position more precisely.

The position and equipment performance information is collected by an IBM compatible, environmentally hardened personal computer. For the system currently under development, the program running on the PC collects the position information from the receiver and the performance information from the datalogger and combines the two sets of data in a single file. The data are time stamped so that the file contains performance information as a function of position and time. Position and performance information are collected every one second. For a tractor running at 8 km/hr (5 mph) that corresponds to a resolution of about 2.3 meters (7.5 feet). The combined information is then transformed into a format compatible with the GIS database software. The database is then used to create archival records, make year by year comparisons, and make comparisons with other position based information collected at the same location.

POSITION INFORMATION SYSTEM - CURRENT STATUS

At the time this article was completed (6/15/91), the status of the National Soil Tilth Laboratory satellite based positioning system was as follows. The stand alone system became operational in early May and was put to use on the ISU research farms collecting information for an ongoing tillage and herbicide experiment. The differential system operation was verified in the lab and a preliminary calibration run made around the ISU campus to estimate the position accuracy in the differential mode. The preliminary test indicated a worst case error of 12 meters. Integration with the GIS database was also begun.
FIGURE 2. SATELLITE BASED POSITIONING SYSTEM
SUMMARY

The position sensing system for farm equipment is being developed using the Global Positioning System. When operational, the system will permit automatic combination of vehicle position information with research data (yield, vehicle performance, soil properties). Data and position will be combined in a common database and used for analysis. The database can also be used to control future farm operations to optimize the amount of fertilizers and chemicals applied, thereby maximizing profits and protecting the environment.

REFERENCES


CHAPTER 12: VARIABLE RATE CONTROL

Variable rate control technology is a second major component of prescription farming. It is in this area where the technology is most easily and economically adapted to meet the needs of prescription farming. There are a number of companies providing control systems for sprayers, liquid applicators and planters. Mid-Tech (1994) markets an electronically controlled injection spray system. The controller can be purchased with an RS-232 interface and can be remotely programmed by a computer. Remote programming capability makes it suitable for use in a prescription farming operation. Microtrak (1993) also provides sprayer and NH3 electronic controllers. This company is developing products for a complete prescription farming system. All products can be connected together and controlled over their proprietary Trak-Net communications system. Hiniker (1993) provides an electronic control system for crop spraying and anhydrous ammonia injection. The controller monitors a flow meter and ground speed sensor while controlling the sprayer or injection control valve. It does not have an RS232 interface, however, one could easily be added in future models for use in prescription farming. Raven Industries (1993) also provides an electronic sprayer control system similar to the Hiniker. With an attachment to the anhydrous line, it can also be used to control NH3 application rates. As with the Hiniker controller, a communication port is not provided, but could easily be incorporated in future models. Rawson Control Systems (1993) produces an electronic unit for control of planter seeding rates. This unit interfaces to a radar ground speed sensor and hydraulic motor drive. The microprocessor in the controller sets the motor speed to maintain the programmed seeding rate at the current planter speed. The company provides controllers with RS-232 ports for use in prescription farming systems. Ag-Chem (1993) provides a complete variable rate application system including the application truck, GPS receiver, ground speed sensor, application controller, and application software. Ag-Chem received a patent for application of materials based on digital maps. At the time of publication, this patent had been challenged in the courts.

Testing of commercial products and research into new control products is also taking place in Universities. Mid-Tech sprayer controllers were tested in a prescription herbicide application experiment at Texas A&M (Rudolf and Searcy, 1994). The study found that one of the main problems that needs to be
addressed is reduction in the response time from injection to application. More basic research into pump and valve control is also being performed. A control system for a servovalve controlled centrifugal pump was designed and tested (Xu et al., 1992). The results indicated that the response of the pump to electronic control was sufficiently fast and accurate for use in variable rate application of chemicals in prescription farming.

The emerging technologies described above are not those best suited for prescription farming. These technologies are constrained by the prevailing design of seed, fertilizer and chemical application equipment. This equipment has been designed based on the current farming practices where one rate is used for an entire field. As a consequence, as the farms have grown in size, the equipment has also increased in size so that 16 row planters, fertilizers and chemical applicators are available and in use on the larger farms. This equipment was designed for fixed rate application over the width of the equipment. This is not the best design for a prescription farming system, because as was shown above, significant yield variations can take place over distances shorter than the width of the larger pieces of equipment.

In an ideal prescription farming system, application rates could be varied row by row. This would permit the operator to vary application rates on closely spaced narrow strips of land located within an area with a nearly uniform yield history (or yield potential). This, of course would require that the harvesting system be capable of measuring yields in each individual strip, so that the effects of the variable seeding and/or fertilization rates in adjacent strips can be compared. This ideal system is not possible to implement with today's equipment. Both the equipment for applying materials and the harvesting equipment are not designed to treat single rows or small strips of land.

References


CHAPTER 13: YIELD MEASUREMENT

An essential part of any prescription farming system is a method to measure results. This generally means some equipment to measure yield. This measurement provides the feedback necessary to evaluate and optimize the performance of the prescription farming systems. Yield measurements combined with yield and weather histories can be used to adjust rates of application.

While both academic research and commercial development of yield measuring devices are proceeding at a rapid pace, the yield measuring process is constrained by the current design of harvesting equipment, just as control of the application processes is constrained by current design of material application equipment. In the case of corn, soybeans and small grains, harvesting on large mechanized farms is done by combine. The combine represents an even larger constraint to implementation of a well designed prescription farming system than does the equipment for the application of fertilizers, seeds and chemicals.

There are at least three major problems with using a combine to measure grain yield in a prescription farming system. The first problem is that commercial combines are designed to harvest grain from multiple rows of grain. Typically 5 to 6 rows for corn and more for soybeans and small grain. This makes it difficult or impossible to do variable application rate experiments on areas of the field with similar yield potential as described in the preceding chapter. It is not possible to measure the grain yield in individual rows or narrow strips of grain with combines currently in use. Therefore, even if it were possible to apply variable rates of fertilizer or plant variable populations of seeds in adjacent narrow strips on a commercial farm, it would not be possible to measure the yields separately in each of those strips to evaluate the variable application rates. This would require specialized harvest equipment that would not be economically viable.

A second problem with current combine technology, is that there is mixing or diffusion of the grain in the combine from sections of higher grain yields to sections of lower grain yields. This effect tends to smooth out the peaks and the valleys in the harvested grain before the grain enters the grain tank. This is a problem with current yield monitor technology. Most yield monitors available at this time are placed at the end of the clean grain auger and measure the grain flow by sensing force or sensing mass in some manner. The grain at
this point has gone through the combine and the mixing process. The accuracy of the measurement when
made in the clean grain path, therefore, is reduced by this mixing process.

A third problem with current combine technology is the time delay from the time of cutting to the
actual entry of the grain into the tank. This problem is not as severe as the other two since that time can be
measured. However, it is another source of inaccuracy.

While there are problems with current yield monitoring technology, it is not known how significant
these problems are. Future research and application may show that farms can be made more profitable with
the application of prescription farming techniques without extensive modifications to existing equipment, in
spite of the fact that application and yield monitoring equipment design is not ideal. Research should
therefore proceed in multiple directions. Research should be conducted with specialized equipment designed
to measure grain yields in narrow strips. This equipment should be used to harvest research plots on which
adjacent narrow strips of grain are planted with varying seeding and fertilization rates. Side by side
comparisons of variable rate application on adjacent narrow small strips within areas of similar yield history or
potential could be performed. These tests would give the most accurate measure of the value of prescription
farming.

Research should also be conducted on commercial farms using conventional harvesting and planting
equipment with added yield monitoring and rate control technology. In this research, variable application
rates would be varied over larger areas and the results would be less precise. Comparison of these results with
the more accurate research farm results could then be used to answer two questions:

1. Can prescription farming techniques be used profitably with existing equipment?
2. Would modification of existing farm equipment to make prescription farming more accurate
   be economically justifiable.

At the present time, most effort in industry and in academia is directed toward using existing
harvesting equipment in prescription farming by adding on yield monitors. A variety of techniques for yield
measurement have been investigated. A pivoted auger flow sensor placed between the clean grain elevator
output and the grain tank was shown to have ± 3% accuracy when measuring the yield of small plots (0.1 ha) of Kansas wheat and grain sorghum (Wagner and Schrock, 1987). This device was installed on a commercial combine. A flow sensor using a triangular paddle elevator installed on a commercial combine was evaluated and found to have total grain readings that were ± 5% within actual readings for the harvested field (Howard et al., 1993). This accuracy was achieved without masking the in field variations in yield. An impulse type flow sensor was used to measure grain flow in the clean grain elevator of a small grain combine (Vansichen and DeBaerdemaeker, 1991). This device had a demonstrated accuracy of ± 2.6% with 95% confidence. Another approach to yield measuring was the use of a weigh tank and weigh bars suspended between the output of the clean grain auger and the main grain tank (Colvin, 1990). This technique has only been developed for stop and go measurements and not continuous yield measurements. It's accuracy is only limited by the accuracy of the weigh bar strain gages.

Commercial yield monitors are also available (Walter, 1994). Both Ag Leader Technology and Micro-Trak offer impulse type flow meters to measure instantaneous grain yield. The sensor for each is mounted in the clean grain elevator. Dronningborg sells a clean grain elevator monitor that measures grain flow using a gamma ray source and detector. The German combine manufacturer Claas also offers a yield monitor as an option.

Research and development of yield monitors is relatively new. A number of problems need to be resolved, even with those monitors designed to work with existing harvest equipment. Most of the monitors, for example, require calibration each time they are used. This is an inconvenient procedure and one that is prone to error. A number of monitors also have problems operating at low combine speeds or low grain flow rates.
References


SUMMARY AND CONCLUSIONS

In this dissertation the requirements for a variable rate material application system in prescription farming have been described. The four major aspects of a variable rate application system were identified: a positioning system; methods of determining application rates; application control systems; and yield measurement devices for analyzing results. It was shown that three of these components - positioning, control and yield measurement - have some remaining economic and technical problems, but that technologies suitable for prescription farming are already commercially available. The fourth aspect of a variable rate application system, however, is not well understood. Determining how much material to apply as a function of position, especially for fertilizer and seeding application rates, will require years of additional research. It was this component of variable rate material application systems that was the main focus of this dissertation.

Three strategies for determination of fertilizer application rates were identified:

1. Adaptation of field scale Extension Service recommendations.
2. Real time on-the-go soil property sensing.
3. Development of prescriptions based on yield histories and weather data.

Field scale recommendations are based on average values for a large number of fertility studies conducted on each major soil association. But there are large variations in soil characteristics within soil associations, so that application rates based on average results for an entire association would not be optimal at most locations within a field. Furthermore, fertility levels by themselves are often poor predictors of yield at specific locations and additional information on the soil physical characteristics is required to determine optimum fertilizer application rates. The second method - real time sensing of soil properties - is experimental and untested. Only a few sensors have been developed and the relationship between the soil properties they measure and the optimum rate of fertilizer application has not been established. The third strategy of using yield histories and weather data combined with expert systems to determine application rates is also untested.
and would require a long time to optimize application rates. However, of the three, the yield history approach is probably the least expensive, and the most likely to have success, at least in the near term.

The recommended procedure for implementation of the yield history approach is as follows:

1. Acquire a detailed yield history for each field in the farm. At least 3 years for each crop.

2. Look for areas of consistent yield or areas of consistent yield under specific weather conditions (e.g. dry years, normal years, and wet years).

3. Determine causes of consistently low and consistently high yielding areas by observation and chemical and physical testing of the soil if necessary.

4. Conduct multiyear tests of variable rate application of fertilizer in the areas of consistent yield, fertilizing narrow adjacent strips at variable rates in each of the areas of consistent yield. Measure the yields in each of the strips.

5. Based on test results, estimate if variable rate application will be profitable on those areas of consistent yield.

6. If profitable, then develop an expert system model to serve as decision support system combined with long range weather forecasts to determine variable application rates for the entire field.

7. If the areas of consistent yield are found to be insignificant, then steps 3 through 5 are not applicable and the only alternative is to conduct variable rate application tests on the entire field to determine if site specific farming for the field will be profitable. Again, if profitability is indicated, then a yield model can be developed to assist in future application rate decisions.

In this dissertation, one expert system using fuzzy rules was developed and tested on a commercial farm for which there were three years of corn yield data collected in adjacent 12 meter (40 foot) long by 3 row wide plots along eight different 340 meter (1320 feet) long transects. The rules were adjusted manually on two of the transects and then applied to a third transect and reasonably good agreement between measured and predicted yields at each position was obtained. This type of model could be combined with variable rate fertilizer application results and used as a decision support system for variable rate application. However, the model described in this dissertation should not be used for that purpose without further refinements. In particular, some method of automatically tuning the rule sets should be incorporated. This could be accomplished by developing a neural network to adjust the rules, or by using a neural network in place of the fuzzy yield model to predict yields.
This type of expert system should not be developed and applied when first converting a farming operation to prescription application of materials. As noted in the recommended procedure above, a number of preliminary steps should be taken before developing and using a yield model to assist in rate application decisions. Initially only those areas where yields are relatively consistent from year to year are considered. Concentrating on the areas of consistent yield has at least two advantages. One advantage is that by concentrating on areas of consistent yield it should be much easier to determine the effects of variable rate application of fertilizer or seeds. The second advantage is that this approach should be able to give a good indication on whether prescription application will be profitable. If the areas of consistent yield do not show a benefit when farmed on a prescription basis, then it is unlikely that those areas with more complicated yield patterns will benefit. If, however, there are few areas of consistent yield on a field, then variable rate application tests must be performed for the entire field.