Improving Nitrogen Management in Corn-Wheat-Soybean Rotations Using Site-Specific Management in Eastern Virginia

Wei Peng

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Darrell Bosch, Chair
Marcus Alley
Anya McGuirk
Saied Mostaghimi
James Pease
Patrick Phipps
Daniel Taylor

November 2, 2001
Blacksburg, Virginia

Keywords: Site-Specific Management, Nitrogen, Cluster Analysis, Markov Chain, Information, Variable Application

Improving Nitrogen Management in Corn-Wheat-Soybean Rotations Using Site-Specific Management in Eastern Virginia

Wei Peng

(ABSTRACT)

Nitrogen (N) is a key nutrient input to crops and one of the major pollutants to the environment from agriculture in the United States. Recent developments in site-specific management (SSM) technology have the potential to reduce both N overapplication and underapplication and increase farmers' net returns. In Virginia, due to the high variability of within-field yield-limiting factors such as soil physical properties and fertility, the adoption of SSM is hindered by high grid-sampling cost. Many Virginia corn-wheat-soybean farms have practiced generating yield maps using yield monitors for several years even though few variable applications based on yield maps were reported. It is unknown if the information generated by yield monitors under actual production situations can be used to direct N management for increased net returns in this area.

The overall objective of the study is to analyze the economic and environmental impact of alternative management strategies for N in corn and wheat production based on site-specific information in eastern Virginia. Specifically, evaluations were made of three levels of site-specific information regarding crop N requirements combined with variable and uniform N application. The three levels of information are information about the yield potential of the predominant soil type within the field, information about yield potentials of all soils within the field (soil zones), information about yield potentials of smaller sub-field units which are aggregated into functional zones. Effects of information on expected net returns and net N (applied N that is not removed by the crop) were evaluated for corn-wheat-soybean fields in eastern Virginia. Ex post and ex ante evaluations of information were carried out.

Historical weather data and farm-level yield data were used to generate yield sequences for individual fields. A Markov chain model was used to describe both temporal and spatial yield variation. Soil maps were used to divide a field into several soil management units. Cluster analysis was used to group sub-field units into functional zones based on yield monitor data. Yield monitor data were used to evaluate ex post information and variable application values for 1995-1999, and ex ante information and variable application values for 1999.

Ex post analysis results show that soil zone information increased N input but decreased net return, while functional zone information decreased N input and increased net returns. Variable application decreased N input compared with uniform application. Variable application based on soil zone information reduced net return due to cost of overapplication or underapplication. Variable application based on functional information increased net return.

Ex ante results show that information on spatial variability was not able to increase farmers' net return due to the cost of variable N application and information. Variable rate application decreases N input relative to uniform application. However, imprecision in the spatial predictor makes the variable application unprofitable due to an imbalance between costs of under- and over-application of N. Sensitivity analysis showed that value of information was positive when temporal uncertainty was eliminated.

The ex post results of this study suggest there is potential to improve efficiency of N use and farmers' net returns with site specific management techniques. The ex ante results suggest that site specific management improvements should be tested under conditions faced by farmers including imperfect information about temporal and spatial yield variability.

Acknowledgements

I think the true value of my whole dissertation research is distilled in this terribly short section in which I would like to say THANK YOU to all those who had helped me through out this study.

First of all, thanks are to my major advisor, Dr. Darrell Bosch. If it were not because of his encouragement and mentoring, I would have given up long time ago. He spent so much time on my academic achievements in the past seven years. He never complained or criticized me of anything even though I failed him so many times at so many places in my research assistant functionality. Instead, he helped me to overcome many obstacles in search for a researchable topic, in shaping up the study schedule, the study objectives, and the study methods. He helped me to identify and get contact with key helpers for this research, and took me to field travels to get me to know about agriculture in Virginia. He read and marked every sentence in my dissertation for many times (note: any further errors are of course mine) and I think he has known better about the whole dissertation than I do! I will here say THANK YOU again to Dr. Bosch. In the world of professors, he deserve my highest respect.

Many thanks also are to other committee members: Dr. Marcus Alley, Dr. Anya McGuirk, Dr. Saied Mostaghimi, Dr. James Pease, Dr. Patrick Phipps, and Dr. Daniel Taylor. Together, they provided me with an excellent environment to seek academic understanding and empirical soundness of the research. Special thanks are to Dr. Pease and Dr. Taylor who also had served in my MS degree committee, and to Dr. Phipps who had helped me a lot in my MS program before serving in my PhD degree committee.

Many thanks are to Phil McClallen, the CEO of MapTech Inc., Blacksburg, Virginia, who spent so many hours helping me to understand the collection and processing of the yield monitor data, and provided valuable advices as how to evaluate the site-specific technology in the practical situation. His experience in site-specific management is very important for this study.

Many thanks are to Courtney Price, manager of the Brandon Plantation, Prince George County, Virginia, one of the earliest practitioners of modern site-specific management in Virginia. He helped me to understand many successes and challenges facing the Virginian farmers in adopting site-specific management technology.

Thanks are to Professor James Baker of Crop and Soil Environmental Science Department of Virginia Tech who provided soil data and advice in simulating the yield responses to weather and soil situation in Virginia.

Lastly and most importantly, my whole-hearted thanks are to my dearest wife Mary, my dearest son Rick, and my dearest daughter Nancy. – Well, anyway, this dissertation is theirs. And I did it!

Table of Contents

Chapter 3. Empirical Model..56

Chapter 4. Results and Discussion..113

Chapter One. Introduction

Public concerns over environmental problems from agriculture are rising. Nitrogen (N) in crop production is not only one of the primary factors to manage in terms of its effect on yields for non-legume crops, but also one of the most important pollutants to the environment. Over 90% of the corn, cotton, potatoes, and rice acres and over 60% of the wheat acres in the United States receive commercial N fertilizers (Kellogg et al.). N loss from fertilizers contributes a large share of soil nitrate-nitrogen (NO3-) contamination of groundwater, especially under irrigated cropping systems (Madison and Brunett). Up to 30 to 50% of applied N may be not taken up by the crop and much of it is lost to the environment (Keeney, Goolsby et al.).[1] Wasted N reduces farmers' income, while N pollution to water bodies accelerates eutrophication of lakes and estuaries (USEPA), threatens human health (Cantor), and degrades ecosystems (National Research Council (NRC), 1978). Large programs to improve environmental quality like the Chesapeake Bay Program (Magnien et al.) are targeting N and phosphorus pollution.

Because N fertilizer is important to crop production, any attempt by the public to reduce N pollution from agriculture should carefully evaluate the effects of these programs on farmers and their reactions to these effects (McSweeny and Shortle). Two logical

---

[1] The loss of N here happens even when N is applied according to extension recommendations. In other words, when crop needs of N are met by commercial N fertilizers and/or carryover N from previous legume crops, then an "excess" of N has to be applied in order to achieve the desired yield goal. Here excess N application is defined as the difference between the amount of N applied and the N taken up by crop and removed from field (Kellogg et al.). This net N applied forms part of the residual N which is lost to leaching particularly after crop harvest. Some researchers (Groffman; Addiscott et al.) point out that the major nutrient loss from agriculture happens when the crop is harvested and the plants lose control over nutrient cycling. A large share of this N pool that is no longer controlled by crops may leave the root zone as runoff or leaching (National Research Council, 1993). To compensate for the total N loss, fertilizers are needed. Groffman says that "the need to add nutrients to agricultural production systems is a fundamental constraint on their environmental performance" (page 56).

approaches to reducing N pollution from a farm are to reduce application rates and to alter application methods to increase uptake efficiency (Legg).

1.1. Reducing application rates versus altering application practices

Other things being equal, higher rates of N application and a longer interval between the time of application and the expected time of heavy uptake by crops mean more opportunity for N loss to the environment and thus lower N use efficiency and higher input costs to producers. Efficiency of N use (and efficiency of N application) is defined in terms of high utilization by the crop and low potential for surface and groundwater pollution as expressed by Simpson et al. To compensate for the loss of applied N to leaching and other pathways, N application rates have to be higher than crop use of applied N. When only pre-plant N is applied, the loss of N to leaching may cause an N deficiency when crop uptake of N is heavy in the middle of the growing season (corn) or late in the growing season (wheat). Studies show that farmers may apply much more N than is recommended (Norris and Shabman; Bosch et al., 1992). Improved information management such as soil testing can reduce N application (Bosch et al., 1994).

A survey in Virginia by Bosch et al. (1992) shows that in Rockingham County, 22% of crop land received N at 50 lb/ac or greater above the recommended rate, while about 25% of cropland in the Northern Neck region of Virginia received 10 lb/ac or more above the recommended rate. Due to low correlation between net N applied (where net N is defined as the difference between N applied and N removed by the harvested crop) and farmers' views of leaching potential from cropping practices on a site, Bosch et al. (1992) suggest that this overapplication may happen because farmers may not know the recommended rates for these sites, or they may not appreciate the extent of water quality damage possible from excess nutrient applications. Intended overapplication of N may also occur with risk-taking farmers who would like to apply more than the crop needs in an average year in order to take advantage of a possibly good year (Bosch et al. 1992).[2]

Traditionally, the recommended N application rates, based on uniform rate and field level yield projection, do not change with the method and timing of applications (Blackmer and White). If N is applied only before planting as is common for corn, then it is probably necessary for many farmers to apply more N than recommended if large N leaching is a problem.[3]

Sidedress (split-application) of N is an attempt to increase the efficiency of N application (Simpson et al.). Studies show that sidedress (split-application) N management can improve profitability as well as environmental quality, even though risk of being unable to sidedress during late spring period is a concern (Bosch et al., 1994; Feinerman et al.). When N is sidedressed, much less pre-plant N is needed to compensate for the low efficiency of early application (Feinerman et al). For example, Alley et al. suggest that for

---

[2] Mulla and Schepers say (p.5): "invariably, producers must place values on the relative importance of yield improvement, product quality, fertilizer savings, time requirements, equipment costs, environmental protection, and public perception (of environmental stewardship). All producers weigh these considerations differently, which is why no one has ever developed a cook-book procedure to optimize them. In fact, the relative importance of these considerations changes during the growing season because of climate, plant growth patterns, and the prospect of generating a profit." For example, a 10 lb/ac N fertilizer saving reduces cost by about $3 and reduces potential pollution. But if an N deficiency does happen and yields are reduced by a few bushels, then farmers' loss of revenues exceeds the gain from fertilizer saving. Thus actual decision making is farm(er)-specific (Alley, personal communication).

[3] Current N recommendations are based on crop and soil budgeting ("N-balance formulas") which uses yield potential as the primary determinant of N rate, adjusted by credits for N from previous legume crops or manure (Mulla and Schepers; Hergert et al.; Blackmer and White; Meisinger et al.; Simpson et al.). Though N mineralization from soil organic matter and residual soil NO_3-N levels are estimated, the recommendations do not account for N losses by leaching, denitrification, erosion, nitrification, and volatilization (Mulla and Schepers). In fact, Everett and Pierce suggest that in humid regions, residual NO_3-N should not be used as a factor in soil N fertilizer recommendation in the spring since residual NO_3-N is leached rather completely over-winter. Furthermore, the amount of fertilizer N lost between pre-planting and the time of rapid uptake of N by plants is not reflected in current recommendations (Blackmer and White). As such, higher-than-recommended pre-plant application rates are an insurance against both N deficiency and the risk that late application will become infeasible due to weather (Feinerman et al.).

wheat only a small amount of N is needed to establish the crop, while most of the N is needed five months later.

Even though sidedressing can improve timing of N application, the current approach of treating the field as a homogeneous whole and applying a uniform rate of N may prevent the realization of application according to plant needs. To improve N management with pre-plant and sidedress N application, the nutrient status of the crop and the soil need to be estimated.[4] Normally farmers do composite soil sample testing to represent the whole field. But farmland is generally rather heterogeneous in terms of soil properties, landscapes, fertility, yield potential, pollution potential, pest distribution, and crop quality. These heterogeneities make crop production distinctive from other industrial sectors (Wolf and Wood). Conventional farming applies inputs like fertilizers uniformly to each field ignoring these heterogeneous factors. As a result, N is over-applied in some places but under-applied in other places (Carr et al.; Mostaghimi et al., 1997). Over-application of N reduces profits because marginal cost exceeds marginal returns. Excessive input may even reduce yield directly (Blackmer and White). Under-application, on the other hand, reduces profit because marginal cost is less than marginal returns (Wolf and Wood, and Nowak, 1998). Thus, efficient N management and control of N pollution may call for site-specific management, also known as precision agriculture.

---

[4] Recent research with the pre-sidedress N test (PSNT) improves N recommendation in that more factors such as soil organic matter, residual NO_3-N, and N credits from manures and legume crops are considered. Soil N test has become a standard management tool for N application (Hergert et al.). Generally, pre-plant soil N testing emphasizes organic N content which indicates future N availability for the coming crop, while pre-sidedress N testing emphasizes N mineralization intensity which indicates the current N availability for the growing crop.

1.2. Site-specific management (SSM)

Definition of SSM

Site-specific management (SSM) or precision agriculture, is defined by NRC (1997, p.2) as a farm "management strategy that uses information technologies to bring data from multiple sources to bear on decisions associated with crop production." SSM identifies and treats the spatial and temporal variability in agricultural production processes in a timely and efficient manner. Specific to crop production, SSM monitors and responds to large in-field variation on as fine a scale as allowed by available and economical means (Whelan and McBratney). The adoption of SSM, which is "simply a more dis-aggregated version of the kinds of best management practices (BMP's) already recommended at the field level" (Heimlich, p.20), can increase efficiency of conventional inputs like fertilizers.[5] Increase in efficiency comes from improved information management and the information generated in SSM becomes an output with value (Nowak, 1997). Although profitability and environmental performance of current precision technology are not clear-cut (Lowenberg-DeBoer and Swinton, Lowenberg-DeBoer and Boehlje), some SSM practices may become standard in the United States (Lowenberg-DeBoer). Technical components of SSM, profitability, environmental impact, and farmers' adoption behavior are discussed below, while a short discussion of economic risk of adopting SSM is deferred to Appendix A.

---

[5] Several commonly cited driving forces to the increasing popularity of precision agriculture are the demand from society for accountability in food safety and environmental health, the need to reduce production risk in an era of fewer government subsidies, the regulatory requirements for reporting and monitoring fertilizer and pesticide use, and the desire to increase profit and capture the opportunities in the industrialization of agriculture (Sonka and Coaldrake, Wolf and Wood).

Components of SSM

Whelan and McBratney summarized five conceptual components in a SSM system. They are i) spatial referencing system, ii) crop-soil-climate monitoring for pest, disease and crop growth indications, iii) attribute mapping of field variations, iv) decision support systems to provide information about economically and agronomically suitable treatment strategies, and v) differential action which allows plant populations and fertilizer and pesticide application to vary in real-time across a field.

From a technical perspective, seven components are involved in SSM: (i) global positioning system (GPS), (ii) geographic information system (GIS) for generating soil attribute maps, micro-meteorological condition maps, and crop yield maps, (iii) remote sensing using aerial photographs to observe important crop production parameters such as soil variability and plant density, (iv) grid sampling techniques, (v) yield monitors to allow farmers to identify factors that contribute to yield variability, (vi) variable rate applicators including controllers, liquid sprayers, granular applicators, air sprayers and spreaders, and drills and planters, and (vii) crop simulation models (Lu et al.). Reliable yield monitors and mapping instruments for grain combines are now available, while yield monitors for cotton, peanut, and potatoes are yet to become commercially available (Smith).[6]

Profitability of SSM

Theoretically SSM matches plant need with input use and thus has potential to increase profit. Empirical studies on corn and small grains and soybeans in Virginia show

---

[6] As of 1999, a complete package of equipment for SSM information analysis and decision-making includes a yield monitor, a GPS unit, data logging capabilities and software for data storage, manipulation, and mapping, costing about $20,000 at purchase (Shatar). A new full-range, variable-rate dry fertilizer spreader, which is a blend on-the-go N-P-K-micronutrient variable product-variable rate applicator, will cost upwards of $250,000, while variable rate single product applicators are much less expensive (Alley; Lu et al.). However, such expensive machinery and electronic equipment are not essential for each farm to own. Small farms can use simple, manually operated systems or purchase analysis, advice, and/or application services from service providers and thus share in the economies of size in spatial data analysis and capital investment (Lu et al.; Tweeten).

that SSM has potential to increase returns (Gupta et al; Mostaghimi et al, 1997). In Virginia, SSM is mainly viewed as a fertilizer management practice rather than a conservation practice and profitability is the major consideration in deciding to adopt (Mostaghimi; McClellan). Furthermore, experience shows that unless farmers see dramatic economic returns from SSM, they will not adopt it (Brann). Farmers have limited time to learn and operate SSM systems (Phipps).

Upon adopting SSM, farmers incur several kinds of costs. Initial investment costs include a yield monitor, GPS, computer and software, and possible modification of crop machinery. Other out-of-pocket costs include training costs to learn how to use the system, increased costs for soil and crop samples, and increased cost for help from consultants to collect and interpret data (Shatar). Non-cash costs also include increased management time since the operation of SSM increases the complexity of the daily workload. A large share of these costs may decrease as farmers become more skilled in managing the SSM system and as equipment and machine cost are spread over several years. Nevertheless, potential profitability of site-specific input management depends on the spatial variability of profitability associated with input use (Malzer et al.).[7]

Environmental impacts of SSM

Some studies in Virginia show that N loss as well as phosphorus (P) and potassium (K) losses were reduced by adopting SSM on farms with corn as the principal crop, even though mean application requirements across the crop rotation did not necessarily change (Mostaghimi et al. 1997, 1998; Gupta, et al.). These results suggest two points. First, the

---

[7] Economic value of SSM will also be enhanced by new developments of biotechnology. In such applications SSM can be used for tracking and managing both the quantity and quality of identity-preserved and specialty crops (for food safety needs for example) (Nowak, 1998). Adoption of SSM itself has some value because adoption may allow farmers to take advantage of future innovations in SSM (Sonka and Coaldrake). As long as comparable income and risk level can be maintained, farmers should be willing to adopt SSM for its own sake. Further values can be derived from possible environment benefits. Society may decide to subsidize SSM adoption to achieve such benefits (Brann). Many farmers can also gain non-monetary social or psychological value when SSM reduces pollution (Lowenberg-DeBoer and Swinton).

adoption of SSM itself does not necessarily reduce the input of fertilizers. SSM, by helping to reduce excess applications of inputs through targeting application to plant requirements, is consistent with the demand for reducing environmental damage from agriculture (Wallace; Wolf and Wood; and Harris). But SSM is not necessarily low-input sustainable agriculture. That is, if the need in preserving soil integrity in terms of richer organic matter, less soil erosion, higher yields and more sustainable agriculture requires higher input levels, then SSM would imply higher input use (Wallace).

Second, for a given level of input, SSM can enhance yields and reduce pollution. N pollution reduction with SSM happens directly when net N application is reduced. Wallace points out that when used in combination with conservation practices such as reduced tillage and cover crops SSM will enhance conservation performance by reducing excess soil N after harvest. It is expected that SSM can enhance the performance of pre-plant-only as well as split N application and reduce unnecessary N pollution. Thus, though some researchers (e.g. Larson et al.) are expecting SSM to greatly reduce N pollution, the magnitude of the reduction in N pollution remains site-specific.[8]

SSM Implementation

SSM implementation consists of generation of field information, input recommendations and variable rate applications. Currently, growers in Virginia are not likely to invest much in machinery for variable rate application. Rather, they prefer to hire companies like Southern States to do the job and thus avoid large fixed machinery and equipment investments (Alley). Variable application of composite combinations of N, P, and K are expensive. Currently, variable rate applications have the most potential for

---

[8] Lowenberg-DeBoer and Boehlje note that environmental considerations need to be incorporated explicitly into the monitoring and decision making process, otherwise SSM can even increase pollution. So for SSM to be environmentally friendly, the need of the crop should be balanced by the need to reduce damage to the environment (Nowak, 1998; Leiva et al.). However, SSM will facilitate record-keeping for input applications, which will not only be beneficial for producer decision making but may also be required for environmental compliance (Smith).

single products such as nitrogen and lime (Alley).

Current SSM development emphasizes producing either yield maps to improve management as in Europe or fertility maps to measure variability in soil fertility and adjust fertilizer use accordingly as in North America (Earl et al.). This difference on emphases between European and North America may reflect the fact that SSM is constrained by balancing information acquisition and related cost and risk (Whelan et al.). An integration of European and North American emphases is needed since input decisions should be based on observed yield data and supporting soil and crop data (Earl et al.).[9]

Accurate yield maps show crop yield potential so that producers can increase inputs such as nitrogen to portions of the field where large gains in yield are achievable, while reducing inputs to areas where yield potential is low. Soil fertility maps, on the other hand, help producers to adjust inputs based on crop nutrient reserves in the soil. Thus, the yield map emphasis overlooks nutrient reserves in soil especially P and K, while the soil fertility maps emphasis overlooks other factors contributing to varying yield potential.[10]

Yield maps are very likely the first thing farmers will utilize in managing input use with available SSM technologies (Alley; Lu et al.; Hergert el al.). However, development of accurate yield maps that show spatial patterns differing in response to N application is difficult. Accuracy refers to the accurate spatial patterns revealed by observed yield data. For example, with increased N application, observed yield variability will decrease (Davis

---

[9] Ferguson et al. argue that due to constraints on economic feasibility, accurate N application rate maps should be produced with information from a variety of sources of spatial data which can be obtained at greater densities and at less cost than grid soil samples. These information sources include yield maps and remotely sensed images, as well as data from selective soil sampling.

[10] For example, moisture availability may affect yields more than variability of soil fertility so variability of water-holding capacity of the soil or the drainage condition of the field may be important (Redulla et al.). Crops very close to woods may have very low yield potential due to insufficient sunshine and excessive pest problems (McClellan; Everett and Pierce). In the latter case, yield maps are more indicative in deciding variable rate input than data about soil texture properties, soil fertility, and water-holding capacity.

et al.).[11] Thus yield maps obtained from well-fertilized fields, as compared with poorly fertilized fields, will be more accurate in indicating areas within the field having different N requirements since effects of other yield-limiting factors are revealed. Yield maps obtained from several years are subject to dynamic interactions among weather, water, pest, and nutrient effects (Davis et al., Redulla et al.). Differences in elevation tend to alter yield levels during a dry year and during a wet year, rendering average yield over the years less variable than yields within the field (Stein et al.). Davis et al. point out that for areas where spatial patterns of N need are not clearly revealed by yield maps, other information sources should be used.[12] Soil maps and soil fertility maps thus come into play (Verhagen and Bouma; Thrikawala et al.).

When the different functionality (in terms of yield and pollution potentials) of soil types within each field is taken into consideration (Simpson et al.), information provided by observed yield data would be enriched in generating input recommendations. In fact, Verhagen and Bouma urge adding data to existing soil survey data that are relevant for modern applications. For example, incorporating data in existing soil survey databases for defined soil series, since traditional soil surveys have been finished in many countries. The point made by Verhagen and Bouma is supported by Atherton et al. who cite (p. 456) that some studies (Steinwand et al.) show that NRCS soil maps are suitable for field-level management decision making, while other studies (Salchow et al.) show that these maps are not adequate for sub-field scale management. Wollenhaupt et al. point out that though valuable in defining important limitations to crop production such as variations in soil

---

[11] This does not mean that the field-level variability of net income will necessarily decrease.

[12] Pocknee et al. list several sources of information that reveal patterns in certain soil parameters relating to the N application decision. They are NRCS soil maps, management history maps, past yield maps, digital elevation maps, ground penetrating radar maps, aerial crop images, past grid sampling maps, aerial soil color images, electromagnetic induction maps, and expert information maps. Effective use of background information will ensure that the spatial variability of several important parameters of the soil, such as yield potential, N mineralization potential, and residual NO3-N before planting are estimable, non-random, of significant magnitude, and manageable (Everett and Pierce).

texture or organic matter content, soil maps alone could be very poor in accounting for spatial patterns in soil N availability. In order to better characterize variations in soil N availability and need, a field can be divided into smaller zones by functionality in terms of yield potential using soil maps and other site characteristics.

Use of functional zones

Soil maps as reported in county surveys are basically pedological, not functional (Verhagen and Bouma). That is, different soil types are determined by differences in certain soil parameters. But in fact, soils with significantly different parameters may have almost the same yield potential, response to N application, and/or pollution potential, that is, they may be functionally the same type (Verhagen and Bouma; Simpson et al.). As such, they can be grouped into one more general soil types.

Static grouping of soils functionally in a field cannot explain many observed differences in actual yields and responses to N applied. For example, Hollands states that elevation alone can be the largest factor to explain the variation of N residual level in the Red River Valley. In the study Hollands presents, when ignoring the elevation factor, even intensive grid sampling can be misleading in generating the spatial pattern of N fertility for the fields. Drainage condition can also explain some variation in yields (Stein et al.; Redulla et al.). These kinds of phenomena can be indicated largely by yield maps, as an empirical study by Stein et al. shows. Wollenhaupt et al. point out that yield maps are very useful in indicating the portions of the field which need to be sampled in order to find the causes of observed yield variability. The pattern of yield variability within a field may vary greatly over time (Stein et al.; Redulla et al.). Thus, a dynamic grouping of functional sub-field zones can utilize yield maps as well as other available information for given (or expected) weather conditions in determining the N application rate for each group. This grouping approach is dynamic because two sub-field zones which are very alike (in yield potential) in one year may be very different in another as indicated by yield maps and explained by other factors like soil type, elevation, drainage situation, as well as closeness to woods.

1.3. Using SSM to improve N management

SSM as a resource management strategy

Proposed conceptual frameworks for evaluating the adoption decision (e.g. Lowenberg-DeBoer and Swinton) often give the impression that adopting SSM is an all-or-none decision. But in fact, SSM is essentially a resource management strategy. One of the most important and costly managed resources is the information about spatial variability of farm land that can be used to increase economic returns and reduce economic and environmental risks.[13] As a management strategy, SSM is similar to other BMPs including integrated pest management (Wolf and Wood). Depending on their farm sizes and their experience, farmers can buy different levels and forms of SSM equipment and services to take advantage of advancing information technologies (Lu et al.).

When viewed as an information management question, the adoption of SSM becomes a problem of deciding how much information is needed with current available SSM technology and related costs and risks. When farmers are risk neutral, they may maximize expected net income. When farmers are risk averse, as many of them are (Wilson and Eidman), they would be willing to give up a certain amount of expected

---

[13] Temporal information is concerned with changes in input prices (including costs of capital investments in data collecting and analysis and decision support service), output prices, and weather conditions within season or over the years. SSM is not established to provide these kinds of temporal information, but it is well suited to utilize these kinds of information for more efficient farm planning and managing. For example, the soil and nutrient dynamics over the years with regard to input use and rotation practices are important considerations in SSM.

income in order to reduce risk. Thus, farmers will choose appropriate investments in information management to derive economic returns and risk control. Differences in information levels can also affect pollution potential. For example, different information about the spatial variability of soil fertility and crop yields on a field will result in different N application rates and application methods.

N management using site-specific information in corn-wheat/soybean production

Nitrogen is the most important fertilizer to be managed in SSM for crops like corn and wheat due to the quantity of usage and potential for pollution (Huang; Blackmer and White). Soybeans are often double-cropped with wheat in rotation with corn in the Coastal Plain of Virginia. Soybean production usually does not need supplemental N, and the N fixed by soybeans can provide additional N to a subsequent corn crop (Simpson et al.).

Traditional N recommendations use only part of the information available or potentially available. N fertilization recommendations in many states are based on average results over years on experimental plots of best or above average fertility, ignoring especially poor soils (Mulla and Schepers). Farmers may have special information in regard to N management which is ignored by researchers (Bosch et al. 1994). The crop yield response to N application, the N content of manure, and the N credited from previous legume crops are all uncertain (Blackmer and White; Bosch et al. 1994). Ignoring spatial variability within a field potentially can make N applications non-optimal. Non-optimality occurs when the farmer ignores the opportunity to obtain a higher level of information about his field in order to improve N management, even if information costs are less than benefits (McSweeny and Shortle).

SSM improves two aspects of N management in corn-wheat/soybean production. The first aspect is the development of better uniform application rates by remote sensing of soil, soil testing, remote sensing of crop, tissue testing, and yield monitoring (Blackmer and White). For example, information on soil types provides soil yield potentials, yield maps provide information on actual yield variation of the fields, and growing season information on N status of crop such as taking pre-sidedress N tests reveals current crop N needs. These types of information can be combined in assessing crop needs and determining N application rates. The value of this aspect is positive when different uniform N application rates with improved information result in higher net income as compared with the conventional uniform N application practices based on standard N recommendations for the dominant soil in the field.

The second aspect of improvement with SSM is variable rate applications that tailor N applications based on the information generated for each site-specific situation. After improved information is generated about N requirements in the field, a variable rate application may result in higher net income as compared to a uniform rate of N application. Whether variable or uniform applications are preferred, depends on economic and environmental considerations (Bock and Hergert; Hergert et al.).

When a field is rather homogeneous, the uniform application based on dominant soil group is appropriate. When a field is rather heterogeneous and the farmer spends more to get more detailed site-specific information, he still can choose between a uniform and variable rate application. If he chooses a uniform application, he can gain over the conventional (no-information) approach by coming up with a better-informed N application decision that offers an optimal (net returns maximizing) result for the whole field under a uniform application method. Variable rate application may not be carried out due to other considerations such as the additional cost involved or the limited benefits. For example, some fields may be statistically heterogeneous in terms of soil sampling results yet corresponding fertilizer recommendations may differ very little (Anderson-Cook et al.). Only with large variability of N need within a field does variable rate application have potential to increase the farmer's income.

Information collection for yield maps and soil sampling has a cost. Generated information has value when it brings about improved N decisions which are translated into savings on N or increased yields. While there has been rapid adoption of precision agricultural technology, it is uncertain how profitable the technology is. Future adoptions are likely to be slower and more uneven (Bullock et al.). There is a need for additional information to help farmers use spatial data to make fertilizer management decisions (Bullock et al.). It is unknown whether and how the additional savings on N use or additional income from increased yields will offset the increased cost in information generation based on yield and soil maps. Even if information is worth collecting, it is also unknown if variable rate application will increase farm net returns. Also unknown is the relative potential to reduce net N using yield and soil map information to guide N applications.

Work is underway to estimate the value of soil test, topographical, and remote sensing information for N application (Hurley et al.) and to develop site specific production functions to guide N applications (Swinton et al.). However, these studies typically evaluate yield patterns ex post and are applied to cases where a large amount of yield, soil, and other information are available. Less is known about the value of yield monitor information collected under typical farm operating conditions for guiding future N management decisions.

Ex ante valuation of SSM[14]

McBratney el al. point out that "it is clear that annual temporal variation (of crop yields) is much larger than the spatial variation within single fields. This leads to the conclusion that if precision agriculture is to have a sound scientific basis and ultimately a practical outcome, then the null hypothesis that still remains to be seriously researched is: given the large temporal variation in yield relative to the scale of a single field, the optimal risk aversion strategy is uniform management" (p.141). They further argue that if spatial variability is small as compared to temporal variability, then it does not pay to treat the field differentially according to spatial variability. An *ex post* study of observed yield variability by Peng and Bosch shows that for a single year's data, variable rate application on a seemingly homogeneous field (in terms of the farmer's opinion and soil distribution) can improve yield and net income, and reduce N pollution. The challenge is to predict the spatial pattern of N application for the coming year (*ex ante*) that maximizes expected return.

When long-term temporal yield variability dominates, it is important to have a long-term series of yield data, since it is possible to predict crop yields over time (Eghball et al.) using historical data or simulated data and statistical procedures (McBratney et al.). After controlling for temporal variation, then it is possible to concentrate on predicting spatial variation using further site specific information including soil maps, sampling results, and geostatistical procedures (Stein et al., 1997b).

In Virginia, some enterprising farmers have used a yield monitor for several years and have a good appreciation of the within-field variability of yield on their farms. Yet most of these farmers still apply N uniformly. Several obstacles to variable rate application

---

[14] Ex ante means before actual outcomes (yields) are observed.

are present. One is that farmers would like to make sure that they apply "enough" fertilizer on the field in order to capitalize on good growing years so they may be less concerned with treating within field variability. Another is the cost of variable rate application. The fertilizer companies would raise the application fee from \$6 per acre for uniform application to \$12 per acre for precision application in eastern Virginia. Even though this increase in cost is not high for fields with enough variability in N needs, it may not be economically viable for the fields without high variability of N needs. Currently, fertilizer companies will not vary application methods (uniform and variable rate) within a farm (i.e. they insist on either uniform or variable rate application on the whole farm). Another obstacle to variable rate application is that yield maps may change from year to year and from crop to crop, making the prediction of yield patterns for next year difficult.

1.4. Objectives

Objectives of the study

The overall objective of the study is to analyze the economic and environmental impacts of alternative management strategies for N in corn and wheat/soybean production in eastern Virginia based on site-specific information. Specifically,

i) To evaluate *ex post* the effects of three levels of site-specific information regarding crop N requirements combined with variable and uniform N application on expected net returns and net N (applied N that is not removed by the crop) for corn-wheat/soybean fields in eastern Virginia. *Ex post* means after actual outcomes (yields) have been observed.

ii) To evaluate ex ante the effects of the three levels of site-specific information combined with variable and uniform application on expected net returns and net

N for corn-wheat/soybean fields in eastern Virginia. *Ex ante* means before the actual outcomes (yields) are observed.

SSM information strategies

Based on the previous discussion, five strategies were developed (Table 1-1). Each strategy is briefly described below while the implementation of these strategies is described further in Chapter 2 and Chapter 3.

Table 1-1. N management strategies to be evaluated for this study[a]

Name	N application rate	Information source
Conventional	Uniform	Dominant soil + extension recommendation[b]
Soil zone-uniform	Uniform	Soil zone + extension recommendation
Functional zone-uniform	Uniform	Functional zone + extension recommendation
Soil zone-variable	Variable	Soil zone + extension recommendation
Functional zone-variable	Variable	Functional zone + extension recommendation

a. Each strategy is a combination of an information level and an application method.
b. In the conventional strategy the farmer regards the whole field as functionally homogeneous with one soil type. In this study, this soil type is the soil type which occupies the most area.

The conventional strategy treats the whole field as if it were homogeneous in terms of soil type, fertility level, and yield potential, ignoring the spatial variability of N needs within each field. The farmer is assumed to take the dominant soil type of the field as the soil type for the whole field and decide the application rate according to general extension recommendation guidelines concerning that dominant soil type.

The soil zone-uniform strategy divides a field into several soil zones by soil type, estimates potential N use for each soil zone and then applies a uniform rate of N over the whole field. Each soil zone is assigned a desirable N rate based on extension recommendation guidelines such as VALUES (Simpson et al.). Then the distributions of soil type are used to determine an optimal uniform N application rate.

The functional zone-uniform strategy estimates potential N use for each functional zone in the field and applies a uniform rate of N over the whole field, taking variability of N needs of this field into consideration. Each functional zone is assigned a desirable N application rate according to extension recommendation guidelines such as VALUES (Simpson et al.), adjusted by the zone's observed yields. An economically optimal uniform N application rate is then determined.

The soil zone-variable strategy differs from the soil zone-uniform strategy in that the soil zone-variable strategy applies a variable rate application by soil zones. Similarly, the functional zone-variable strategy differs from functional zone-uniform strategy in that functional zone-variable strategy applies a variable rate application by functional zones. The conceptual and the empirical bases for these five strategies can be found in Chapter 2 and Chapter 3.

1.5. Assumptions and hypotheses

Assumptions

1) Soil maps exist or can be provided free. Yield maps and variable rate applications can be done by hiring service-providing firms. In other words, there are no long-term SSM investment costs on the farmer's part.
2) Adequate P and K and lime are present for the fields under study.
3) Yield potential for a field does not change when N application strategy changes.
4) It is always feasible to do N side-dress (topdress) application at the proper time.
5) Farmers attempt to maximize expected net income.
6) Farmers' behavior in adopting SSM is not affected by whether the land is rented or owned.
7) Wheat and corn yields respond to nitrogen according to a linear response-plateau function (discussed in Chapter Two).

Hypotheses

1) The soil zone-uniform strategy and the functional zone-uniform strategy have higher expected net returns than the conventional strategy on all fields.
2) The functional zone-uniform strategy has a higher expected net return than the soil zone-uniform strategy on all fields.
3) The soil zone-variable strategy has a higher expected net return than the soil zone-uniform strategy on all fields.
4) The functional zone-variable strategy has higher expected net return than the functional zone-uniform strategy on all fields.
5) Uniform N application with increased information does not generally reduce net N.
6) Variable N application will have lower net N than uniform N application.
7) The functional zone variable strategy will have the smallest level of net N of all N application strategies on all fields.

1.6. Study area and farm type

The study area is a row-crop farm in eastern Virginia. The study crops are corn and wheat with wheat being double-cropped with soybeans.

1.7. Organization of the dissertation

The remainder of this dissertation is organized as follows: Chapter 2, The Conceptual Framework, will discuss the mathematical and graphical models to demonstrate basic ideas and methods used in this study. Chapter 3, The Empirical Model, will lay out the empirical farm model and the statistical model for predicting N reductions and net returns from N application strategies. Chapter 4, Results and Discussion, will report and discuss results from the empirical model, and Chapter 5, Summary and Conclusions, will summarize the whole study including motivation, objectives, models, results, conclusions, implications, and suggestions for further study.

Chapter Two. Conceptual Framework

2.0. Overview

This chapter presents the conceptual framework for the study. Section 2.1 describes mathematically each of the five strategies the farmer can adopt for a field and calculation of information values and values of variable application. Section 2.2 discusses the method to estimate N pollution potential for the strategies. Section 2.3 describes the procedures to generate yield maps and define functional zones. Section 2.4 describes how temporal and spatial yield variations are modeled. Section 2.5 describes how the temporal and spatial yield information can be used to determine the optimal N application rates. Section 2.6 is a brief summary of the chapter.

2.1. Mathematical layout of the management strategies

Smallest management unit (SMU) and production function

Consider only one field on the farm. Suppose the acreage of the field is A and the field is divided into I zones by soil type and J zones by functionality (J could be larger or smaller than I depending on the criterion to define distinctive functional zones). Let the acreage of the i^{th} soil zone be A_i and the acreage of the j^{th} functional zone be B_j. Furthermore, suppose the smallest management unit (SMU) within a field is m acres and the field is divided into K SMUs with each unit entirely belonging to only one soil zone as well as only one functional zone (K is generally larger than I and J). The size of SMU is so chosen that the farmer will regard each unit as homogeneous, ignoring any variability within. Thus $A = \sum_{i=1}^{I} A_i = \sum_{j=1}^{J} B_j = K \cdot m$, ignoring the fact that because of the irregularities of the field boundary actual areas of some of the SMUs can be less than m acres.

Suppose the farmer is planting a single rotation with several crops involved (e.g. corn- wheat/soybean rotation) on the field and only one crop is planted on the whole field at one time, i.e. the whole field is in one phase of the rotation in any given year. The farmer can adopt any N application strategy on a specific field for any crop in any year. Thus, in the discussion that follows, only one crop is considered to simplify notations.

The crop yield response to N application can be expressed by a production function. The variables to be included are those that are measurable and believed to spatially affect yield, such as field characteristics, nutrient status, and inputs (e.g. N) (Snyder et al. (1996)). The crop production function can be expressed as:

$$y_k = f_k(N_k; W_k, M_k, R) \tag{2.1}$$

where y_k is crop yield on the k^{th} SMU. $N_k = x_k + s_k$ where x_k is the total N application rate,[15] and s_k is the estimated soil N carried over from previous crops. R represents the weather effect (rainfall and temperature) during the growing season; W_k denotes further within-field locational descriptors not expressed by k (or in other words, the vector of spatial indicators for the SMU k); and M_k is the vector of other management inputs such as pest control, plant population, and plant variety. The semicolon ";" in the production function denotes that W_k, M_k, R are exogenous. It is assumed that W_k and M_k are not random but weather condition R is random.

---

[15] In this study, N is split-applied for all strategies, since split N application is a rather established practice in the study area (Phipps).

The production function for the j^{th} functional zone is $y_j = g_j(N_j; M_j, R)$. There are no additional location descriptors W_j for this function since the farmer treats all SMUs within the same functional zone as identical. Similarly, the production function for the i^{th} soil zone is $y_i = h_i(N_i; M_i, R)$. The difference between $g_j(\cdot)$ and $h_i(\cdot)$ is that $h_i(\cdot)$ will not change for given soil type over the years, while $g_j(\cdot)$ changes from year to year, reflecting the fact that functional zones are results of dynamic groupings of the SMUs.

It is expected that the production function for each functional zone satisfies $\frac{\partial()}{\partial t} \geq 0$, and $\frac{\partial^2()}{\partial t^2} \leq 0$, where () represents a production function ($f_k(\cdot)$, $g_j(\cdot)$, and $h_i(\cdot)$) and t represents any input (e.g. N, P, K, and pesticides). That is, adding N (and/or other inputs) to crops will not reduce the crop yields, while the increase of yield from one additional unit of input will be less than or equal to that from the previous unit of the same input. The random variable R is expressed as a “state of nature” which provides an exogenous constraint on the production function (e.g. R can be expressed as “good year”, “normal year”, and “bad year” with corresponding probabilities). There are no explicit analytical properties needed at this stage for the spatial indicators vector M for the production function.

The nonlinear relationship between N level and yield can be approximated by a linear or quadratic form, assuming the existence of a well behaved “true” functional form and the fact that the range of economic N application rate is narrow (Snyder et al. (1996)). Many researchers use a quadratic form to quantify yield response to N (e.g. Snyder et al. (1996), Malzer et al., Bundy and Andraski, Cerrato and Blackmer, and Thrikawala et al.). A proper quadratic production function with the range of N application properly controlled does satisfy the expected properties as stated above. Similarly, linear functional forms can also be chosen, based on the argument that the narrow range of N application rates may make the nonlinearity of the true production function visibly linear. Other researchers such as Babcock, Babcock and Blackmer, Mallarino and Blackmer, and Cox use a linear response and plateau function.

The form of a deterministic linear response and plateau function can be written as $y = \min\left\{\boldsymbol{a} + \boldsymbol{b}N, y_p\right\}$ where y_p is the plateau value. Agronomists found that this functional form fits most field data well (Cox). Mallarino and Blackmer note that a linear response and plateau function is effective in determining the economic concentrations of soil test P for corn (cited by Cox). Babcock comments that this functional form is appropriate when a nutrient limiting concentration is considered, meaning that virtually no yield gains result when nutrient concentration exceeds a certain level. This function is appropriate for this study especially since extension recommendations guidelines used for this study (VALUES by Simpson et al.) give optimal expected yield for each soil type, assuming the yield response to N before the optimal yield is reached is linear, while there is no expected yield gain after N application reaches the level required to achieve the optimal yield. Also, this function is straightforward to use to calculate the amount of net N (more discussion of net N can be found later in this chapter).

Conventional strategy

The profit-maximizing farmer’s objective for the field is to

$$
\begin{aligned}
\underset{\bar{x}}{Max}\ \boldsymbol{p}^c &= E_R\left\{A\left[Pf^c(\bar{x};\bar{s},M,R) - P_N\bar{x} - C\right]\right\} \\
&= A\int_R \left[Pf^c(\bar{x};\bar{s},M,R) - P_N\bar{x} - C\right]dG(R) \qquad (2.2) \\
&= A\left\{P\int_R f^c(\bar{x};\bar{s},M,R)dG(R) - P_N\bar{x} - C\right\}
\end{aligned}
$$

where P is output price for the crop; $\bar{x} + \bar{s} = \overline{N}$ where $\bar{x}$ is the uniform N application rate and $\bar{s}$ is the average N fertility level for the whole field, respectively; f^c is the production function for the crop for the whole field assumed by the farmer following the conventional strategy (see more discussion below); P_N is the price of fertilizer N; C is the per acre fixed cost related to production of the crop less cost of N fertilizer; and $G(R)$ is the distribution function of R. In this study, the $G(R)$ is discrete ("states of nature").

The farmer uses f^c, the production function for the dominant soil type in the field, to represent the whole field. He also assumes that the whole field is of the same fertility, $\bar{s}$, which can be estimated by means such as one composite soil sample test. With assumptions that one production function applies to the entire field and that one average fertility applies to the entire field, the farmer comes up with an N application decision $\bar{x}^{c*}$ by solving the first-order condition:

$$P\int_R \left(\frac{d}{d\bar{x}} f^c(\bar{x};\bar{s},M,R) \right) dG(R) - P_N = 0 \qquad (2.3)$$

He expects a net income of

$$\boldsymbol{p}^{\mathrm{C*}}_{\text{expected}} = A\left\{P\int_R f^c(\bar{x}^{c*};\bar{s},M,R)dG(R) - P_N\bar{x}^{c*} - C\right\} \qquad (2.2')$$

But actually, he gets

$$\boldsymbol{p}^{C*}_{\text{actual}} = m\left\{\int_R \sum_{k=1}^{K} Pf_k(\bar{x}^{c*} + s_k; W_k, M_k, R)dG(R) - \sum_{k=1}^{K} P_N\bar{x}^{c*} - C\right\} \qquad (2.2'')$$

Here $\boldsymbol{p}^{\mathrm{C*}}_{\text{expected}}$ and $\boldsymbol{p}^{\mathrm{C*}}_{\text{actual}}$ are probably different and the difference is site-specific.

Soil zone-uniform strategy

In this strategy, the farmer regards each soil zone i as a distinctive homogeneous sub-field and his decision is to

$$
\begin{aligned}
\underset{\bar{x}}{Max}\,\boldsymbol{p}^{su} &= \int_R \sum_{i=1}^{I} A_i\left[Ph_i(\bar{x};\bar{s}_i, M_i, R) - P_N\,\bar{x} - C - C_S\right]dG(R) \\
&= \int_R \left(\sum_{i=1}^{I} A_i Ph_i(\bar{x};\bar{s}_i, M_i, R)\right)dG(R) - A\left(P_N\,\bar{x} + C + C_S\right)
\end{aligned}
\tag{2.4}
$$

where superscript *su* denotes the information strategy in soil-zone approach, and C_S is the per acre cost to delineate the soil zones in the field. The first-order condition that gives the optimal N application rate $\bar{x}^{su*}$ is

$$
\int_R \left(\sum_{i=1}^{I} A_i P \frac{d}{d\bar{x}} h_i(\bar{x};\bar{s}_i, M, R)\right)dG(R) - AP_N = 0 \tag{2.5}
$$

Similar to (2.2') and (2.2"), the $\boldsymbol{p}^{su*}_{\text{expected}}$ is

$$
\int_R \left(\sum_{i=1}^{I} A_i Ph_i(\bar{x}^{su*};\bar{s}_i, M_i, R)\right)dG(R) - A\left(P_N\,\bar{x}^{su*} + C + C_S\right) \tag{2.4'}
$$

And $\boldsymbol{p}^{su*}_{\text{actual}}$ is

$$
m\left\{\int_R \sum_{i=1}^{I} Ph_i(\bar{x}^{su*} + s_i; W_i, M_i, R)dG(R) - \sum_{i=1}^{I} P_N\,\bar{x}^{su*} - C - C_S\right\} \tag{2.4"}
$$

Generally $\boldsymbol{p}^{su*}_{\text{expected}}$ and $\boldsymbol{p}^{su*}_{\text{actual}}$ are different, though the difference should be smaller than that for the conventional strategy because more field information is used.

Functional zone-uniform strategy

Functional zones are formed by grouping SMUs into several groups according to their similarities in yield level and N requirements based on historical observation and expected weather situation. Because of the way functional zones are formed, the SMUs that belong to the same group in one year do not necessarily belong to the same group for the next year, and the expected yield for a functional zone may also change from year to year (as later discussion makes more clear). The farmer's objective in the functional zone-uniform strategy is

$$\begin{aligned} \underset{\bar{x}}{Max}\,\boldsymbol{p}^{fu} &= \int_R \sum_{j=1}^{J} A_j \left[Pg_j(\bar{x};\bar{s}_j, M_j, R) - P_N\bar{x} - C - C_F\right] dG(R) \\ &= \int_R \left(\sum_{j=1}^{J} A_j Pg_j(\bar{x};\bar{s}_j, M_j, R) \right) dG(R) - A\left(P_N\bar{x} + C + C_F\right) \end{aligned} \tag{2.6}$$

where superscript *fu* denotes the information strategy in functional zone-uniform approach, and C_F is the total cost to delineate the functional zones in the field. The functional zone-uniform strategy problem is very similar in mathematical expression to the soil zone-uniform strategy except that $\bar{s}_j$ refers to an average soil fertility for functional zone j. So expressions corresponding to (2.4'), (2.4"), and (2.5) are omitted here. The net income expected by the farmer in this strategy, $\boldsymbol{p}^{fu*}_{\text{expected}}$, and the correct expected net income, $\boldsymbol{p}^{fu*}_{\text{actual}}$, are also different generally.

Variable rate strategies

Each soil zone is applied a different rate of N in the soil zone-variable strategy and each functional zone is applied a different rate of N in the functional zone-variable strategy. In the soil zone-variable strategy, the farmer's objective is to

$$\begin{aligned} \underset{x_1,\dots,x_I}{Max}\,\boldsymbol{p}^{sv} &= \int_R \sum_{i=1}^{I} A_i \left[Ph_i(x_i;\bar{s}_i, M_i, R) - P_N x_i - C - C_S - C_{SV}\right] dG(R) \\ &= \int_R \left(\sum_{i=1}^{I} A_i Ph_i(x_i;\bar{s}_i, M_i, R) \right) dG(R) - \sum_{i=1}^{I} A_i P_N x_i - A\left(C + C_S + C_{SV}\right) \end{aligned} \tag{2.7}$$

where C_{SV} is the additional cost for variable application. The optimal N rate for each soil zone is determined by solving the first-order condition for each soil zone. As with other strategies, the net income expected by the farmer, $\boldsymbol{p}^{sv*}_{\text{expected}}$, is different from the actual expected net income, $\boldsymbol{p}^{sv*}_{\text{actual}}$.

In the functional zone-variable strategy, the farmer's objective is to

$$\begin{aligned}\underset{x_1,\ldots,x_J}{Max}\,\boldsymbol{p}^{fv} &= \int_R \sum_{j=1}^{J} A_j\left[Pg_j(x_j;\bar{s}_j,M_j,R) - P_N x_j - C - C_F - C_{FV}\right]dG(R) \\ &= \int_R \left(\sum_{j=1}^{J} A_j Pg_j(x_j;\bar{s}_j,M_j,R)\right)dG(R) - \sum_{j=1}^{J} A_j P_N x_j - A\left(C + C_F + C_{FV}\right)\end{aligned} \tag{2.8}$$

where C_{FV} is the additional cost for variable rate application for functional zones. C_{FV} could be different from C_{SV}. The distinctive N rate for each functional zone is determined from the first order condition for the above problem. Still, $\boldsymbol{p}_{\text{expected}}^{fv*}$ is not necessarily equal to $\boldsymbol{p}_{\text{actual}}^{fv*}$ because variability still exists within each functional zone and is beyond the farmer’s ability to delineate or control.

Above discussion is for only one crop on one field on the farm. The farm may have several fields. Since it is assumed that information collecting and analysis and the variable rate application will be contracted to service companies, this layout suffices for the whole farm. The farmer can maximize total net income for the whole farm by maximizing net income on each of the individual fields which may involve choosing different strategies for different fields on his farm. Thus, this setting allows the farmer to look at individual fields instead of the whole farm to evaluate the adoption of SSM and the consequent environmental impact.

Information values and values of variable rate application

In SSM, the information value can be separated from the value of variable rate application. Information has value when the production decision with information available differs from that without information (Schnitkey et al.). Under a uniform application method, if the application rate of N changes and/or yield changes after information is obtained, then information has value. Similarly, for given information level, if the application rate of N changes and/or yield changes with variable rate application, then variable rate application has value also. Using the conventional strategy as the baseline, a summary of the information values, variable rate application values, and profitability of the strategies is presented in Table 2-1. In Table 2-1, gross values are calculated by ignoring the extra costs for information generation and additional costs for carrying out variable application, while the net values are calculated as gross values minus the extra costs for information generation and additional cost for carrying out variable application wherever applicable.

Table 2-1. Values of information and variable application for different strategies[ab]

Strategy	Change of profit from baseline[c]	Information value		Value of variable rate application	
		Gross	Net	Gross	Net
Conventional	0	0	0	0	0
Soil zone-uniform	$p^{su*}_{\text{actual}} - p^{c*}_{\text{actual}}$	$p^{su*}_{\text{actual}} - p^{c*}_{\text{actual}} + A \times C_S$	$p^{su*}_{\text{actual}} - p^{c*}_{\text{actual}}$	0	0
Functional zone-uniform	$p^{fu*}_{\text{actual}} - p^{c*}_{\text{actual}}$	$p^{fu*}_{\text{actual}} - p^{c*}_{\text{actual}} + A \times C_F$	$p^{fu*}_{\text{actual}} - p^{c*}_{\text{actual}}$	0	0
Soil zone-variable	$p^{sv*}_{\text{actual}} - p^{c*}_{\text{actual}}$	$p^{su*}_{\text{actual}} - p^{c*}_{\text{actual}} + A \times C_S$	$p^{su*}_{\text{actual}} - p^{c*}_{\text{actual}}$	$p^{sv*}_{\text{actual}} - p^{su*}_{\text{actual}} + A \times C_{SV}$	$p^{sv*}_{\text{actual}} - p^{su*}_{\text{actual}}$
Functional zone-variable	$p^{fv*}_{\text{actual}} - p^{c*}_{\text{actual}}$	$p^{fu*}_{\text{actual}} - p^{c*}_{\text{actual}} + A \times C_F$	$p^{fu*}_{\text{actual}} - p^{c*}_{\text{actual}}$	$p^{fv*}_{\text{actual}} - p^{fu*}_{\text{actual}} + A \times C_{FV}$	$p^{fv*}_{\text{actual}} - p^{fu*}_{\text{actual}}$

a. All terms used in this table have been formerly defined from equations (2.1) through (2.8). Thus, values presented in this table are the total values for a field and not per acre values.

b. Superscript "*" indicates the optimal solution, "*su*" indicates soil zone-uniform strategy, "*fu*" indicates functional zone-uniform strategy, "*sv*" indicates soil zone-variable strategy, and "*fv*" indicates functional zone-variable strategy.

c. Conventional strategy is the baseline case throughout this study. Its profit is referred to as p^{c*}_{actual}

From Table 2-1, it is seen that the information value does not necessarily increase with information level. The same is true for values of variable application. As can be seen in the table, information values do not change with the change of application method, while total profits may change with the change of either information level or application method.

Increased information levels and variable rate application bring higher costs. It is possible that even though a certain information value is not positive, yet the variable application based on this level of information may achieve positive income over the baseline strategy, the conventional strategy. For example, if $p_{\text{actual}}^{fu*} - p_{\text{actual}}^{c*}$ (net information value) is negative, but $p_{\text{actual}}^{fv*} - p_{\text{actual}}^{c*}$ (net information value plus net value of variable application) is positive, then precision application with functional-zone information is still preferred over the conventional strategy. Of course, to carry out the functional-zone variable rate application, the functional-zone information has to be generated.

2.2. Potential N pollution assessment

Net N as an indicator of pollution potential

The major pathway for N removed from soil is crop uptake of inorganic N in the NH_4-N or NO_3-N form (Mulla and Schepers). Variations in N uptake efficiency across a landscape can be very large (Fiez et al.). There is a significant risk of NO_3-N leaching for pre-plant N application in corn production since corn takes up most N during the middle one-half of the growing season (Mulla and Schepers). Simpson et al. strongly recommend split application of N whenever possible to minimize the potential for surface and groundwater pollution. Split-application of N is well recognized as an efficient and economic N management practice in eastern Virginia (Phipps).

N mineralization, immobilization, and N fixed from leguminous plants all influence the availability of inorganic N in the soil. The quantity of available mineral N for the crop from soil reserves depends on rainfall, soil type, and residues from the previous crop, soil organic matter content, soil temperature and moisture content. Major pathways for N loss from soil are denitrification and leaching of NO_3-N, volatilization, and nitrification (Mulla and Schepers). Organic N can also be lost with sediment.

Site-specific N management is not intended to reduce the overall N application rate, rather it is intended to better match applications to crop needs (Wallace). Even if site-specific management results in higher levels of N application, total N pollution can still be reduced compared with the conventional approach. In fact, Bosch et al. (1994) point out that improved information may have greater potential for reducing N loss than for reducing application rates. If more information indicates higher N need, that means much of the increased N application will be utilized by crop; but if information indicates lower N rates, that means much of the formerly over-applied N was potentially available for loss to pollution (Bosch et al., 1994). Thus, "(f)rom the perspective of water quality protection, negative excess nitrogen applications do not cancel out the damaging effects of positive excesses that may reach water supplies." (Bosch et al. 1992, p. 33). This viewpoint is consistent with the concept of "unrecovered N" (Kitchen et al.) which equals the maximum of zero and the difference between the amount of N applied and the amount of N in crop yield removed from field. This difference is unrecovered and may eventually end up in pollution to the environment, especially for the field vulnerable to leaching (Meisinger). Larson et al. review a conceptual model by Stevenson which "indicates that once the amount of N applied exceeds the amount needed for economic plant production, NO3-N in the soil increases exponentially and is vulnerable to leaching by fall and spring rains" (p.339).

Thus, the amount of net N (Kitchen et al.'s "unrecovered N") is a good indicator (index) of the N pollution potential from a field, especially in evaluating the environmental impact of site-specific management at the field level. For example, using this approach, Thrikawala et al. demonstrated that the absolute levels of net N are highest under fields with high average fertility and high variability in fertility levels. Under all fertility distributions, the site-specific management results in significantly less net N than the constant-rate application method.

Assuming a fixed N content in yield for each crop and by equating the unrecovered N to net N, a simple formula can be used to express the net N concept for this study:

$$EN_k = \max\left\{x_k - \boldsymbol{r} f_k(N_k, W_k, M_k, R)\ , 0\right\} \tag{2.9}$$

where EN_k is the net N applied for SMU k for the crop, x_k is applied N, and $\boldsymbol{r}$ is the decimal fraction of N content in yield.[16] The expression EN_k can be estimated using available data.

A graphical illustration

A graphical illustration is presented in Figure 2-1 that is adopted from Huang and Lantin to compare the net N under different strategies in this study. Only the conventional strategy, the functional zone uniform strategy, and the functional zone variable strategy are compared.

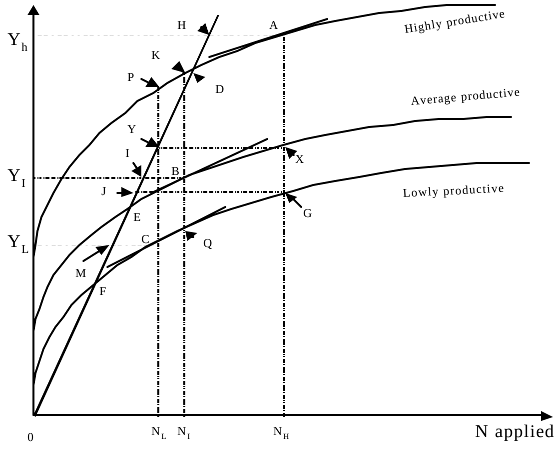


Figure 2-1. Net N under different strategies and field conditions for a functional zone (Adopted from Huang and Lantin)

---

[16] Huang and Lantin include the N credited from other sources like previous legume crop, manure, crop residues, and irrigation water, in short, s_k, as denoted in this study. For the purpose of this study, N credited from other sources is not included except for corn when the previous crop is soybean. Irrigation and animal manure application are not common in the study area.

The field has three zones, namely, the high productive zone, the average productive zone, and the low productivity zone. Each zone has its own production function as displayed as a concave curve on the above graph. To simplify the illustration, a simple profit equation is used:

$$\text{Profit} = P^{*}Y - P_N^{*}(\text{N applied}) \qquad (2.10)$$

Thus, the profit maximizing solutions for the high productivity zone, the low productivity zone, and the average productivity function, respectively, are A, C, and B in the above graph where the slope of the tangent lines to these points equals the price ratio P_N/P. The assumption that the N content in crop yield is a fixed percentage is represented by an upward-sloped straight line. The positive intercepts of the production functions represent the fact that N inherently contained in soil can produce some yield even if no N fertilizer is applied. To the left of the N content line (the straight line OH), any N application decision for any production function is not only under-applying N in the economic sense, but may also be mining soil N in the long run, which is unsustainable. The ideal of "no-net N" can only be realized by operating at points D, E, and F where N removed in yields equals N applied for corresponding functional zones. The operation points D, E, and F are generally sub-optimal as shown in the graph. To the right of the N content line, production is sustainable but net N always exists, reflecting the fact that this aspect of the agricultural crop production is "a fundamental constraint to the long-term environmental performance and sustainability" (Groffman, p.57; Wallace).

If applied in an economically optimal way, N rates for high productivity, average productivity, and low productivity functional zones, respectively, will be N_H, N_I, and N_L, respectively. The optimal operation points are A, B, and C, respectively, and result in corresponding net N represented by the lengths of AH, BI, and CM, respectively. As a general observation from the above particular drawing, it can be seen that net N is not necessarily higher for higher application rates when soil productivity is also variable.

Under the conventional strategy, when the dominant portion of the field is highly productive, then uniform N application rate N_H to the entire field results in large net N, GJ, for the lowly productive portion of the field and XY for the average productive zone. Total net N for the field reaches its maximum and at least part of the net N in GJ and XY is economically unproductive. On the other hand, if the dominant portion of the field is lowly productive or average productive, then the highly productive portion of the field is under-applied at point P or at point K in the figure.

The functional zone-uniform strategy takes into account all zones within the field to increase net return from the conventional strategy for the farmer. An optimal functional zone-uniform strategy would apply more than N_I but less than N_H to maximize expected profits. Thus, when the predominant portion of the field is highly productive, the functional zone-uniform strategy results in less total net N as compared with the conventional strategy. When the predominant portion of the field is lowly productive or average productive, the functional zone-uniform strategy results in larger total net N as compared with the conventional strategy. This result indicates that without variable rate application, functional zone information alone may not be able to reduce net N for a field substantially and sustainably.

Finally, with variable rate application, uneconomic over- and under-application and unsustainable under-application do not occur. Even though it is not clear whether or not total net N from variable application is less than that of uniform application when mining of soil N is permissible, it is very likely that total net N from variable application will never be greater than that from uniform application given that soil N is not mined. Thus it is likely that the functional zone-variable strategy is superior to both conventional and functional zone-uniform strategies in reducing net N and N pollution potential.

2.3. Yield map and the functional zones

SSM technology allows the farmer to manage a field as a collection of many SMUs, e.g. 500 SMUs. That is, instead having a field, the farmer now has 500 sub-fields to manage. To manage each SMU separately, it is important for the farmer to predict yield level and thus N need for each SMU. Then the farmer can practically group these 500 SMUs into several distinct zones with respect to N needs.[17] This section develops the procedure to predict yield for each SMU, and describes the establishment of the functional zones. Major steps involved in this section are illustrated in the following figure.

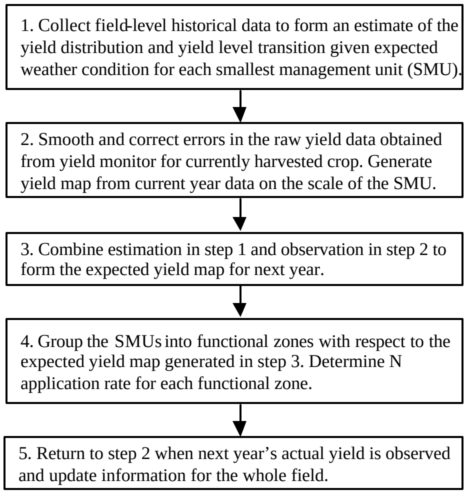
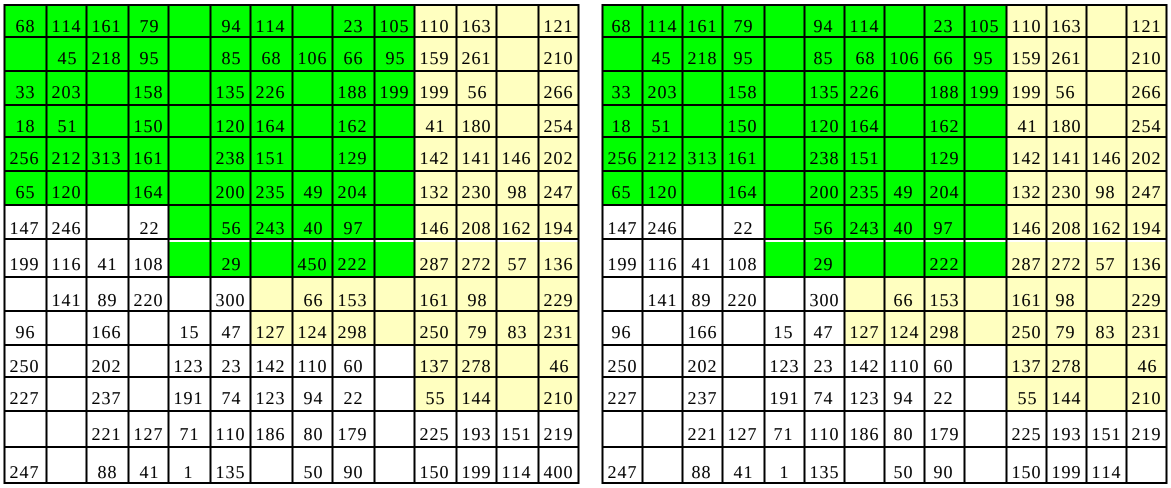


Figure 2-2. The steps to generate the functional zones for a field

---

[17] The 500 SMUs are not treated as 500 zones but *J* zones (e.g. 5 zones) for two reasons: one is that predicted yield for a SMU is subject to error so two SMUs with similar but not equal predicted yields can receive same amount of N; two is that fertilizer recommendations are usually given

Figure 2-3 further illustrates the steps listed above. Figure 2-3(a) shows the observed raw yield data for a field; Figure 2-3(b) shows the edited yield data which correct certain errors in the observed yield data; Figure 2-3(c) shows the observed yield generated at the scale of SMU; and Figure 2-3(d) shows the functional zones generated by cluster analysis. The numbers in Figure 2-3 are all hypothetical. More detailed discussion is in the following subsections.

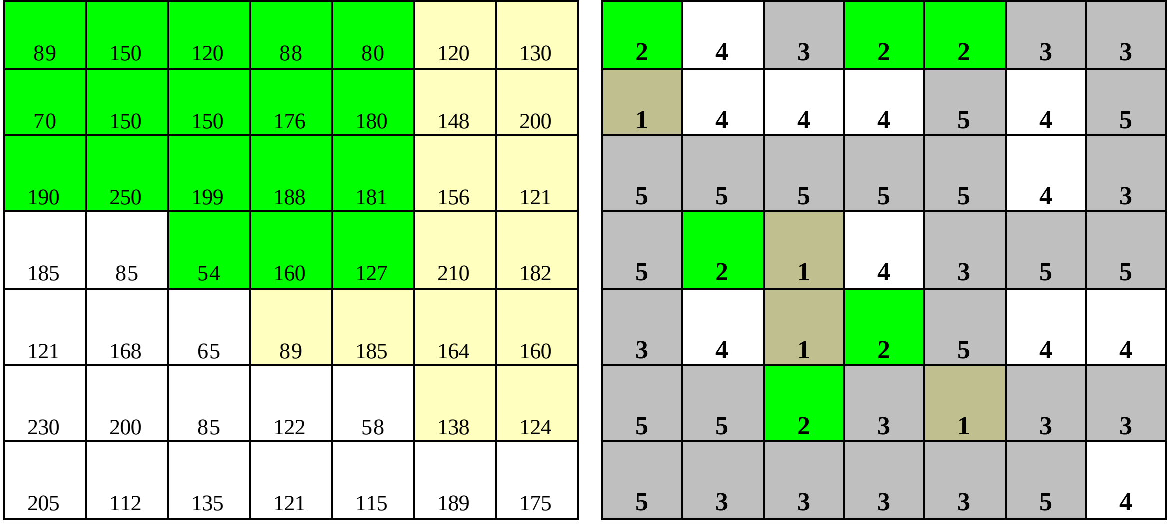


68	114	161	79		94	114		23	105	110	163		121
	45	218	95		85	68	106	66	95	159	261		210
33	203		158		135	226		188	199	199	56		266
18	51		150		120	164		162		41	180		254
256	212	313	161		238	151		129		142	141	146	202
65	120		164		200	235	49	204		132	230	98	247
147	246		22		56	243	40	97		146	208	162	194
199	116	41	108		29		450	222		287	272	57	136
	141	89	220		300		66	153		161	98		229
96		166		15	47	127	124	298		250	79	83	231
250		202		123	23	142	110	60		137	278		46
227		237		191	74	123	94	22		55	144		210
		221	127	71	110	186	80	179		225	193	151	219
247		88	41	1	135		50	90		150	199	114	400

(a). Raw yield data from yield monitor

68	114	161	79		94	114		23	105	110	163		121
	45	218	95		85	68	106	66	95	159	261		210
33	203		158		135	226		188	199	199	56		266
18	51		150		120	164		162		41	180		254
256	212	313	161		238	151		129		142	141	146	202
65	120		164		200	235	49	204		132	230	98	247
147	246		22		56	243	40	97		146	208	162	194
199	116	41	108		29			222		287	272	57	136
	141	89	220		300		66	153		161	98		229
96		166		15	47	127	124	298		250	79	83	231
250		202		123	23	142	110	60		137	278		46
227		237		191	74	123	94	22		55	144		210
		221	127	71	110	186	80	179		225	193	151	219
247		88	41	1	135		50	90		150	199	114	

(b). Edited yield data from raw yield data



89	150	120	88	80	120	130
70	150	150	176	180	148	200
190	250	199	188	181	156	121
185	85	54	160	127	210	182
121	168	65	89	185	164	160
230	200	85	122	58	138	124
205	112	135	121	115	189	175

(c). Yield map generated from observed yield

2	**4**	**3**	**2**	**2**	**3**	**3**
1	**4**	**4**	**4**	**5**	**4**	**5**
5	**5**	**5**	**5**	**5**	**4**	**3**
5	**2**	**1**	**4**	**3**	**5**	**5**
3	**4**	**1**	**2**	**5**	**4**	**4**
5	**5**	**2**	**3**	**1**	**3**	**3**
5	**3**	**3**	**3**	**3**	**5**	**4**

(d). Functional zones form by cluster analysis

Figure 2-3. The generation of functional zones from observed yield data

---

by intervals (e.g. "50 pounds K and 30 pounds P if yield is 80 to 100 bushels") and several SMUs may fall in the same interval.

Observed yieldfor current year

Raw data from yield monitor. A yield monitor gives automated measurements of the production harvested at intervals as a combine travels over a field, adjusted by vehicle speed, head position, and crop moisture levels derived from separate sensors (Heimlich). Data collected by real-time sensors such as a yield monitor can be seen as intensive samples providing rich information on small-scale spatial variability. With as many as 1,400 observations per hectare from the yield monitor, it is rarely necessary to estimate missing observations (Whelan et al.)[18].

Data directly recorded are subject to many errors. As described by Burrough and Swindell, the first type is errors in GPS measurements that cause errors in locations, so raw data have to be adjusted to the common coordinate base used for the digital elevation model. The second is the error of extreme values resulting from short-range variation due to harvesting and locational errors. Raw yield data need to be interpolated as averages for larger blocks in order to remove effects of these extreme errors. Furthermore, certain irregularities both in yield data or locational data need to be smoothed out. For example, grain flow through the combine may be temporarily interrupted due to a mechanical problem. Sometimes the yield monitor records a yield even when there is no crop to harvest such as at the end of the field (McClellan). As an example, in Figure 2-5(b), two extreme values are removed (see highlighted boxes).

Generally, the data points reported by the yield monitor cover different areas and need to be geo-referenced (i.e. the four corners of the area represented by the data point need to be calculated explicitly). A simpler approach is to treat each yield recording as a

---

[18] The yield monitor attached to a global positioning system, or GPS-located combine harvester referenced to a local base station (or differential correction provided via satellite signal) records data on a smart card every one to two seconds during which the combine travels over the field some 2 meters (Burrough and Swindell; McClellan).

data point. As it is shown later, a SMU defined for this study can contain up to 100 such points, so the each-recording-as-a-point approach is precise enough in generating yield maps. The adjusted data then become manageable and are ready to be used to generate yield maps for given grid size (Figure 2-5(c)). In this study, the adjusted data will be calculated by a simple average method.

Yield for each SMU. There are three things to consider in generating a yield map of given grid size from corrected yield monitor data. The first is to choose the right grid size. The right size is one which is convenient to manage as a distinctive field unit and one for which the within-grid variability is not a concern to the farmer. The correct-sized grid is illustrated in Figure 2-5 (c) and has been referred to previously in this study as the smallest management unit (SMU).

The second consideration is the determination of the yield level for each grid when at least part of the grid has yield data reported by yield monitor. Since the grid size is so chosen that the within-grid variability is negligible, a simple average yield of the data points contained in the grid will suffice.

The third consideration is how to handle grids with no yield data. When data points are totally missing from a SMU that is inside the field, then the yield for this grid needs to be estimated from neighboring grids. Geo-statistical methods such as kriging or other spatial interpolating methods such as the inverse distance interpolator method can be used (Wollenhaupt et al.; Cressie). Geostatistical methods such as kriging are used widely in SSM studies (Webster; Whelan et al.; Webster and Oliver).

The functional zones (cluster analysis)

After an expected or observed yield map at the scale of SMU is established for a field, the farmer then can divide the SMUs into several functional zones according to their yields and N needs. Within each functional zone, the predicted yield for the SMUs, which are not necessarily next to each other, is statistically homogeneous (or “similar”), while two different zones have yields that are statistically different from each other. In this study, both ex post analysis and ex ante analysis use cluster analysis to generate functional zones.

The fundamental clustering criterion is from the matrix equation (Everitt):

$$\mathbf{T} = \mathbf{W} + \mathbf{B} \tag{2.11}$$

where $\mathbf{T}$ is the total dispersion matrix, $\mathbf{W}$ is the matrix of within-groups dispersion, and $\mathbf{B}$ is the between-groups dispersion matrix. Because $\mathbf{T}$ is fixed for a given data set, clustering criteria should be functions of $\mathbf{B}$ and $\mathbf{W}$. When only one variable is involved in grouping (as in the case for the farmer who only uses yield data to do the grouping), then it is straightforward to minimize $\mathbf{W}$ or maximize $\mathbf{B}$. When more than one variable is involved in grouping, then several criteria are suggested, such as minimizing the trace of $\mathbf{W}$, minimizing the determinant of $\mathbf{W}$, and maximizing the trace of $\mathbf{BW}^{-1}$ (Everitt). These types of analysis and algorithms are described in Everitt, and Massart and Kaufman, as well as in *SAS User’s Manual* (SAS Institute) and can be carried out by using statistical software like SAS.

The desired number of groups, K, needs to be determined by certain criteria. For example, statistically, one of the most operational methods is based on Marriot’s criterion, namely choosing the K such that $K^2|\mathbf{W}|$ is a minimum (SAS Institute). In this study, the desired number of groups will be chosen not based on the Marriot’s criterion, but based on the observed range of yield within the field and agronomic recommendations as to the minimum yield range for varying N applications. For example, when wheat yield varies from 40 to 120 bu/ac, eight groups can be chosen, with a interval of 10 bushels, which is usually the minimum difference of yield for recommending a change of N application rate. After the desired number of groups is decided, then the functional zones can be established by using an appropriate cluster algorithm. As information increases, e.g. with more years of data observations, the farmer will have more yield maps and will have more accurate expected yield patterns, which may result in a change in the functional zone grouping.

When grouping number, K, is decided, then for each group, a "centroid" of the group is chosen (a centroid is a vector and in one-dimensional data is the mean). An observation's membership in the group depends on its distance from the centroid. The methods to classify observations (objects) thus are called nearest centroid sorting methods (Massart and Kaufman). One straightforward method is by minimizing the sum of the squares (when the distance is Euclidean) of the distances to the centroid:

$$E = \sum_{k=1}^{K} \sum_{i \in C_k} |\vec{x}_{i.} - \bar{x}_k|^2 \qquad (2.12)$$

where $\bar{x}_k$ is the pattern of the centroid of cluster C_k with m_k members, i.e.,

$$\bar{x}_k = \frac{1}{m_k} \sum_{i \in C_k} \vec{x}_{i.} \text{ and } m = \sum m_k \qquad (2.13)$$

The procedure to carry out the actual cluster analysis described above is (Massart and Kaufman, p. 105-107):

1. Select an initial set of K centroids and the corresponding clustering of m objects into K clusters.
2. Compute the centroids of the clusters and the distance of all objects to all centroids.
3. Reallocate all objects that were incorrectly located. That is, after the new centroid for each group is calculated (step 2), each object will be compared with

the centroid of its own cluster and those of the neighboring clusters. If an object is found closer to the centroid of one of its neighboring clusters than to that of its own cluster, then this object is incorrectly located and should be moved to that neighboring cluster.

4. If any reallocation happens, return to 2, otherwise, stop.

2.4. Modeling temporal and spatial variability

Before functional zones for next year can be generated, it is necessary to predict the yield level for each SMU. It is the within-field variability for each year that is important to the farmer in SSM, but the pattern of within-field variability itself varies greatly from year to year. The farmer needs to predict the state of the spatial variability for the coming season. Several approaches used by researchers are discussed briefly here.

A brief review of modeling approaches

Spatiotemporal models. One approach is to decompose yield data at location (x, y) at year t, $q(x, y, t)$, into:

$$q(x, y, t) = m(x, y, t) + r(x, y, t) \tag{2.14}$$

where $m(x, y, t)$ is mean term and $r(x, y, t)$ is residual term. The basic idea of this decomposition is to fit first the mean term using generalized linear models, generalized additive models, regression-tree models, or a deterministic function describing a physical process, and then model the residuals using spatial-temporal semivariance models (McBratney et al.). Empirical applications of this approach often show that large spatial-temporal variation or lack of data make the predictions from this approach problematic (McBratney et al.). Lack of data over years makes it difficult to determine the effects of temporal change on within-field spatial variability (Stein et al., 1997b).

Pattern-recognition using explanatory variables. Another approach as exemplified by Stein et al. (1997a) is using the spatial pattern of certain measurable and/or stable variables (such as elevation, slope, and selected soil variables) to explain the observed spatial variability of yield within a field. Methods such as distance (such as Euclidean distance) between maps (spatial patterns of a selected variable versus spatial patterns of observed yields) are used. Though it is possible to use this approach to guide site-specific recommendations, empirical results show that variability explained by the models is rather small. Furthermore, in order to utilize proper geostatistical analysis, adequate sampling and accurate soil maps at fine resolution are essential. As indicated in Chapter One of this study, soil maps with high resolution do not exist in the study area, and the cost of adequate sampling is high.

Defining management zones using aerial observation. Some researchers such as Asim et al. use aerial graphs over several years to identify management units within a field for wheat. In Asim et al., based on interpretation of the wheat growth pattern from the aerial observation over seven years, cluster analysis is used to divide a field into several management zones. This approach may ignore the fact that the management zones thus identified are still subject to temporal variation.

Markov chain models for both temporal and spatial transition. In this study, a Markov chain model is used to predict both field-level temporal variation (year to year variation due to weather and/or pests) and within-field spatial variation based on several considerations. First, long-term field-level yield variation that dominates total yield variation for a field[19] can be described by a Markov chain model (discussed later). Second,

---

[19] McBratney et al. commenting on research by Eghball et al. and by Eghball and Power, point out that average yields tend to be dominated by long-term temporal variation. When annual correlation of yields is low as compared with spatial

sub-fields (SMUs) now can be viewed as individual fields with yield monitor and variable rate application technology available so within-field variation, which determines the spatial variability of a field in given year, can be traced by looking at the temporal variation of individual sub-fields (i.e. the SMUs). An advantage of this approach is that the required data are often available for a farm. Long-term field-level historical data may be available or can be generated with a simulation model. Short-term sub-field yield data are available from the yield monitor.

The Markov chain approach is an attractive alternative to regression analysis for crop yield forecasting (Matis et al. 1989, 1985). Matis et al. (1985) use a Markov chain model to forecast a crop yield distribution at intermediate times in the growing season relating crop and soil moisture condition to final yield. They used regression analysis as a preliminary step to select the classification variable for defining the states of nature. They argue that the single required assumption of a Markov process is statistically testable and biologically reasonable; that the approach is nonparametric; and that it is easy to implement. One disadvantage of the Markov chain model is that it requires many years of yield data. However, this disadvantage could be overcome by using a simulation model with historical weather data (Matis et al. 1985) or by observing many samples over several years (Matis et al. 1989). The long-term variation of detrended crop yields is determined mostly by weather variation and historical weather data for many years are available for the study area. In fact, Markov chain models are often used to model weather-related or biological processes (e.g. Martell).

---

correlation, it is said that temporal variation dominates the yield. Clearly, long-term temporal variation of yield, when detrended to reflect the technological progress, is mainly determined by weather. A field-level yield is a weighted average of the sub-fields within the field.

The Markov chain model

The following layout of a Markov chain model relies heavily on Anderson and Goodman. Suppose a stochastic phenomenon (e.g. yield of a SMU) is classified by a finite set of states of nature (e.g. “above average”, “average”, and “below average” where “average” is the mean yield of the SMUs within a field for a given year) and is measured at equi-distance in time (e.g. once every two years). If the variable of interest (such as crop yield) follows a Markov process, then the probability distribution of the variable at future time $t + 1$ depends only on its realization at time t. Denote the states as $i = 1, 2, \ldots, s$, and the time as $t = 0, 1, \ldots, T$. Let $p_{ij}(t)$ be the transition probability which is the probability of state j at time t, given state i at time $t - 1$. Let $n_i(0)$ be the number of individuals (e.g. SMUs in a field) in state i at time 0 and $n_i(t)$ be the number of individuals in state i at time t. Assume that there are many observations in each of the initial states and the same set of transition probabilities operates at each time step.[20] Thus, the Markov chain model is characterized by the Markov property and a transition matrix:

(1). The transition matrix:

$$\mathbf{P} = \begin{array}{r|cccc|} \text{States} & 1 & 2 & \ldots & s \\ \hline 1 & p_{11} & p_{21} & \cdots & p_{s1} \\ 2 & p_{12} & p_{22} & \cdots & p_{s2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ s & p_{1s} & p_{2s} & \cdots & p_{ss} \end{array}$$

where $p_{ij} \geq 0$ and $\sum_{j=1}^{s} p_{ij} = 1$, all i. The probability of p_{ij} is the probability that given current state of nature is i, the state of nature for next period is j. Note the transition probabilities

---

[20] This is one of the two approaches to study Markov chain models. In the second approach, only one individual is studied and the sequence of observations is large (e.g. observed for many years). Anderson and Goodman show that these two settings have the same asymptotic properties.

above are not functions of time, meaning that the same transition probability pattern holds for all time.

(2). Homogeneous Markov property:

$$P(s(t+1)=i_{t+1} \mid s(t)=i_t,\ldots,s(1)=i_1,s(0)=i_0)=P(s(t+1)=i_{t+1} \mid s(t)=i_t) \qquad (2.15)$$

where $s(t)$ is state of nature at time t. . This property says that given the current state of nature is known, the probability distribution of the state of nature for future time t + 1 is not related to the history of the state of nature before time t.

Anderson and Goodman show that the maximum likelihood estimate for p_{ij} is

$$\hat{p}_{ij}=n_{ij} / n_{i\cdot}=\left(\sum_{t=1}^{T} n_{ij}(t)\right) / \left(\sum_{j=1}^{S}\sum_{t=1}^{T} n_{ij}(t)\right)=\left(\sum_{t=1}^{T} n_{ij}(t)\right) / \left(\sum_{t=0}^{T-1} n_{i}(t)\right) \qquad (2.16)$$

where $n_i(t)=\sum_{j=1}^{S} n_{ij}(t)$. To calculate $\hat{p}_{ij}$, T should be at least 1, that is, there should be at least an initial observation time and one follow-up observation time. Based on asymptotic properties, Anderson and Goodman also developed test statistics for testing of homogeneity (that p_{ij}'s are constants over time and across individuals), for testing of independence (that a first order Markov transition does not exist), and for testing of hypotheses about specific probabilities and constructing confidence regions (for details see Anderson and Goodman, pp.96-99).

Once the transition matrix is estimated, the future distribution of the states of nature given current realization can be predicted. For example, if the current state of nature is i, then j periods later, the distribution of the states of nature is $\boldsymbol{p}_i' \cdot \mathbf{P}^j$, where $\boldsymbol{p}_i'=(0,\ldots,1(i^{th}),0,\ldots,0)$ and $\mathbf{P}^j$ is the $\mathbf{P}$ raised to the power of j. For example, let the transition probability matrix for all SMUs in the field be

$$\mathbf{P} = \begin{array}{r} \\ > Average \\ Average \\ < Average \end{array} \begin{array}{ccc} > Average & Average & < Average \\ \end{array}$$

$$\mathbf{P} = \begin{array}{r|ccc|} & > \mathit{Average} & \mathit{Average} & < \mathit{Average} \\ > \mathit{Average} & 1/4 & 2/4 & 1/4 \\ \mathit{Average} & 1/3 & 1/3 & 1/3 \\ < \mathit{Average} & 2/5 & 2/5 & 1/5 \end{array}$$

The row index represents current state and the column index represents the future state. For example, the cell at third row and first column is “the probability that the future state is above average, given that the current state is below average”. Suppose for a certain SMU, the current state is “average” (second row), then the transition probability matrix for two periods from now is:

$$\mathbf{P}^2 = \begin{array}{r|ccc|} & > \mathit{Average} & \mathit{Average} & < \mathit{Average} \\ > \mathit{Average} & 0.329 & 0.391 & 0.279 \\ \mathit{Average} & 0.328 & 0.411 & 0.261 \\ < \mathit{Average} & 0.313 & 0.43 & 0.273 \end{array}$$

$p_2' = (0 \;\; 1 \;\; 0)$, so $p_2' \cdot \mathbf{P}^2 = (0.328 \;\; 0.411 \;\; 0.261)$. That is, for this certain SMU, the probabilities that two periods from now the state of nature is “above average”, “average”, and “below average”, respectively, are 0.328, 0.411, and 0.261, respectively, given that the current state of nature is “average”. These states of nature describe the relative yield levels among the SMUs in a field and are used to represent the within-field variability for a given year.

The above layout of the Markov chain model is for the situation where many individuals are observed over several periods, i.e. $\sum_{i=1}^{s} n_i(0)$ is large as well as $n_i(t)$, and $T \geq 1$. This model will be used to study the sub-fields variation after field-level temporal variation is explained. For the field-level temporal variation, the Markov chain model is specified differently where only one individual (i.e. one field) is observed but the observation is taken over many years. That is, $\sum_{i=1}^{s} n_i(0) = 1$ and T is large. Models of this kind are discussed elsewhere (see for example, Bartlett; Iosifescu; Kemeny and Snell). The asymptotic theories about these two approaches are similar (Anderson and Goodman). While the states of nature for a SMU are defined in terms of relative yield levels among the SMUs, the states of nature for a field are defined in terms of absolute yield level.

For both models, the estimates of the transition probabilities can be updated as more years' data become available to the farmer. For example, suppose the farmer used the past 20 years' historical data to predict the 21st year's yield. After the 21st year becomes history, he can use 21 years' data to estimate the transition probability matrix and then use the new transition matrix and the 21st year's state of nature to predict the 22nd year's yield. Once the distribution of the future yield for a SMU or a field is determined, the expected yield or optimal yield as well as corresponding N application rates for alternative strategies can be determined.

2.5. Optimal target yield and optimal N rate

After the probability distribution of the states of nature for next period has been determined for a SMU or a field, the next step is to choose an optimal target yield that maximizes expected net return for the SMU or the field. Determining the optimal target yield is equivalent to determining the optimal N application rate for the SMU or field for the coming season as following discussion shows. For uniform application strategies (the conventional strategy, the soil zone-uniform strategy, and the functional zone-uniform strategy), a field-level target yield needs to be determined. For variable application strategies (the soil zone-variable strategy and the functional zone-variable strategy), an optimal target yield needs to be determined for each management unit within the field. The optimal target yield is so chosen that the expected cost of inaccurate N application will be minimized. The following discussion presents the procedure to determine the optimal N rate for uniform application.

General procedure for determining the optimal N rate. Suppose that based on the soil zone or functional zone information, the farmer knows that zone *i* should receive N at a rate of N_i in order to achieve the expected yield that maximizes expected profit for the zone. Denote the uniform N application rate as x. Then when x is larger than N_i, overapplication of N happens and $O_i = \max(x - N_i, 0)$ is the amount of net N. If x is smaller than N_i, then underapplication of N happens and a yield penalty of $A_i \cdot \boldsymbol{d} \cdot U_i$ results, where $\boldsymbol{d}$ is the yield penalty in bushels for each pound of N underapplied and $U_i = -\min(x - N_i, 0)$ is the amount of N underapplied per acre. So the total cost of N misapplication for zone i is

$$A_i \cdot \left(P_N \cdot O_i + (P - P_h) \cdot \boldsymbol{d} \cdot U_i - P_N \cdot U_i\right) \tag{2.17}$$

where P_N is the price of N, P is the price of crop, and P_h is the variable harvest cost per bushel. That is, when N is overapplied, there is no yield penalty but all overapplied N becomes part of the net N (applied but not removed by the crop). When N is underapplied, there is cost from yield penalty as well as savings from less harvest cost and less N cost. Collecting terms, the above expression becomes

$$\left(A_i \cdot P_N\right) \cdot O_i + \left(A_i \cdot \left((P - P_h) \cdot \boldsymbol{d} - P_N\right)\right) \cdot U_i \tag{2.17'}$$

Since $(P - P_h) \cdot \boldsymbol{d} - P_N \geq 0$, the farmer loses money when N is underapplied for any zone, other things equal. It is assumed $P_N \neq (P - P_h)\boldsymbol{d} - P_N$ because in the very unlikely case where $P_N = (P - P_h)\delta - P_N$, the optimization problem below cannot be solved.[21] Thus, the farmer's objective is to choose a uniform N application rate in order to minimize the total cost of misapplication for the whole field. When cost of misapplication is minimized, the expected net return is maximized. Mathematically, his objective is

$$\min_{x} \sum_{i=1}^{I} \left\{ \left(A_i \cdot P_N\right) \cdot O_i + \left(A_i \cdot \left((P - P_h) \cdot \delta - P_N\right)\right) \cdot U_i \right\} \qquad (2.17.1)$$

subject to

$$x + U_i - O_i = N_i,\ i = 1, 2, \ldots, I \qquad (2.17.2)$$

$$U_i \cdot O_i = 0,\ i = 1, 2, \ldots, I \qquad (2.17.3)$$

$$x, U_i, O_i \geq 0,\ i = 1, 2, \ldots, I \qquad (2.17.4)$$

where constraint (2.17.3) highlights the fact that a zone cannot be simultaneously over- and under-applied. The above formulation is a standard nonpreemptive goal-programming model after the redundant non-linear constraint (2.17.3) is dropped (Hillier and Lieberman). The constraint (2.17.3) can be dropped because with any optimal solution x determined, if O_i and U_i were both positive, then assuming $P_N \neq (P - P_h)\delta - P_N$ the objective function could be further improved if all O_i is given to U_i when $P_N > (P - P_h)\delta - P_N$ or the opposite when $P_N < (P - P_h)\delta - P_N$, leaving constraint (2.17.3) redundant. The above procedure is applicable directly to the situation where there is a field with a probability distribution of yield over the years and an optimal yield is to be determined in making uniform N application decision (the conventional strategy). This

---

[21] In fact, $(P-P_h)\delta$ - P_N is much larger than P_N in this study, meaning that farmers suffer much larger loss from underapplying N than from overapplying N. Alley comments that "this is a point that growers understand very well."

procedure is the basis for the later discussion of determining optimal yield for next year when transition probabilities are involved.

Above layout describes the rules to determine the optimal uniform N rate over a field when soil zone or functional zone information is available. The optimal variable N application rate for a given information level is the amount that results in zero overapplication and zero underapplication and the cost of misapplication is thus zero.

After N rates are determined by rules discussed above, the actual costs of misapplication can be calculated for each strategy when actual yields are observed. Based on actual costs of misapplication for the strategies, the values of information and variable application can be calculated. Suppose the actual costs of misapplication for the conventional strategy, the soil zone-uniform strategy, the soil zone-variable strategy, the functional zone-uniform strategy, and the functional zone-variable strategy are, respectively, L_C, L_{SU}, L_{SV}, L_{FU}, and L_{FV}, respectively. Then the gross values and net values of information and variable application in Table 2-1 can be more specifically expressed in Table 2-2.

Table 2-2. Values of information and variable application for different strategies using costs of misapplication[ab]

Strategy	Change of profit from baseline[c]	Information value		Value of variable rate application	
		Gross	Net	Gross	Net
Conventional	0	0	0	0	0
Soil zone-uniform	$L_C - L_{SU}$	$L_C - L_{SU} + AC_S$	$L_C - L_{SU}$	0	0
Functional zone-uniform	$L_C - L_{FU}$	$L_C - L_{FU} + AC_F$	$L_C - L_{FU}$	0	0
Soil zone-variable	$L_C - L_{SV}$	$L_C - L_{SU} + AC_S$	$L_C - L_{SU}$	$L_{SU} - L_{SV} + AC_{SV}$	$L_{SU} - L_{SV}$
Functional zone-variable	$L_C - L_{FV}$	$L_C - L_{FU} + AC_F$	$L_C - L_{FU}$	$L_{FU} - L_{FV} + AC_{FV}$	$L_{FU} - L_{FV}$

a. Values presented in this table are the total values for a field and not per acre values.

b. Subscript “C” indicates conventional strategy, “*SU*” indicates soil zone-uniform strategy, “*FU*” indicates functional zone-uniform strategy, “*SV*” indicates soil zone-variable strategy, and “*FV*” indicates functional zone-variable strategy.

c. Conventional strategy is the baseline case throughout this study.

It should be noted that determining an optimal N rate is equivalent to determining an optimal crop yield goal from which the optimal N rate is calculated. The following discussion will translate the procedure for determining the optimal N rate described above into the procedure for determining an optimal yield (target yield) because both temporal and spatial variability of yield thus can be incorporated explicitly into the decision making process.

Field-level target yield. Suppose the current year is in the state *i*, then the probability distribution for next year is (p_{i1}, p_{i2}, …, p_{is}). Because the farmer is assumed to maximize expected profit, the target yield should be that which gives him the maximal expected net return from the field. Denote state of nature *j* as a yield interval of (y_{j1}, y_{j2}) where $y_{j1} < y_{j2}$. Suppose the conditional distribution of yield *Y* in the state of nature *j* is $f(y|S = j)$, e.g., $f(y|S = j)$ is uniform, i.e.

$$f(y \mid S = j) = \begin{cases} \dfrac{1}{y_{j1} - y_{j2}} & \forall\, y \in (y_{j1}, y_{j2}) \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

Suppose that in order to achieve yield level *y* the N application needed is x. Denote the cost of misapplication as

$$C_y = P_N O_y + (P - P_h) d U_y - P_N U_y \tag{2.19}$$

which is similar to (2.17) but with slightly different subscripts. Thus, if the state of nature for a future time is given, the farmer's objective is to choose an application rate of *x* to

$$\min_y \int f(y \mid S = j) C_y \, dy \tag{2.20}$$

and from this an optimal target yield can be solved for state of nature j. Since the state of nature for a future time is expressed as a probability distribution, the farmer's actual objective becomes:

$$\min_y \sum_{k=1}^{s} p_{jk} \int f(y|S = j) C_y \, dy \tag{2.21}$$

In this study, the probability distribution $f(y|S=j)$ is represented by assigning equal probability to each observed yield. For example, if in state nature of j, there are observations of 70, 79, 78, 74, and 81 bu/ac, a probability of 1/5 can be assigned to each of these observations. Then the algorithm used to solve (2.21) is basically the same as that for solving the mathematical programming model presented by (2.17.1) to (2.17.4).

Target yield for a SMU. The preceding discussion described a procedure to determine an optimal yield when field level yield probabilities are available. When yields are available for each SMU, target yields must be derived for each SMU in order to derive uniform and variable application amounts. The procedure to determine optimal SMU yield is discussed below.

Based on the field-level transition matrix above, the expected field-level yield for next year can be calculated as

$$EY = \sum_{k=1}^{K} p_{jk} \int yf(y|S = j)dy \tag{2.22}$$

assuming N is not limiting. Suppose there are K states to describe the yield of a SMU relative to EY, e.g. state $k = 1$ is "more than 25% lower than the field-level yield". Given the current year's state of nature i, the distribution for the future state for this SMU is (p_{i1}, p_{i2}, …, p_{iK}). Based on information from EY and the definition of states of nature for a SMU, the distribution (p_{i1}, p_{i2}, …, p_{iK}) is translated into a distribution for yield in terms of absolute (not relative) yields.

Thus, for a variable rate application strategy, the determination of the optimal yield for a SMU is identical in procedure to the determination of field-level target yield as described in (2.21). For uniform N application with information for each SMU available, the field-level optimal yields are derived based on the SMUs. In order to determine a field-level target yield (as reflected in the uniform N rate) for the whole field when SMU information is available, suppose the total number of SMUs that are in state of nature k in the current year is n_k and the area for a SMU is b, that is, $A = b\sum_{k=1}^{K} n_k$. Suppose that the distribution of yield for a given state of nature k is $f(y \mid S = k)$ and the cost of misapplication for the SMU is $C(y)$ per acre. Then based on (2.21), the farmer's problem becomes

$$\min_{y} \; b\sum_{j=1}^{K}\sum_{k=1}^{K} p_{jk} n_k \int f(y|S = j)C(y)dy \tag{2.23}$$

The optimization problem (2.23) can be understood more easily in thinking that each group of SMUs that is in the same state in current year is a sub-field and there are K such sub-fields. That is, the farmer is determining a single optimal N rate (target yield) for all the sub-fields and thus obtains the optimal result for the whole field. This setting is applicable to the soil zone-uniform strategy and the functional zone-uniform strategy when the yield transition probabilities are considered. Since different soil zones have different areas, (2.23) needs to be modified slightly by replacing bn_k with a single symbol B_k, the area of the k^{th} soil zone.

2.6. Summary of this chapter

This chapter described the theoretical framework for the procedures to achieve the objectives of the study. Theoretical procedures for describing farmers' decision making, evaluation of information and variable rate application are developed, and an index for potential N pollution is developed. This chapter also described the procedure to set up functional zones using cluster analysis and the procedure to model temporal and spatial yield variability using Markov chain models. A procedure for optimizing N application rates (and thus target yield level) conditional on the level of information and N application strategy is described. In Chapter 3, the empirical model will be described.

Chapter 3. Empirical Model

3.0. An overview

The farmer's assumed objective is to select the N management information and application strategy that maximizes expected net return for the field. The farmer's expected net return from a field is

$$\pi = P_y \cdot y - C_0 - C_S - C_{SS} \tag{3.1}$$

where P_y is the crop price, y is the expected crop yield for a given N application, C_0 is the part of production cost which is not related to either the N application strategy adopted or the site-specific situation of the field (such as planting the crop), C_S is the cost of following an N application strategy that does not depend on the site-specific situation of the field (such as cost for yield delineation and record keeping by soil or functional zone), and C_{SS} is the additional cost which results from the site-specific situation of the field under a given strategy (e.g. cost of total N fertilizer applied). In the following discussion, C_S is referred to as "information and application-related cost", and C_{SS} is referred to as "yield and fertilizer-related cost".

There are five N management strategies identified: conventional, soil zone-uniform, functional zone-uniform, soil zone-variable, and functional zone-variable. The additional net returns from a strategy are subdivided into an information value and a variable rate application value. Information value is calculated as the difference in expected net returns between a uniform application strategy (i.e. soil zone-uniform strategy and functional zone-uniform strategy) and the conventional strategy. Value of variable rate application is calculated as the difference of expected net returns between a uniform application information strategy and its corresponding variable application strategy.

In this chapter, empirical procedures are described for estimating the value of information and variable application for N when the farmer's objective is to maximize net returns. Section 3.1 describes the study area and the case farm. Sections 3.2 through 3.6 describe the empirical model for ex post analysis. Sections 3.7 through 3.10 described the empirical model for ex ante analysis. Section 3.11 gives a summary and limitations of the empirical models developed in this chapter.

3.1. Study area and case farm

The study area: Prince George County, Virginia

Prince George County is in the Tidewater area of southeastern Virginia and covers a total area of about 188,992 acres (Jones et al.). Thirty-seven percent of the total area, or 66,100 acres, was classified as prime farmland in 1985.[22] Major land uses were farming and wood production. Most farms produce corn, peanuts, soybeans and wheat. According to the Virginia Agricultural Statistical Service (VASS), in 1997 the average farm size was 338 acres in this county. Total cropland was 21,402 acres, including acreages for corn, wheat, and soybeans, of 4,425 acres, 3,848 acres, and 9,224 acres, respectively.[23]

"Prince George County lies entirely within the Coastal Plain physiographic province" (Jones et al. p.2). Most of the area is about 90 to 175 feet above sea level and is

---

[22] Prime farmland is defined as the land "that is best suited to producing food, feed, forage, fiber, and oilseed crops. ... Prime farmland produces the highest yields with minimal inputs of energy and economic resources, and farming it results in the least damage to the environment." (Jones et al. p.37) A field being classified as prime farmland should be currently used for producing food or fiber or be available for those uses. Current total acreage of prime farmland for the whole county may be different from that of 1985.

nearly level to slightly sloping. Soils are well drained or moderately well drained by the James, Appomattox, Nottoway, and Blackwater Rivers and their tributaries. Most of the soils in Prince George County have high acidity and low fertility, so fertilization and liming are needed for sustainable crop production. Environmental damages such as loss of chemicals and nutrients and wet soil with drainage from agriculture activity are major concerns in this area (Jones et al.).

Field description

Eight fields on a large crop farm in Prince George County were chosen for the study. The dominant soil series on this farm is Pamunkey loam and fine sandy loam. The slopes of the fields range from 1 to 4%. The fields are well drained to moderately well drained (Mostaghmi et al., 1997). The total acreage of these eight fields is 409.2 acres and the shapes and locations are displayed in Figure 3-1.

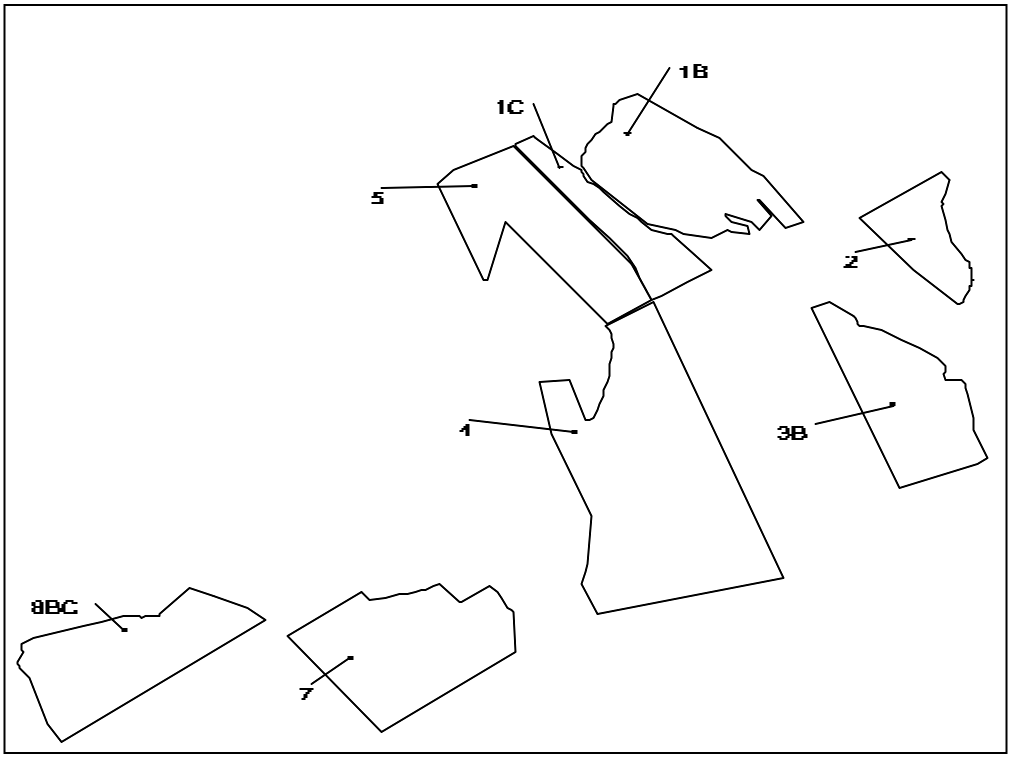


Figure 3-1. The eight fields on the case farm

---

[23] Peanut acreage decreased from 2,446 acres in 1995 to 1,557 acres in 1997 (VASS).

These fields were chosen because for the past several years, corn, wheat and soybean were planted in rotation. Good historical records are available for each field and the farm manager has practiced generating yield maps for several years (1995-1999). Among these eight fields, four were planted as corn-wheat/soybean-corn-wheat/soybean-corn from 1995 to 1999, while the other four were planted as wheat/soybean-corn-wheat/soybean-corn-wheat/soybean from 1995 to 1999. Table 3-1 reports information about these fields in terms of soil distribution, acreage, and expected yields based on soil distributions. From the table, it is seen that differences of expected yields of both crops among the fields are small due to the large proportion of Pamunkey soil in all fields.[24]

Table 3-1. Soil distributions and crop yield potential for case study fields

Soil	Acres by soil type[a]						Field acres		Exp. yield (bu/ac)	
name	ARG	BOL	EMA	MUC	PA1	PA2	Ac1[b]	Ac2[c]	Wheat[d]	Corn[d]
F1B	4.37	9.47		0.24	6.96	29.98	51.5	50.8	83	146
F1C	4.03				6.19	11.57	22.2	21.8	79	142
F2[e]				0.69	16.49	2.40	17.8	19.6	88	156
F3B		11.76		0.70	24.32	8.48	44.7	45.3	86	150
F4	9.32				84.76	23.63	118.8	117.7	85	152
F5	3.58	3.44			28.24	10.63	46.5	45.9	85	150
F7		2.00	2.26		37.55	16.61	59.2	58.4	89	157
F8B	0.68	22.93			22.46	3.71	50.4	49.8	85	145

a. Abbreviations for soil types are:

ARG: Argent (#=148) BOL: Bolling (#=151) EMA: Emporia (#=161)
MUC: Muckalee (#=168) PA1: Pamunkey (#=171) PA2: Pamunkey (#=172)

b. Total acreage calculated from 3x3 m^2 grid resolution.

c. Total acreage calculated from 30x30 m^2 grid resolution (this number will be used in this study).

d. Weighted average from VALUES (Simpson et al.) estimation for the soil types in the fields.

e. About 3 acres of the field is classified as "water" by the soil map due to coarse resolution. In this study, the "water" portion is deleted.

The N application method is uniform over the whole field. All fields were grid sampled for P, K, and pH and the variations of P, K, and pH are not large (Mostaghimi et al. 1997). Farm-level average yields for the past 10 years are about 135 bu/ac for corn and

---

[24] Pamumkey soils are deep, nearly level, and well drained with slopes from 0 to 2%. Pamunkey soils are similar to Rappahannock, Rumford, Tetorum, and Tomotley soils and are well suited to grow crops (Jones et al.),

86 bu/ac for wheat. A yield monitor has been used since 1995 to record crop yields, but no site-specific management practices were used (personal communication with the farm manager). Corn and soybeans are planted no-till, while both conventional-till and no-till are practiced for wheat. In no-till planting of corn and soybean, an in-row ripper operation is carried out. No-till planting reduced labor and machine cost, but increased herbicide cost.

Soil test or tissue tests are not done to determine the N application rate for corn. Generally 30 lb/ac N are applied pre-plant and an additional 125 lb/ac are applied as side-dress when corn is 12 inches high by dribbling the material between rows. For wheat, the rate of pre-plant N application is 30 to 40 lb/ac in November. At growth stage 25 (January), about 40 lb/ac N are applied. During late March to early April (growth stage 30), another N application of 40 to 60 lb/ac is carried out (actual amount depends on tissue test). Thus, for average 86-bushel yield, total N application averages about 125 lb/ac, which is consistent with the recommended rate (discussed below in Section 3.4). The observed field-level yields for the period of 1992 to 1998 for these fields are reported in Table 3-2.

Table 3-2. Historical yields for the case fields during 1992-1998

Field	1992	1993	1994	1995	1996	1997	1998	1992	1993	1994	1995	1996	1997	1998
	------------------- wheat (bu/ac) -------------------							------------------- corn (bu/ac) -------------------						
F1B		74		85		74		183		104		160		125
F1C		88		86		68		175		106		191		119
F2	66		89		44		42		81		118		105	
F3B		89		93		88		186		111		162		124
F4		89		101		92		196		123		179		145
F5			86		62		69	180	98		177		87	
F7	78		83		66		76		119		188		125	
F8B			87		65		64		131		182		111	
Farm-level[a]	84	88	83	96	58	91	54	186	107	125	178	170	107	126

a. Farm-level yields are from whole farm average for given crops. Acreage is around 550 acres in each year for each crop.

The fields on this farm are well-fertilized by the uniform application method, so variations in yield due to N, P, K, and pH level are suppressed. The data reveal the variability of yield potential related to factors such as soil type, erosion, elevation, closeness to woods, and insect and disease pests. For example, crop yield close to woods may drop by half even though no fertility difference is detected (McClellan). Thus, the actual data from this farm are very useful to reveal the patterns of spatial variability and potential gains from SSM information on N requirements.

3.2. Overview of ex post evaluation procedures

As stated in Chapter One, both ex post and ex ante analyses are carried out in this study to estimate values of information and variable application. Ex post analysis provides assessment on the information values and variable application values for alternative strategies when actual temporal and spatial variability is known. Ex post analyses use the conventional strategy as the baseline. Information level is specified as follows for each strategy.

The conventional strategy. The conventional strategy (baseline) in ex post analysis is based on expected yield suggested by VALUES. The conventional strategy uses the information about the predominant soil type in the field and the N application rate is the VALUES recommended rate for that soil type. This information can be obtained by consulting soil maps for the area and using VALUES recommendations. The conventional strategy ignores the minor soil types within the field even though this information may be already available.

Soil zone strategies. Soil zone-uniform and -variable strategies use VALUES to estimate yields by soil type as does the conventional strategy but include the minor soil types and soil type distributions within the field. By consulting VALUES, an N application rate is assigned to each soil zone within the field and an optimal uniform rate for the whole field (for uniform application) or an individual N application rate for each soil type within the field (for variable application) is determined.

Functional zone strategies. To generate functional zone information for ex post analysis, the soil type and distribution within the field and actual yield variability within each soil zone are used in deciding the N application rate. As discussed before, each distinct functional zone is a group of smallest management units (SMUs) that are similar in yield potential and N requirement for each crop for a given year. For ex post analysis, current yield records from the yield monitor are used to generate functional zones for the current year. The functional zone-uniform strategy, similar to the soil zone-uniform strategy, applies N at a uniform rate which is optimal for the whole field. The variable application strategy applies an individual application rate for each functional zone.

The baseline for ex post analysis is the conventional strategy without consideration of temporal and spatial transition. The soil zone strategies do not consider temporal yield variability. The functional zone strategies are based on the actual observed yields for all fields from 1995 to 1999. The generation of functional zones for ex post analyses is described in the following.

For each field in each year, data recorded by the yield monitor were mapped into 98.425x98.425 ft^2 (30x30 m^2) grids and the yield for each grid was calculated. These grids were then grouped according to their similarity by using cluster techniques as described in

Chapter 2, equations (2.11), (2.12), and (2.13). Then treating all grids in the same group as homogeneous, N application rate was determined according to the observed mean yield of each group. The number of groups is eight for wheat and fifteen for corn for each field. Eight groups for wheat were selected because the difference between the VALUES yields for the poorest soil and best soil on the study fields for wheat is 60 bu/ac. Fifteen groups for corn were determined because observed yield range for the case farm is from around 50 to around 200 bu/ac. The numbers of groups for wheat and corn reflect a reasonable tradeoff between a manageable number of groups and an accurate representation of a site's potential yield.

The ex post evaluation of information values and values of variable application in was carried out on all fields for crops for each year (1995-1999) in the following steps:

- For a field in each year, N applied to each 30x30 m^2 grid in the baseline strategy (the conventional strategy) was compared with the N needed to grow the observed crop yield as recorded by yield monitor.[25] Based on this comparison, yield penalty and net N were calculated for the baseline strategy in the given year for the given crop on the given 30x30 m^2 grids. Yield penalty and overapplication of N then were translated into cost of inaccurate N application for the grid. Field-level yield penalty, net N, and cost of inaccurate N application then were obtained by summing up all grids in the fields for the baseline strategy.
- Similar comparisons were carried out for the soil zone strategies and functional zone strategies for each field in each year (1995-1999). For example, with soil zone information, N applied to each SMU of a given soil type was compared with the N

---

[25] As stated before, it is assumed that the actual yields as recorded by yield monitor were obtained from adequate N application so extra N applied above what is needed to achieve the observed yield becomes part of net N.

needed to grow the observed crop yield as recorded by the yield monitor. Based on this comparison, yield penalty and net N were calculated for each SMU based on the given yield for the given crop. Total yield loss and net N thus can be calculated by summing up all SMUs.

- For a given field, crop, and year, the soil information value is the cost of inaccurate N application for the conventional strategy minus that for the soil zone uniform strategy. The value of variable application for soil zone strategies is the cost of inaccurate N application for the soil zone uniform strategy minus that for the soil zone variable application minus the extra cost of variable N application. The information value and value of variable application for functional zone strategies are similarly calculated.

Production costs, input and output prices, and guidelines for N application, yield potentials, and calculation of net N are described in Sections 3.3 and 3.4. The smallest management units (SMUs) are described in Section 3.5. The procedure to determine the optimal N application rates for each strategy is described in Section 3.6.

3.3. Production costs and crop prices

Production cost for corn and wheat

For the purpose of this study, all costs not related to site-specific situation and N application strategy (C_0 in equation (1)) were excluded since they do not affect the estimated information values and values of variable rate application. Costs thus excluded are fixed machine costs, some pre-harvest costs such as lime, chemicals, and chemical application, fuel, oil, repairs, and related production interest, and fixed harvest costs that do not vary by yield such as fuel, oil, repairs, and harvest labor.

The remaining crop production cost is divided into two parts. The information and N application-related cost (C_S in equation (3.1)) includes fertilizer application cost, information cost (for field delineation and record keeping), additional cost for variable rate application, and production interest related to each N application strategy. The yield and fertilizer-related cost (C_{SS} in equation 3.1) includes N fertilizer, related production interest, and hauling cost. In the following discussion, the sum $C_S + C_{SS}$ is frequently referred to as total cost.

Corn production cost. The information and N application-related cost, C_S, is decomposed in Table 3-3. All strategies are assumed to split the N application. Application costs are based on a $5.00/ac custom rate current for the area (Alley) since fertilizer application is assumed to be done by the fertilizer dealer. Information cost occurs when delineating the field using soil maps or generating yield maps and establishing functional zones. Additional costs for variable rate application differ between soil zone and functional zone strategies.

Table 3-3. Information and N application-related costs for corn

Strategy	Fertilizer application ($)[a]	Information cost ($)[b]	Variable rate application ($)[c]	Production interest ($)[d]	Total ($/acre)
conventional	10.00	0	0	0.45	10.45
soil zone-uniform	10.00	0	0	0.45	10.45
func. zone-uniform	10.00	2.50	0	0.56	13.06
soil zone-variable	10.00	0	2.00	0.54	12.54
func. zone-variable	10.00	2.50	2.50	0.68	15.68

a. All strategies split the N application.
b. Information cost is incurred for yield maps and generating functional zones.
c. This application cost is the cost in addition to the uniform application cost.
d. Calculated as 0.09*(sum of all the costs to the left in the table)/2 where 0.09 is the annual interest rate.

Information cost includes that which is needed for hiring custom-services for yield map making and record keeping. In a study on P and K site-specific management, Lowenberg-Deboer et al. estimated that per acre cost for yield map making is $2.50/ac, record keeping is $1.00/ac per year, and grid sampling is $0.75/ac. This is similar to the cost in Virginia wherein service companies such as Southern States are doing grid sampling and making fertility (P, K, and pH) maps for farmers at about $8/ac which can be used for three years (Alley). Snyder et al (1998) estimate based on surveys that per-acre cost associated with yield monitor/GPS receiver use, yield map generating, and record keeping is $2.75/ac for corn production in central Kansas. In this study, there is no grid sampling cost and yield maps are assumed to be updated every year with good record keeping. Thus, it is assumed that the per-acre cost for yield map making and record keeping is $2.50/ac per year.

The cost of variable application is estimated at $3.00/ac above uniform application in reports by Lowenberg-DeBoer et al. Snyder et al (1998) estimate (based on surveys) that per acre additional spreading cost for variable rate N application is $2.00/ac for corn production in central Kansas. English et al. report that the price to purchase SSM service (map generating plus variable N application) for corn is $4.75/ac for Tennessee farmers. In another study by Roberts et al., the additional cost for variable rate application over uniform application is estimated to be $2.00/ac for the soil-map based N application strategy. When a field is homogeneous or all SMUs in each zone within the field are connected with each other as in soil-zone maps, variable rate application could be less costly. For example, a manually-adjusted three-rate variable applicator can be used which needs no GPS facility (Thrikawala et al.; Alley). On the other hand, due to large within-field variability in terms of soil type and other soil attributes, it is expected that the variable application cost for soil-zone strategy will not be much less than that for functional zone strategy (McClellan). Thus in this study, it is assumed that the additional per acre cost incurred by variable rate application is \$2/ac for soil zone-variable application, and is \$2.50/ac for functional zone-variable application, as shown in the above table.

For the yield and fertilizer-related cost, C_{SS}, only N cost and crop hauling cost were considered in this study because variable application is only for N and yield response is expressed in terms of N applied. The price of N at purchase is \$0.24/lb, and the related interest is 0.5*0.09*\$0.24 or \$0.01 where 0.09 is the annual interest rate and 0.5 indicates the funds are used for one-half year. Thus the total cost of N fertilizer is fixed at \$0.25/lb. The yield-related crop hauling cost is \$0.15/bu (Virginia Cooperative Extension Farm Management Staff, 1999). Actual costs of N fertilizer and hauling can be determined only when expected yield is determined and actual yields are observed, respectively. For example, if the farmer applies 125 pounds of N in conventional strategy and the actual yield achieved is 115 bushels, then C_S = \$10.45, and C_{SS} = \$0.25*125 + \$0.15*115 = \$48.50. Thus, the total cost, $C_S + C_{SS}$, is \$58.95/ac.

Wheat production cost. The estimation of information and application-related cost, C_S, and the yield and fertilizer-related cost, C_{SS}, are similar to that of corn. It is assumed that all strategies apply N three times. The information and application-related cost is decomposed in Table 3.4.

Table 3-4. Information and N application-related costs for wheat

Strategy	Fertilizer application (\$) [a]	Information cost (\$) [b]	Variable application (\$) [c]	Production interest (\$) [d]	Total (\$/ac)[e]
Conventional	15.00	0	0	0.68	15.68
Soil zone-uniform	15.00	0	0	0.68	15.68
Func. zone-uniform	15.00	2.50	0	0.79	18.29
Soil zone-variable	15.00	0	2.00	0.77	17.77
Func. zone-variable	15.00	2.50	2.50	0.90	20.90

a) All strategies apply N three times including at planting.
b) Information cost is for making yield maps and generating functional zones.
c) This application cost is in addition to the uniform application cost.
d) Calculated as 0.09*(sum of all the costs to the left in the table)/2 where 0.09 is the annual interest rate.
e) It is assumed that only the last application is variable for the variable application strategies.

The yield and fertilizer-related cost for wheat is only for N fertilizer ($0.25/lb) and wheat crop hauling (assumed to be $0.15/bu). Based on above information, the farmer's cost for wheat production can be calculated. For example, if the farmer decides to apply 70 lb/ac N and achieves a yield of 60 bu/ac under the conventional strategy, then his total cost is $15.68 + 70*$0.25 + 60*$0.15, or $42.18/ac. In this study, it is assumed that only the last of the split applications is variable for the variable application strategies.

Crop prices

Crop prices are estimated from Food and Agricultural Policy Research Institute (FAPRI) forecasts from 2000 to 2007, deflated by the GDP deflators projected by FAPRI. After forecast prices for each future year are deflated, expected national prices are obtained by averaging the deflated price for each of the years. Then based on the historical price differences between Virginia and national level from 1986 to 1995 (USDA Agricultural Statistics Board) the expected national prices are further adjusted to Virginia levels. As seen from Table 3.5, the expected crop prices in this study are $2.10/bu (corn), and $2.87/bu (wheat). Since it is assumed that the farmer is risk neutral, the expected crop prices as obtained above will be used in this study.

Table 3.5. Expected crop prices for the study area[a]

Item		1999	2000	2001	2002	2003	2004	2005	2006	2007	National average[b]	Adjustment Ratio[c]	Expected price[d]
GDP	Change from last year	0	2.7	2.8	2.7	2.6	2.6	2.6	2.6	2.6			
	Deflator[e]	1.00	0.97	0.95	0.92	0.90	0.88	0.85	0.83	0.81			
Corn price	Nominal	1.94	2.00	2.06	2.10	2.17	2.24	2.29	2.35	2.39			
($/bu)	Real[f]	1.94	1.95	1.95	1.94	1.95	1.96	1.96	1.96	1.94	1.95	1.08	2.10
Wheat price	Nominal	2.66	2.98	3.15	3.25	3.34	3.46	3.54	3.64	3.66			
($/bu)	Real	2.66	2.90	2.98	3.00	3.00	3.03	3.02	3.03	2.97	2.96	0.97	2.87

a. Information sources: *FAPRI 1998, 1999 U.S Agricultural Outlook* (FAPRI, 1998, 1999); *Agricultural Prices (1985-1996 Summary)* (USDA-NASS).

b. Simple average of deflated (real) prices. The values are in 1999 dollars.

c. Adjustment ratio is calculated as average of (Historical Virginia average price)/(Historical national price).

d. Expected price is calculated as (Adjustment ratio)*(National average)

e. The deflator is calculated as 1/[(1+chang_year_1/100)*...*(1+change_year_t)]. For example, the deflator for the year of 2001 is calculated as 1/[(1+2.7/100)*(1+2.8/100)].

f. Real crop price is calculated as (nominal price)*deflator.

3.4. N applications and yield potentials

N application guidelines

The Virginia Agronomic Land Use Evaluation System (VALUES) restructures and reorients soil test recommendations to include the best currently available scientific recommendations for nutrient applications to corn, soybeans, wheat, cotton, peanuts, and other crops (Simpson et al.). "Nitrogen, a key nutrient plants required for optimum crop production, as well as a potential contaminant, has been given particular treatment with respect to both rate of application and time of application (in VALUES). Information on specific rates of P and K fertilizer to use, as well as N, is provided for all major soils in Virginia." (Simpson et al., p.1)[26] This study will determine N application rates and yield potentials for the fields in the study area according to the methods and concepts developed by VALUES. Some key concepts and procedures in VALUES important to this study are briefly explained in the following paragraphs.

Soil productivity groups. VALUES acknowledges that the ability of the soil to retain water that is accessible to plants for well drained and moderately well drained soil, or the effective drainage system to remove excess water for the poorly drained soil is one major determining factor for yield in Virginia. Soils with quite different physical or

---

[26] The minor and especially poor soils are ignored. For example, Simpson et al. state on page 4 that "The data received was from university variety trials, research plots (maximum yield), small research plots, test demonstrations (field size), five acre clubs, actual farmers, maximum yield clubs, seed companies, theses and dissertations." And there are some standards in accepting the data such as "(yield data) were obtained under the use of high crop management practices; In the case of a low yield, it was determined whether it could be attributed to the failure to use good crop management. If so, the data were omitted. If not attributed to bad management, it was assumed that the poor yield was due to the actual interaction of the crop, the soil and the rainfall/temperature pattern for that growing season, and the data were included in the study." Simpson et al. admit (p.2), "However, the best indication of the yield to expect and plan for in the coming year is soil specific yields that have been obtained under good management in the past." After grouping soil types into three groups, they use some straightforward assumption to come up with N application requirements such as "the amount of nitrogen needed to produce a given yield of corn is one pound per bushel" (p.6).

chemical properties may have essentially the same yield potential and because of this, “soil productivity groups” are developed by categorizing soil types with similar yield potential into one group.

Soil management groups. Due to differences in drainage and profile textural and depth properties, soil types within the same productivity group may have to be managed differently. The concept of soil management groups is used to categorize soils with similar profile characteristics which affect soil and crop management practices into the same soil management group. The concept of soil management group addresses the similarity of physical and chemical properties of the soils, while soil productivity group acknowledges the similarity of yield potentials. Both crop productivity groups and soil management groups for all the major soils in Virginia are developed by VALUES.

In determining fertilizer N application rates, soil management groups are first placed into larger soil productivity groups. Then for each soil productivity group, a realistic yield level for each crop is developed. This yield is not necessarily either a maximum yield or an optimal yield. Rather it represents a realistic yield that is close to a maximum yield for many if not most cropping situations with good management. For example, wheat yield potential for Pamunkey soil series (productivity group) is 64 bu/ac for standard wheat and 90 bu/ac with intensive management of wheat. But in the case farm where Pamunkey soil is one of the major soil types, no-till wheat double-cropped with soybean achieves an average yield around 86 bu/ac, well above the standard yield potential estimated by VALUES. Thus, yield with intensive management of wheat will be used for this study on the case farm.

N application rate for corn

Based on experimental and historical field data, VALUES develops a realistic yield expectation ("yield potential") for each soil management group (Simpson et al.). Simpson et al. assume that the amount of N needed to produce a given yield is 1 lb/bu (i.e., linear response)[27] and that sufficient residual soil N from the previous soybean crop is present in any given situation to produce at least 20 b/ac of corn. In this study, it is assumed that the net N over that needed to achieve the potential yield does not bring about a higher yield on average (thus the use of the linear response plateau model).

To calculate the net N applied to the corn crop (applied N that is not removed by the crop), the information in *Agronomy Handbook* (Virginia Cooperative Extension Service (VCES), 2000) is used. The N contained in every 150 bushels of corn grain at a moisture level of 15.5% is 135 pounds, or 0.90 of a pound of N in every bushel of grain. Every 100 pounds of corn grain contains 1.6 pounds of N. At a seeding rate of 83 pounds (VCES, 1999), the N input from seed is 1.3 lb/ac. Including N carryover from soybean (Simpson et al.) but not considering other N sources such as N from rainfall, the net N in corn production can be calculated as:

$$\max\left\{x-(yield\ \text{-}\ 20)*0.9+1.3,0\right\}=\max\left\{x+19.3-yield*0.9,0\right\} \qquad (3.2)$$

where x is the N applied and (yield - 20) is the yield gained from N application. For example, if the farmer applies 90 lb/ac N and achieves a yield of 110 bu/ac, then net N is equal to 10.3 lb/ac. When N is underapplied, it is assumed that a yield penalty of 1 bu/ac will result from 1 lb/ac N underapplied.

---

[27] This relationship, termed "N use efficiency", actually may vary greatly (e.g. 0.8 to 1.2 pound N applied per bushel of crop yield) depending on various factors such as weather, soil, and crop variety (Alley). One pound N for one bushel crop yield is an average value.

N application rate for wheat

The N application rate for wheat in Virginia is less related to soil productivity potential than is corn since the plant available water is often adequate to excessive during the fall and spring periods (Simpson et al.). The timing of the application is more important for wheat compared to corn especially for soils with high leaching potential (Simpson et al.). Current N application recommendations are mainly based on the nitrate soil test plus plant tissue N concentration and crop growth conditions (Simpson et al.).

According to Simpson et al., the N application for wheat is divided into three stages: at planting, midwinter, and spring. At planting, N is applied at 25 to 30 lb/ac when there is no previous manure application. In midwinter the crop is topdressed with N at 30 lb/ac. However, if only a single application of N is used, 80 lb/ac would be recommended if plants have less than 100 tillers/ft^2 or 30-40 lb/ac N if plants have more than 100 tillers/ft^2. When split applications are made, the first application may apply N at 40 to 60 lb/ac based on the number of tillers (60 lb/ac N for tillers more than 100 per square foot). The second application should be based on tissue test results.

In this study, other sources of N such as manure or carryover from previous crops are not considered in determining the rate of N application for wheat. Since every 80 bushels of wheat contains 100 pounds of N or 1.25 lb/bu (VCES, 2000), the sustainable rate of N application should be at least 100 pounds if the expected yield is 80 bu/ac. Assuming a seeding rate of 135 lb/ac (VCES, 1999), the N input from seed is 2.8 pounds. The formula to calculate net N (lb/ac) from wheat is as follows

$$\max\{x - yield * 1.25 + 2.8, \quad 0\} \tag{3.3}$$

Thus, if the target yield level is 86 bushels as on the case farm, the N application rate should be 1.25*86 = 107.5 pounds. Thus, when N is accurately applied according to the expected yield of 86 bushels, the inherent net N is 2.8 lb/ac, which equals that from seeds alone. The yield reduction from under-applying N is calculated as 1/1.25 or 0.8 of a bushel. The expected yields and N application required for the major soil types on the case farm are reported in Table 3-6.

Table 3-6. Expected yields and N requirements for the major soil types on the case farm[a]

Soil name	Soil #	Expected wheat yield (bu/ac)	N application for wheat (lb/ac)	Expected corn yield (bu/ac)	N application for corn (lb/ac)
Argent	148	30	38	65	45
Bolling	151	80	100	130	110
Emporia	161	70	88	120	100
Muckalee	168	30	38	65	45
Pamunkey	171, 172	90	113	160	140

a. Expected yields are based on VALUES (Simpson et al.) estimations of soil potential yields.

3.5. Delineation of smallest management units (SMUs)

The combine with a yield monitor covers a maximum width of 210 inches meters during harvest. The yield monitor generally records the quantity of grains harvested every second, covering a distance from 54 to 90 inches. Thus recorded yield data correspond to rectangular blocks of various sizes in the field. Blocks of these sizes are too small and irregular for meaningful SSM operations. Based on expert opinion (McClellan), experience of some researchers (e.g. Malzer et al., Snyder et al.), and the fact that the soil data resolution for the case study area is 1/9 hectare or larger, the size for a SMU is selected to be 98.425x98.425 ft^2 (30x30 m^2). A field is thus divided into 30×30m^2 (98.425x98.425 ft^2) grids, starting from the lower-left corner of the field.

The actual raw yield data obtained for this study were already corrected and converted into yields of the same moisture level for each crop (McClellan). After conversion of geographic projection and datum were done[28], raw yield data points were

---

[28] A freeware Windows program called CORPSCON (v5.11) was downloaded from www.tec.army.mil/TD/corpscon.html. This program was used by the U.S. military for conversion among several geographical projection and datum systems. The original data in the format *Geographic Coordinates, NAD27* (horizontal) in *decimal degrees* and *NGVD29* (vertical) in *the U.S. Survey feet* are converted into *UTM*

geo-referenced as easting-northing coordinates expressed in metric meters. Each point represents an area around 10 m^2 and represents the most precise yield information achievable for this study.

The field boundary dBASE file containing the points coordinates needed to generate the polygon of the field shape was imported into SAS. Then $x_1 = \min(X)$, $x_2 = \max(X)$, $y_1 = \min(Y)$, and $y_2 = \max(Y)$ were identified where X is the set of X-coordinates of the boundary points and Y is the set of Y-coordinates of the boundary points. Then starting from point (x_1, y_1), a matrix representing all 3×3 m^2 grids bounded within (x_1, y_1), (x_1, y_2), (x_2, y_2), (x_2, y_1) was generated. Each 3×3m^2 grid was represented by its lower-left corner coordinates and, being small enough, it is assumed to be wholly within the field or totally out of the field boundary. The matrix was saved as a dBASE file.

Using ArcView (ESRI), the matrix of 3×3 m^2 grids was imported and generated as a “theme” (in ArcView’s term). Then by “clipping” operation, all the 3×3 m^2 grids that are outside the field boundary were deleted from the field shape theme which is digitized from the dBASE file containing all the field boundary points. Soil type provided by the Virginia Department of Conservation and Recreation was assigned to all the 3×3 m^2 grids that were inside the field by a spatial data assignment. The matrix of 3×3m^2 grids that are inside the field boundary was saved as a dBASE file.

The next step was to assign observed yield data (all years and all crops) to each 3×3m^2 grid that is inside the field boundary. A yield was assigned to a 3×3m^2 grid if the yield point was located within a 3×3m^2 grid. This operation resulted in a data file that

---

Coordinates, NAD27 for *Zone 18* in metric meters (horizontal), and *NGVD29* (vertical) in metric meters. After conversion, longitude and latitude coordinates (degrees) are expressed in easting and northing X-Y coordinates (meters) and thus can be manipulated easily.

contains the most precise and rich information for the field. Each 3×3m^2 grid in the field contains soil type and crop yields from 1995 to 1999.

The next step was to generate the 30×30m^2 SMUs for the field from the 3×3m^2 grids. The procedures to generate the coordinates are similar to that for 3×3m^2 grids. That is, based on the minimal and maximal X and Y coordinates, a 30×30m^2 blank grid matrix was generated, starting from the lower-left corner. Then each 3×3m^2 grid was assigned to the 30×30m^2 which contains it. A 30×30m^2 grid was marked as outside the field and was deleted if no 3×3 m^2 grid was contained in it. The soil type for a 30×30m^2 grid was that of the largest group of 3×3m^2 grids that have the same soil type and are contained in the 30×30m^2 grid. Yield data for all crops in all years for a 30×30m^2 grid were obtained by simple average of that for the 3×3 m^2 grids contained in the 30×30m^2 grid. In the rare instances that a 30×30m^2 grid contains no 3×3m^2 grids that have observed yield data, missing yield data are replaced with that of a neighboring 30×30m^2 grid. All 30×30m^2 grids that have an area of less than 225m^2 (= 15×15) were deleted.[29] Thus, the complete raw data set for the field was transformed from 3×3m^2 grids to 30×30m^2 grids. The data manipulation described here was carried out using SAS programming.

Figure 3-2 gives a hypothetical portion of a field. As can be seen in the figure, SMU A is within the field but recorded yield are missing within A (this occurred for less than 1% of the data in this study). Yields for A were replaced by that of the grid to the left (F) or, if there are no grids to the left, the grid above it. The SMUs G, J, and K were

---

[29] These grids are among those that are located on the edge of the field. After deletion of these grids, the field acreage retained for further study could be slightly smaller than that indicated by field boundary data. The deletion simplifies further analysis since these grids are more likely to have missing yield data and estimating missing yields provides very little more information for the analysis.

deleted from further analysis for having too little area in the field. As shown in Table 3-1, the field acreage measured at 3×3 m^2 grid level and that measured at 30×30 m^2 grid level are very close. The size of SMU selected in this study is small enough not to distort the field shape.

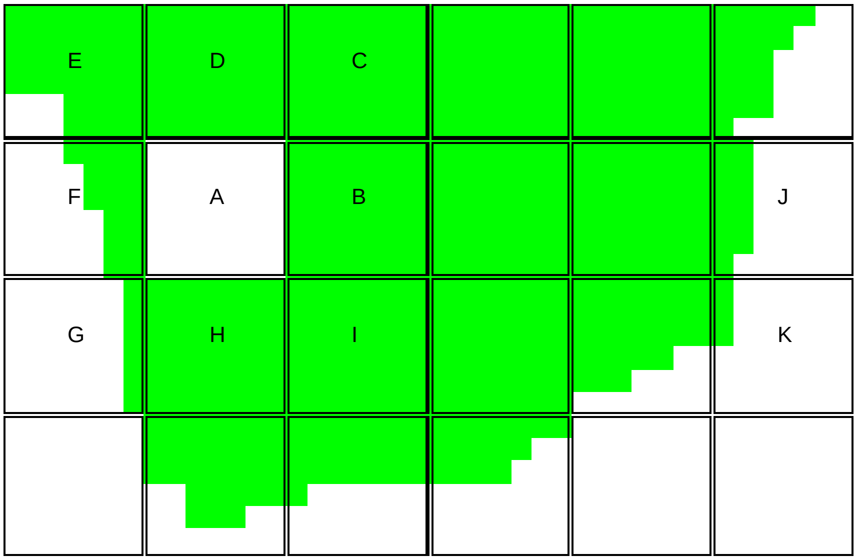


Figure 3.2. A hypothetical portion of the field location near the boundary

3.6. Optimizing ex post N applications

At each information level for each N application strategy, the optimal N rates are determined to give the farmer the highest expected net income. Assuming that the yield potentials of the fields in a specific year are not affected by N input, the optimal N rates are those that result in the lowest cost of N misapplication. The cost of N misapplication is the cost resulting from either overapplication of N or underapplication of N or both as in uniform application strategies. With the linear response and plateau production function used for this study, an optimal N rate is uniquely related to an optimal target yield and is uniquely related to the highest expected net return for the farmer.

In this section, optimal target yields and corresponding N rates for each N application strategy and crop are determined and expected cost of N misapplication to each field is estimated. This section also summarizes how the information values and the values of variable application are evaluated given the estimated optimal target yields.

Target yields and expected costs

Conventional strategy. For conventional strategy, the farmer uses only the VALUES recommendation (Simpson et al.) because Pamunkey is the dominant soil type in all fields. Therefore, the farmer always applies N at 113 lb/ac for wheat and 140 lb/ac for corn in all years on all fields. The optimal yields for the fields are equal to expected yields of Pamunkey (i.e. wheat, 90 bu/ac; corn, 160 bu/ac). For this strategy, expected cost from N misapplication is 0 because minor soils in the fields are not considered.[30]

Soil zone strategies. Soil zone N rates for variable application are listed in Table 3-6. The field-level total N applied for soil zone variable application is less than that of the conventional strategy for all fields because the minor soils have lower N application rates than the dominant soil. The expected cost from inaccurate N application is also 0 in this approach.

For soil zone uniform application, a search procedure for optimal (target) yield was carried out by using SAS DATA steps and PROC IML. Weighted average costs of inaccurate N application based on soil zone information were calculated for each suggested optimal yield. The yield having the minimal cost is the optimal target yield for the field in any year. Suggested target yields were evaluated from 1 bu/ac to 90 bu/ac for wheat and from 1 to 160 bu/ac for corn (VALUES yields for Pamunkey soil). Optimal yields are

---

[30] Expected cost of N misapplication is based on the farmer's belief about temporal and spatial yield variability while actual cost of N misapplication is based on yield monitor data. The expected cost and actual cost of N misapplication are generally different.

reported in Table 3-7. The optimal strategy always applies N according to the good predominant soils, Pamunkey soils, even though this strategy results in overapplication to the minor poor soils within the field. Thus, the conventional strategy and soil zone uniform strategy are identical. In contrast to the conventional strategy, there are positive costs expected from N overapplication in each field because soil variability is considered in estimating N misapplication costs.

Table 3-7. Optimal yields, N rate, and expected costs of N misapplication for soil zone uniform application with ex post analysis

Field	Crop	Optimal yield in any year (bu/ac)	Optimal N rate in any year (lb/ac)	Expected cost of N misapplication in any year ($/ac)
F1B	wheat	90	113	2.29
F1C	wheat	90	113	3.47
F2	wheat	90	113	0.66
F3B	wheat	90	113	1.10
F4	wheat	90	113	1.49
F5	wheat	90	113	1.70
F7	wheat	90	113	0.35
F8B	wheat	90	113	1.69
F1B	corn	160	140	3.56
F1C	corn	160	140	4.40
F2	corn	160	140	0.83
F3B	corn	160	140	2.32
F4	corn	160	140	1.88
F5	corn	160	140	2.42
F7	corn	160	140	0.64
F8B	corn	160	140	3.78

Functional zone strategies. When temporal and spatial transitions are not considered, functional zones can be generated each year based on observed yield recorded by the yield monitor. Thus, optimal yields and expected cost of N misapplication for functional zone strategies in ex post analysis are true yields and costs as reported in Chapter Four, Results and Discussion. For this reason, further discussion of the generation of functional zone and optimal N application for a field in each year is deferred to Chapter Four.

3.7. Overview of ex ante evaluation procedures

Ex ante analysis uses the Markov chain method to assess information values and variable application values of alternative strategies when observed temporal and spatial variability is utilized to improve N application for the future. Due to data availability, only the year 1999 is used to evaluate alternative strategies. The conventional strategy is the baseline for the ex ante analysis.

As discussed in Chapter Two, temporal variation of yield for a field can be obtained from historical yields or a simulation model based on weather data. Prediction of field-level yields can be based on the temporal variation pattern. The farmer can use predicted yield revealed by the field-level temporal variation pattern in combination with spatial variations to derive N applications for each information level. Information level is specified as following for each strategy.

Conventional strategy. The conventional strategy uses a Markov chain model to determine the future field-level yield distribution and thus the optimal N application rate based on the predominant soil. The Markov model was used rather than VALUES to estimate target yields to keep the conventional strategy consistent with other information levels so the differences in returns and net N would be due to differences in spatial information alone. The conventional strategy applies N according to the need of the predominant soil type (Pamunkey) in the field based on the Markov prediction of yield for that soil and the VALUES recommendation of 1 lb/bu minus carryover from soybean for corn and 1.25 lb/bu for wheat.

Soil zone strategies. Predicted field-level optimal yields for the coming year were used to predict the optimal yields for individual soil zones within the field. In order to estimate the soil zone optimal yield, it was assumed that relative yield variations from year to year are the same across soil types. If yield of soil A in 1998 in the field increases 10% from that of 1997, the yield of soil B in 1998 in the same field also increases by 10% from that of 1997. More discussion follows in the description of modeling field-level temporal variations.

Functional zone strategies. While ex post analysis uses current yield data to determine the optimal N application rate for the current year, ex ante analysis predicts future functional zone patterns based on the modeling of temporal and spatial variation for a given field and crop. The field-level Markov chain model predicts yields for the field and the within-field Markov chain model uses yield monitor data over several years to predict yields for the functional zones within a field.

Temporal variability was modeled first by generating a farm-level crop yield sequence using a crop yield simulation model with actual soil data and historical weather data. Then the farm-level yield sequence was translated into field-level yield sequences based on observed yields from 1992 to 1998. From the field-level yield sequence, the soil zone-level yield sequences were generated by rescaling the field-level sequences according to the soil distribution and relative yield potential for each soil type. Based on generated yield sequences, the temporal transitions of the field-level and soil zone-level yield were estimated by using the Markov model described in Chapter Two. The conventional strategy and the soil zone strategies then were evaluated.

For functional zone strategies, field-level temporal transition yield probabilities were used to predict the field-level yield for 1999. Yield monitor data in 1995 and 1997 were used to describe the within-field yield transition probabilities using another Markov chain model. Cluster analysis was used to group the SMUs into functional zones for 1999. After functional zones were established for each field, an optimization procedure was used to determine the optimal N application for each field in 1999.

Procedures for modeling ex ante temporal yield variability are described in Section 3.8. Procedures for modeling ex ante spatial yield variability are described in Section 3.9. Procedures for determining optimal N applications are described in Section 3.10.

3.8. Ex ante temporal variability

Field-level temporal variation

Several steps are involved to model field-level temporal variation. 1) Farm-level yields were simulated by using historical weather data, farm soil data, and farm management practices data. 2) Simulated farm-level yields were rescaled to the individual field-level. A different yield sequence was not simulated for each field due to lack of field-specific soil parameters required to run the crop simulation model. Rescaling was possible because several years' observed field-level yield data are available. 3) Based on the simulated yield sequence for a field, a transition probability matrix was estimated which was used to derive an optimal yield and associated N application for 1999. The weather year 1999 was used for the ex ante analysis.

EPIC model. EPIC, the Erosion-Productivity Impact Calculator (Williams and Renard)[31], was developed as a result of the Soil and Water Resources Conservation Act of 1977 and was used in this study to simulate the field-level temporal variation of yield. This model is a bio-physical-process-based calculator with weather, hydrology, erosion, nutrient cycling, pesticide fate, soil temperature, tillage, crop growth, and crop and soil management components (Williams and Renard). The input data include a management operation description for the site, crop parameters, tillage parameters, and pesticide parameters. Weather data, wind erosion data, and soil data for the site are needed input. The output data include crop yields, N loss in runoff, sub-lateral flow, leaching and with sediment, and other nutrient, pesticide, and sediment variables. The model has been tested widely with good performance (Williams and Renard).

For this study, daily weather data (maximum temperature, minimum temperature, and rainfall) were obtained from the National Climatic Data Center (www.ncdc.noaa.gov) for the period of 1922 to 1998. Since there is no national weather station in Prince George County, data from the nearest station to Prince George County at the International Airport of Richmond, Virginia were used. Long-term wind speed and direction data are contained in EPIC for Richmond station.

Soil data used for the EPIC simulation are based on the Pamunkey soil series since Pamunkey is the predominant soil on the farm. Soil data were provided by James Baker (Professor, Crop and Soil Environmental Science Department, Virginia Tech). Field management schedules (dates of planting and harvesting, and dates and amounts of N applications) are based on information provided by the farmer. Verification of simulation results was mainly based on comparison of simulated yields and observed farm-level yield for the years 1992-1998 and long term average yields.

Wheat yield sequences at farm-level. The simulated results for wheat are reported in Table B-1. A comparison of simulated wheat yields for the year 1992 to 1998 with farm-level actual wheat yield is reported in Table 3-8. Simulated yields are similar in year-to-year variation to the observed farm-level yields. Except for the year 1998, simulated yields

---

[31] EPIC is also called the Environmental Policy Integrated Climate in its recent version 8120.

are within 13% of the observed yields and on average the simulated yield is within 5% of the observed yield. Overall variation reflected in simulated yields is smaller than that in observed yields.

Table 3-8. Observed and simulated farm-level wheat yields

Year	Actual yield (bu/ac)[a]	Simulated yield (bu/ac)	Ratio (actual/simulated)
1992	84	97	0.87
1993	88	88	1.00
1994	83	75	1.11
1995	96	90	1.07
1996	58	62	0.94
1997	91	94	0.97
1998	54	80	0.68
Average	79	84	0.95
Std. dev.[b]	16.44	12.26	
CV[c]	20.8	14.6	

a. Based on total farm acreage of the crop for the given year which is larger than the totality of the acreage of the eight case fields. See Table 3-2.
b. Standard deviation.
c. Coefficient of variation expressed in percentage values.

The next step is to rescale the simulated farm-level yield data by multiplying 0.95 by each of the simulated yields of 1923-1998. The resultant yield sequence is plotted in Figure 3-3 with 1992-1998 yields being the observed farm-level yields.

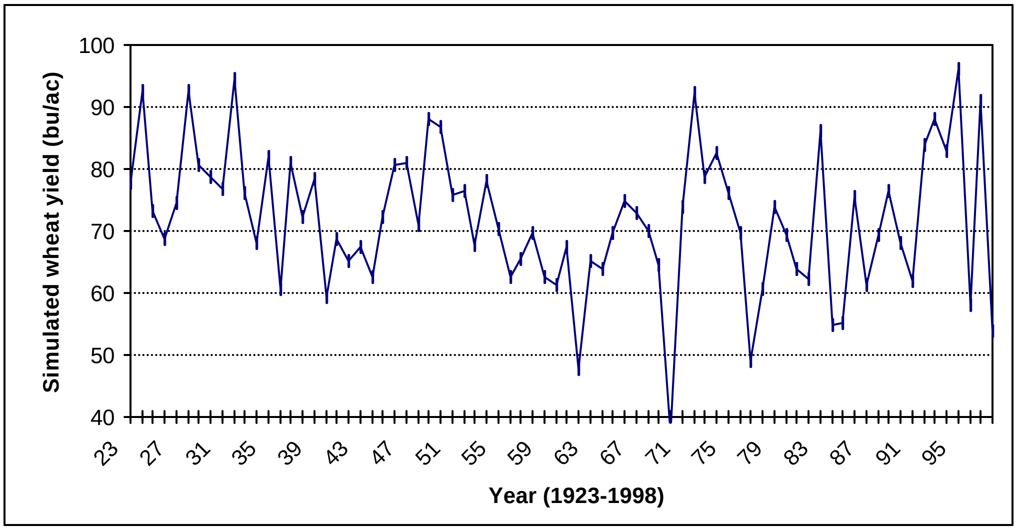


Figure 3-3. Simulated and farm-level wheat yield sequence (1923-1998)[a]

a. Yields from 1992 to 1998 are observed farm-level yields.

From the Figure 3-3, some temporal transition patterns can be seen at the farm-level. For example, if yield for current year is above 80 bu/ac, then in 8 out of 17 cases yield will be between 70 and 80 bu/ac next year.

Wheat yields sequences for individual fields. The simulated farm-level yields for 1923 to 1998 then are translated into a yield sequence for each individual field using information from Table 3-2. With limited information on historical field-level yields, the simple rule of scaling by ratio of average field-level yield and farm-level yield for the year 1992 to 1998 is used[32] when observed field yields vary similarly to that of observed farm-level yields. When an obvious dissimilarity is observed between field-level temporal yield variation and the farm-level temporal yield variation, refined scaling is used. For example, wheat yield for field F1B is 15% lower than the farm-level average according to observations on 1993, 1995, and 1997 and the temporal variation pattern is similar to that of farm-level yield, so the simulated sequence is multiplied by 0.85 to get the yield sequence for field F1B. However, for field F2 in bad years (1996 and 1998), average yield is 23% less than farm-level yield while in good years (1992 and 1994), it is only 7% lower. Thus for field F2, the simulated farm-level yield sequence was multiplied by 0.77 for simulated yields below 60 bu/ac and by 0.93 for other years.[33] The detailed scaling standards for each field are listed in Table B-2 of Appendix B.

Table 3-9 compares the scaled simulated yields and observed field-level yields for 1992-1998. After rescaling, the simulated yields for each field are close to observed field-

---

[32] This procedure implies that yields for all fields vary in the same direction temporally. This is a consistent assumption for this study and, of course, a farmer using this procedure would be able to make adjustment for individual fields which do not vary the same as the rest of the farm. For example, a field with low elevation may achieve higher than average yield when the rest of the fields on the farm are affected by drought and have lower than average yield.

[33] As can be seen, this procedure of translating simulated farm-level yield sequence into field-level is subjective and may over-use the observed data. On the other hand, the farmer is able to improve the accuracy of the sequence when more information is accumulated over the future years.

level yields from 1992 to 1998. With the exception of field F-2, average simulated and rescaled yields are within 7% of actual average yields. The complete field-level wheat yield sequences for individual fields are listed in Table B-1 in which simulated (and rescaled) yields are replaced by observed yield wherever available.

Table 3-9. Field comparisons: rescaled simulated versus actual wheat yields (bu/ac)[a]

	F1B		F1C		F2		F3B		F4		F5		F7		F8B	
Year	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated
1992					66	86							78	88		
1993	74	71	88	73			89	83	89	85						
1994					89	66					86	79	83	71	87	79
1995	85	73	86	76			93	86	101	89						
1996					44	45					62	65	66	75	65	65
1997	74	77	68	79			88	90	92	93						
1998					42	71					69	84	76	76	64	84
Average	78	73	81	76	60	67	90	86	94	89	72	76	76	78	72	76

a. For each field, comparisons are done only for the years when wheat was grown on the field. The numbers are rounded to the nearest integers.

Temporal wheat yield transition for individual fields. Good, normal, and bad states of nature were defined for each field. Descriptive statistics of a yield sequence were used to determine the cutoff values for each state. The descriptive statistics are listed in Table 3-10 and were generated by using SAS PROC UNIVARIATE. From Table 3-10, it is seen that the mean and the median for a field are very close. In all cases, yields between the 25% quantile and the 75% quantile are within one standard deviation from the mean. Thus, it was decided that a yield that is within 0.70 of one standard deviation (bu/ac) around the average is defined as the normal state.[34] The mean and standard deviation instead of median and quantile values are used because the former is more computationally convenient. Above the normal state is the good state and below the normal state is the bad state. The intervals for normal states were restricted to be narrow in order to increase the number of observations (simulated) in good and bad states so the estimation of the transition probabilities is more statistically sound. For example, if there is only one point in a good state, it will be hard to estimation the transition from good to other states of nature.

Table 3-10. Definition of states of nature and optimal yields for wheat at field-level[a]

Field	Yield (bu/ac)					Bad (bu/ac)		Normal (bu/ac)		Good (bu/ac)	
Name	Mean	Stc	25% quantile	Median	75% quantile	Range	% of observation	Range	% of observation	Range	% of observation
F1B	61	9.6	55	61	67	54 ≤	25	[55, 68]	54	≥ 69	21
F1C	64	10.1	57	63	70	56 ≤	25	[57, 71]	55	≥ 72	20
F2	66	12.2	60	66	73	59 ≤	22	[60, 74]	57	≥ 75	21
F3B	72	11.2	65	72	79	63 ≤	25	[64, 80]	53	≥ 81	22
F4	74	11.7	67	74	81	65 ≤	25	[66, 82]	54	≥ 83	21
F5	80	12.3	71	80	88	70 ≤	26	[71, 88]	51	≥ 89	22
F7	72	8.2	67	73	78	≤ 66	25	[67, 78]	53	≥ 79	22
F8B	80	12.4	71	80	88	70 ≤	26	[71, 88]	51	≥ 89	22

a. Yields are rounded to the nearest bushels. The descriptive statistics are based on the simulated yield sequence as listed in Table B-1.

---

[34] The number of 0.7 is chosen because for a normal distribution, about 50% of the values are within 0.7 standard deviation from the mean. A test for normality using PROC UNIVARIATE was carried out and results show that except for F2, normality was not rejected. This does not mean that normality is meaningful in this situation, but it does support the decision to choose the number 0.7. In cases where normality does not hold, 0.7 was used arbitrarily.

Based on the definitions of states of nature and the yield sequences generated for individual fields, the transition probability matrices were estimated and are listed in Table 3-11. The rows are for states of the current year, while the columns are for states of next year. A chi-square test was carried out using SAS PROC FREQ (SAS Inc.) to test if the probability of a future state of nature is independent of the current state of nature. This test for independence is equal to a test for the existence of a first-order Markov chain (Anderson and Goodman; Gbur and Steelman). The test results clearly fail to reject the hypothesis that there is no first-order Markov chain for any individual field. This result is not surprising since it had been pointed out by Simpson et al. that for wheat production in the study area, water availability is generally not a limiting factor. Instead, the timeliness of N application based on tissue-test is more related to yield levels. Thus, the estimated transition probabilities as listed in Table 3-11 were not used to calculate the optimal field-level yield for 1999. Instead, equal probability (=1/76) was assigned to each observed year from 1923 to 1998 in order to determine the field-level optimal yield for 1999.

Table 3-11. Estimated field-level transition matrices for wheat[a]

Field		Transition matrix			Field		Transition matrix		
F1B		Bad	Normal	Good	F4		Bad	Normal	Good
	Bad	0.263	0.526	0.211		Bad	0.263	0.526	0.211
	Normal	0.275	0.500	0.225		Normal	0.275	0.500	0.225
	Good	0.188	0.625	0.187		Good	0.188	0.625	0.188
		Test for independence: $X^2 = 0.7685$, p-value = 0.9426					Test for independence: $X^2 = 0.7685$, p-value= 0.9426		
F1C		Bad	Normal	Good	F5		Bad	Normal	Good
	Bad	0.263	0.579	0.158		Bad	0.263	0.526	0.211
	Normal	0.268	0.512	0.220		Normal	0.282	0.487	0.231
	Good	0.200	0.600	0.200		Good	0.235	0.529	0.235
		Test for independence: $X^2 = 0.6617$, p-value = 0.9560					Test for independence: $X^2 = 0.1896$, p-value = 0.9958		
F2		Bad	Normal	Good	F7		Bad	Normal	Good
	Bad	0.188	0.625	0.188		Bad	0.368	0.526	0.105
	Normal	0.233	0.558	0.209		Normal	0.256	0.538	0.205
	Good	0.250	0.500	0.250		Good	0.118	0.471	0.412
		Test for independence: $X^2 = 0.5313$, p-value = 0.9704					Test for independence: $X^2 =6.2220$, p-value = 0.1832		
F3B		Bad	Normal	Good	F8B		Bad	Normal	Good
	Bad	0.263	0.526	0.211		Bad	0.263	0.526	0.211
	Normal	0.282	0.487	0.231		Normal	0.282	0.487	0.231
	Good	0.176	0.588	0.235		Good	0.235	0.529	0.235
		Test for independence: $X^2 = 0.7954$, p-value = 0.9391					Test for independence: $X^2 = 0.1896$, p-value = 0.9958		

a. For each field, row headings correspond to yield states of nature in the current year and column headings show states of nature for next year. The chi-square test here is equal to testing the existence of the first-order Markov chain (Anderson and Goodman; Gbur and Steelman). Fisher's exact test was also used here due to the fact that more than 30% of the cells in some two-way 3x3 tables have counts less than 5. The results from chi-square test and Fisher's exact test are similar nevertheless. The statistics and p-values from Fisher's exact test are not reported in this table. Testing is carried out by using SAS PROC FREQ (SAS Inc.).

Corn yield sequences for individual fields. Table 3-12 reports a comparison of simulated corn yields for 1992 to 1998 with farm-level actual corn yields. The simulated yields tend to smooth out the high yields (e.g. in 1992 and 1995) and overpredict the low yields (1997 and 1998). However, the averages of the two sequences over 1992-1998 are very close. The comparison shows that while corn yields are sensitive to factors not considered in the simulation model, the simulation model predicts well the long run yield variation.

Table 3-12. Observed and simulated farm-level corn yields

Year	Actual yield (bu/ac)[a]	Simulated yield (bu/ac)	Ratio
1992	186	146	1.27
1993	107	103	1.04
1994	125	134	0.93
1995	178	128	1.39
1996	170	152	1.12
1997	107	140	0.76
1998	126	138	0.91
Average	135	134	1.007
Std. dev.[b]	34	16	
CV[c]	25	12	

a. Based on total farm acreage on the crop for the given year which is larger than the totality of the acreage of the eight case fields. See Table 3-2.
b. Standard deviation.
c. Coefficient of variation expression in percentage values.

The simulated farm-level corn yield sequence was rescaled to form a corn yield sequence for individual fields based on the comparison in Table 3-12. Simulated and rescaled farm and field-level yields are reported in Table B-3. Because the simulated farm-level corn yield sequence understates variability shown by actual farm-level yields from 1992 to 1998, the simulated farm-level yield sequence was first rescaled to farm-level before being rescaled to generate individual field sequences. For the first one or two years in which a yield above 145 was simulated, the simulated yield was multiplied by 1.15. If the third or fourth year was still over 140, the simulated yields were multiplied by 0.75. After the farm-level sequence was rescaled, a sequence then was rescaled from farm-level for each individual field. Rescaling for individual fields was based on the percent difference between the average yield for the field from 1992-1998 and the farm-level average yield (see Table 3-2). The rescaling is explained in detail in Table B-4. The generated field-level yield sequences are reported in Table 3-13 for the years when actual yields were reported (1992-1998). As shown by Table 3-13, the farm-level rescaling combined with rescaling for individual field results in very close field-level yield sequences compared with actual observations from 1992 to 1999. Actual and simulated yields were in the same state of nature 23 out of 29 times (79%). States of nature are defined in Table 3-14. When available, actual farm-level and field-level yields were used to replace the simulated as well as rescaled yields listed in Table B-3.

Table 3-13. Field comparisons: simulated versus actual corn yields (bu/ac)[a]

	F1B		F1C		F2		F3B		F4		F5		F7		F8B	
Year	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated
1992	183	175	175	180			186	179	196	197	180	175				
	(good)	(good)	(good)	(good)			(good)	(good)	(good)	(good)	(good)	(good)				
1993					81	93					98	101	119	118	131	116
					(bad)	(bad)					(bad)	(bad)	(bad)	(bad)	(normal)	(bad)
1994	104	118	106	121			111	120	123	133						
	(bad)	(normal)	(bad)	(normal)			(bad)	(normal)	(bad)	(normal)						
1995					118	121					177	167	188	196	182	192
					(normal)	(normal)					(good)	(good)	(good)	(good)	(good)	(good)
1996	160	160	191	165			162	163	179	180						
	(good)	(good)	(good)	(good)			(good)	(good)	(good)	(good)						
1997					105	93					87	101	125	118	111	116
					(normal)	(bad)					(bad)	(bad)	(bad)	(bad)	(bad)	(bad)
1998	125	118	119	122			124	121	145	134						
	(normal)	(normal)	(normal)	(normal)			(normal)	(normal)	(normal)	(normal)						
Average	143	143	148	147	101	102	146	146	161	161	136	136	144	144	141	141

a. For each field, comparisons were done only for the years when corn was grown on the field. The numbers are rounded to the nearest integers. “Good”, “normal”, and “bad” refer to the states of nature which are discussed later.

Temporal corn yield transition for individual fields. The rescaled farm-level corn yield sequence is plotted in Figure 3-4 with actual yields used for 1992 to 1998. The temporal dependence pattern is clear in this plot. A high yield is likely followed by a large drop of yields, low yields are likely followed by a large yield increase, while average yields may stay average, go up, or go down the following year.[35]

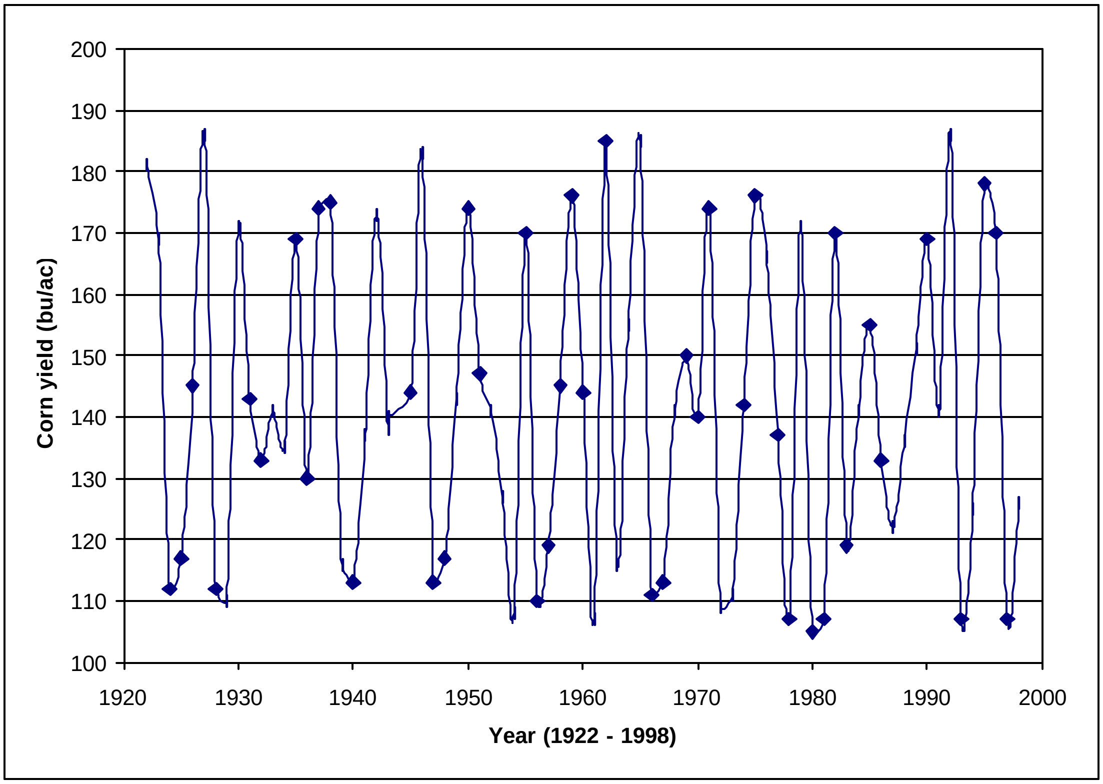


Figure 3-4. Simulated farm-level corn yield sequence (1923-1998)[a]

a. Simulated yields from 1992 to 1998 were replaced with observed yields.

Similar to wheat, 3 states of nature of corn yields are defined for each field. Based on the descriptive statistics reported in Table 3-14, the normal state is a yield equal to the mean plus/minus 0.8 of the standard deviation, which includes about 50% of the yields. The definition of normal state includes a larger deviation from the mean for corn compared to wheat because yield variation in corn is much larger than in wheat. Above the normal

---

[35] This time-dependent pattern is somewhat forced in the rescaling process. However, it is reasonable in the sense that the farmer does observe the time pattern from 1992-1998 by which he must calibrate the EPIC simulated yield sequence.

state is the good state, and below the normal state is the bad state. Table 3-14 lists the state definitions for each field. In Table 3-15, estimated field-level transition matrices are listed and results of the test for independence (equal to the test for the existence of first-order Markov chain) also are listed. The statistical test is carried out on the original 3×3 contingency table and the results support the use of Markov chain model for the field-level temporal variation for fields.

Table 3-14. Definition of states of nature for corn at the field-level[a]

Field	Yield (bu/ac)					Bad (bu/ac)		Normal (bu/ac)		Good (bu/ac)	
Name	Mean	Std	25% quantile	Median	75% quantile	Range	% of obs[b]	Range	% of obs[b]	Range	% of obs[b]
F1B	133	24.5	109	133	159	≤ 113	31.2	[114, 153]	39.0	≥ 154	29.9
F1C	138	25.5	113	137	164	≤ 116	31.2	[117, 158]	39.0	≥ 159	29.9
F2	111	10.4	102	109	116	≤ 94	28.6	[95, 126]	42.9	≥ 127	28.6
F3B	136	24.9	111	135	162	≤ 115	31.2	[116, 156]	39.0	≥ 157	29.9
F4	150	27.2	123	149	179	≤ 128	31.2	[129, 172]	39.0	≥ 173	29.9
F5	133	24.9	110	133	159	≤ 113	29.9	[114, 153]	40.3	≥ 154	29.9
F7	156	28.0	129	155	186	≤ 133	29.9	[134, 179]	40.3	≥ 180	29.9
F8B	153	27.5	126	152	183	≤ 130	29.9	[131, 175]	40.3	≥ 176	.29.9

a. The data are rounded to the nearest integers.

Table 3-15. Estimated field-level transition matrices for corn[a]

Field	Transition matrix				Field	Transition matrix			
F1B		Bad	Normal	Good	F4		Bad	Normal	Good
	Bad	0.375	0.375	0.250		Bad	0.375	0.375	0.250
	Normal	0.103	0.483	0.414		Normal	0.103	0.483	0.414
	Good	0.522	0.304	0.174		Good	0.522	0.304	0.174
	Test for independence: $X^2 = 11.304$, p-value =0.023					Test for independence: $X^2 = 11.304$, p-value=0.023			
F1C		Bad	Normal	Good	F5		Bad	Normal	Good
	Bad	0.375	0.375	0.250		Bad	0.348	0.435	0.217
	Normal	0.103	0.483	0.414		Normal	0.100	0.467	0.433
	Good	0.522	0.304	0.174		Good	0.522	0.304	0.174
	Test for independence: $X^2 = 11.304$, p-value = 0.023					Test for independence: $X^2 = 12.392$, p-value = 0.015			
F2		Bad	Normal	Good	F7		Bad	Normal	Good
	Bad	0.364	0.409	0.227		Bad	0.348	0.435	0.217
	Normal	0.094	0.500	0.406		Normal	0.100	0.467	0.433
	Good	0.500	0.364	0.136		Good	0.522	0.304	0.174
	Test for independence: $X^2 = 12.331$, p-value = 0.015					Test for independence: $X^2 = 12.392$, p-value = 0.015			
F3B		Bad	Normal	Good	F8B		Bad	Normal	Good
	Bad	0.375	0.375	0.250		Bad	0.348	0.435	0.217
	Normal	0.103	0.483	0.414		Normal	0.100	0.467	0.433
	Good	0.522	0.304	0.174		Good	0.522	0.304	0.174
	Test for independence: $X^2 = 11.304$, p-value = 0.023					Test for independence: $X^2 = 12.392$, p-value = 0.015			

a. The chi-square test on the 3×3 contingency table here is testing the existence of the first-order Markov chain (Anderson and Goodman; Gbur and Steelman). Testing is carried out using SAS PROC FREQ (SAS Inc.).

Temporal transition yields for soil zones

A temporal transition for soil zones is needed to evaluate the soil zone information strategy. Temporal variation was assumed to be the same for all soil types within each field. Given a temporal yield sequence at the field-level, a separate yield sequence was formed for a soil type simply by rescaling the field-level yield sequence based on relative expected yields suggested by VALUES. This scaling factor is calculated as

$$\frac{\text{Soil type expected yield by VALUES}}{\text{Field expected yield calculated from VALUES}}$$

where field expected yields calculated from VALUES can be found in Table 3-1, and soil type expected yields by VALUES can be found in Table 3-6. For example, F1B has an expected corn yield of 146 bu/ac (Table 3-1), and for Argent soil, the expected yield is 65 bu/ac (Table 3-6). Thus for Argent soil in field F1B, the rescaling factor is 65/146 = 0.445. Because field-level wheat yield transition does not follow the Markov chain model, soil zone yields were assumed not to follow a Markov Pattern. Equal probability was assigned to all points in the soil zone yield sequences. For corn yield, the same transition matrix applies to the field-level as well the soil zone level within the field. The soil zone rescaling factors are found in Section 3.10 where optimal yields for soil zone strategies are discussed.

Temporal transition yields for functional zones

For functional zones, the temporal transition at the field-level used for the conventional strategy was used to determine the expected field-level yield for 1999. However, the temporal transition at the field-level was not used to determine the optimal N application for the functional zone strategies. Instead, the observed SMU-level yields in 1995 and 1997 were used to determine the optimal N application in 1999 for the functional zone strategies. More discussion is found in the following sections.

3.9. Ex ante spatial variability

Data. The Markov chain model was used also to describe the yield transitions among the SMUs in a field. These SMU yield transitions were used to evaluate the functional zone uniform and variable application strategies. In this second Markov chain model (many objects observed over some small number of time periods), a minimum sample of two periods is needed. Because the equi-distant observations are needed to estimate the transition probabilities and because of the fact that corn and wheat were planted in rotation during 1995-1999 on the case farm, only 1995 and 1997 yield monitor data were used to estimate the within-field transition matrices. The estimated transitions were used to make N application decisions for each crop in 1999. That is, fields F1B, F1C, F3B, and F4 were used to evaluate SSM of N in 1999 for wheat because these fields were planted wheat in 1995, 1997, and 1999, while fields F2, F5, F7, and F8B were used to evaluate SSM of N on 1999 for corn because these fields were planted in corn in 1995, 1997, and 1999.

All observed SMU yield data in 1995 and 1997 were grouped into five clusters by using SAS PROC FASTCLUS (SAS Inc.). As discussed in Chapter Two, memberships of the SMUs are based on the closeness of observed yields among the SMUs within a group and differences between two groups.[36] Five clusters were chosen because the largest number of soil zones for a field is five (F1B). Each SMU within each cluster was assigned to one of the three states of nature according to its yield relative to field-level average. For wheat, a SMU is in good state if its yield was 0.7 times standard deviation above the field average yield, in bad state if its yield was 0.7 times standard deviation below the field

---

[36] In the terminology of clustering analysis, each object (SMU) has two observations (observed yields in 1995 and 1997).

average yield, and in normal state otherwise. For corn, 0.8 times standard deviation was used similarly to define the three states of nature.[37] In order to reflect the differences among the 5 clusters, SMUs with same state of nature in different clusters were treated differently as long as the within-field probability transition exists. Thus, there are 15 functional zones within each field if within-field probability transition exists, or five functional zones within each field if within-field probability transition does not exist. The area and mean yields in 1995 and 1997 of the five clusters are listed in Table 3-16. The statistical test results in Table 3-16 show that the clusters have significantly different means in each year as expected. For F7, the clusters 2 and 3 have similar yields in 1995, yet their 1997 yields are quite different.

---

[37] Note 0.7 and 0.8 times standard deviation used here were corresponding to those in temporal transition modeling. The uses of these two numbers were crudely justified in temporal transition modeling.

Table 3-16. Observed yield differences by functional zones for each individual field[a]

	Cluster	Wheat yield (bu/ac)			Cluster	Corn yields (bu/ac)	
Field	#	1995	1997	Field	#	1995	1997
F1B	1	41	23	F2	1	48	60
	2	65	32		2	107	56
	3	75	56		3	139	117
	4	85	74		4	161	138
	5	93	91		5	74	101
	p-value =	0.0001	0.0001		p-value =	0.0001	0.0001
F1C	1	84	65	F5	1	96	31
	2	66	24		2	168	41
	3	77	49		3	195	79
	4	89	76		4	182	114
	5	93	82		5	165	94
	p-value =	0.0001	0.0001		p-value =	0.0001	0.0001
F3B	1	93	88	F7	1	114	31
	2	64	50		2	164	83
	3	88	44		3	165	126
	4	86	72		4	191	122
	5	97	105		5	109	141
	p-value =	0.0001	0.0001		p-value =	0.0001	0.0001
F4	1	50	17	F8	1	112	33
	2	61	47		2	178	50
	3	75	72		3	184	98
	4	96	79		4	194	124
	5	106	98		5	180	138
	p-value =	0.0001	0.0001		p-value =	0.0001	0.0001

a. P-values are from ANOVA by using SAS PROC GLM (SAS Inc.). The p-values indicate the significance of differences among the clusters. Cluster numbers are directly from output of SAS PROC FASTCLUS and are not ordered in terms of cluster means.

Within-field yield transition. Similar to temporal transition states of nature, three states of nature for within-field yield transitions among the SMUs are defined for wheat: yield that is at least 0.7 times the standard deviation above the cluster mean yield is the good state, yield that is at least 0.7 times the standard deviation below the cluster mean yield is the bad state, and the rest is the normal state. For corn, three states of nature are defined similarly using 0.8 (instead of 0.7) times the standard deviation. The numbers 0.7 and 0.8 are adopted directly from those used in temporal transition estimation.

After states of nature was defined for every cluster and thus every SMU within a given field, the transition probabilities were estimated by assuming that the same transition matrix applies to all the SMUs within a field. For example, if *n* SMUs were in state A in 1995 across all clusters and *m* of these SMUs change to state B in 1997 across all clusters, then the estimated transition probability from state A to B is *m/n*.

Before the estimation of within-field transition, the observed yield data of all SMUs in 1995 and 1997 for these eight fields were rescaled to the same respective expected yield levels of 1999 as predicted by the field-level temporal Markov chain model. This rescaling is necessary since the within-field Markov chain describes only relative yields. In addition, the rescaling was necessary to make the functional zone strategies consistent with the soil zone strategies and the conventional strategy because the rescaled yields in 1995 and 1997 were used both in estimating within field transition matrices and the prediction of the optimal yields for 1999. Table 3-17 reports the rescaling standards at the field-level. The yields of the SMUs were multiplied by the rescaling factors listed in Table 3-17. For example, if a SMU in F1B has wheat yields of 85 bu/ac in 1995 and 74 bu/ac in 1997, rescaled yields are 85*0.72 bu/ac for year 1 (1995) and 74*0.82 bu/ac for year 2 (1997).

Table 3-17. Rescaling factors for individual fields

	1999 expected yield	1995 yield		1997 yield	
Field	(bu/ac)	Actual (bu/ac)	Rescaling factor[a]	Actual (bu/ac)	Rescaling factor
	---------- Wheat ----------				
F1B	61	85	0.72	74	0.82
F1C	64	86	0.74	68	0.94
F3B	72	93	0.77	88	0.82
F4	74	101	0.73	92	0.80
	---------- Corn ----------				
F2	118	118	1.00	105	1.12
F5	143	177	0.81	87	1.64
F7	167	188	0.89	125	1.34
F8B	164	182	0.90	111	1.48

a. The yield of each SMU in the field will be multiplied by the rescaling factor to estimate the transition probabilities.

The estimated within-field transition matrices are listed in Table 3-18. The Markov transition is significant in only two fields, namely, F4 and F5. Thus, Markov transition matrices will be estimated only for these two fields. For these two fields, the normal state has the highest probability of occurring in the next period. For F4 (wheat), given the current year state of nature is bad, the probability of the next period being bad is larger than the probability that the next state will be good. For F5 (corn), it is the opposite. This difference between F4 and F5 probably means that wheat yield within-field variability pattern is less affected by weather than that of corn.

For the fields with no within-field transition pattern, the number of functional zones equals the number of clusters. Each cluster of SMUs contains SMUs with similar temporal variation patterns. For these fields, the distribution of future states of nature is the same regardless of the current state of nature for a SMU. In these fields, the distribution of future states of nature was estimated by pooling all SMUs together. For example, if there are a total of N_1 and N_2 SMUs that are in normal state in 1995 and 1997, respectively, and the total number of SMUs in the field is M, then the probability that a SMU in 1999 is in the normal state is $(N_1+N_2)/2M$.

One possible explanation for the non-existence of a within-field Markov transition for most of the fields is that relative yields are rather stable for a SMU in a field. This may happen when the factors that affect yield interact to keep yield variation patterns stable. For example, two SMUs with different elevations may have different soil N levels with the higher SMU having higher N (Hollands). Thus, in a dry year, the lower elevation SMU has potential for higher yield but N is limiting as compared with the higher SMU. Another possible explanation is that there are too few years of data available to reveal trends for individual sites within fields.

Table 3-18. Estimated within-field transition matrices for each field[ab]

Wheat field		Transition matrix			Corn field		Transition matrix		
F4		Bad	Normal	Good	F5		Bad	Normal	Good
	Bad	0.205	0.638	0.157		Bad	0.108	0.595	0.297
	Normal	0.196	0.477	0.327		Normal	0.215	0.511	0.274
	Good	0.132	0.629	0.240		Good	0.264	0.642	0.094
		p-value < 0.001					p-value =0.029		
F1C		p-value = 0.433			F2		p-value = 0.833		
F3B		p-value =0.745			F7		p-value =0.064		
F1B		p-value = 0.811			F8B		p-value = 0.642		

a. Fields F1B, F1C, F3B, and F4 are for wheat. The Fields F2, F5, F7, and F8B are for corn.

b. p-value is from Fisher's exact test. The Fisher's exact test on the 3×3 contingency table here is testing the existence of the first-order Markov chain (Anderson and Goodman; Gbur and Steelman). Testing is carried out using SAS PROC FREQ (SAS Inc.).

3.10. Optimizing ex ante N applications

In this section, optimal target yields and corresponding N rate for each ex ante N application strategy are described. Procedures for estimating the expected cost of inaccurate N application to each field are shown. This section also summarizes how the information values and the values of variable application are evaluated given the estimated optimal target yields. The expected costs of N misapplication were based on the knowledge available to the farmer at each given information level. Expected costs of N misapplication may be different from the actual costs. This difference is seen in the comparison of expected values presented below with the actual results presented in Chapter Four. SAS PROC IML (SAS Inc.) was used to program the search procedure to estimate the optimal field-level target yield for 1999 which minimizes costs of N misapplication for each strategy as described in Chapter 2 (equation 2.17.1). The SAS codes can be found in Appendix D.

Target yields and expected costs with ex ante analysis

Conventional strategy. The field-level target yield, expected yield, and expected cost from N application for the target yield for wheat in 1999 for the conventional strategy are listed in Table 3-19.

Table 3-19. Field-level optimal yields, expected yields, N rates, and expected costs of N misapplication in 1999 for conventional strategy

Field	Crop	Expected yield 1999 (bu/ac)	1999 optimal yield (bu/ac)	N rate for 1999 (lb/ac)	Expected cost of N misapplication ($/ac)
F1B	Wheat	61	74	93	5.21
F1C	Wheat	64	77	96	5.76
F3B	Wheat	72	88	110	6.03
F4	Wheat	74	90	113	6.45
F2	Corn	118	137	117	6.14
F5	Corn	143	170	150	7.82
F7	Corn	167	194	174	8.62
F8B	Corn	164	190	170	8.43

Because equal probability was assigned to each wheat yield simulated 1923-1998, the field-level expected yields for 1999 are identical to the long-term average yields. Because transition probabilities were assigned to corn yields, expected yields are different and higher than long-term averages. For every field the optimal yield for 1999 is clearly higher than the respective expected yield for a profit-maximizing farmer because the forgone yield of one bushel from underapplication of N results in much larger income loss than the cost from one pound N overapplication. This result is consistent with the observed farmer's behavior for the case farm. For wheat, only on F4 is N applied at the level recommended by VALUES (Simpson et al.) for Pamunkey soil (113 lb/ac) while other fields apply N at lower rates. For corn, optimal yields are higher in F5, F7 and F8B than the VALUES assumed yields for Pamunkey soil while the optimal yield for F2 is lower than the VALUES assumed yield for Pamunkey soil, which is 160 bu/ac.

For example, the expected cost of N misapplication was calculated for F1B as follows. F1B has a yield sequence for wheat from 1923 to 1998 shown in Table B-2. If N is applied at 93 lb/ac for 1999 and the yield predicted for 1923 (66 bu/ac) occurs in 1999, then N will be overapplied by (optimal yield – yield in 1923)*1.25 = (74 – 66)*1.25 = 10 lb/ac and there will be a cost of 10*0.25 = $2.50/ac for wasted N input. If the yield for 1995 occurs in 1999, then N will be underapplied by (yield in 1995 – optimal yield)*1.25 = (85 – 74)*1.25 = 13.75 lb/ac and the resulting cost from yield loss = (85 - 74)*2.87 = $31.57/acre while the saving from less N input =13.75*0.25 = $3.44/acre, and the saving from less crop hauling = (85 – 74)*0.15 = $1.65/acre. So if the year 1995 repeats in 1999 the cost from N misapplication will be 31.57 – 3.44 – 1.65 =$26.48/acre. The costs are calculated similarly for other years and the simple average of these costs gives the value of $5.21/acre as listed in Table 3-19. The simple average is used because there is no transition probability for wheat in F1B.

Soil zone variable strategy.[38] Because field-level wheat yields do not follow a Markov chain, for each field soil zone expected yields and optimal yields as well as expected costs of inaccurate N application for 1999 can be directly derived from Table 3-19 by using soil rescaling factors discussed in Section 3.8 and shown in Table 3-20.[39] For example, Argent soil has a rescaling factor of 0.361 in field F1B for wheat. Consequently, its expected wheat yield is 22 bu/ac and the optimal yield is 27 bu/ac for F1B. The results are listed in Table 3-20. As expected due to the influence of other site factors, the same soil in different fields may have different expected yields, optimal yields, N rates, and expected costs of N misapplication. Costs of N misapplication are positive due to under- and over-application.

---

[38] Variable application strategy at a given information level is discussed before the uniform strategy because the algorithm of the latter is based on that of the former, as clearly shown in Chapter Two.

[39] The results are verified by using the same SAS program that generated results in Tables 3-19.

Table 3-20. Soil zone expected yields, optimal yields, N rates, and expected costs of N misapplication for wheat in 1999 for each field under variable application

Field	Soil zone[a]	Soil zone rescaling factor	Expected yield 1999 (bu/ac)	1999 optimal yield (bu/ac)	Optimal N rate for 1999 (lb/ac)	Expected cost of N misapplication ($/ac)
F1B	Argent	0.361	22	27	34	1.88
	Bolling	0.964	59	71	90	5.02
	Muckalee	0.361	22	27	34	1.88
	Pamunkey1	1.084	66	80	101	5.65
	Pamunkey2	1.084	66	80	101	5.65
F1C	Argent	0.380	24	29	36	2.19
	Pamunkey1	1.139	73	88	109	6.56
	Pamunkey2	1.139	73	88	109	6.56
F3B	Bolling	0.930	67	82	102	5.61
	Muckalee	0.349	25	31	38	2.10
	Pamunkey1	1.047	75	92	115	6.31
	Pamunkey2	1.047	75	92	115	6.31
F4	Argent	0.353	26	32	40	2.28
	Pamunkey1	1.059	78	95	120	6.83
	Pamunkey2	1.059	78	95	120	6.83

a. Soil identification numbers are Argent (148), Bolling (151), Emporia (161), Muckalee (168), Pamunkey1 (171), and Pamunkey2 (172).

The procedure that determined field-level corn optimal yield and expected cost is used for each soil zone for corn since it is assumed that all soil zones within a field follow the same field-level transition. The results in Table 3-21 show that the same soil type in different fields may have different expected yields, optimal yields, N rates, and expected costs of inaccurate N application. Generally for both wheat and corn, better soil zones have higher costs of inaccurate N application from temporal variation. Even though good soil zones can achieve much higher yields than poor soils zones in normal and good years, their yields are similar to the poor zones in bad years.

Table 3-21. Soil zone expected yields, optimal yields, N rates, and expected costs of N misapplication for corn in 1999 for each field under variable application

Field	Soil zone[a]	Soil zone rescaling factor	Expected yield 1999 (bu/ac)	1999 optimal yield (bu/ac)	Optimal N rate for 1999 (lb/ac)	Expected cost of N misapplication ($/ac)
F2	Muckalee	0.417	49	57	37	2.56
	Pamunkey1	1.026	121	141	121	6.30
	Pamunkey2	1.026	121	141	121	6.30
F5	Argent	0.433	62	74	54	3.39
	Bolling	0.867	124	147	127	6.78
	Pamunkey1	1.067	153	181	161	8.35
	Pamunkey2	1.067	153	181	161	8.35
F7	Bolling	0.828	139	161	141	7.14
	Emporia	0.764	128	148	128	6.59
	Pamunkey1	1.019	171	198	178	8.79
	Pamunkey2	1.019	171	198	178	8.79
F8B	Argent	0.448	74	85	65	3.78
	Bolling	0.897	147	170	150	7.56
	Pamunkey1	1.103	181	210	190	9.30
	Pamunkey2	1.103	181	210	190	9.30

a. Soil identification numbers are Argent (148), Bolling (151), Emporia (161), Muckalee (168), Pamunkey1 (171), and Pamunkey2 (172).

The field-level results for expected yields, optimal yields, N application rates, and expected cost of N misapplication for the soil zone variable strategy were the same as for the conventional strategy because of the way the soil zone yield sequences were generated. Thus, no additional table is given here for the field-level results for the soil zone variable strategy.

Soil zone uniform strategy. Given above derived yield sequences and estimated transition matrices, a search procedure was developed using SAS PROC IML to calculate the field-level predicted optimal yields and related costs of N misapplication in 1999. This procedure was similar to that used for the conventional strategy. A search was conducted over all yields between 30 to 160 bu/ac at one-bushel intervals and the yield which resulted in the smallest weighted average cost of N misapplication was the optimal yield for the field. The predicted optimal yields and expected costs are reported in Table 3-22. It should be noted that expected field-level yields in 1999 are the same as listed in Table 3-19.

Table 3-22. Soil zone, optimal yields, N rates, and expected costs of N misapplication for corn and wheat in 1999 for each field under uniform application

Field	Crop	1999 optimal yield (bu/ac)	N rate for 1999 (lb/ac)	Expected cost of N misapplication ($/ac)
F1B	Wheat	76	95	6.87
F1C	Wheat	82	103	8.86
F3B	Wheat	87	109	6.65
F4	Wheat	94	118	7.99
F2	Corn	141	121	6.96
F5	Corn	176	156	10.29
F7	Corn	198	178	9.34
F8B	Corn	206	186	12.32

The optimal target yields for 1999 are higher than that of the conventional strategy for the same field (see Table 3-19) indicating that focusing on good and dominating soil zones is more profitable even though large overapplication of N can happen to the poor soil zone within the field.[40] However, for each field, the optimal yield under uniform application is less than or equal to that of best soil zone optimal yield under variable application (Tables 3-20, and 3-21) because of the cost from large overapplication on poor soil zones.

Functional zone variable strategy. After the within-field transition matrix was estimated, a distribution of states of nature for 1999 was determined based on the actual state for 1997 of a given SMU.[41] Then the observed yields for all the SMUs of the cluster to which the specific SMU belongs in 1995 and 1997, after being rescaled to the level of expected yield in 1999, were pooled together. The search procedure will determine optimal

---

[40] The rescaling procedure used to recover soil zone yields from field yields led to higher yields for Pamunkey under the soil zone strategy than under the conventional strategy.

[41] For fields where Markov transition is not significant, the distribution of the future states of nature are the same for all SMUs in all functional zones that are in a given state of nature in the current year.

yield in 1999 for a SMU in this cluster using the pooled yields given the state of nature of this SMU in 1997.

The predicted optimal yields and related costs of N misapplication are listed in Table 3-23. In Table 3-23, each row in a field is a functional zone. From Table 3-23, it is seen that for a given crop, different functional zones have different expected yields, optimal yields, and expected costs. For F4 and F5 where within-field transition is present, the within field transition has only minor effects on optimal yields and N rate in F4. F5 has lower expected and optimal yields in 1999 for SMUs which were in good state in 1997 than those of SMUs which were in bad or normal states in 1997 because, according to transition probabilities for F5 reported in Table 3-18, the SMUs which are currently in good states are much more likely shift to a bad state next year than staying in a good state.

Table 3-23. Optimal yield and expected cost for functional zone variable strategy in each field in 1999[a]

Wheat field	Cluster #	State of nature in 1997[b]	Exp. wheat yield in 1999	Optimal wheat yield in 1999	N rate needed	Exp. cost of N misappl.	Corn field	Cluster #	State of nature in 1997[b]	Exp. corn yield in 1999	Optimal corn yield in 1999	N rate needed	Exp. cost of N misappl.
F1B	1	1, 2, or 3	19	33	41	2.45	F2	1	1, 2, or 3	55	81	61	8.96
	2	1, 2, or 3	37	52	65	5.71		2	1, 2, or 3	84	120	100	12.39
	3	1, 2, or 3	51	60	75	3.75		3	1, 2, or 3	135	148	128	4.54
	4	1, 2, or 3	61	67	84	2.24		4	1, 2, or 3	159	166	146	2.43
	5	1, 2, or 3	71	78	98	3.00		5	1, 2, or 3	95	119	99	7.72
F1C	1	1, 2, or 3	62	67	84	1.89	F7	1	1, 2, or 3	73	126	106	13.96
	2	1, 2, or 3	35	51	64	6.19		2	1, 2, or 3	128	163	143	11.28
	3	1, 2, or 3	51	63	79	3.93		3	1, 2, or 3	157	175	155	5.48
	4	1, 2, or 3	69	73	91	1.56		4	1, 2, or 3	166	176	156	3.65
	5	1, 2, or 3	73	79	99	2.75		5	1, 2, or 3	187	198	178	3.64
F3B	1	1, 2, or 3	72	76	95	1.52	F8B	1	1, 2, or 3	75	124	104	13.22
	2	1, 2, or 3	45	55	69	3.44		2	1, 2, or 3	119	169	149	14.03
	3	1, 2, or 3	52	72	90	6.97		3	1, 2, or 3	156	175	155	5.79
	4	1, 2, or 3	63	74	93	4.06		4	1, 2, or 3	179	191	171	4.62
	5	1, 2, or 3	81	89	111	3.56		5	1, 2, or 3	183	210	190	7.85
F4	1	1	24	40	50	6.56	F5	1	1	66	107	87	20.71
		2	25	40	50	6.22			2	63	104	84	20.71
		3	25	40	50	6.55			3	53	98	78	16.17
	2	1	40	50	63	3.52		2	1	104	148	128	12.21
		2	41	50	63	3.20			2	101	148	128	12.81
		3	41	50	63	3.26			3	97	149	129	13.98
	3	1	56	62	78	2.84		3	1	148	163	143	4.15
		2	57	63	79	3.26			2	145	163	143	4.72
		3	57	63	79	2.95			3	142	162	142	5.45
	4	1	67	75	94	3.32		4	1	169	204	184	11.95
		2	67	74	93	3.02			2	168	204	184	11.89
		3	67	75	94	3.07			3	166	189	169	8.84
	5	1	78	83	104	2.02		5	1	145	168	148	6.52
		2	78	84	105	2.08			2	143	168	148	6.84
		3	78	83	104	1.99			3	142	156	136	5.36

a. Unit of measure for yield is bushels per acre; for N rate is pounds per acre; and for cost is dollars per acre.

b. State of nature 1 is bad, 2 is normal, and 3 is good.

The expected and optimal yields for each field with the functional zone variable strategy are listed in Table 3-24. The expected yields listed in Table 3-24 are weighted averages for each functional zone listed in Table 3-23. As compared with the conventional strategy (Table 3-19), the expected yields and optimal yields with functional zone information for wheat and corn are lower. Expected costs are lower in most cases as compared with the conventional strategy, indicating that the functional zone variable strategy is expected to be more able to reduce inaccurate N applications. Yet, the expected costs are still positive due to the stochasticity of future yield levels.

Table 3-24. Field-level 1999 expected and optimal yield under functional zone variable strategy

Field	Crop	Expected yield (bu/ac)	Optimal yield (bu/ac)	Expected cost of N misapplication ($/ac)
F1B	Wheat	61	69	3.17
F1C	Wheat	64	70	2.54
F3B	Wheat	72	81	3.44
F4	Wheat	74	80	2.34
F2	Corn	117	134	5.92
F5	Corn	143	174	10.11
F7	Corn	167	181	4.68
F8B	Corn	164	190	7.70

a. All weighted averages by area for each soil type in each state of nature in 1997.

Functional zone uniform strategy. The same procedure that determined optimal yield and expected cost for uniform soil zone application is adopted to determine the optimal yields and expected costs for uniform functional zone application. The predicted 1999 optimal yields and expected costs of inaccurate N application for each field based on functional zone information are listed in Table 3-25.

In most cases (F1B, F1C, F2, F5, F7, and F8B), optimal yields for uniform application are equal to or higher than those for variable application, indicating that the farmer will benefit by applying higher N rate on good zones even though this brings larger overapplication for poor zones. This pattern is somewhat different from that of soil zone strategies where optimal yield for uniform application is less or equal to that of variable application. This change of pattern probably indicates that the yield monitor reveals some areas where consistently very high yields are observed. In all cases, uniform application with functional zone information brings higher expected cost of inaccurate N application than the variable application strategy.

Table 3-25. Field-level 1999 optimal wheat and corn yield and N rate under functional zone uniform application[a]

Field	Crop in 1999	Field-level optimal yield for 1999 (bu/ac)	Field-level expected cost of N misapplication ($/ac)
F1B	Wheat	75	5.94
F1C	Wheat	74	4.63
F3B	Wheat	86	6.34
F4	Wheat	81	3.52
F2	Corn	159	12.13
F5	Corn	182	16.29
F7	Corn	187	8.71
F8B	Corn	201	13.41

a. All weighted averages by area for each soil type in each state of nature in 1997.

3.11. Summary of the chapter and limitations

Summary. This chapter described the study area, the case farm, the case fields, the extension recommendation related to N management in corn and wheat, and related prices and production costs. It also described how the soil-zone information and the functional-zone information were generated for the ex ante and ex post evaluation procedures and what the costs of the information were for the case farm. The chapter described how the Markov chain models were used to characterize the field-level temporal yield variation and the within-field yield variability for the case farm.

With VALUES recommendations and the information of predominant soil types within the case fields, the conventional strategy can be established to evaluate information values and the values of variable rate application of the alternative strategies, as presented in Table 3-26. Based on the empirical model developed in this chapter, both ex post analyses and ex ante analyses can be carried out. Ex post analyses were carried out based on a known pattern of yields for all fields in each year from 1995 to 1999. Ex ante analyses consider both temporal and spatial yield transitions and were carried out for 1999.

Table 3-26. Description of strategies, information levels, and purpose of analyses of the empirical model

Strategy	Information	Purpose
Ex post Conventional strategy	Production cost, VALUES,[a] predominant soil types for a field	Baseline strategy to evaluate information values and the values of variable rate application in alternative strategies when yield patterns are known
Soil zone strategies	Production cost, soil map, VALUES	To see if information in current soil maps can have positive values in SSM of N when yield patterns are known.
Functional zone strategies	Production cost, soil map, VALUES, and yield monitor recordings.	To estimate the value of information regarding spatial yield patterns within field when yield patterns are known
Ex ante: Conventional with field-level transition pattern identified	Information for conventional strategy plus farm-level yield transition information	Baseline strategy to evaluate information values and the values of variable rate application under uncertainty
Soil zone strategies	Soil information described under ex post strategy plus soil zone-level temporal transition information	To estimate the value of soil map and historical field yield information when yield patterns are unknown
Functional zone strategies	Functional zone information described under ex post strategy plus functional zone temporal transition information	To estimate the value of information provided by yield-monitor plus soil map plus historical field yield when within-field yield patterns are unknown

a. VALUES refers to yield goals and N application rates by soil type as recommended by Virginia Cooperative Extension Service (Simpson et al.)

Study limitations. First, the EPIC-simulated farm-level yield sequence over 1922-1998 does not fully reflect farm yield variability for corn and wheat. Second, errors were introduced in translating simulated farm level yields to field yields which may have lowered the accuracy of the field yield transition probabilities. Third, the errors in constructing field transition probabilities may have affected the within-field yield transition probabilities and optimal N application amounts. The assumption that all SMUs within the same field follow the same transition matrix may not always hold. However, the objective of the study is to rank SSM strategies for N applications in terms of net N and farm profits. Thus, the errors described above do not bias the results in favor of one strategy vs. another since all levels of information should be affected similarly. The information values and the values of variable application for the alternative N management strategies will be presented in Chapter Four.

Chapter 4. Results and Discussion

4.0 Overview

This chapter reports study results in the following order: Section 4.1 reports the observed yield variability by field, soil zone, and functional zone. Section 4.2 and Section 4.3 report results of the ex post analyses on net N and values of information and variable application. Yield data recorded by yield monitors from 1995 to 1999 are used to evaluate N application strategies. Section 4.4 and Section 4.5 report results of ex-ante analyses of net N and net returns from N application strategies on wheat (fields F1B, F1C, F3B, and F4) and corn (field F2, F5, F7, and F8B) in 1999. Section 4.6 reports results of a sensitivity analysis which removes the temporal variation from the model and concentrates on spatial information in 1999. Section 4.7 summarizes the major results of this chapter.

4.1 Variability of yield across fields, soil zones, and functional zones

Field-level yield variability

Field-level yields for each field from 1995 to 1999 are reported in Table 4-1. The yield level of wheat and yield level of corn seem to be unrelated over years. For example, in 1995, both corn and wheat were high for most fields, but in 1996, corn yields were high but wheat yields were low. Most wheat yields were high in 1997 and 1999 but corn yields were low. And in 1998, both wheat and corn yields were low. This observation indicates that it is reasonable to separate the study of yield variability in the wheat and corn.

Table 4-1. Observed within-field wheat and corn yield variability for the case fields[a]

Field	Mean	St.dev	CV[b]	Mean	St.dev	CV[b]	Mean	St.dev	CV[b]	Mean	St.dev	CV[b]	Mean	St.dev	CV[b]
	--- wheat in 1995 ---			--- wheat in 1996 ---			--- wheat in 1997 ---			--- wheat in 1998 ---			--- wheat in 1999 ---		
F1B	84	13	15				72	21	29				44	18	41
F1C	85	10	11				66	18	27				91	21	23
F3B	91	10	11				86	22	26				93	11	12
F4	100	14	14				90	17	19				86	19	22
F2				44	15	34				41	13	32			
F5				61	14	23				68	9	14			
F7				66	13	20				75	14	19			
F8B				64	13	21				63	15	23			
	--- corn in 1995 ---			--- corn in 1996 ---			--- corn in 1997 ---			--- corn in 1998 ---			--- corn in 1999 ---		
F1B				156	34	22				120	35	29			
F1C				187	25	14				118	28	24			
F3B				160	21	13				122	24	20			
F4				177	25	14				142	30	21			
F2	112	45	41				101	34	34				92	32	36
F5	173	30	18				85	35	41				123	22	18
F7	186	24	13				123	23	19				104	28	27
F8B	181	21	12				109	35	32				118	25	22

a. Yields were calculated from the 30x30 m^2 grids (SMUs) within the field ignoring the fact that some SMU are not fully 30x30 m^2 (e.g. for SMUs on the field boundary). Note the means presented here are nevertheless almost identical to those presented in Table 3-2. In this table, all SMUs are treated equally, regardless of their soil types and the functional zone to which they belong. All data are rounded to the nearest integers.

b. The CV is calculated as standard deviation divided by mean and then expressed as a percentage.

Wheat. Soil, elevation and other site factors are important in explaining yield variability. Of the fields planted with wheat in 1996 and 1998, F2 has the highest coefficient of variation (CV) although F2 (Table 3-1) is the most homogeneous field in terms of soil type. F2 has low-lying areas which may be subject to water-logging which lowers yields. In this case, elevation may dominate in explaining yield variability. However, soil variability is also important. Of the fields planted with wheat in 1995, 1997, and 1999, F1B has the highest CV. According to Table 3-1, F1B has the largest share of very poor Argent soil and medium Bolling soil.

The same soil type in different fields in different years may have quite different yields and different soil types may have similar yields. Fields F1B and F1C have lower wheat yields in 1995 and 1997 than fields F3B and F4 because they have more of the poor Argent soil. But in 1999, F1B has very low yield while all other fields including F1C achieved very high yields. A closer look at the mean yields by soil type on F1B (see Table 4-2) reveals that in 1999, good Pamunkey soils (171 and 172) which are the predominant soils of the fields actually achieved yields as low as those on the poorest Argent soil (148) while the medium Bolling soil (151) achieved the highest yield. For fields planted with wheat in 1996 and 1998, the field F2 in which the best soil is most prevalent consistently yielded the lowest. Thus, field-specific factors other than soil types are also important in determining yield levels.

Table 4-2. Within-field corn and wheat yield variation by soil types for case fields 1995-1999[a]

Field	Soil	Mean	Std	CV	Mean	Std	CV	Mean	Std	CV	Mean	Std	CV	Mean	Std	CV
		---- 1995: wheat ----			---- 1996: corn ----			---- 1997: wheat ----			---- 1998: corn ----			---- 1999: wheat ----		
F1B	Argent	71	12	17	106	29	28	54	18	34	80	31	38	35	18	52
	Bolling	85	14	17	154	30	20	71	23	32	115	36	32	60	12	19
	Muckalee	63	16	24	95	29	30	57	6	10	50	31	62	27	2	6
	Pamunkey1	85	14	16	170	28	16	77	20	26	134	32	24	30	10	33
	Pamunkey2	86	10	12	163	29	18	75	19	25	128	29	22	43	17	40
F1C	Argent	79	10	13	172	28	16	49	19	40	121	37	31	81	26	32
	Pamunkey1	89	7	8	190	25	13	76	7	10	118	20	17	94	19	20
	Pamunkey2	85	9	11	191	23	12	67	17	25	117	29	25	92	19	20
F3B	Bolling	94	5	6	162	16	10	81	23	28	126	22	18	92	10	11
	Muckalee	77	13	17	135	32	24	66	23	36	80	24	31	90	12	14
	Pamunkey1	92	10	11	166	17	10	94	16	17	129	17	13	98	10	10
	Pamunkey2	87	12	14	143	26	18	71	25	35	103	27	26	85	11	13
F4	Argent	90	27	30	151	46	30	75	30	40	120	42	35	61	28	46
	Pamunkey1	103	10	10	184	16	9	95	11	11	150	22	15	92	13	14
	Pamunkey2	95	16	17	165	26	16	81	18	23	123	35	28	75	20	26
		---- 1995: corn ----			---- 1996: wheat ----			---- 1997: corn ----			---- 1998: wheat ----			---- 1999: corn ----		
F2	Muckalee	91	31	34	19	9	49	59	38	66	24	14	58	71	15	21
	Pamunkey1	112	45	40	45	15	33	101	33	33	42	12	28	90	32	35
	Pamunkey2	121	50	41	43	8	19	109	32	29	39	16	42	82	37	46
F5	Argent	149	34	23	42	18	41	38	22	59	64	16	26	123	20	16
	Bolling	134	40	30	45	14	32	21	20	97	62	10	17	120	15	13
	Pamunkey1	179	25	14	67	10	15	97	26	27	69	8	12	123	23	19
	Pamunkey2	182	23	13	59	12	20	93	22	24	68	8	11	125	22	17
F7	Bolling	146	22	15	42	18	44	104	45	43	47	10	21	48	30	64
	Emporia	179	46	26	66	10	15	91	42	46	65	9	14	85	19	23
	Pamunkey1	195	18	9	71	6	9	130	15	12	82	8	10	110	24	22
	Pamunkey2	173	21	12	59	15	26	116	22	19	66	14	21	104	27	26
F8B	Argent	114	51	45	30	13	43	29	18	62	24	10	42	78	31	40
	Bolling	182	24	13	58	15	25	88	36	40	57	16	28	113	25	22
	Pamunkey1	182	10	5	72	5	7	132	13	10	70	8	11	128	17	13
	Pamunkey2	182	12	7	65	8	12	126	15	12	67	9	14	96	40	41

a. Soil identification numbers are Argent (148), Bolling (151), Emporia (161), Muckalee (168), Pamunkey1 (171), and Pamunkey2 (172).

All fields except F2 have their lowest CVs in 1995 which is also the best year for farm-level yields from 1992 to 1999, indicating that within-field wheat yield variability is smaller in very good years. In other more normal years at the farm-level (e.g. 1997), the within-field yield variability varies from field to field. For example, F1B has larger wheat variability in 1999 than in 1997 while F1C has larger variability in 1997.

Corn. Weather may have been the most important factor in determining corn yield levels. For fields planted in corn in 1995, 1997, and 1999, within-field corn yield variability as measured by CV was the highest for F2, the most homogeneous field of all eight fields in terms of soil type. Except for F2 which has consistently lower corn yields, corn yields for other fields (F5, F7, and F8B) vary greatly from year to year. For other fields planted to corn in 1996 and 1998 (F1B, F1C, F3B, and F4), CVs were consistently higher for F1B which had lower corn yields than the other three fields. The differences of corn yield CVs among F1B, F1C, F3B, and F4 within a given year are similar over the years, indicating that relative variability among these four fields may be stable over the years.

Across crops, field-level variations differ across fields. From Table 4-1, the eight fields can be roughly divided into three groups. Group one includes F3B, F4, F7, and F8B. In group one, yields were high in good years and are consistently highest of all fields. For F3B and F4, 1995, 1996, 1997, and 1999 were good years. For F8B, 1995 was good year. In normal or bad years, the yields for group one were stable and similar among the fields. Group two includes F1C and F5. In group two, yields are high in good years (e.g. F1C in 1995, 1996, and 1999, F5 in 1995), but much lower as compared with other fields in bad years (e.g. F1C in 1997 and 1998, F5 in 1997). Group three includes F1B and F2 which had consistently lower to much lower yields as compared with group one fields. This grouping is not consistent with soil variability of the fields indicating the importance of other factors such as elevation in determining yield variability. However, as described in the following paragraphs, soil zone variability is an important factor in determining yields.

Soil zone yield variability

Wheat. Generally, within each field, different soil zones have different average yield levels. An analysis of variance (ANOVA) using SAS PROC GLM (SAS Inc.) on wheat in 1995 and 1997 for fields F1B, F1C, F3B, and F4 shows that wheat yields are significantly different by soil types on every field (p-values are less than 0.0001 for all fields for data shown in Table 4-2), indicating that soil zones are effective to indicate the existence of within-field yield variability for wheat.

Yields are more variable across soils within a field than across fields. For example, in 1995, wheat yield CVs for soil zones in F1B range from 12 to 24 (Table 4-2) while wheat yield CVs for the fields range from 11 to 15 (Table 4-1). Within each field, there are some soil zones that have larger CVs than that of the field-level CV, and there are some soil zones that have smaller CVs than that of the field-level CV.

The soil zone with highest CVs is either Argent soil (148) or Muckalee soil (168) which are the poorest soils according to VALUES (Simpson et al.) for the study fields. For example, for F1C in 1997, the CV for Argent soil is 40% while CVs for other soils are 25% or less, indicating that some SMUs (smallest management units) within Argent soil zone can achieve rather high yields while some other SMUs have very low yields.

In some years, poor soil zones in terms of VALUES-rated yield capacity can achieve yields higher than good soil zones. For example, in F1B in 1999, Argent soil has higher zone-level yield than Pamunkey1 (35 bu/ac vs. 30 bu/ac). In some years, probably other factors than soil type may dominate in determining yield levels within a field.

Pamunkey soils (171 and 172) differ in yields in some fields. For example, in F4, wheat yields for Pamunkey1 soil are consistently higher than that of Pamunkey2 soil and the CVs for Pamunkey1 soil are consistently lower than that of Pamukey2 soil. VALUES estimation does not differentiate between these two soils and on the whole, the general pattern is consistent with VALUES rankings. In most cases, the poorest Argent and Muckalee soils have much lower yields as compared with other soil zones in all fields (e.g., Muckalee soil in F2 in 1996 and 1998). The middle soils (Bolling and Emporia) may achieve very high yields as compared with good Pamunkey soils (e.g., Bolling soil in F3B in 1999 wheat, and Emporia soil in F7 in 1996 wheat). However, the good Pamunkey soils achieve high and stable yields in all cases.

Corn. Soil zone variability is important in describing within-field corn yield variability. An analysis of variance (ANOVA) using SAS PROC GLM (SAS Inc.) on corn in 1995 and 1997 for fields F2, F5, F7, and F8B shows that, except for F2 in 1995, crop yields are significantly different by soil types on each field (p-values are less than 0.01 for individual fields for data shown in Table 4-2).

Corn yields are more sensitive to soil type than are wheat yields. A poor soil (in terms of VALUES-rated yield capacity) is less likely to produce a comparable corn yield to that of a good soil in a good year, even though in a bad year, poor and good soils are more comparable. For example, the Argent soil in 1995 (good year) in fields F5 and F8B had much lower yields than that of Pamunkey soils, while in F5 in 1999 (bad year), comparable yields were recorded for Argent and Pamunkey soils. Except for F2, which has clearly lower yields of corn as well as wheat, the fields that have a higher proportion of good Pamunkey soil achieve higher yields of corn as well as wheat (e.g. F7 and F4).

Corn soil zone yield CVs for good soils are generally lower than those of poor soils but could be much higher than those of poor soils in bad years. Corn yields in some SMUs in good soil zones are more sensitive to poor weather than other SMUs in the same soil zones within a field. For example, good Pamunkey soils (171 and 172) have smaller CVs in F1B in 1996, F4 in 1996, and F5 in 1995 than those of poor Argent soil (148).

Functional zone yield variability

Crop yield variation by functional zones is reported in Table 4-3. Only CVs are reported in Table 4-3 and other statistics are reported in Table C-1 (in Appendix C). Table 4-3 shows that except for the functional zone 1 which has the lowest yields, the within functional zone variability generally is small, indicating that a sufficient number of functional zones were identified to establish homogeneous zones (8 for wheat and 15 for corn) for this study.

The CVs for functional zone 1 are mostly high. Yields for SMUs in this zone are lower than other SMUs in the field while the standard deviation of the group is similar to SMUs in other functional zones.

Corn and wheat yield CVs across functional zones are not different. For example, CVs for wheat in F1B ranged from 2-10 in all years while those for corn ranged from 0-12. If within field variation can be accurately described by this ex post functional zoning, then perhaps N application rate can be determined accurately to give the farmer the highest net income.

Table 4-3. Within-field yield variation by functional zones for case fields 1995-1999

		Functional	CV of crop yields[a]				
Field	Crop	zone # [b]	1995	1996	1997	1998	1999
F1B	Wheat	#1	None[c]		22		30
		#2 to #8	2 ~ 8		2 ~ 9		2 ~ 10
	Corn	#1		20		32	
		#2 - #15		1 ~ 5		0 ~ 12	
F1C	Wheat	#1	None[c]		34		41
		#2 to #8	1 ~ 4		2 ~ 7		2 ~ 8
	Corn	#1		None[c]		65	
		#2 - #15		0 ~ 3		0 ~ 11	
F3B	Wheat	#1	None[c]		13		5
		#2 to #8	1 ~ 6		3 ~ 7		2 ~ 3
	Corn	#1		4		13	
		#2 - #15		1 ~ 3		0 ~ 6	
F4	Wheat	#1	18		33		37
		#2 to #8	2 ~ 5		2 ~ 16		2 ~ 10
	Corn	#1		24		23	
		#2 - #15		1 ~ 6		1 ~ 7	
F2	Wheat	#1		47		38	
		#2 to #8		2 ~ 13		3 ~ 7	
	Corn	#1	25		38		25
		#2 - #15	1 ~ 12		1 ~ 9		1 ~ 6
F5	Wheat	#1		45		19	
		#2 to #8		2 ~ 13		1 ~ 6	
	Corn	#1	27		45		None[c]
		#2 - #15	1 ~ 6		1 ~ 17		0 ~ 5
F7	Wheat	#1		45		23	
		#2 to #8		1 ~ 12		2 ~ 6	
	Corn	#1	6		50		68
		#2 - #15	1 ~ 2		0 ~ 11		1 ~ 12
F8B	Wheat	#1		27		25	
		#2 to #8		1 ~ 6		1 ~ 8	
	Corn	#1	51		33		30
		#2 - #15	1 ~ 2		0 ~ 11		0 ~ 9

a. Values in the table indicate range in yield CVs within a given functional zone.
b. #1 functional zone has the lowest yield level.
c. Only one SMU is included in the functional zone.

4.2. Ex post analysis: Reduction of N rates, yield loss, and net N

Table 4-4 reports the field-level weighted average changes in N application rates, yields, and net N from the baseline for each year, each field, and each alternative strategy. All data are summed across SMUs and then divided by the acreage of the field. Negative values indicate reductions from the baseline and positive values indicate increases from the baseline. The soil zone uniform strategy is identical to the conventional strategy for all fields and all crops, meaning that the profit-maximizing strategy is to apply N based on recommendations for the best Pamunkey soils which is the predominant part of all the case fields. Minor soils are ignored in the N rate determinations in the soil zone uniform strategy because of the imbalance between the cost of one pound N underapplication and the cost of one pound N overapplication. Because the soil zone uniform strategy is equal to the conventional strategy, these two strategies are not compared.

Table 4-4. Ex post change in N rates, yields, and net N from the baseline for each N application strategy[a]

Field	Strategy	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	Average net N change (lb/ac)
		------- 1995: wheat -------			------- 1996: corn -------			------- 1997: wheat -------			------- 1998: corn -------			------- 1999: wheat -------			
F1B	Soil zone-uniform	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-10	-5.3	-3	-14	-8.3	-8	-10	-3.2	-5	-14	-3.6	-12	-10	-0.9	-8	-7.2
	Func. zone-uniform	9	1.7	7	30	12	27	8	0.9	7	-9	-1	-9	-30	-1.2	-28	0.8
	Func. zone-variable	-7	1.1	-8	0	11.1	-3	-20	-0.1	-20	-35	-1	-35	-58	-1.3	-56	-24.4
F1C	Soil zone-uniform	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-14	-9.6	-2	-18	-16.8	-1	-14	-4.6	-9	-18	-10.7	-9	-14	-9.3	-3	-4.8
	Func. zone-uniform	1	0.2	1	53	32.7	36	-14	-0.7	-14	-20	-1.5	-20	24	8.8	13	3.2
	Func. zone-variable	-6	0.3	-5	31	32.1	14	-28	-1	-27	-41	-0.9	-41	2	8	-8	-13.4
F3B	Soil zone-uniform	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-5	-3.4	-1	-9	-8	-3	-5	-2.3	-2	-9	-1.9	-9	-5	-3	-2	-3.4
	Func. zone-uniform	9	4.2	4	18	7.4	17	24	6.8	15	-16	-0.4	-16	21	6.5	12	6.4
	Func. zone-variable	2	4	-2	2	7.2	1	-3	5.8	-11	-36	-0.8	-36	4	5.9	-3	-10.2
F4	Soil zone-uniform	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-6	-4.5	0	-8	-6.3	-2	-6	-4.1	-2	-8	-5	-3	-6	-3.3	-2	-1.8
	Func. zone-uniform	26	12.3	11	42	20.5	35	17	5.7	9	7	1.8	7	11	3.5	6	13.6
	Func. zone-variable	13	11.6	-1	19	19.9	11	2	4.9	-5	-15	1.6	-15	-5	3.3	-9	-3.8

a. All numbers are in per acre units. A positive value indicates an increase from the baseline in yield loss, or N rates, or net N. Average net N change is field-level average over 1995-1999.

Table 4-4 continued

Field	Strategy	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	N rate change (lb/ac)	Yield change (bu/ac)	Net N change (lb/ac)	Average net N change (lb/ac)
		------- 1995: corn -------			------- 1996: wheat -------			------- 1997: corn -------			------- 1998: wheat -------			------- 1999: corn -------			
F2	Soil zone-uniform	0	0	0	0	0	-2.8	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-4	-1.1	-2	-3	0	-27.8	-4	-0.5	-3	-3	-0.1	-3	-4	-0.4	-3	-2.8
	Func. zone-uniform	-1	-0.2	-1	-41	-0.7	-55.8	-21	-0.6	-21	-44	-0.3	-43	-34	-0.9	-34	-27.8
	Func. zone-variable	-42	-0.6	-42	-58	-0.9	0	-55	-1.7	-55	-61	-0.8	-59	-67	-1.3	-67	-55.8
F5	Soil zone-uniform	0	0	0	0	0	-5.8	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-10	-7.3	-4	-7	-1.3	-11.4	-10	-0.1	-10	-7	-2.8	-4	-10	-4.7	-6	-5.8
	Func. zone-uniform	36	21.2	28	-18	-0.2	-31.8	-40	-1	-40	-17	-0.2	-17	-11	-0.3	-11	-11.4
	Func. zone-variable	17	21.1	9	-36	-0.9	0	-73	-1.5	-73	-27	-0.6	-26	-35	-0.9	-35	-31.8
F7	Soil zone-uniform	0	0	0	0	0	-1.8	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-2	-1.9	-1	-2	0	-6	-2	-0.3	-2	-2	-0.1	-2	-2	0	-2	-1.8
	Func. zone-uniform	46	28.2	32	-17	-0.2	-24.4	-19	-0.9	-19	-4	-0.4	-4	-22	-0.9	-22	-6
	Func. zone-variable	28	28.8	14	-30	-0.8	0	-35	-0.8	-35	-18	-0.6	-18	-54	-1	-54	-24.4
F8B	Soil zone-uniform	0	0	0	0	0	-9.4	0	0	0	0	0	0	0	0	0	0
	Soil zone-variable	-15	-13.9	-3	-7	-0.1	-9.8	-15	-0.2	-15	-7	0	-7	-15	-1.7	-15	-9.4
	Func. zone-uniform	38	23.4	29	-18	-0.1	-27.6	-22	-1	-22	-21	-0.8	-20	-19	-1.1	-19	-9.8
	Func. zone-variable	22	23	14	-32	-0.8	-30	-49	-1.3	-49	-34	-1	-32	-41	-1	-41	-27.6

a. All numbers are in per acre units. A positive value indicates an increase from the baseline in yield loss, or N rates, or net N. Average net N change is field-level average over 1995-1999.

Change in N application rates

Corn. The functional zone uniform strategy applies higher N rates than the conventional strategy in good years (e.g. 1996 for F1B, F1C, F3B, and F4, and 1995 for F5, F7, and F8B) and lower N rates in the bad years (e.g. 1997 and 1999 for F2, F5, F7, and F8B). More information does not necessarily decrease N application but does increase the efficiency of N use. The functional zone uniform strategy can apply over 50 lb/ac more N than does the conventional strategy as in F1C in 1996, or about 40 lb/ac less N than does the conventional strategy as in F5 in 1997. For a given field and year, the functional zone uniform strategy may call for much higher N than the conventional strategy in good years (e.g. F5 in 1995), or much lower N than the conventional strategy in bad years (e.g. F5 in 1997) indicating the importance of temporal variation in yields.

In all cases, for given information levels, the variable application strategies apply less N as compared to the uniform strategies as expected. However, functional zone variable strategy may apply more than 40 pounds less N than the functional zone uniform strategy (e.g. F2 in 1995). Reductions for the soil zone variable strategy are much smaller, indicating that actual within-field variability is revealed more in the functional zone information than in the soil zone information.

Wheat. The functional zone uniform strategy may call for more N than the conventional strategy in some cases (1995, 1997, and 1999), and less N than the conventional strategy in other cases (1996, 1998). The effects on N input of more information about the within field variability depend on weather conditions.

Similar to corn, variable N application can reduce field-level N input at a given information level. The difference of N input between the soil zone uniform strategy and the soil zone variable strategy is smaller than that between the functional zone uniform strategy and the functional zone variable strategy for a given field in a given year. The smaller differences imply that the soil zone strategy is less able to differentiate areas of field by N requirement than is the functional zone strategy.

For both wheat and corn, the uniform N rate with soil zone information level is higher than the weighted average variable application rate for different soil zones within the field. It is reasonable for the profit-maximizing farmer to apply N rates higher than the recommended field-level N rate when several soil types are present in a field because income losses per unit of N underapplication are higher than losses per unit of N overapplication.

Change in yield

Corn. Table 4-4 indicates changes in yield relative to the conventional strategy. Positive values show yield increases relative to the baseline (reduction in yield losses) which reduce the cost of N misapplication and increase the values of information. The functional zone uniform strategy increases yields relative to the conventional strategy when it applies higher N (e.g. F1B in 1996). The functional zone uniform strategy reduces yields slightly compared to the conventional strategy when it applies much lower N (e.g. F2 in 1997). Functional zone information is very effective in enhancing yields per unit of N application.

The soil zone variable strategy in most cases reduces yields relative to the conventional strategy without large reduction of N input (e.g. F1C in 1996). This is because soil zone variable application resulted in higher N applications than the conventional strategy to the predominant good soil zone and lower N applications to the minor soils. Because some SMUs in predominant good Pamunkey soils had low yields, the variable application resulted in further N overapplication to these SMUs. Because some

SMUs in the minor soils had high yields, there was N underapplication and thus yield losses increased for these SMUs under variable application. By contrast, the functional zone variable and uniform strategies maintain similar yields which are considerably higher than the conventional strategy. The functional zone variable strategy clearly decreases N input as compared with the functional zone uniform strategy (e.g. F1B in 1996). In fact, only the functional zone variable strategy both reduces N input and increases yields relative to the baseline (e.g. F4 in 1998) indicating that with functional zone information, variable application is able to reduce total N input while maintaining yield levels.

Wheat. The functional zone uniform strategy was able to increase yields by increasing N input (e.g. F4 in 1997) or slightly decrease yield by clearly larger decreases of N input (e.g. F1B in 1999). Thus, functional zone information is able to increase crop yield relative to N input (i.e. increase N efficiency).

Variable application based on soil zone information may not increase N use efficiency. In most cases, soil zone variable application results in a reduction in N application but also a reduction in yields. By contrast, in all cases, variable application based on the functional zone information clearly reduces N input relative to the functional zone uniform strategy while maintaining yield levels equal to those obtained by the functional zone uniform strategy. Thus, with functional zone information, variable application clearly increases yield obtained per unit of applied N compared to uniform application.

Change in net N

Corn. Reductions in N input by the functional zone uniform strategy result in comparable reductions in net N relative to the conventional strategy (e.g. F2 in 1999). This result occurs when the conventional strategy overapplies N. Increases in N input by the functional zone uniform strategy relative to the conventional strategy result in increased net N (e.g. F5 in 1995). However, the increase of net N in most of these cases is less than the increases of applied N (e.g. F8B in 1995) indicating that functional information can increase N use efficiency.

Variable application based on soil zone information always slightly reduces net N relative to the conventional strategy, and in most cases reduced yield (e.g. F1C in 1998) indicating that reduction of net N could be costly. By contrast, variable application based on the functional zone information always reduces net N from the uniform application while maintaining yield levels (e.g. F1B in 1998). With functional zone information, the reduction of net N for variable application relative to uniform application is usually over 20 lb/ac indicating that it is very effective to use variable application to reduce net N.

Wheat. The functional zone uniform strategy increases net N relative to the conventional strategy when the conventional strategy underapplies N (e.g. F1C in 1999). Underapplication by the conventional strategy is implied because increased N applications with better information resulted in higher yields. The functional zone uniform strategy decreases net N relative to the conventional strategy when the conventional strategy overapplies N (e.g. F1B in 1999). Overapplication by the conventional strategy is implied because reduced N applications with better information do not reduce yields. Generally, when the conventional strategy overapplies N, the functional zone uniform strategy can reduce net N by an amount equal to the reduced N input (e.g. F1C in 1997). When the conventional strategy underapplies N, the functional zone uniform strategy increases net N by amounts less than the increased N input (e.g. F4 in 1995). These results show that functional zone information can improve N use efficiency, but does not always reduce net N.

Variable application based on soil zone information in all cases except F4 in 1995 reduces net N relative to the uniform strategy because of reduced N input. For F4 in 1995, even though N input is reduced by 6 lb/ac, net N is not reduced. Only when the soil zone uniform strategy clearly overapplies N (e.g. F8B in 1998) is variable application able to reduce net N by the same amount as N input.[42] In most cases of soil zone information, the reduction of net N in variable application occurs with increased yield losses (e.g. F5 in 1998). By contrast with the functional zone information, variable application always reduces net N while maintaining similar yield levels as compared with the uniform strategy indicating that variable application is very effective in reducing net N.

Average net N over 1995-1999. For fields F1B, F1C, F3B, and F4, the functional zone uniform strategy increases average net N, while the soil zone variable strategy and the functional zone variable application strategy reduce average net N compared to the conventional strategy. For these four fields, the functional zone variable strategy reduces net N from the baseline by the largest amount. For fields F2, F5, F7, and F8B, all strategies except the soil zone uniform strategy reduce average net N compared to the baseline. However, the soil zone variable strategy achieves much smaller average reductions than the functional zone strategies. For the functional zone variable strategy, the average net N reduction is always greater than 20 lb/ac. Generally, for all eight fields, functional zone information combined with variable application reduces average net N from the baseline.

---

[42] Clearly overapplied N means that the field-level N input rates are clearly higher than the N needed to achieve the actual field-level yield. Such clear overapplication also can be seen in Table 4-4. For example, for F8B in 1998, the functional zone variable strategy reduces N rates relative to the conventional strategy by 34 lb/ac with yield loss increase of only 1 bu/ac indicating that it is clear the conventional strategy overapplies N.

4.3 Ex post values of information and variable N application

As discussed in Chapter Two and Chapter Three, the gross information value is the reduction of the cost of N misapplication relative to the conventional strategy (e.g. the gross soil information value is reduced cost of N misapplication of the soil zone uniform strategy relative to the conventional strategy). The net information value is the gross value minus the additional cost of information generation. The gross value of variable application is the reduced cost of N misapplication relative to uniform application for that information level (e.g. gross value of soil variable application is the reduced cost of N misapplication relative to the soil zone uniform application). The net value of variable application is the gross variable application value minus the additional cost of carrying out variable application. The sums of information values reported for variable applications in Table 4-5 represent the sums of net information values for a given information level plus the values of variable application.

Table 4-5. Ex post values of information and variable application for alternative N application strategies[a]

Field	Value	1995 value ($/ac)			1996 value ($/ac)			1997 value ($/ac)			1998 value ($/ac)			1999 value ($/ac)			Five year average ($/ac)		
		Gross	Net	Sum	Gross	Net	Sum	Gross	Net	Sum	Gross	Net	Sum	Gross	Net	Sum	Gross	Net	Sum
		--------- wheat ----------			--------- corn ----------			--------- wheat ----------			--------- corn ----------			--------- wheat ----------					
F1B	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-11.90	-13.99	-13.99	-12.66	-14.75	-14.75	-6.34	-8.43	-8.43	-3.30	-5.39	-5.39	-0.15	-2.24	-2.24	-6.87	-8.96	-8.96
	Functional information	2.57	-0.04		15.76	13.15		0.53	-2.08		0.37	-2.24		4.29	1.68		4.71	2.09	
	Functional variable application	2.11	-0.50	-0.54	5.89	3.28	16.43	4.35	1.74	-0.34	6.49	3.88	1.64	6.60	3.99	5.67	5.08	2.48	4.57
F1C	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-22.53	-24.62	-24.62	-28.41	-30.50	-30.50	-8.81	-10.90	-10.90	-16.50	-18.59	-18.59	-21.60	-23.69	-23.69	-19.57	-21.66	-21.66
	Functional information	0.46	-2.15		50.46	47.85		1.68	-0.93		2.05	-0.56		18.04	15.43		14.54	11.93	
	Functional variable application	1.79	-0.82	-2.97	4.24	1.63	49.48	2.73	0.12	-0.81	6.25	3.64	3.08	3.30	0.69	16.12	3.66	1.05	12.98
F3B	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-8.15	-10.24	-10.24	-13.23	-15.32	-15.32	-4.89	-6.98	-6.98	-1.40	-3.49	-3.49	-6.94	-9.03	-9.03	-6.92	-9.01	-9.01
	Functional information	9.23	6.62		9.87	7.26		12.66	10.05		3.19	0.58		12.37	9.76		9.46	6.85	
	Functional variable application	0.95	-1.66	4.96	3.69	1.08	8.34	3.97	1.36	11.41	4.29	1.68	2.26	2.43	-0.18	9.58	3.06	0.46	7.31
F4	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-10.59	-12.68	-12.68	-10.42	-12.51	-12.51	-9.56	-11.65	-11.65	-7.95	-10.04	-10.04	-7.34	-9.43	-9.43	-9.17	-11.26	-11.26
	Functional information	27.02	24.41		29.44	26.83		11.48	8.87		1.59	-1.02		6.99	4.38		15.31	12.69	
	Functional variable application	1.34	-1.27	23.14	4.66	2.05	28.88	1.68	-0.93	7.94	5.18	2.57	1.55	3.23	0.62	5.00	3.22	0.61	13.30
		--------- corn ----------			--------- wheat ----------			--------- corn ----------			--------- wheat ----------			---------corn ----------					
F2	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-1.28	-3.37	-3.37	0.79	-1.30	-1.30	-0.11	-2.20	-2.20	0.37	-1.72	-1.72	0.09	-2.00	-2.00	-0.02	-2.12	-2.12
	Functional information	-0.04	-2.65		8.43	5.82		4.01	1.40		10.12	7.51		6.81	4.20		5.87	3.26	
	Functional variable application	9.39	6.78	4.13	3.48	0.87	6.69	6.56	3.95	5.35	2.72	0.11	7.62	7.61	5.00	9.20	5.95	3.34	6.60
F5	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-11.90	-13.99	-13.99	-1.76	-3.85	-3.85	2.18	0.09	0.09	-5.83	-7.92	-7.92	-6.73	-8.82	-8.82	-4.80	-6.90	-6.90
	Functional information	32.28	29.67		4.08	1.47		8.14	5.53		3.77	1.16		2.19	-0.42		10.10	7.48	
	Functional variable application	4.51	1.90	31.57	2.44	-0.17	1.30	7.13	4.52	10.05	1.36	-1.25	-0.09	4.91	2.30	1.88	4.07	1.46	8.94
F7	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-3.02	-5.11	-5.11	0.33	-1.76	-1.76	0.02	-2.07	-2.07	0.26	-1.83	-1.83	0.60	-1.49	-1.49	-0.36	-2.45	-2.45
	Functional information	43.50	40.89		3.80	1.19		2.92	0.31		-0.08	-2.69		3.78	1.17		10.79	8.17	
	Functional variable application	5.75	3.14	44.03	1.66	-0.95	0.24	4.36	1.75	2.06	3.14	0.53	-2.16	7.73	5.12	6.29	4.52	1.92	10.09
F8B	Soil information	0	0		0	0		0	0		0	0		0	0		0	0	
	Soil variable application	-23.24	-25.33	-25.33	1.67	-0.42	-0.42	3.44	1.35	1.35	1.72	-0.37	-0.37	0.50	-1.59	-1.59	-3.19	-5.27	-5.27
	Functional information	36.15	33.54		4.12	1.51		3.58	0.97		2.97	0.36		2.69	0.08		9.90	7.29	
	Functional variable application	3.23	0.62	34.16	1.52	-1.09	0.42	6.22	3.61	4.58	2.83	0.22	0.58	5.58	2.97	3.05	3.87	1.27	8.56

a. The net values of information are also the values of uniform strategies (the soil zone uniform strategy and the functional zone uniform strategy). Sum refers to the sum of net information value and net variable application value. Sum is also the value of the variable strategy for given information level relative to the baseline.

Gross information values

Corn. Gross soil zone information values are all zero because the soil zone uniform strategy is identical to the conventional strategy. Except for F2, gross functional zone information values for corn are positive. Gross information values are high in years when N is underapplied in the conventional strategy (e.g. F7 in 1995) and low in years where N is overapplied in the conventional strategy (e.g. F8B in 1999). As stated earlier, underapplication is implied when better information levels increase yields. Overapplication is implied when better information levels reduce N applications but not yields. In years when the conventional strategy underapplies N, gross information values increase with the amount of underapplication[43] (e.g. F7 and F8B in 1995). In years when the conventional strategy overapplies N, the gross information values increase with the amount of overapplication by the conventional strategy (e.g. F7 and F8B in 1999). The gross information value for corn reaches a maximum of $50.46 (F1C in 1996). Functional information can reduce overapplication and underapplication of N. Temporal yield variation is important in affecting the level of the values of information.

Wheat. Gross functional zone information values for wheat are all positive except F7 in 1998. Similar to corn, high gross information values occur when the conventional strategy would underapply N (e.g. F4 in 1995), lower gross information values occur when the conventional strategy overapplies N (e.g. F1C in 1997). Since the conventional strategy generally did not underapply N very much, the gross information values for wheat are

---

[43] The degrees of overapplication or underapplication by the conventional strategy can be obtained easily by comparing actual yields in Table 4-1 with the target yields for the conventional strategy which are 90 and 160 bu/ac for wheat and corn respectively (Table 3-6). This means application rate was based on target yield which was 160 bu/ac while actual N need was only that for roughly 90 bu/ac yield. It can also seen in Table 4-4 by comparing N rate with yield loss for the functional zone variable strategy relative to the conventional strategy.

generally lower than those for corn. The highest gross information values are less than $30.00 in all cases.

Net information values

Corn. Net functional information values for corn are positive in most cases. In the cases where the conventional strategy clearly underapplies N, the net information values are clearly positive (e.g. F5 in 1995, F1C in 1996). In the cases where the conventional strategy overapplies N, the net information values are small or negative (e.g. F5 in 1999, F1B in 1998). Functional information increases the farmer's net income mainly through reducing the degree of N underapplication.

Wheat. Similar to corn, net information values are positive in most cases. The net information values for wheat are less variable than those for corn. The high net information values come from the cases where the conventional strategy underapplies N and better information levels increase yields (Table 4-4). The low or negative net information values come from the cases where the conventional strategy overapplies N and better information levels reduce N applications but not yields.

Field-level average over 1995-1999 shows that all net functional zone information values are positive, with the highest being $12.69/ac (F4) and the lowest being $2.09/ac (F1B). The high returns to functional zone information indicate the benefits of anticipating variations in yield potential due to weather. The results also show that a one-unit overapplication of N is less costly than a similar underapplication of N.

Gross values of variable application

Corn. In most cases, the gross values of variable application based on soil zone information are negative (e.g. the gross value of variable application for F1C in 1996 is - $28.41/ac) because of the underapplication of N to SMUs with poor soil types. Soil zone information does a poor job of directing variable application for corn. Soil variable application values are positive when the conventional strategy overapplies N to a large degree (e.g. F5 in 1997). Positive variable application values come mainly from decreased field-level overapplication. In all cases, the gross values of variable application based on functional zone information are positive, showing that variable application reduces both overapplication and underapplication of N for corn. If within-field variability can be correctly described by the functional zones then variable application can increase the farmer's net income.

Wheat. In most cases, the gross values of variable application based on soil information are negative. The large negative values of variable application based on soil information come from increased underapplication relative to the conventional strategy (e.g. F1C in 1999) while the positive values of variable application come from the decreased overapplication relative to the conventional strategy (e.g. F7 in 1998). The net return increased from reducing overappliction is rather small while the net return decreased from larger underapplication is rather large.

In all cases, the gross values of variable application based on functional zone information are positive. Given that field-level N application from the functional zone variable strategy is lower than the functional zone uniform strategy for each field, variable application greatly increased N use efficiency. Since variable application increases net income mainly by reducing overapplication, it is as expected that the values of variable application are not large for any field (the largest value is $5.95 on F2). The results show that variable application based on functional zone information not only reduces total N and net N compared with the uniform strategy, but also has a positive gross value to the farmer.

Net values of variable application

Corn. In all cases except F5 and F8B in 1997, net values of variable application based on soil zone information are negative. The saving from reduced overapplication does not compensate for the increased operation cost of variable application. In the cases where the soil zone variable strategy underapplied N, net income losses from variable application are very large (e.g. F1C in 1998, and F8B in 1995).

In all cases, net values of variable application based on functional zone information are positive, indicating that with accurate delineation of within-field yield variability, the variable application can increase net return either by reducing overapplication or by reducing underapplication of N.

Wheat. In all cases, net values of variable application based on soil zone information are negative, indicating that variable application reduces N application but does not increase N use efficiency. In all cases, net values of variable application based on functional zone information are positive or slightly negative. Some values are slightly negative because of the increased operation costs for the variable application.

Averages over five years show that the soil zone variable strategy is not profitable for any of the eight fields. The soil zone variable strategy reduces the farmer's net income relative to the conventional strategy because the sum of increased operation cost and increased cost of underapplication in some parts of the field is larger than savings from reduced N overapplication. Soil zone information alone is not likely to increase net farm income when used to direct variable application for wheat and corn. Averages over five years show that the functional soil variable strategy increases net income relative to the functional zone uniform strategy in all fields, indicating that variable application not only reduces the field-level total N application and net N, but also reduces income losses from underapplication and overapplication.

Summary

Results from this case study suggest that soil zone information alone will not increase net income when used to determine either uniform N rate or variable application rates for any field. In Chapter Three, it was shown that observed yield levels differ significantly across soil zones in each year and field. However, some SMUs in the poor soil zones had high yields and some SMUs in the good soil zones had low yields. Because the cost from underapplication of N greatly exceeds the cost from overapplication of N, soil zone information led to large losses of net income by underapplying N to SMUs that are in poor soil zones but nevertheless achieved high yield levels. Net income losses due to underapplication to poor soil zones are especially large in corn in good years when poor soils achieved yields comparable to the good soil zones.

In the ex post analysis, functional zone information raises net income when used to direct both uniform and variable application of N. For all fields, the variable application achieves higher net income than the uniform application based on 5-year averages in spite of the additional cost of variable application. The results establish the potential to increase N use efficiency and net income by dividing a field into several functional zones. However, when future field-level yields and within-field yield variability can only be described by conditional probability distributions, it is not clear if information and variable application values will be ranked similarly. In the following ex ante analysis, this question will be investigated using historical data to establish the conditional probability of both field-level and within-field yields.

4.4. Ex ante reduction of N rates, yield loss, and net N from the baseline for alternative strategies in 1999

The year 1999 is used to examine the information values, values of variable application, and net N reduction for alternative strategies when both temporal and spatial information is utilized. N applications, yields, and net N are reported for each field in Table 4-6. The baseline is the conventional strategy. As discussed in Chapter Three, the baseline optimal yields for 1999 were determined based on the historical farm-level yields and individual field yields. N application rates for baseline were based on optimal yields in 1999, and net N was calculated based on the optimal N application rates corresponding to the optimal yields compared with N required for the actual yields.

Table 4-6. Ex ante change in N rate, yield, and net N from the baseline in 1999 for alternative strategies

Strategy	Wheat field	Change in N applied (lb/ac)	Change in yield (bu/ac)	Change in net N (lb/ac)	Corn Field	Change in N applied (lb/ac)	Change in yield (bu/ac)	Change in net N (lb/ac)
Soil zone-uniform	F1B	2	0	1	F2	4	0	4
Soil zone-variable		0	-1	0		1	-1	2
Func. zone-uniform		1	0	1		22	0	22
Func. zone-variable		-7	0	-7		-2	0	-2
Soil zone-uniform	F1C	7	5	2	F5	6	0	6
Soil zone-variable		-1	-1	1		0	-4	3
Func. zone-uniform		-3	-1	-1		12	0	12
Func. zone-variable		-8	-4	-3		4	-1	5
Soil zone-uniform	F3B	-1	-1	-1	F7	4	0	4
Soil zone-variable		0	0	0		1	0	1
Func. zone-uniform		-2	-2	-1		-7	0	-7
Func. zone-variable		-9	-6	-3		-13	0	-13
Soil zone-uniform	F4	5	1	2	F8B	16	0	16
Soil zone-variable		1	-1	1		0	0	0
Func. zone-uniform		-12	-6	-6		11	0	11
Func. zone-variable		-13	-5	-8		0	0	0

In the following discussion, changes in yields relative to the conventional baseline will be used to describe the effects of alternative strategies on yields. An increased yield with better information indicates that the better N management based on better information raises yields relative to the baseline. As discussed before, yield loss comes from N underapplication and good management of N should reduce yield losses, other things equal. A summary table (Table 4-7) is derived from tables in Chapter Three and previous tables in this chapter to display the expected yields, the actual yield, and the optimal yields for alternative strategies for each field in 1999. This summary table will be used to discuss field-level underapplication/overapplication for alternative strategies.

Table 4-7. Expected yields, actual yields, optimal yields of all strategies in 1999 for study fields

	Wheat fields in 1999				Corn fields in 1999			
Term	F1B (bu/ac)	F1C (bu/ac)	F3B (bu/ac)	F4 (bu/ac)	F2 (bu/ac)	F5 (bu/ac)	F7 (bu/ac)	F8B (bu/ac)
Expected yields	61	64	72	74	118	143	167	164
Actual yields	44	91	93	86	92	123	104	118
Opt. yld. for the conventional strategy	74	77	88	90	137	170	194	190
Opt. yld. for the Soil zone-uniform strategy	76	82	87	94	141	176	198	206
Opt. yld. for the Soil zone-variable strategy	74	77	88	90	137	170	194	190
Opt. yld. for the Func. Zone-uniform strategy	75	74	86	81	159	182	187	201
Opt. yld. for the Func. Zone-variable strategy	69	70	81	80	134	174	181	190

Change in N rates

Corn. Soil zone information results in higher N rates for all fields, indicating that the optimal yield in 1999 for the good and predominant soil zone (Pamunkey) is higher than that of the conventional strategy in all four fields. However, except F8B, the increased N rates are less than 10 lb/ac, indicating that when the field is rather homogeneous in terms of soil type and the dominant soil type is a good soil, then there is a small difference between the soil zone uniform strategy and the conventional strategy. This result is consistent with the ex post analysis.

Functional zone information increased the N rate in three fields and decreased the N rate in one field (F7) in 1999. In F2, F5, and F8B, the functional zone uniform strategy applied more than 10 lb/ac over the conventional strategy, indicating that the functional zone information results in different optimal yield predictions than that of the conventional strategy or the soil zone strategy.

Variable application based on both soil zone information and functional zone information decreased the N rate compared to the uniform application in most cases. This suggests that variable application is able to efficiently utilize soil or functional zone information to direct N use and thus reduce overapplication. Comparatively, variable application based on functional zone information is able to reduce N rate more than that based on soil zone information (e.g. F2). However variable application reduced N application relative to the baseline in only two cases (F2 and F7).

Wheat. Soil zone information led to higher N application rates in three cases and lower N application in one case (F3B). But the N rate differences between the soil zone uniform strategy and the conventional strategy are generally small (< 7 lb/ac). Differences are small because the temporal variability for wheat as described by the simulated yield sequences is small.

Functional zone information led to an increased N rate only in one case (F1B). In F4 in 1999, the functional zone uniform strategy applied 12 lb/ac less N than the conventional strategy, indicating that the observed within-field variability is less than that indicated by soil zone information.

Variable application based on soil zone information or functional zone information led to the same or less N application than the uniform application strategy in all but one case. Field-level N rate for the soil zone variable strategy is basically equal to that of the conventional strategy, while the functional zone variable strategy applied less N than the conventional strategy (e.g. F4).

Higher information levels do not necessarily reduce N application, but variable applications based on higher levels of information lead to lower N application compared to uniform application for both corn and wheat in most cases.

Change in yields

Corn. For both soil zone information and functional zone information, there is no change in yield relative to the baseline in all four fields. The conventional strategy overapplied N in all corn fields in 1999, while the soil zone uniform strategy and the functional zone uniform strategy further overapplied N (except the functional zone uniform strategy in F7). Because of the large temporal variability in corn yields on the case farm, net incomes are increased by applying N for a higher corn yield which may result in high overapplication in a bad year for corn such as 1999.

Variable application based on soil zone information and functional zone information resulted in a slight yield loss in three cases. For F5, soil zone variable strategy resulted in 4 bu/ac yield loss while its field-level average N rate did not change from the conventional strategy, indicating that there were some SMUs in the poor soil zone that actually achieved yields much higher than would be indicated by the soil zone information alone. For F5, the functional zone variable strategy applied 8 lb/ac less than the functional zone uniform strategy but reduced yield by only 1 bu/ac relative to the uniform application. These results suggest that functional zone information properly described the within-field variability for this case.

Wheat. Soil information led to reduced yield in one case (F3B) due to further underapplication compared with the conventional strategy, increased yields in two cases (F1C and F4) due to decreased underapplication, and the same yield in one case (F1B) due to further overapplication compared to the conventional strategy. For functional information, in three cases, yields were reduced relative to the conventional strategy (F1C, F3B, and F4) due to further underapplication in some SMUs within the field. In most cases, functional zone information led to lower yields compared to soil zone information.

Variable application based on soil information reduced yields relative to the uniform strategy in three cases (F1B, F1C, and F4) and increased yield in one case (F3B). In F1B where large overapplication occurred for all strategies, variable application based on soil zone information resulted in a slight yield loss due to underapplication to some SMUs even though the field-level N rate did not change from the conventional strategy. A similar situation occurred on F4 where N is overapplied at the field-level but where a slight decrease of N application in the soil zone variable application compared with the conventional strategy resulted in a small yield reduction. For F1C and F3B where optimal yields are close to actual yields, reduced yields from variable applications with soil zone information reflected lower N applications. Yield losses from underapplication on the SMUs in the poor soil zones, which achieved high actual yields in 1999, equal the yield gains from increased N application for the SMUs that are in good soil zones where higher yields were achieved in 1999. Soil zone information in F1C and F3B resulted in underapplication of N in some poor soil zones while reducing the underapplication of N on the good soil zones. These results indicate the inadequacy of directing variable application based on soil zone information alone.

Variable application based on functional zone information reduced yields relative to uniform application in two cases (F1C and F3B). For F1B where all strategies clearly overapplied N in 1999, variable application based on functional information was able to maintain yields while reducing the N rate relative to the uniform application. For the other three fields where N was clearly under applied at the field-level, the reduction in yields is less than what would be reflected by the reduced N rates.[44] For F4, variable application reduced N input compared to the uniform application, yet wheat yields from variable application were also higher compared to the yields from uniform application. These results indicate that functional zone information does increase the efficiency of N use.[45]

Change in net N

Corn. Because of the clear overapplication of N at the field-level by all strategies,[46] changes in net N are equal to the change of N rates in all cases for both the soil zone uniform strategy and the functional zone uniform strategy. Net N from the soil zone uniform strategy can be smaller (e.g. F2 and F5) or larger (F7 and F8B) than that from the functional zone uniform strategy.

Variable application reduced net N compared to the uniform strategy in all cases. The reduction of net N by variable application is much larger for functional zone information than for soil zone information. In F2 and F5, the soil zone variable strategy reduced net N by 2 to 3 lb/ac as compared with the soil zone uniform strategy but reduced yields as well. The functional zone variable strategy reduced net N by 6 to 24 lb/ac as compared with the functional zone uniform strategy with smaller yield losses. Variable application based on functional zone information is more effective in reducing net N than that based on soil zone information.

---

[44] As stated in Chapter 3, it is assumed that with one pound N underapplied, yield loss for corn is 1 bushel, and for wheat is 0.8 bushel.

[45] Note, as it was defined in Chapter 1, N efficiency is the percent uptake by crop of the N applied.

[46] Overapplication or underapplication can be seen by comparing target yields as calculated in Chapter 3 with the actual yields as reported in Table 4-1. It is clear that corn fields are greatly overapplied in 1999 by all strategies while both overapplication and underapplication happened in wheat fields with different strategies. For example, the target yield for the soil zone uniform strategy is F8B in 1999 is 206 bu/ac (Table 3-22) so N rate would be 186 lb/ac. But actual yield for F8B in 1999 was 118 bu/ac (Table 4-1) which required an N rate of 98 lb/ac. Thus the field-level overapplication by the soil zone uniform strategy would be 186 – 98 = 96 lb/ac. Field-level underapplication is similarly estimated.

Wheat. The reduction of net N from the baseline is similar between the soil zone uniform strategy and the functional zone uniform strategy with the exception of F4 where functional zone information did better. In all wheat fields, there are only small changes in net N from the baseline for the alternative uniform strategies.

Variable application based on soil zone and functional zone information is able to reduce net N compared to uniform application. In all cases, variable application based on functional zone information is able to reduce more net N relative to the uniform application compared with that based on soil zone information (e.g. F1B). Functional zone information is able to describe more within field variability than does soil zone information.

4.5. Ex ante values of information and variable application in 1999

Based on actual yield, N application, yield loss, and production costs for each strategy, the gross and net values of information and variable applications were calculated for each field in 1999 when both temporal information and spatial information were used. The results are reported in Table 4-8. As discussed in Chapter Two and Chapter Three, the gross value of information is the reduction relative to the baseline of N misapplication cost for the uniform strategy with a given information level. Gross variable application value is the reduction of N misapplication cost relative to the uniform strategy for that information level. The net value is gross value minus the additional operation cost of generating information and carrying out variable application. Net information value plus net variable application value for a given information level is also the increased net return for an information variable strategy over the conventional strategy. In Table 4-8, the strategy sum is the sum of net information value plus net variable application value.

Table 4-8. Ex ante information values and values of variable rate application in 1999 for alternative N application strategies

Strategy	Field	Information value ($/ac)		Variable appl. value ($/ac)		Strategy
		Gross	Net	Gross	Net	sum ($/ac)
		Wheat				
Soil zone-uniform	F1B	-0.25	-0.25			-0.25
Soil zone-variable				-2.33	-4.42	-4.67
Func. zone-uniform		-0.12	-2.73			-2.73
Func. zone-variable				1.53	-1.08	-3.81
Soil zone-uniform	F1C	9.54	9.54			9.54
Soil zone-variable				-13.66	-15.75	-6.21
Func. zone-uniform		-4.33	-6.94			-6.94
Func. zone-variable				-5.28	-7.89	-14.83
Soil zone-uniform	F3B	-1.35	-1.35			-1.35
Soil zone-variable				1.25	-0.84	-2.19
Func. zone-uniform		-2.75	-5.36			-5.36
Func. zone-variable				-9.89	-12.50	-17.86
Soil zone-uniform	F4	3.87	3.87			3.87
Soil zone-variable				-5.19	-7.28	-3.41
Func. zone-uniform		-12.99	-15.60			-15.60
Func. zone-variable				2.89	0.28	-15.32
		Corn				
Soil zone-uniform	F2	-1.00	-1.00			-1.00
Soil zone-variable				-0.41	-2.50	-3.50
Func. zone-uniform		-5.50	-8.11			-8.11
Func. zone-variable				6.06	3.45	-4.66
Soil zone-uniform	F5	-1.50	-1.50			-1.50
Soil zone-variable				-6.06	-8.15	-9.65
Func. zone-uniform		-3.00	-5.61			-5.61
Func. zone-variable				0.22	-2.39	-8.00
Soil zone-uniform	F7	-1.00	-1.00			-1.00
Soil zone-variable				0.80	-1.29	-2.29
Func. zone-uniform		1.75	-0.86			-0.86
Func. zone-variable				1.52	-1.09	-1.95
Soil zone-uniform	F8B	-4.00	-4.00			-4.00
Soil zone-variable				3.76	1.67	-2.33
Func. zone-uniform		-2.75	-5.36			-5.36
Func. zone-variable				2.78	0.17	-5.19

It should be noted that in ex ante analysis, both gross and net information values and variable application values can be negative for any field because the farmer's N application decision was assumed to be based on his belief or his modeling of the temporal/spatial variability for each field while the evaluation of information values were based on actual yield data. Thus, gross and net variable application values can be negative because of the errors in the farmer's belief or the error terms not captured in the farmer's model which induce a sub-optimal allocation of N to SMUs within the field.

Gross information values

Corn. In all cases except for the functional zone uniform strategy in F7, gross information values (soil zone or functional zone) are negative due to further overapplication of N on these fields. For example, for F5, the field-level expected yield is 143 bu/ac, the actual yield in 1999 was 123 bu/ac, the conventional strategy's as well as the soil zone variable strategy's optimal yield was 170 bu/ac, the soil zone variable strategy's optimal yield was 176 bu/ac, the functional zone variable strategy's optimal yield was 174 bu/ac, and the functional zone uniform strategy's optimal yield was 182 bu/ac (Table 4-7). Thus, for F5 in 1999, N overapplications from all strategies were 50 lb/ac or higher at the field-level. Gross information value for the functional zone on F7 is positive due to decreased N application compared to the conventional strategy. When temporal prediction of field-level yield is not accurate, soil zone information or functional zone information does not necessarily increase the farmer's net return.

Wheat. Gross soil information values are positive in F1C and F4 due to increased N application because the conventional strategy underapplied N on these two fields. For example, for F1C, all strategies underapplied N for F1C in 1999. Gross soil information value is slightly negative on F1B due to increased overapplication and is negative on F3B due to increased underapplication on some SMUs in the field (Table 4-7).

Gross functional zone information values are negative for all cases. On F1B, the slight negative information value comes from slight further overapplication relative to the conventional strategy. On other fields, gross information values are negative because of increased underapplication of N on these fields (see Table 4-7).

For all fields across crops, it is seen that neither soil zone nor functional zone information necessarily increases the precision of temporal field-level yield predictions. However, in the case where the predicted yield level is lower (as reflected by the optimal yield for the conventional strategy) than actual, soil information may lead to higher N application relative to the case when predicted yield is higher than actual and thus decrease underapplication and increase net return while functional zone information may lead to further overall underapplication of N and decreased net return (see Table 4-7).

When yields are underestimated, the soil zone information does a better job than the functional zone information because the predominant Pamunkey soil yield sequence is no longer scaled which results in a higher target for Pamunkey soils compared to the conventional strategy. By contrast, the better functional zones in the observed yield sequences were not scaled down in the previous analysis in determining target yields.

Net information values

Corn. In all cases, net information values are negative. Functional zone information values are negative because of the additional costs of information generation for functional zone delineation. For both soil zone and functional zone information, the main reason for negative values of information is that information results in further overapplication of N to corn in 1999 for all four study fields.

Wheat. Net soil information values are positive in fields where soil information led to decreased underapplication (F1C and F4) and negative in fields where soil information led to increased overapplication (F1B) or increased underapplication (F3B).

Net values of functional zone information are negative in all cases with additional cost for functional zone delineation added. Functional zone information values are negative either because of increased overapplication as in F1B or increased underapplication as in other cases.

In most cases across crops, the net value for functional zone information is smaller than the net value for soil zone information for a given field because of larger underapplication (e.g. F1C), or larger overapplication (e.g. F2), or higher additional cost for information generation (e.g. F1B and F8B). However, large errors in predicting field-level yields in most cases make it difficult to compare information values. The functional information only utilized yield monitor data in 1995 and 1997 and as a result, in some cases, the functional zone strategies applied higher N relative to the conventional strategy, and in other cases, the functional zone strategies applied lower N relative to the conventional strategy. This situation also makes the evaluation of functional information difficult.

Gross values of variable application

Corn. Gross variable application values based on soil zone information are either negative because of increased underapplication in some SMUs (e.g. F2 and F5) or positive because of decreased overapplication (e.g. F7 and F8B). For example, for F5, even though soil zone variable application strategy applied the same amount of N at the field-level as the conventional strategy, yield decreased 4 bu/ac for the soil zone variable strategy relative to the conventional strategy (Table 4-6). The soil zone variable application underapplied N for some SMUs in the minor soil zones even though at the field-level, N was overapplied by 47 lb/ac (Table 4-7). It is not clear if variable application improves N efficiency for information levels with large N overapplication (e.g. F7 and F8B) (Table 4-6). However, in cases (F2 and F5) where overapplication is smaller, variable application by soil zone increased yield loss and thus decreased gross variable application values.

Gross values of variable application based on functional zone information are positive in all cases due to decreased N overapplication. Variable N application based on functional zone information improves N use efficiency. Functional zone variable application also performed better than soil zone variable application as clearly seen in F2. In F2, functional zone variable application reduced per acre N application by 24 pounds relative to uniform application while maintaining yields. Soil zone variable application reduced N application by only 3 lb/ac relative to soil zone uniform application while reducing yield by 1 bu/ac.

Wheat. In this case study, variable application based on soil zone information does not increase net returns when yields are uncertain. In all cases except F3B, values of variable application based on soil zone information are negative due to increased yield loss relative to uniform application. In F3B, variable application value is positive because slightly increased N rates increase yields relative to uniform application.

Gross value of variable application based on functional information can be positive either through reducing overapplication of N (e.g. F1B) or reducing both N application rate and yield loss relative to uniform application (e.g. F4). However, in fields where the uniform strategy (including the conventional strategy) already underapplied N (e.g. F1C and F3B, see Tables 4-6 and 4-7), further underapplication on some of the SMUs by variable application results in a negative gross value. For example, for F1C, variable application further underapplied N compared to the uniform strategy (Table 4-7).

Net values of variable application

Corn. Net values of variable application based on soil zone information are mostly negative except for F8B due to the additional operation cost for variable application (there is no cost for soil zone information generation). There is a clear field-level overapplication with the soil zone uniform strategy for all fields. However, variable application based on soil zone information does not reduce the overapplication of N when using the soil zone uniform strategy, probably due to the homogeneity of these fields in terms of soil type distribution.

Net values of variable application based on functional zone information are positive for F2 and F8B and negative for F5 and F7. Some values are negative because of the additional operation cost for variable application. Where uniform application clearly over applied N, the saving on reduced N overapplication with variable application may not be enough to compensate for the increased operation cost for variable application.

Wheat. Net values of variable application based on soil zone information are all negative, indicating that when field-level overapplication was relatively small (e.g. F1B) or field-level underapplication occurs (F1C, F3B, and F4), variable application based on soil zone information reduced the farmer's net income. Soil zones were not very homogenous. There are costs of N underapplication for some SMUs which were in a poor soil zone but can achieve high yield. There are also costs from N overapplication to some SMUs which were in good soil zones but achieved low yields. These types of costs made the variable application by soil zone not profitable. Soil zone information was not adequate to direct variable application in order to increase net incomes.

Net values of variable application based on functional zone information are all negative except for F4. In F1B, saving on reduced overapplication from variable application did not compensate for the additional operation cost for variable application. Functional zone variable application may improve N efficiency as shown in F4, but because of the large errors in predicted field-level yields in 1999, variable application did not increase net income.

Strategy rankings

Soil zone uniform strategy had the highest returns in six of the eight fields mainly through less N overapplication (e.g. wheat in F1B and all corn fields) or less underapplication (e.g. wheat F1C, F3B, and F4) and because of the zero cost of soil zone information. However, the returns from the soil zone uniform strategy were higher than those of the conventional strategy on only 2 fields. Except F8B, the soil zone variable strategy had lower returns than the soil zone uniform strategy due to increased costs of N misapplication and/or increased operation cost. The functional zone variable strategy enhanced N efficiency in most cases over the functional zone uniform strategy but did not increase net return due to additional operation costs. In some cases where the functional zone uniform strategy clearly underapplied N at the field-level (e.g. F1C and F3B), the functional zone variable strategy resulted in additional loss in net return by further N underapplication on many SMUs within the field.

Discussion

Due to large errors in predicting field-level yields in 1999, evaluation of soil zone and functional zone information is very difficult. Where field-level yields were underestimated in the baseline, the functional zone uniform strategy further underapplied N. In most cases, the functional zone variable strategy shows improved N efficiency but may further increase underapplication of N in terms of net returns maximization. Uniform application generally did not underapply the high yielding SMUs, but variable application did underapply some SMUs classified as low yielding. Thus, if N was underapplied by uniform application, the variable application further underapplied N resulting in a larger loss in profit.

Accurate estimates of optimal yield are very important to optimizing net returns from N management. However, as shown in Table 4-7, estimated optimal yields in 1999 are much higher than actual yields in F1B, F2, F5, F7, and F8B, and lower than actual in

F1C, F3B, and F4. Inaccurate estimates of optimal yields for 1999 makes it difficult to determine the values of spatial information and variable application based on spatial information. In the next section a sensitivity analysis was carried out to more clearly describe the performance of spatial information and variable N application for improved N management.

4.6. The value of spatial information without temporal uncertainty: a sensitivity analysis

As seen from Sections 4.4 and 4.5, large errors occurred in predicting yields for 1999, making the comparison of the effects of spatial information across fields very difficult. In this section, a sensitivity analysis is described which assumes that field-level yields in 1999 for all fields are certain and are equal to the actual 1999 yields as recorded by the yield monitor. The sensitivity analysis eliminates temporal uncertainty and analyzes the effect of spatial variability on N application rates, yields, net N reduction potential, and values of information and variable application for alternative strategies. This sensitivity analysis is different from the ex post analysis. First, the sensitivity analysis applies N by actual field-level yields for the conventional strategy and the soil zone strategies while the ex post analysis applied N with yield level recommended by VALUES (Simpson et al.). Second, the sensitivity analysis models 1999 within-field spatial variability at the functional zone level with yield monitor data in 1995 and 1997, while the ex post analysis used the observed SMU yield data for 1999.

Some modifications are made to the procedure for generating the model parameters in Chapter Three and conducting the sensitivity analysis. For the conventional strategy, because the field-level yields in 1999 are certain, optimal N application rates are calculated as yield*1.25 lb/ac for wheat and (yield – 20) lb/ac for corn. As discussed in Chapter 3, when field-level yield is certain, the conventional strategy applies N as if the whole field were homogeneous. Again the conventional strategy is used to determine a baseline.

For the soil zone variable strategy, the yield for each soil zone is the field-level yield times a soil zone rescaling factor as discussed in Section 3.4 and N rates are determined for each soil zone in the same way as for the field level. For the soil zone uniform strategy, the optimization procedure was used to determine the optimal target yields at the field level. The optimal yields for the whole fields are those for the good and predominant Pamunkey soils (171 and 172) in these fields as discussed in Section 4.2 and Section 4.3. For example, for F1B, the wheat yield in 1999 is 44 (Table 4-1), the rescaling factor for Pamunkey soil is 1.084 (Table 3-20), thus N rate for Pamunkey soil is based on a yield level of 44*1.084 = 47.7 bu/ac.

For the functional strategies, the same procedure described in Chapter 3 was used, with field-level rescaling factors as listed in Table 3-17 modified by using the actual observed field-level yields in 1999 to replace the expected yield in 1999 as listed in Table 3-17. Actual field-level yields in 1999 are listed in Table 4-1. For example, the scaling factor for F1B in 1995 now changes from 61/85 = 0.72 (Table 3-17) to 44/85 = 0.52. After field-level rescaling, the Markov chain model for within-field SMUs' transition was used to describe the spatial variation based on observed data. In short, the difference between previous ex ante analysis and current sensitivity analysis in the functional zone strategy procedures are the rescaling factors. Sensitivity analysis results for changes in N application rates, yields, and net N on the study fields in 1999 when field-level yield are certain are reported in Table 4-9.

Table 4-9. Sensitivity analysis of ex ante changes in N rate, yields, and net N relative to the baseline in 1999 when field-level yields are certain

Strategy	Wheat field	Change in N applied (lb/ac)	Change in yield (bu/ac)	Change in net N (lb/ac)	Corn Field	Change in N applied (lb/ac)	Change in yield (bu/ac)	Change in net N (lb/ac)
Soil zone-uniform	F1B	4	1	2	F2	2	1	1
Soil zone-variable		-1	-1	0		0	0	0
Func. zone-uniform		12	3	8		27	11	20
Func. zone-variable		6	3	3		14	11	7
Soil zone-uniform	F1C	16	6	9	F5	8	4	6
Soil zone-variable		0	-3	3		0	-2	2
Func. zone-uniform		19	7	11		30	9	27
Func. zone-variable		12	7	5		25	6	23
Soil zone-uniform	F3B	5	1	3	F7	2	1	1
Soil zone-variable		0	-1	1		0	1	0
Func. zone-uniform		21	4	17		11	5	8
Func. zone-variable		14	3	12		9	5	5
Soil zone-uniform	F4	7	3	4	F8B	12	6	9
Soil zone-variable		1	0	1		0	0	1
Func. zone-uniform		11	5	6		26	9	23
Func. zone-variable		8	5	2		19	9	16

N application rates

Corn. Except for the soil zone variable strategy, all strategies increased N rates from the baseline. Soil information alone increased slightly the N application rate (2 to 12 lb/ac), while functional zone information increased N application by 11 to 30 lb/ac from the baseline. N application increases from soil zone information alone are not large because all study fields are rather homogeneous in terms of soil distribution. With field-level yield certain, increased information on spatial variability brought about higher N application rates in all corn fields.

In all cases, variable applications reduced N rates relative to uniform application for given information levels. The soil zone variable strategy is almost identical to the baseline because the slightly increased N application rate in good soil zones is balanced out by the large decrease of N applications in the poor soil zones. The functional zone variable strategy still applied higher N rates than the baseline, but applied 15 to 50% less than its corresponding functional zone uniform strategy.

Wheat. Similar to corn, all strategies except the soil zone variable strategy increased N rates from the baseline, indicating that increased spatial information increased N rates when field-level yields are known. N application rate increases based on soil zone information alone range from 4 to 16 lb/ac, while N rate increases based on functional zone information range from 11 to 21 lb/ac. The range of increased N rates for soil zone information is smaller because the areas of poor soils in the fields are smaller than the areas of low-yielding functional zones.

In all cases, variable applications reduced N rates compared to corresponding uniform applications for a given information level. The soil zone variable strategy is very close to the conventional strategy at the field-level. The functional zone variable strategy still applied higher N rates than the conventional baseline indicating that with the absence of temporal variability more spatial information can increase N application from the baseline. Possibly with further research this result might be generalized to other cases where the predominant soils are lower-yielding rather than higher-yielding as in this study.

Yield changes

Corn. Yields were increased relative to the baseline through higher N input with more spatial information in all cases. The soil zone uniform strategy increased yields relative to the conventional strategy from 1 to 6 bu/ac, while the functional zone uniform strategy increased yields by 5 to 11 bu/ac relative to the conventional baseline.

Variable application based on soil zone information reduced yields relative to those with the uniform application in all cases except F7, indicating that soil zone information alone was not adequate to direct variable application without reducing yields in this study. By contrast, variable application based on functional zone information was able to reduce N application rates while not reducing yields (except F5).

Wheat. In all cases, increased spatial information increased yields relative to the baseline by increasing N applications when only uniform application was used. The soil zone information alone was able to increase yields by 1 to 6 bu/ac relative to the baseline. Functional zone information alone increased yields by 3 to 7 bu/ac relative to the baseline.

Variable application based on soil zone information in all cases reduced yields relative to uniform application. By contrast, variable application based on functional zone information did not decrease yield relative to uniform application except by 1 bushel on F3B.

Net N

Corn. Even without temporal variability, spatial information with uniform application increased net N in all cases. Net N increased from 1 to 9 lb/ac for the soil zone uniform strategy and from 8 to 27 lb/ac for the functional zone uniform strategy relative to the baseline.

Variable application was able to reduce net N relative to the corresponding uniform application in all cases. The soil zone variable strategy reduced net N by 1 to 8 lb/ac relative to the soil zone uniform strategy. The functional zone variable strategy reduced net N by 3 to 13 lb/ac relative to the functional zone uniform strategy. When yield losses are considered, the net N reduction for variable application based on soil zone information does not indicate an improvement of N use efficiency. However, the net N reduction for variable application based on functional zone information indicates an improvement of N use efficiency because yields were generally the same for the uniform and variable applications.

Wheat. Spatial information with uniform application increased net N in all cases. Net N increased from 2 to 9 lb/ac for the soil zone uniform strategy and from 6 to 17 lb/ac for the functional zone uniform strategy relative to the baseline.

Variable application reduced net N relative to the corresponding uniform application in all cases. The soil zone variable strategy reduced net N by 2 to 6 lb/ac relative to the soil zone uniform strategy. The functional zone variable strategy reduced net N by 4 to 6 lb/ac relative to the functional zone uniform strategy. When yield losses are considered, the net N reduction for variable application based on soil zone information does not indicate an improvement of N use efficiency, but the net N reduction for variable application based on functional zone information indicates an improvement of N efficiency.

The results of the sensitivity analysis for ex ante information values and values of variable rate application in 1999 when field-level yields are certain are listed in Table 4-10.

Table 4-10. Results of sensitivity analysis showing values of information and variable rate application in 1999 when field-level yields are certain

		Information value ($/ac)		Variable appl. value ($/ac)		Strategy
Strategy	Field	Gross	Net	Gross	Net	Sum ($/ac)
		Wheat				
Soil zone-uniform	F1B	2.63	2.63			2.63
Soil zone-variable				-5.66	-7.75	-5.12
Func. zone-uniform		6.41	3.80			3.80
Func. zone-variable				-0.80	-3.42	0.38
Soil zone-uniform	F1C	13.58	13.58			13.58
Soil zone-variable				-20.45	-22.54	-8.96
Func. zone-uniform		15.15	12.54			12.54
Func. zone-variable				0.56	-2.06	10.48
Soil zone-uniform	F3B	3.77	3.77			3.77
Soil zone-variable				-4.73	-6.82	-3.05
Func. zone-uniform		6.44	3.83			3.83
Func. zone-variable				-1.32	-3.94	-0.11
Soil zone-uniform	F4	6.46	6.46			6.46
Soil zone-variable				-6.80	-8.89	-2.43
Func. zone-uniform		9.31	6.70			6.70
Func. zone-variable				2.15	-0.47	6.23
		Corn				
Soil zone-uniform	F2	1.50	1.50			1.50
Soil zone-variable				-1.84	-3.93	-2.43
Func. zone-uniform		14.01	11.40			11.40
Func. zone-variable				4.52	1.90	13.30
Soil zone-uniform	F5	5.93	5.93			5.93
Soil zone-variable				-10.02	-12.11	-6.18
Func. zone-uniform		9.96	7.35			7.35
Func. zone-variable				-4.84	-7.46	-0.11
Soil zone-uniform	F7	1.39	1.39			1.39
Soil zone-variable				-0.38	-2.47	-1.08
Func. zone-uniform		6.03	3.42			3.42
Func. zone-variable				0.21	-2.41	1.01
Soil zone-uniform	F8B	7.94	7.94			7.94
Soil zone-variable				-7.92	-10.01	-2.07
Func. zone-uniform		10.95	8.34			8.34
Func. zone-variable				1.27	-1.35	6.99

Information values

Corn. Gross and net information values are positive in all cases. Without temporal variability, increased spatial information improved N management and increased net return. Gross functional zone information values (ranging from $6.03 per acre to $14.01 per acre) are higher than the corresponding gross soil zone information values (ranging from $1.50 per acre to $7.94 per acre) in all cases. Net functional zone information values are higher than the corresponding net soil zone information values in all cases. In this case study when yield uncertainty was eliminated at the field level, functional zone information generated from yield monitor observations alone was more useful than soil zone information generated from soil maps in terms of increasing returns to N applications.

Wheat. Information values are positive in all cases. Without temporal variability, increased spatial information improved N management and increased net returns. Gross functional zone information values (ranging from $6.41 per acre to $15.15 per acre) are higher than the corresponding gross soil zone information values (ranging from $2.63 per acre to $13.58 per acre) in all cases. Net functional zone information values are higher than the corresponding net soil zone information values except on field F1C. In this case study when yield uncertainty was eliminated at the field level, functional zone information generated from yield monitor observation alone was more useful than soil zone information generated from soil maps.

Variable application values

Corn. In all cases, the gross and net values of variable application based on soil zone information are negative. Soil zone information was not sufficient to direct variable application to further improve net return when temporal yield uncertainty was removed. In all cases except for F5, gross values of variable application based on functional information are positive. But except for F2, net values of variable application based on functional zone information are negative due to additional cost of variable application. In this study, variable application based on functional zone information improved N use efficiency but generally did not increase net returns.

Wheat. In all cases, gross and net values of variable application based on soil zone information are negative. Soil zone information was not sufficient to vary N applications and increase net returns. Gross values of variable application based on functional information are positive for F1C and F4 and negative for F1B and F3B. Net values of variable application based on functional information are negative because of the additional operation cost. Variable application based on functional zone information may improve N use efficiency but did not increase the farmer's net return in this study.

In summary, variable application based on soil zone information reduces net return because soil zone information did not adequately predict yield variations relative to the field average. Variable application based on soil zone information did not improve N use efficiency in all fields. Variable application based on functional zone information improved N use efficiency in most cases. However, variable application reduced net return relative to the uniform application. Net returns were reduced due to the cost of variable N application and the imbalance between the cost of one pound N overapplied and the cost of one pound N underapplied.

Strategy rankings

In all fields except F2, the soil zone uniform strategy and the functional zone uniform strategy resulted in higher net returns relative to variable applications. In this case study, spatial information improves net returns when temporal variability is absent. In wheat fields, the performance of the soil zone uniform strategy and that of the functional zone uniform strategy are comparable, while in the corn fields, the functional zone uniform strategy performed better than the soil zone uniform strategy.

In all cases, the soil zone variable strategy reduced the farmer's net return because of the increased yield loss from the baseline and the additional operation cost, indicating that for this case study soil zone information is not adequate to direct variable application even when no temporal variability is considered. With the exception of F2, the functional zone variable strategy has lower net return than the corresponding functional zone uniform strategy. The combination of additional operation cost for variable application and the imbalance between the cost of one pound N overapplied and that of one pound N underapplied makes variable application unprofitable even if temporal variability is not considered.

Major results from sensitivity analysis

Conditioning on given temporal observation (i.e. when field-level yields are certain), results from this sensitivity analysis show:

a). Information about spatial variability increased N input, yields, net N, and net return relative to the conventional strategy. Higher level information (functional zone information) resulted in higher N input, yields, net N, and net return than lower level information (soil zone information).

b). Variable application does not have higher values relative to uniform application, indicating lack of precision in predicting spatial variation. Variable application based on soil zone information decreased yields and net return due to yield loss from underapplication to some SMUs which were in low-yielding soil zones as well as wasted N input to some SMUs which were in high-yielding soil zones. Variable application based on functional zone information increased N use efficiency but generally did increase net return due to the imbalance between the costs of underapplication and overapplication. Underpredictions of site yield potential have high costs.

4.7. Summary of the chapter

Temporal and spatial variability in crop yields

In this chapter, observed yield variability for 1995-1999 on the study farm is discussed for eight fields. Yield variations are described by field, by soil type within each field, and by functional zone in each field. Observed temporal variability is large for both wheat and corn. Temporal variability for wheat fields is smaller than for corn fields. Observed spatial variability is also big and can be described in several ways. First, differences between fields are large even though all fields are rather homogeneous in terms of soil distribution. Some fields achieved consistently higher yields than other fields while some fields achieved much higher yields in some years and achieved much lower yields in other years relative to other fields. Second, the differences of yields by soil type are significant for each field in each year. Soil types with similar yield potentials according to VALUES also differ from each other in observed yield level and variability among the SMUs. Third, each field can be divided into several functional zones according to the observed yields for each SMU in a field for a given year (as in ex post analysis) or over several years (as in ex ante analysis). Fourth, the variation among the SMUs within each soil type for a given field is often large indicating the soil zones may not be precise enough to direct N application. By contrast, the variation among the SMUs in a functional zone generated from observed yield for each year (ex post) is generally very small.

Ex post analysis results

In this case study, the profit maximizing strategy is to apply N according to the needs of the SMUs which are of high yield potential. The profit maximizing strategy applies N according to the predominant good soil type within a field and disregards minor poor soils because savings from any lower N rate to reduce overapplication in the minor (and poor) soil zones would be more than offset by the cost from N underapplication to the predominant (and good) Pamunkey soils in any given field. As a result, the soil zone uniform strategy is identical (i.e. N application rates are equal) to the conventional strategy.

In this case study, soil zone information was not adequate to direct variable N application to increase net returns. On average, variable application based on soil zone information slightly reduced net N from the field but clearly reduced the farmer's net return in each field. Variable application underapplied N for SMUs which were in poor soil zones but achieved high yields and overpplied N for SMUs which were in good soil zones but achieved low yields.

Functional zone information alone did not always reduce net N from each field. However, on average, uniform variable rate application based on functional zone information both reduced net N from each field and increased net return for each field. The functional zone variable application strategy reduced net N and increased net returns compared to the conventional strategy and the functional zone uniform strategy. The performance of the functional zone variable strategy was helped by the fact that functional zones were precisely identified by using the actual yields of the SMUs (as reflected in part by the small CVs) so both overapplication and underapplication were controlled for each SMU.

Ex ante analysis results

In this case study, soil zone or functional zone information with uniform application either increased or decreased N rates and net N for a field from the baseline. Variable application always reduced N rate and net N for a field compared to the uniform application for a given information level.

Soil zone variable application did not always improve N efficiency. Soil zone information alone was not adequate to direct variable application in order to increase the farmer's net return.

Variable application based on functional zone information improved N use efficiency, but did not increase the farmer's net return. Returns declined due to higher yield loss from underapplication of N to some SMUs and additional operation cost as compared with the uniform application based on functional zone information. In some cases, functional zone information with uniform application led to optimal target yields lower than the actual yields. Variable application further underapplied N, resulting in larger losses in net returns. In some other cases, functional zone information with uniform application led to target yields that were higher than the actual yields. However, in these cases, the savings from less wasted N with variable application were offset by the increased application cost. Because the cost of underapplication was much larger than the cost of overapplication, the performance of variable application was heavily affected by errors in temporal as well as spatial predictors of yields. As a result, returns from variable application were lower than returns from uniform application. Errors in the temporal predictor were reflected by the fact that the expected yields and target yields were either over- or under-estimated for 1999. This type of error was caused by the annual yield simulation procedure. Errors in the spatial predictor were reflected by the fact that the relative yields among the SMUs in 1999 were estimated with insufficient precision. This type of error was caused by the lack of site-specific information which describes the dynamic interactions among various yield-determining factors.

In general, large errors in predicting yields made it very difficult to compare the effects of soil zone information and functional zone information and their corresponding variable applications. In ex ante analysis, functional zone strategies depended on the prediction of field-level yields before utilizing the observed spatial information concerning relative yields among the SMUs. Yields for some fields were over-estimated while for others the yields were underestimated and thus the effects on net returns of information and variable application were mixed.

Sensitivity analysis results

With certain field-level yields, higher information levels led to higher N input. Spatial variability described by functional zone information led to higher N application rates than application rates with soil zone information.

Both variable application based on soil zone information and variable application based on functional zone information reduced N rates and net N as compared with corresponding uniform application. However, yields were lower for soil zone variable application compared to uniform application and as a result N use efficiency did not improve.

Values of soil zone and functional zone information with uniform application were positive in all cases, indicating that increased information was able to increase the farmer’s net return in this case study. The awareness of heterogeneity within the field itself led to a higher N application rate which increased net returns due to the large imbalance between the cost of one pound N under applied and the cost of one pound N over applied.

Soil zone information was not adequate to direct variable application in order to increase the farmer’s net return. Uncertainty about each zone’s yield relative to the overall yield led to losses from underapplication and overapplication of which underapplication losses were higher. In all cases, the soil zone variable strategy had lower net return as compared with the baseline even though the soil zone uniform strategy had higher net return compared with the baseline.

The functional zone variable strategy maintained or increased net return relative to the baseline in all cases. However, the functional zone variable strategy did not generally increase net return relative to the functional zone uniform strategy due to the additional cost for variable application as well as the errors in functional zoning.

The effects of information and variable application in terms of reduction of N rates, reduction of yield loss and reduction of net N from the baseline are similar for wheat and corn. There are no clear differences between wheat and corn in terms of information values and variable application values.

Just as the awareness of temporal yield variability resulted in an optimal field-level target yield which was higher than the expected field-level yield, the awareness of spatial yield variability within a field also resulted in an optimal field-level target yield which was higher than the actual field-level yields. Thus, it is concluded that larger awareness of spatial variability led to higher N input, other things equal, for a profit maximizing farmer in this case study.

Chapter Five. Summary and Conclusions

This chapter reviews the problem statement, objectives, study area, the theoretical framework, and empirical model developed (Section 5.1) and the major findings of the research (Section 5.2). Limitations of the study and suggestions for further study are discussed (Section 5.3), and the policy implications from the study are also presented (Section 5.4).

5.1. Study objectives and procedures

Nitrogen (N) is both the most important input to grain crops and one of the major pollutants to the environment from agriculture in the United States. Recent developments in site-specific management (SSM) technology utilizing global positioning systems (GPS), geographic information systems (GIS), remote-sensing, grid sampling, yield monitors, and variable rate applicators have the potential to decrease both overapplication and underapplication of N and thus increase N use efficiency and farmers' net returns. In Virginia, due to the high variability of within-field yield-limiting factors such as soil physical properties and fertility, the adoption of SSM is hindered by the costs of grid-sampling (Alley; Anderson-Cook et al.). However, many big Virginia corn-wheat/soybean farms have been generating yield maps using yield monitors for several years even though few variable applications based on yield maps were reported. It is unknown if the information generated by yield monitors under actual production situations can be used to direct N management for increased net returns in Virginia.

The overall objective of the study is to analyze the economic and environmental impacts of alternative management strategies for N for corn and wheat/soybean production in eastern Virginia based on site-specific information. Specifically, evaluations of expected net returns and net N (applied N that is not removed by the crop) are made at three levels of site-specific information regarding crop N requirements combined with variable and uniform N application for corn-wheat/soybean fields in eastern Virginia. Ex post and ex ante evaluations of information are carried out.

Five N application strategies were identified for evaluation: the conventional strategy which treats the field as homogeneous in yield potential and N needs, the soil zone uniform strategy which treats different soil zones within a field as different management zone while applying N uniformly over the whole field, the soil zone variable strategy which applies different N rates to different soil zones, the functional zone uniform strategy which groups the smallest management units (SMUs) within a field into different functional zones according to observed yield patterns of these SMUs while applying N uniformly over the whole field, and the functional zone variable strategy which applies different N rates to different functional zones. The farmer was assumed to maximize the expected net return by determining the optimal N application rate for each given strategy. The value of information is evaluated based on the differences in expected net returns between returns obtained with that level of information (uniform application) and the conventional strategy. The value of variable application is evaluated based on the differences of expected net returns obtained with uniform application and with variable application at given information level. In both ex post and ex ante analyses, the conventional strategy was used as the baseline.

In Chapter 2, the theoretical framework was discussed. The farmer was assumed to choose a N application rate to maximize the expected net return from each field for a given N application strategy. A linear response-plateau production function was used to describe the crop yield from N application. In this setting, for a given strategy, the expected net return was maximized when the cost of N misapplication was minimized. The cost of N misapplication came from wasted N input with overapplication and yield loss with underapplication. A standard nonpreemptive goal-programming model was used as justification for a search procedure to minimize the cost of N misapplication. The search procedure was programmed using (SAS) (SAS Institute). A linear response-plateau production function was used to estimate the relationship between N application and crop yield.

Cluster analysis was introduced to group the smallest management units (SMUs) of a field into functional zones based on their yield potentials and N application needs. It was assumed that the farmer treats individual functional zones as homogeneous. Cluster analysis was used in both ex post and ex ante analyses.

For ex ante analysis, Markov chain models were introduced to model both the temporal and spatial transition of crop yields within fields. The Markov chain model for temporal yield transition describes yield transition of one unit of land over many periods of time. This model used historical farm-level data. The Markov chain model for spatial transition describes yield transition of many units of land within a field over a few periods. This model used yield monitor data for each field over 1995-1997. Based on the generated distributions of future yield for a field or a smallest management unit (SMU), the search procedure was able to determine the ex ante optimal N rates for each strategy.

In Chapter 3, a case study farm including eight case fields totaling 409.2 acres in the Prince George County, Virginia was described. The case study farm had been generating yield maps using a yield monitor on the combine since 1995. The yield monitor data from 1995 to 1999 were used.

Based on a literature review, the input and output prices and production costs were determined for the study. For each strategy, production cost was decomposed into the strategy-specific cost and site-specific cost. Strategy-specific cost included the cost to generate information and the additional cost for variable application. Site-specific cost included the cost from N input, from yield loss, and from crop hauling. Empirical formulas for N application rate, yield loss due to underapplication, and net N were developed.

A SMU was described as a 98.425x98.425 ft^2 (30x30 m^2) grid composed of 100 9.8425x9.8425 ft^2 (3x3 m^2) grids which were generated from yield monitor data. Based on yield monitor data, yield maps were generated for each field in each year on the scale of SMU. These SMUs were used to develop functional zones.

In the ex post analysis, the optimal procedure applied N for the conventional strategy at the plateau yield levels for the predominant Pamunkey soils as suggested by the Virginia Agronomic Land Use Evaluation System (VALUES). The soil zone strategies searched for optimal target yields ranging from the lowest plateau yield levels for the poor soils in the field to the highest plateau yield levels for the good soils in the field. The conventional strategy and the soil zone strategies applied N according to the VALUES recommendation. The conventional strategy was identical to the soil zone uniform strategy and both strategies applied N according to the predominant Pamunkey soils on all fields.

In ex post analysis, the functional zone strategies searched for optimal target yields ranging from the lowest yields of the functional zones as recorded by the yield monitors to the highest yields of the functional zones as recorded by the yield monitor. The functional zone strategies applied N according to observed yield monitor data on the SMU scale. In uniform strategies (the soil zone uniform strategy and the functional zone uniform strategy), highest net return for the whole field was based on the area-weighted returns for the sub-zones (soil zones or functional zones) at a uniform N application rate.

In the variable strategies (the soil zone variable strategy and the functional zone variable strategy), the highest net return for the whole field was the sum of highest returns for the sub-zones (soil zones or functional zones) with each sub-zone possibly employing different N application, to achieve the highest net return for that sub-zone. Assuming the actual yield data observed by yield monitors were the true yield potentials for observed areas in the fields in a given year for a given crop, the optimization procedure determined N application rate by minimizing costs of N misapplication (due to overapplication and/or underapplication).

In the ex ante analysis, the optimization procedure in the ex post analysis was modified to accommodate the yield transition probabilities. Field-level yield sequences were generated at the farm level by Erosion-Productivity Impact Calculator (EPIC) using historical weather data from 1922-1998. Farm yields were rescaled by observed field-level yield data over 1992-1998. The conventional strategy searched for the optimal yields among the simulated yield sequences with estimated temporal Markov chain transition probabilities. Similarly, the soil zone strategies also searched for optimal yields for soil zones with estimated temporal Markov chain transition probabilities. The functional zone strategies searched for the optimal yields among observed SMU yields in 1995 and 1997 with estimated spatial Markov chain transition probabilities attached. The optimization procedures for both ex post and ex ante analyses were programmed in the SAS PROC IML.

In the ex ante analysis, optimal yields were always higher than expected yields when temporal and/or spatial variability were considered. Increased information led to higher expected costs of N misapplication.

5.2 Results and discussion

In Chapter 4, the actual temporal and spatial yield variability for each field over 1995-1999 were discussed at field-level, within-field soil zone level, and within-field functional zone level by crops. Ex post results were presented for the N application rates, yield loss due to N misapplication, and net N for each strategy. The gross and net ex post and ex ante information values and gross and net values of variable application were presented as well. Due to large errors in predicting the field-level yields for 1999, a sensitivity analysis was carried out to remove the temporal uncertainty in order to evaluate the spatial information and variable application values.

Temporal and spatial variability in crop yields

Observed temporal variability is large for both wheat and corn. Temporal variability for wheat fields is smaller than for corn fields. Observed spatial variability is large even though all fields are very homogeneous in terms of soil types. First, differences between fields are large even though all fields are rather homogeneous in terms of soil distribution. Some fields achieved consistently higher yields than other fields while some fields achieved much higher yields in some years and much lower yields in other years relative to other fields. Second, the differences in yields by soil type are significant for each field in each year. Soil types with similar yield potentials according to VALUES also differ from each other in observed yield level and variability among the SMUs. Third, each field can be divided into several functional zones according to the observed yields for each SMU in a field for a given year (as in ex post analysis) or over several years (as in ex ante analysis). Fourth, the variation among the SMUs within each soil type for a given field is often large indicating the soil zones may not be precise enough to direct N application. By contrast, the variation among the SMUs in a functional zone generated from observed yield for each year (ex post) is generally very small.

Ex post analysis results

In this case study, when spatial variability is present, the profit maximizing strategy is to apply N according to the needs of the SMUs which are of high yield potential. The profit maximizing strategy applies N according to the predominant good soil type within a field and disregards minor and low-yielding soil zones because savings from any lower N rate to reduce overapplication in the minor and low-yielding soil zones would be more than offset by the cost from N underapplication to the predominant and high-yielding Pamunkey soils in any given field. As a result, the soil zone uniform strategy is identical to the conventional strategy.

In this case study, soil zone information was not adequate to direct variable N application to increase net returns. On average, variable application based on soil zone information slightly reduced net N but clearly reduced the farmer’s net return in each field. Variable application underapplied N for SMUs which were in poor soil zones but achieved high yields and overapplied N for SMUs which were in good soil zones but achieved low yields, resulting in both yield loss and N waste.

Functional zone information alone did not always reduce net N from each field. However, on average, uniform application based on functional zone information reduced net N from each corn field and increased net returns for all fields. The functional zone variable application strategy reduced net N and increased net returns compared to the conventional strategy and the functional zone uniform strategy. The better performance of the functional zone variable strategy was due to the fact that functional zones were precisely identified based on the actual yields of the SMUs so both N overapplication and underapplication were minimized for each SMU.

Ex ante analysis results

In this case study, soil zone and functional zone information with uniform application could either increase or decrease N rates and net N for a field relative to the baseline. Variable application always reduced the N rate and net N for a field as compared with the uniform application for a given information level. However, in general, large errors in predicting yields made it very difficult to compare the effects of soil zone information and functional information and their corresponding variable applications.

Soil zone variable application did not always improve N efficiency. Soil zone information alone was not adequate to direct variable application in order to increase the farmer's net return.

Variable application based on functional zone information improved N use efficiency, but did not increase the farmer's net return relative to the baseline or in most cases relative to uniform application. Returns declined due to higher yield loss from underapplication of N to some SMUs and additional operation cost as compared with the uniform application based on functional zone information. In some cases where functional zone information led to target yields lower than the actual yields for uniform application, variable application further underapplied N resulting in larger losses in net return. In other cases where functional zone information with uniform application led to target yields that were higher than actual yields. However, the savings from reduced N overapplication with variable application were offset by the increased application cost. Because the cost of underapplication was much larger than the cost of overapplication, the performance of variable application was sensitive to errors in predicted site (SMU) yields.

Large errors in predicting site yields caused returns from variable application to be lower than returns from uniform application. Site yield errors were affected by errors in the simulation procedure in predicting farm and field-level yield potential. Spatial prediction errors caused relative yields among the SMUs in 1999 to be estimated with insufficient precision. Spatial prediction errors resulted from the lack of sufficient site-specific information which describes the dynamic interactions among various yield-determining factors.

Sensitivity analysis results

With certain field-level yields, higher information levels led to higher N input. Spatial variability described by functional zone information led to higher N application rates than application rates with soil zone information. The effects of information and variable application in terms of reduction of N rates, reduction of yield loss and reduction of net N from the baseline are similar for wheat and corn.

Both variable application based on soil zone information and variable application based on functional zone information reduced N rates and net N as compared with corresponding uniform application. However, yields were lower for soil zone variable application compared to uniform application, and, as a result, N use efficiency did not improve.

Information values were positive in all cases, indicating that increased information was able to increase the farmer's net return in this case study. The awareness of heterogeneity within the field itself led to a higher N application rate which increased net returns due to the large imbalance between the cost of one pound N under applied and the cost of one pound N over applied.

Soil zone information was not adequate to direct variable application in order to increase the farmer's net return. Uncertainty about each zone's yield relative to the overall yield led to losses from underapplication and overapplication of which underapplication losses were higher. In all cases, the soil zone variable strategy had lower net returns as compared with the baseline while the soil zone uniform strategy had higher net returns compared with the baseline.

The functional zone variable strategy maintained or increased net return relative to the baseline in all cases. However, the functional zone variable strategy did not generally increase net return relative to the functional zone uniform strategy mainly due to the additional cost for variable application.

Just as the awareness of temporal yield variability resulted in an optimal field-level target yield which was higher than the expected field-level yield, the awareness of spatial yield variability within a field also resulted in an optimal field-level target yield which was higher than the actual field-level yields. Thus, it is concluded that awareness of yield variability led to higher N input for a profit maximizing farmer in this case study. This result confirms the findings of Babcock.

5.3. Limitations and suggestions for further research

Limitations

The modeling of temporal yield variability in this study resulted in large errors compared to observed farm yields. First, the EPIC-simulated farm-level yield sequence over 1922-1998 may not fully reflect farm yield variability for corn and wheat. Second, errors were introduced in translating simulated farm level yields to field yield which may have lowered the accuracy of the field yield transition probability. Third, the errors in constructing field transition probabilities may have affected the within-field yield transition probabilities and optimal N application amounts.

The Markov chain model for spatial variability was not statistically significant in describing the yield transition of a SMU. Markov chain transitions existed only for two fields. A possible reason for the lack of significance is that because of the rotation, the crop yields used were two years apart. The underlying assumption was that the within-field Markov transition was caused by a first-order transition of the weather condition which is one-year apart. It is not surprising to see that Markov chain thus modeled did not exist for most fields. This means it is very difficult to model the transition of spatial variability by using yield monitor data on multi-year rotations.

The production function used in this study made the crop yield very sensitive to N underapplication while insensitive to increased N application beyond a certain level. Using the linear response-plateau production function, the ratio of N price to crop price always drives the farmer to apply for the plateau yield. At that point, the cost of one pound N underapplication is several times larger than one pound N overapplication. Thus there is an imbalance between the effects of N applications that are too low and applications that are too high. The form of production function used resulted in high penalties for yield prediction errors which may have contributed to the low values of information in the ex ante analysis.

The ex ante analysis in this study was performed for only one year. Multiple year evaluations may have shown how the effects of learning and updating of information result in increased information values.

Implications for further research

Study results suggest two major areas for further research. First, research is needed on better ways to capture variations in yield potential within fields. Second, strategies for N management must be tested under conditions faced by farmers as they make N management decision. Ex ante as well as ex post information evaluations are needed.

Lack of precision in delineating soil zones may have been a major reason for the inadequacy of soil zone information to direct N application. The resolution of soil maps used in this study was one hectare[47], while the acreage for a SMU is less than one tenth hectare. The N application decision in soil zone strategies was based on the resolution of one hectare while the evaluation of soil information and variable application was based on observed yield on sites 30x30 m^2 in size. Further study should be carried out using finer resolution soil maps. Costs versus benefits of improving soil map resolution will need to be considered.

Lack of precision in the spatial predictor for functional zones is also a major reason for the unsatisfactory performance of variable application based on functional zone information. Further study should be carried out on quantifying the "break-even" spatial

---

[47] Even though the soil type data were provided (derived) at 1/9 hectare scale, the original data was obtained at 1 hectare resolution.

variability which calls for variable application. Both ex post and ex ante analyses are need in this direction.

With improved soil maps available, future study could compare functional zones generated within soil zones and functional zones generated across soil zones. This comparison study may show that functional zone information is more descriptive of the spatial variability with increased accuracy of soil information.

Farmers may have more knowledge about field variability than is captured by soil maps and yield monitor data. Future studies should consider how to incorporate farmers' observations and opinions into descriptions of field-level and within-field yield variability.

Further study may incorporate other information in modeling temporal and spatial yield variability. Yield monitoring information data may not be sufficient to direct SSM. Other information readily available includes elevation, slope, closeness to woods, closeness to water bodies, drainage, and specific pest situations. Of particular importance is research to reveal spatial variations in crop N requirements shortly before N is applied. Aerial photographs of crop condition may be especially helpful for this purpose.

Future study should consider the effects of risk aversion on the net return, N application, and net N as well as values of information and variable application. A risk averse farmer may be more concerned about income losses and thus may be inclined to apply higher N to reduce the risk of large yield losses in certain years. However, increased information about within field yield variability may reduce the N application rates for a risk averse farmer.

Improved procedures for ex ante evaluation of N management strategies from a farmer's perspective are needed to give a better indication of the value of spatial information. These procedures could focus on how farmers form subjective farm and field yield probabilities and how these probabilities would be combined with site specific yield information in making N application decisions.

5.4. Implication for policy

More information about spatial yield variability may reveal soils and farming regions with high potential for N overapplication. Site specific information can be targeted to reduce N input to these areas. With reduced cost and increased capability of the advancing information technology, the public may establish information databases to monitor the variability of crop yield and production inputs. Through education, extension services, and industry support, farmers may be encouraged to adopt variable application to increase efficiency and net returns.

The increased information and awareness of yield variability within a field may bring higher N input and thus increase N pollution potential. Thus the public needs to find ways to encourage farmers to incorporate environmental considerations into their decision making. For example, crop insurance can be designed to compensate farmers for losses from N underapplication based on spatial information.

At any given information level, variable application may lead to less N input and thus less net N as compared to uniform application. But the use of variable application may be hindered by the increased cost of variable application. Thus, the public may need to encourage variable application by providing subsidies to support these costs. The public may monitor N application on farms to identify the fields where SSM has the largest potential to reduce N pollution. Targeting public resources to these areas may be the most cost effective means for promoting SSM.

The imbalance between the cost of N underapplication and overapplication prompts the profit maximizing farmers to prefer overapplication to underapplication. The profit maximizing strategy is to apply N for the highest crop yields even though in most years the highest crop yields are not realized. Thus, the public should encourage the use of crop monitoring tools such as tissue tests or aerial photographs to provide information for within-season N needs.

References

Addiscott, T.M., A.P. Whitemore, and D.S. Powlson. *Farming, Fertilizers and the Nitrate Problem.* CAB International, Wallingford, 1991.

Agricultural Conservation Innovation Center (ACIC). Promoting Conservation Innovation in Agriculture Through Crop Insurance. http://www.agconserv.com/risk.htm. Accessed in June 1998.

Alley, M.M. *Personal Communication*. Professor, Crop and Soil Environmental Sciences Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Alley, M.M., P. Scharf, D.E. Brann, W.E. Baethgen, and J.L. Hammons. *Nitrogen Management for Winter Wheat: Principles and Recommendations.* Publication number 424-026, Virginia Cooperative Extension Service, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1996

Anderberg, M.R. *Cluster Analysis for Applications,* Academic Press, New York, 1973.

Anderson, T.W., and L.A. Goodman. “Statistical Inference About Markov Chains.” *Annals of Mathematical Statistics*, 28(1) (1957): 89-110.

Anderson-Cook, C.M., M.M. Alley, R. Noble, R. Khosla. “Phosphorous and Potassium Fertilizer Recommendation Variability for Two Mid-Atlantic Coastal Plain Fields.” *Soil Sciences Society of America Journal*, 63(1999): 1740-1747.

Asim, M., H. Mapp, and J. Solie. “Using Spatial and Temporal Variability to Identify Site-Specific Management Areas in Wheat.” Paper presented at AAEA annual meeting, Tampa, Florida, 1-4 August 2000.

Atherton, B.C., M.T. Morgan, S.A. Shearer, T.S. Stombaugh, and A.D. Ward. “Site-Specific Farming: A Perspective on Information Needs, Benefits and Limitations.” *Journal of Soil and Water Conservation*, 54(2)(1999): 455-461.

Babcock, B.A. “The Effects of Uncertainty on Optimal Nitrogen Applications.” *Rev. Agric. Econ*. 14 (2) (1992): 271-80.

Babcock, B.A., and A.M. Blackmer. “The Value of Reducing Temporal Input Nonuniformities.” *J. Agric. Res. Econ*. 17(2)(1992): 335-347.

Bartlett, M.S. “The Frequency Goodness of Fit Test for Probability Chains”, *Proc. Cambridge Philos. Soc.*, 47(1951): 86-95.

Bernardo, D.J., H.P. Mapp, G. J. Sabbagh, S. Geleta, K. B. Watkins, R. L. Elliott, and J. F. Stone. “Economic and Environmental Impacts of Water Quality Protection Policies. 1. Framework for Regional Analysis.” *Water Resources Research*. 29(9)(1993): 3069-3079.

Blackmer, A.M., and S.E. White. "Using Precision Farming Technologies to Improve Management of Soil and Fertilizer Nitrogen." *Australian J. Agric. Res.* 49 (1998): 555-64.

Bock, B.R., and G.W. Hergert. "Fertilizer nitrogen management". p.139-164. In *Managing Nitrogen for Ground Water Quality and Farm Profitability*, eds. R.F. Follett et al. SSSA, Madison, WI., 1991.

Bosch, D.J., K.O. Fuglie, and R.W. Keim. *Economic and Environmental Effects of Nitrogen Testing for Fertilizer Management*. Staff Report No. AGES9413. Resources and Technology Division, USDA-ERS, Washington D.C., April 1994.

Bosch, Darrell J., James W. Pease, Sandra S. Batie, and Vernon O. Shanholtz. *Crop Selection, Tillage Practices, and Chemical and Nutrient Applications in Two Regions of the Chesapeake Bay Watershed.* Virginia Water Resources Research Center Bulletin 176. Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1992.

Braden, John, B., Gary V. Johnson, Aziz Bouzaher, and David Miltz. "Optimal Spatial Management of Agricultural Pollution." *Amer.J.Agr.Econ.* 71(1989): 404-413.

Brann, D.E. *Personal Communication*. Professor, Crop and Soil Environmental Sciences Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Bregt, A.K. "GIS Support for Precision Agriculture: Problems and Possibilities." *Ciba Foundation Symposium 210:Precision Agriculture: Spatial and Temporal Variability of Environmental Quality*, John Wiley & Sons, New York, 1997.

Bullock, D.S., S. Swinton, and J. Lowenberg-DeBoer. "Can Precision Agricultural Technology Pay for Itself? – The Complementarity of Precision Agriculture Technology and Information." Paper presented at AAEA 2001 Learning Workshop on Spatial Data Analysis, Chicago, Illinois, August 4, 2001.

Bundy, L.G., and T.W. Andraski. "Soil Yield Potential Effects on Performance of Soil Nitrate Tests." *J. Produc. Agric.* 8(1995): 561-567.

Burrough, P.A., and J. Swindell. "Optimal Mapping of Site-Specific Multivariate Soil Properties." *Ciba Foundation Symposium 210: Precision Agriculture: Spatial and Temporal Variability of Environmental Quality*, John Wiley & Sons, 1997.

Cantor, P. "Health Effects of Agrichemicals in Groundwater: What Do We Know?" *Agricultural Chemicals and Groundwater Protection: Emerging Management and Policy*. Proceedings of a Conference held in St. Paul, MN, October 22-23, 1988.

Carpentier, C. L, D. J. Bosch, and S. S. Batie. "Using Spatial Information to Reduce Costs of Controlling Agricultural Nonpoint Source Pollution." *Agricultural and Resource Economics Review.* 27(April 1998): 72-84.

Carr, P. M., G. R. Carlson, J. S. Jacobsen, G. A. Nielsen, and E. O. Skogley. "Farming Soils, Not Fields: A Strategy for Increasing Fertilizer Profitability." *J.Prod.Agric.*, 4(1)(1991): 57-61.

Cerrato, M.E., and A.M. Blackmer. "Comparison of Models for Describing Corn Yield Response to Nitrogen Fertilizer." *Agronomy Journal*, 82(1990): 138-143.

Cox, F.R. "Economic Phosphorus Fertilization Using a Linear Response and Plateau Function." *Communication in Soil Sciences and Plant Analysis*. 27(3&4)(1996): 531-543.

Cressie, N. "Geostatistical Analysis of Spatial Data." *Spatial Statistics and Digital Image Analysis*. National Research Council, National Academy Press, Washington D.C., 1991.

Davis, J.G., G.L. Malzer, P.J. Copeland, J.A. Lamb, P.C. Robert, and T.W. Bruulsema. "Using Yield Variability to Characterize Spatial Crop Response to Applied N. In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Earl, R., P. N. Wheeler, B. S. Blackmore, and R. J.Godwin. "Precision Farming: the Management of Variability." Landwards, V.51(4)(1996): 18-23.

Eghball, B., and J.F. Power. "Fractal Description of Temporal Yield Variability of 10 Crops in the United States." *Agronomy Journal*, 87(1995): 152-156.

Eghball, B., G.D. Binford, J.F. Power, D.D. Baltensperger, and F.N. Anderson. "Maize Temporal Yield Variability Under Long-term Manure and Fertilizer Application: Fractal Analysis." *Soil Sci. Soc. Am. J.* 59(1995): 1360-1364.

English, B.C., R.K. Roberts, and S.B. Mahajanashetti. "Spatial Break-Even Variability for Variable Rate Technology Adoption." In *Proceedings of the Fourth International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson. Madison, WI: ASS-CSSA-SSSA, 1998.

Everett, M.W., and F.J. Pierce. "Variability of Corn Yield and Soil Profile Nitrates n Relation to Site-Specific N Management." In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Everitt, B. *Cluster Analysis*. London: London: Heinemann Educational Center [for] the Social Science Research Council, 1974.

Feinerman E, E. K. Choi, and S. R. Johnson. "Uncertainty and Split Nitrogen Applications in Corn Production", *Amer. J. Agri. Econ*. 72(November 1990): 975-984.

Ferguson, R.B., C.A. Gotway, G.W. Hergert, and T.A. Peterson. "Soil Sampling for Site Specific Nitrogen Management." In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Fiez, T.E., B.C. Miller, and W.L. Pan. "Assessment of Spatially Variable Nitrogen Fertilizer Management in Winter Wheat." *J. Prod. Agric*. 7(1994): 86-93.

Finke, P.A. "Field Scale Variability of Soil Structure and Its Impact on Crop Growth and Nitrate Leaching in the Analysis of Fertilizing Scenarios." *Geoderma* 60(1993): 89-109.

Food and Agricultural Policy Research Institute. "*FAPRI 1998: U.S. Agricultural Outlook*." Staff Report #1-98, August 1998. Iowa State University, University of Missouri-Columbia.

Franzen, D.W., L.J. Cihacek, and V.L. Hofman. "Variability of Soil Nitrate and Phosphate Under Different Landscapes." In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Fuglie, O. Keith, and Darrell J. Bosch. "Economic and Environmental Implications of Soil Nitrogen Testing: A Switching-Regresssion Analysis". *Amer.J.Agr.Econ.* 77 (November 1995): 981-900.

Gbur, E., and C.D. Steelman. "A Markov Chain Model To Assess Resistance of Cattle to Horn Flies." *Proceedings of the Kansas State University Conference on Applied Statistics in Agriculture*, April 26-28, 1992, Department of Statistics, Kansas State University, Manhattan, Kansas, 1992.

Goh, S., C. Shih, M. J. Cochran, and R. Raskin, "A Generalized Stochastic Dominance Program for the IBM PC". *Southern Journal of Agricultural Economics.* 21(December 1989): 175-182

Goodman, A.S., and K.A. Edwards. "Integrated Water Resources Planning." *Natural Resources Forum.* 16(1)(1992): 65-70.

Goolsby, D.A., W.A. Battaglin, and E.M. Thurman. "Occurrence and Transport of Agricultural Chemicals in the Mississippi River Basin, July Through August 1993. *Floods in the Upper Mississippi River Basin*, U.S. Geological Survey Circular, 20(1993): 1120-, U.S.G.P.O., D.C., 1993.

Gotway, C.A., R.B. Ferguson, and G.W. Hergert. "The Effects of Mapping and Scale on Variable-Rate Fertilizer Recommendations for Corn." In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Groffman, P.M. "Ecological Constraints on the Ability of Precision Agriculture to Improve the Environmental Performance of Agricultural Production Systems." *Ciba Foundation Symposium 210: Precision Agriculture: Spatial and Temporal Variability of Environmental Quality*, John Wiley & Sons, New York, 1997.

Gupta, R.K., S. Mostaghimi, and P.W. McClellan. "Multi-Site Evaluation of Precision Farming Technology for Coastal Plains of Virginia-Implications on Fertilizer Input Applications." Unpublished article, Biological Systems Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1997.

Hall, J., and Ciannat Howett. "Albemarle-Pamlico: Case Study in Pollutant Trading: Most of the Nutrients Came from Nonpoint Sources." *EPA-Journal.* 20(Summer 1994): 27-29.

Harris, D. "Risk Management in Precision Farming." In *Precision Agriculture '97: Proceedings of the First European Conference on Precision Agriculture*, eds. J.V. Stafford, BIOS Scientific Publisher Ltd., 1997.

Heathcote, I.W. *Integrated Watershed Managemen: Principles and Practices*. John Wiley & Sons, Inc., New York, 1998.

Heimlich, R. “Precision Agriculture: Information Technology for Improved Resource Use.” *Agricultural Outlook*, 250(April 1998): 19-23, USDA, ERS, Herndon, Virginia.

Hergert, G. W., W. L. Pan, D. R Huggins, J. H. Grove, and T. R. Peck. “Adequacy of Curent Fertilizer Recommendations for Site-Specific Management.” In *The State of Site Specific Management for Agriculture- Invited papers and the proceedings of a Symposium on Site-specific Management (NCR-1180) in St. Louis, Missouri, 31 October 1995*, F. J. Pierce and E. J. Sadler (eds.), ASA-CSSA-SSSA, Madison, WI.

Hillier, P.S., and G.J. Lieberman. *Introduction to Mathematical Programming*, 2nd edition, McGraw-Hill, Inc., New York, 1995.

Hollands, K.R. “Relationship of Nitrogen and Topography.” In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Huang, W.Y., “Appendix F: An Estimate of Excess Nitrogen Fertilizer for the U.S.” In *Agricultural Chemical Use and Ground Water Quality: Where Are the Potential Problem Areas?* eds. R.L. Kellogg, M.S. Maizel, and D.W. Goss, USDA-ERS, Washington D.C., December 1992.

Huang, W.Y. “Market-Based Incentives for Addressing Non-Point Water Quality Problems: A Residual Nitrogen Tax Approach.” *Review of Agricultural Economics*, 16(1994): 427-440.

Huang, W.Y., R. Heifner, and H. Taylor. “An Analysis of Using Insurance to Enhance Nitrogen Fertilizer Application Timing to Reduce Nitrogen Losses to the Environment.” Paper presented to Southern Agricultural Economics Association Annual Conference, Memphis, Tennessee, February, 1999.

Huang, W.Y., and R.M. Lantin. “A Comparison of Farmers’ Compliance Costs to Reduce Excess Nitrogen Fertilizer Use Under Alternative Policy Options.” *Review of Agricultural Economics*, 15(1)(1993): 51-62.

Hurley, T., B. Kilion, G. Malzer, and M. Dikici. “The Value of Information for Variable Rate Nitrogen Application: A Comparison of Soil Test, Topographical and Remote Sensing Information.” Paper presented at the *AAEA Annual Meeting*, Chicago, Illinois, August 5-8, 2001.

Iosifescu, M. *Finite Markov Processes and Their Applications*. John Wiley and Sons Ltd., New York, 1980.

Jones, D.L., I.A. Rodihan, L.E. Cullipher, J.W. Clay, and M.J. Marks. *Soil Survey of Prince George County, Virginia*. USDA, Soil Conservation Service, in cooperation with the Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1985.

Keeney, D. R. “Nitrogen Management for Maximum Efficiency and Minimum Pollution.” *Nitrogen in Agricultural Soils*, ed. F. J. Stevenson. Madison WI: American Society of Agronomy, Agronomy Monograph No. 22, 1982.

Kellogg, R. L., M. S. Maizel, and D. W. Goss. *Agricultural Chemical Use and Ground Water Quality: Where Are the Potential Problem Areas?* USDA, National Center for Resource Innovations, Washington D.C., December 1992.

Kemeny, J.G., and J.L. Snell. *Finite Markov Chains*. Princeton, N. J.: Van Nostrand, 1960.

King, R., and G. Oamek, "Risk Management by Colorado Dryland Wheat Farmers and the Elimination of the Disaster Assistance Program." *Amer. J. Agri. Econ*. 65(May 1983): 247-255.

Kitchen, N.R., D.F. Kughes, K.A. Sudduth, and S.J. Birrell. "Comparison of Variable Rate to Single Rate Nitrogen Fertilizer Application: Corn Production and Residual Soil NO3-N. p.427-441. In *Site-Specific Management for Agricultural Systems*, eds. P.C. Robert et al., ASA Misc. Publ. ASA, CSSA, and SSSA, Madison, WI.

Knight, F. H., *Risk, Uncertainty, and Profit*. Boston, Houghton, Milton, 1921.

Krause, M.A., and J.R. Black. "Optimal Adoption Strategies for No-till Technology in Michigan." *Review of Agricultural Economics*. 17(1997): 299-310.

Lake, J. V., G. Gregory, and J. A. Goode. *Precision Agriculture: Spatial and Temporal Variability of Environmental Quality*. Ciba Foundation, Wiley & Sons Ltd., New York, 1997.

Larson, E. E., J. A. Labm, B. R. Khakural, R. B. Ferguson, and G. W. Rehm. "Potential of Site-Specific Management for Nonpoint Environmental Protection." In *The State of Site-Specific Management for Agriculture* by F. J. Pierce, and E. J. Sadler (eds.). Pp 337-367. Madision, WI: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America.

Legg, T. D. *Farm Level Effects of Environmental Policies Aimed at Nitrogen Management*. Ph.D. dissertation, University of Minnesota, March 1991.

Leiva, R. R., J. Morris, and S. B. Blackmore. "Precision Farming Techniques For Sustainable Agriculture." In *Precision Agriculture '97: Proceedings of the First European Conference on Precision Agriculture*, eds. J.V. Stafford, Warwick University, UK, 7-10, September 1997. Bios Scientific Pub., Hemdon, VA, 1997.

Loucks, P.D. "Watershed Planning: Change Issues, Processes and Expectations." *Water Resources Update*, 111(Spring 1998): 38-46, Universities Council on Water Resouces.

Lowenberg-DeBoer, J. "Precision Farming and the New Information Technology: Implications for Farm Management, Policy, and Research: Discussion." *Amer. J. Agr. Econ*. 78(December 1996): 1281-1284.

Lowenberg-DeBoer, J., S. Hawkins, and R. Nielsen. "Economics of Precision Farming." Paper presented at the Agron. Research Center Field Day. Purdue University, West Lafayette, IN., 1994.

Lowenberg-DeBoer, J., and S. M. Swinton. "Economics of Site-Specific Management in Agronomic Crops." In *The State of Site Specific Management for Agriculture- Invited papers and the proceedings of a Symposium on Site-specific Management (NCR-1180) in St. Louis, Missouri, 31 October 1995,* F. J. Pierce and E. J. Sadler (eds.), ASA-CSSA-SSSA, Madison, WI.

Lowenberg-DeBoer, J., and M. Boehlje. "Revolution, Evolution or Dead-end: Economic Perspectives on Precision Agriculture." Proceedings of the Third International Conference on Precision Agriculture, ASA-CSSA-SSSA, Madison WI, forthcoming.

Lu, Y. C., C. Daughtry, G. Hart, and B. Watkins. "The Current State of Precision Farming." Food Review International. 13(2) (1997): 141-162.

Madison, R.J., and J.O. Brunett. *Overview of the Occurences of Nitrates in Groundwater of the U.S.* U.S. Geolo. Survey Water Supply Paper 2275. pp. 93-105, Geological Survey, US Department of the Interior, D.C., 1985.

Magnien, R., D. Boward, and S. Bieber. "The State of the Chesapeake Bay 1995." EPA website *http://www.epa.gov/ncepihom/Catalog/EPA903R95001.html*, latest accessed on August 13, 2001.

Mallarino, A.P., and A.B. Blackmer. "Comparison of Methods for Determinig Critical Concentrations of Soil Test Phosphorus for Corn." *Agronomy Journal*, 84(1992): 850-856.

Malzer, G.L., P.J. Copeland, J.G. Davic, J.A. Lamb, P.C. Robert, and T.W. Bruulsema. "Spatial Variability of Profitability in Site Specific N Management." In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Mapp, H. P., D. J. Bernardo, G. J. Sabbagh, S. Geleta, and K. B. Watkins. "Economic and Environmental Impacts of Limiting Nitrogen Use to Protect Water Quality: A Stochastic Regional Analysis." *Amer.J.Agr.Econ*. 76 (November 1994): 889-903.

Martell, D.L. "A Markov Chain Model of a Forest Fire Danger Rating Index." Paper presented at *The 10th Conference on Fire and Forest Meteorology,* April 17-21, 1989, Ottawa, Canada.

Massart, D.L., and L. Kaufman. *The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis.* John Wiley & Sons, New York, 1983.

Matis, J.H., T. Birkett, and D. Boudreaux. "An Application of the Markov Chain Approach to Forecasting Cotton Yields from Survey." *Agricultural Systems*, 29(1989):357-370.

Matis, J.H., T.Saito, W.E. Grant, W.C. Iwig, and J.T. Ritchie. "A Markov Chain Approach to Crop Yield Forecasting." *Agricultural Systems*, 18(1985): 171-187.

McBratney, A.B., B.M. Whelan, and T.M. Shatar. "Variability and Uncertainty in Spatial, Temporal and Spatiotemporal Crop-Yield and Related Data." In *Ciba Foundation Symposium 210: Precision Agriculture: Spatial and Temporal Variability of Environmental Quality*, eds. Lake, J.V., G.R. Bock, and J.A. Goode, John Wiley & Sons, New York, 1997.

McClellan, P. W. Personal Communication. Biological Systems Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1999.

McSweeny, W. T., and J. S. Shortle. "Reducing Nutrient Application Rates for Water Quality Protection in Intensive Livestock Areas: Policy Implications of Alternative Producer Behavior." *Northeastern Journal of Agricultural and Resource Economics*. 18(1)(April 1989): 1-11.

Meisinger, J.J. "Evaluating Plant-Available Nitrogen in Soil-Crop Systems." *Nitrogen in Crop Production,* Chapter 26. Madison, WI: Soil Science Society of America, 1984.

Meisinger, J. J., F. R. Magdoff, and J. S. Schepers. "Predicting N Fertilizer Needs for Corn in Humid Regions: Underlying Principles." Chapter 2 in B. R. Bock and K. R. Kelley (eds.) *Predicting N Fertilizer Needs for Corn in Humid Regions.* Bulletin Y-226. Muscle Shoals, AL: National Fertilizer and Environmental Research Center, Tennessee Valley Authority, 1993.

Meyer, J. "Second Degree Stochastic Dominance With Respect to a Function." *International Economic Review*. 18(1977): 477-487.

Mostaghimi, S. *Personal Communication*. Biological Systems Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1999.

Mostaghimi, S., McClellan, P. W., Gupta, R. K., Vaughan, D. H. and Y. Fu. "Suitability of Precision Farming Technology in the Virginia's Coastal Resources Management Area". Final Report to Commonwealth of Virginia Department of Environmental Quality, Virginia Coastal Resources Management Program, Richmond, Virginia, 1997.

Mostaghimi, S., P. W. McClellan, K. Brannnan, and Y. Fu. "Precision Farming Technology for Reducing Agricultural Chemical Inputs in Virginia's Coastal Resources Management Area." Virginia Division of Soil and Water Conservation, report No.PF-0298, Richmond, Virginia. 1998.

Mulla, D. J., and J. S. Schepers. "Key Processes and Properties for Site-Specific Soil and Crop Management." In F. J. Pierce and E. J. Sadler (eds.) *The State of Site Specific Management for Agriculture- Invited papers and the proceedings of a Symposium on Site-specific Management (NCR-1180) in St. Louis, Missouri, 31 October 1995.* ASA-CSSA-SSSA, 66 S. Segoe Rd., Madison, WI 53711.

National Research Council (NRC). *Precision Agriculture in the 21st Century: Geospatial and Information Technologies in Crop Management*. National Academy Press, Washington, D.C., 1997.

\- *Soil and Water Quality: An Agenda for Agriculture*. National Academy Press, Washington D.C., 1993.

\- *Nitrates: An Environmental Assessment.* National Academy of Sciences, Washington, DC, 1978.

NCDEM (North Carolina Division of Environmental Management). *Water Quality Progress in North Carolina,* 1988-1989 305(B) Report. Report No.90-07. Raleigh, North Carolina, 1990.

Norris, P. E. and L. A. Shabman. "Economic and Environmental Considerations for Nitrogen Management in the Mid-Atlantic Coastal Plain." *American Journal of Alternative Agriculture*. 7(1992): 148-156

Nowak, P. "Agriculture and Change: The Promise and Pitfalls of Precision." *Communications in Soil Science and Plant Analysis*. 29(11-14)(1998): 1537-1541.

Nowak, P. “A Sociological Analysis of Site-Specific Management.” In *The State of Site Specific Management for Agriculture* by F.J. Pierce and E.J. Sadler (eds.). P.397-422. ASA-CSSA-SSSA, Madison, WI., 1997.

Peng, W., and D. Bosch. "Improving Nitrogen Management on A Case Corn-Wheat/Soybean Farm Using Site-Specific Management in Coastal Plains in Virginia". Paper presented at AAEA Annual Meeting, Tampa, Florida, 1-4 August 2000.

Phipps, P. Personal Communication. Professor, Tidewater Agricultural Research and Experiment Center, Virginia Polytechnic Institute and State University, Suffolk, Virginia, 2000.

Pierce, F., D. Warncke, and M. Everett. “Yield and Nutrient Variability in Glacial Soils of Michigan.” In *Site-Specific Management for Agricultural Systems*, eds. P. Robert, R. Rust, and W. Larson. Madison, WI: ASS-CSSA-SSSA, 1998.

Randhir, Timothy O., and John G. Lee. “Designing Spatial Incentives to Manage Agricultural Nonpoint Source Pollution.” Paper presented at AAEA annual meeting in Toronto, Canada, 1997.

Redulla, C.A., J.L. Havlin, G.J. Kluitenberg, N. Zhang, M.D. Schrock. “Variable Nitrogen Management for Improving Groundwater Quality.” In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Roberts, R.K., B.C. English, and S.B. Mahajanashetti. “Evaluating the Returns to Variable Rate Nitrogen Application.” *Journal of Agricultural and Applied Economics*, 32(1)(April 2000): 133-143.

Salchow, E., R. Lal, R.R. Fausey, and A.D. Ward. “Pedotransfer Functions for Variable Allucial Soils in Southern Ohio.” *Geoderma*, 3(1996): 165-181.

SAS Institute Inc. *SAS User’s Guide: Statistics*, SAS Institute Inc., Cary, N.C., 1982.

Schnitkey, G.D., J.W. Hopkins, and L.G. Tweeten. “An Economic Evaluation of Precision Fertilizer Appication on Corn-Soybean Fields.” In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Shanholtz, V.O., and N. Zhang. *Supplementary Information Of Virginia Geographic Information System*. Special Report, VirGIS 88-21. Blacksburg, Virginia: VirGIS Laboratory, Department of Agricultural Engineering, Virginia Polytechnic Institute and State University, 1988.

Shatar T. “Does Precision Agriculture Make Economic Sense?” *Leading Edge -- The Journal of Agricultural Engineering and Precision Agriculture*, 1(1)(1998): 36-37.

Simpson, T. W., S. J. Donohue, G. W. Hawkins, M. M. Monnett, and J. C. Baker. *The Development and Implementation of the Virginia Agronomic Land Use Evaluation System (VALUES)*. Dept. of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, December 1992.

Smith, W. F. “Precision Farming Overview.” *Proceedings of Beltwide Cotton Conferences*, 1996, V.1 (p.179-180), National Cotton Council of America, Memphis, Tenn., 1996.

Snyder, C., J. Havlin, G. Kluitenberg, and T. Schroeder. “An Economic Analysis of Variable Rate Nitrogen Management.” In *Proceedings of the Third International Conference on Precision Agriculture,* eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Sonka S. T., and K. F. Coaldrake. “Cyberfarm: What Does It Look Like? What Does It Mean?” *Amer. J. Agr. Econ*. 78(December 1996): 263-268

Stein, A., J. Brouwer, and J. Bouma. “Methods for omparing Spatial Variability Patterns of Millet Yield and Soil Data.” *Soil Sci. Soc. Am. J.* 61(1997a): 861-870.

Stein, A., M.R. Hoosbeek, and G. Sterk. “Space-Time Statistics for Decision Support to Smart Farming.” *Precision Agriculture: Spatial and Temporal Variability of Environmental Quality.* P.120-130. John Wiley & Sons, New York, 1997b.

Steinwand, A.L., D.L. Karlen, and T.E. Fenton. “An Evaluation of Soil Survey Crop Yield Interpretations for Two Central Iowa Farms.” *J. Soil and Water Cons.* 51(1)(1996): 66-71.

Stevenson, F.G. *Cycles of Soil.* John Wiley and Sons, New York, 1986.

Sturt, Guy. *1999 Crop Enterprise Cost Analysis for Eastern Virginia.* Eastern District Farm Management Staff. Virginia Cooperative Extension Service, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1998.

Swinton, S.M., R. Bongiovanni, J. Lowenberg-DeBoer, and D.S. Bullock. “Assessing the Value of Precision Agriculture Data: On-Farm Nitrogen Response Research in Argentina.” Paper presented at the AAEA Learning Workshop on Spatial Data Analysis, Chicago, Illinois, August 4, 2001.

Tauer, Loren W. “Target MOTAD”. *Amer. J. Agr. Econ.*. 65(1983): 606-610.

Thrikawala, S., A. Weersink, G. Kachanoski, and G. Fox. “Economic Feasibility of Variable-Rate Technology for Nitrogen on Corn.” *Amer. J. Agr. Econ.* 81 (Nov. 1999): 914-927.

Tweeten, L. “Is Precision Farming Good for Society?” *Better Crops*, 80(3)(1996): 1-5.

USDA. “Albemarle-Pamlico Drainage Area Study Links Agricultural Production and Natural Resource DATA.” *RTD Updates: Area Studies*, Number 1. Economic Research Service, Washington D.C., January 1994.

USDA. *Soil Survey of City of Suffolk, Virginia.* Soil Conservation Service, Washington, D.C., June, 1981.

USDA, Soil Conservation Service, and Statistical Laboratory of Iowa State University. National *Resources Inventory (1987): Instruction for Collecting Sample Data.* Washington D.C., Government Printing Office, September 1986.

USDA-NASS. *Agricultural Prices (19xx Summary).* 1998 and earlier issues. NASS, Agricultural Statistics Board, D.C.

USEPA, Office of Water Program Operations. *Report to Congress: Nonpoint Pollution in the U.S.* Washington, D.C.: USEPA, January 1984.

Verhagen, J., and J. Bouma. "Modeling Soil Variability." In *The State of Site Specific Management for Agriculture*, eds. F.J. Pierce and E.J. Sadler, ASA-CSSA-SSA, Madison, WI. 1997.

Virginia Agricultural Statistical Service (VASS). Agriculture Census for Prince George County, Virginia. *http://govinfo.ibrary.orst.edu/cgi-bin/ag-list?01-149.vac*, accessed December, 1999.

Virginia Cooperative Extension Service. "Crop and Livestock Enterprise Budgets 2000." Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 2000.

Virginia Department of Conservation and Recreation. *Nutrient Management Handbook*. The Department of Conservation and Recreation, Division of Soil and Water Conservation, Richmond, Virginia, March 1991.

Wallace, A., "High-Precision Agriculture is An Excellent Tool for Conservation of Natural Resources." *Communications in Soil Science and Plant Analysis*. 25 (1&2)(1994): 45-49.

Webster, R. "Quantitative Spatial Analysis of Soil in the Field." In *Advances in Soil Science, Volume 3*, ed. B.A. Stewart, Springer-Verlag Inc., New York, 1985.

Webster, R., and M.A. Oliver. "Sample Adequately to Estimate Variograms of Soil Properties." *Journal of Soil Science*, 43(1992): 177-192.

Whelan, B., and A. McBratney. "Just What is Precision Agriculture?" *Leading Edge -- The Journal of Agricultural Engineering and Precision Agriculture*, 1(2)(September 1998): 14-15.

Whelan, B.M., A.B. McBratney, and R.A. Rossel. "Spatial Prediction for Precision Agriculture." In *Proceedings of the Third International Conference on Precision Agriculture*, eds., P.C. Robert, R.H. Rust, and W.E. Larson, Madison, WI: ASA-CSSA-SSSA, 1996.

Williams, J., and K. Renard. "Assessment of Soil Erosion and Crop Productivity with Process Model (EPIC)." In *Soil Erosion and Crop Productivity*, eds. Follett, R.F. and B.A. Steward, American Society of Agronomy, Madison, WI, 1985.

Williams, J., S. Jones, and P. Dyke. "The EPIC model." Chapter 2, pp.3-92. In *EPIC – Erosion/Productivity Impact Calculator: 1. Model Documentation*, eds. A.N. Sharpley and J.R. Williams, USDA Tech. Bull. No.1768, 1990.

Wilson, P. N., and V. R. Eidman. "An Empirical Test of the Interval Approach for Estimating Risk Preference." *Western Journal of Agricultural Economics* 8(1993): 170-182.

Wolf, S. A., and F. H. Buttel. "The Political Economy of Precision Farming." *Amer.J. A. Agr. Econ*. 78(December 1996): 1269-1274.

Wolf, S. A., and S. D. Wood. "Precision Farming: Environmental Legitimation, Commodification of Information, and Industrial Coordination." *Rural Sociology*, 62(2)(1997): 180-206.

Wollenhaupt, N.C., D.J. Mulla, and C.A Crawford. “Soil Sampling and Interpolation Techniques for Mapping Spatial Variability of Soil Properties.” In *The State of Site Specific Management for Agriculture,* eds. F.J. Pierce and E.J. Sadler, ASA-CSSA-SSA, Madison, WI. 1997.

Yagow, G., and Shanholtz, V. *PC-VirGIS / AGNPS 5.00 File Builder*. Draft Documentation. MapTech, Inc. Virginia Tech Corporate Research Center, 200 Kraft Drive, Suite 2200, Blacksburg, Virginia, 2000.

Yule, I. J., and E. Crooks. “Precision Farming: The Price of Inperfection: A Case Study Using Fertilizer Distribution.” *Landwards*. V.51(1)(Spring 1996): 5-9.

Appendix A. A short discussion of the economic risk of adopting SSM

Theoretically, SSM is a good risk management tool because it transforms a large amount of uncertainty into risk (Nowak, 1998). Uncertainty is the state in which the probability distributions are unknown, and risk is the state in which probability distributions are known (Knight). By gathering and analyzing soil and yield data spatially and temporally, SSM results in better knowledge of every stage of production, which in turn results in more efficient input use. More efficient input use both enhances net returns and reduces income and environmental risks (Heimlich).

However, the adoption of SSM itself increases farmers' financial risk because of increased capital investment which increases farmers' debt service obligations. Learning costs, defined as costs from lost yield as well as wasted input costs due to mistakes in managing a new technology (Krause and Black), present another type of risk. The risk of loss due to obsolescence of investment is high and very uncertain for information technologies (computer, yield monitor, and software). The values of yield and fertility maps increase for the farmers as they are used and updated for several years. Land value increases as soil integrity is maintained by SSM. However, if the land is to be sold in the future for non-agricultural use, or if the land is rented and the operator loses the lease in the future, the long-term benefit of farmers' investment in SSM may be lost. In addition, human capital risk increases because SSM operation relies more on skilled personnel to operate machines and analyze data. Thus one of the challenges for SSM adopters is risk management even though SSM is often cited as an excellent mechanism to manage risk. This situation can be viewed as a tradeoff between temporal and spatial risk.

Larger farms (usually the adopters of innovative BMPs) tend to be SSM innovators (Heimlich). Adopters of SSM are more likely those who have adopted other relatively sophisticated methods of input management such as intensive soil sampling, pest scouting, and record keeping, indicating that human capital is indeed a major factor in adoption behavior (Wolf and Wood).

Evaluation of SSM tends to compare profitability or risk of SSM with conventional systems where no BMP's or innovation practices are adopted, thus crediting some benefits derived from other practices to SSM (Lowenberg-DeBoer and Swinton). This can introduce two biases in many SSM profitability and risk analyses. For those adopters who have already taken certain steps toward improved innovative management systems, profit and risk levels as well as opportunities to reduce pollution can be overestimated. For those adopters who have been slow to adopt innovative management systems, the risk may be underestimated. Thus, it is warned that biased evaluations of the economic and environmental impact of SSM from industrial or promotional organizations may present one more risk for farmers in deciding whether or when to adopt SSM practices (Heimlich). However, by hiring service providers to carry out some key SSM operations like mapping and variable rate application, small and inexperienced farmers can also adopt SSM because (1) they are no longer constrained by lack of experience, and (2) much of the cost in SSM becomes variable cost to the farmers and thus reduces risks by reducing required investments. Currently in Virginia, soil sampling can be done by fertilizer companies, and companies like Southern States can generate yield maps and do variable rate application for farmers. The most likely investment in SSM now for farmers in Virginia is just a yield monitor at cost of around $2,500 (Mostaghimi).

Appendix B. Simulated field-level yield sequences and rescaling procedures

Table B-1. Simulated farm-level wheat yields and field-level yields

Year	Farm-level (bu/ac)	F1B (bu/ac)	F1C (bu/ac)	F2 (bu/ac)	F3B (bu/ac)	F4 (bu/ac)	F5 (bu/ac)	F7 (bu/ac)	F8B (bu/ac)
1923	78	66	69	73	78	80	87	78	87
1924	93	79	82	86	93	96	103	89	103
1925	73	62	64	68	73	75	81	73	81
1926	69	59	61	64	69	71	77	69	77
1927	75	64	66	70	75	77	83	75	83
1928	92	78	81	86	92	95	102	88	102
1929	80	68	70	74	80	82	89	80	89
1930	79	67	70	73	79	81	88	79	88
1931	77	65	68	72	77	79	85	77	85
1932	95	81	84	88	95	98	105	91	105
1933	76	65	67	71	76	78	84	76	84
1934	68	58	60	63	68	70	75	68	75
1935	82	70	72	76	82	84	91	79	91
1936	61	52	54	57	61	63	68	61	68
1937	81	69	71	75	81	83	90	78	90
1938	72	61	63	67	72	74	80	72	80
1939	79	67	70	73	79	81	88	79	88
1940	59	50	52	45	59	61	65	75	65
1941	69	59	61	64	69	71	77	69	77
1942	65	55	57	60	65	67	72	65	72
1943	67	57	59	62	67	69	74	67	74
1943	63	54	55	59	63	65	70	63	70
1945	72	61	63	67	72	74	80	72	80
1946	81	69	71	75	81	83	90	78	90
1947	81	69	71	75	81	83	90	78	90
1948	71	60	62	66	71	73	79	71	79
1949	88	75	77	82	88	91	98	84	98
1950	87	74	77	81	87	90	97	84	97
1951	76	65	67	71	76	78	84	76	84
1952	77	65	68	72	77	79	85	77	85
1953	68	58	60	63	68	70	75	68	75
1954	78	66	69	73	78	80	87	78	87
1955	70	60	62	65	70	72	78	70	78
1956	63	54	55	59	63	65	70	63	70
1957	65	55	57	60	65	67	72	65	72
1958	70	60	62	65	70	72	78	70	78
1959	62	53	55	58	62	64	69	62	69
1960	61	52	54	57	61	63	68	61	68
1961	67	57	59	62	67	69	74	67	74
1962	48	41	42	37	48	49	53	61	53
1963	65	55	57	60	65	67	72	65	72
1964	64	54	56	60	64	66	71	64	71
1965	70	60	62	65	70	72	78	70	78
1966	75	64	66	70	75	77	83	75	83
1967	73	62	64	68	73	75	81	73	81

Table B-1. Simulated farm-level wheat yields and field-level yields (*continued*)

Year	Farm-level (bu/ac)	F1B (bu/ac)	F1C (bu/ac)	F2 (bu/ac)	F3B (bu/ac)	F4 (bu/ac)	F5 (bu/ac)	F7 (bu/ac)	F8B (bu/ac)
1968	70	60	62	65	70	72	78	70	78
1969	64	54	56	60	64	66	71	64	71
1970	37	31	33	28	37	38	41	47	41
1971	74	63	65	69	74	76	82	74	82
1972	92	78	81	86	92	95	102	88	102
1973	79	67	70	73	79	81	88	79	88
1974	82	70	72	76	82	84	91	79	91
1975	76	65	67	71	76	78	84	76	84
1976	70	60	62	65	70	72	78	70	78
1977	49	42	43	38	49	50	54	62	54
1978	61	52	54	57	61	63	68	61	68
1979	74	63	65	69	74	76	82	74	82
1980	69	59	61	64	69	71	77	69	77
1981	64	54	56	60	64	66	71	64	71
1982	62	53	55	58	62	64	69	62	69
1983	86	73	76	80	86	89	95	83	95
1984	55	47	48	42	55	57	61	70	61
1985	55	47	48	42	55	57	61	70	61
1986	75	64	66	70	75	77	83	75	83
1987	61	52	54	57	61	63	68	61	68
1988	69	59	61	64	69	71	77	69	77
1989	76	65	67	71	76	78	84	76	84
1990	68	58	60	63	68	70	75	68	75
1991	62	53	55	58	62	64	69	62	69
1992	92	78	81	66	92	95	102	78	102
1993	83	74	88	77	89	89	92	80	92
1994	71	60	62	89	71	73	86	83	87
1995	86	85	86	80	93	101	95	83	95
1996	59	50	52	44	59	61	62	66	65
1997	90	74	68	84	88	92	100	86	100
1998	76	65	67	42	76	78	69	76	64
Mean	72	61	64	66	72	74	80	72	80
Std.dev.	11.0	9.6	10.1	12.2	11.2	11.7	12.3	8.2	12.4

Table B-2. Rescaling standards for wheat yield for each field[a]

Field	Pattern	Scaling standard
F1B	Wheat yield for field 1B is 15% lower than farm-level average according to observations on 1993, 1995, and 1997 and the temporal variation pattern is similar to that of farm-level yield.	The simulated farm-level sequence is multiplied by 0.85.
F1C	Wheat yield for field F1C is 12% lower than farm-level average according to observations on 1993, 1995, and 1997 and the temporal variation pattern is similar to that of farm-level yield.	The simulated farm-level sequence is multiplied by 0.88.
F2	In bad years (1996 and 1998), average yield is 23% less than farm-level yield while in good years (1992 and 1994) it is only 7% lower.	The simulated farm-level yield sequence is multiplied by 0.77 for yields below 60 bushels per-acre and by 0.93 for other years
F3B	Identical (in 1993, 1995, and 1997) to the farm-level yields.	No re-scaling is done for this field.
F4	It is 3% higher (in 1993, 1995, 1997) than farm-level yields.	The simulated farm-level yield sequence is multiplied by 1.03.
F5	The yields are 11% higher than the farm-level (1994, 1996, and 1998).	The simulated farm-level yield sequence is multiplied by 1.11.
F7	In above average years (1992 and 1994 where farm-level yields are above 80), the field yield is 4% lower than farm-level, while in bad years (1996 and 1998 when farm-level yields are below 60) the field yields are 27% higher than farm-level.	For farm-level yields below 60, they are multiplied by 1.27 and for farm-level yields above 80, they are multiplied by 0.96. Other years remain intact.
F8B	The yields are 11% higher than the farm-level (1994, 1996, and 1998).	The simulated farm-level yield sequence is multiplied by 1.11.

a. In determining scaling standards, pattern comparison for corn in the years when corn were planted is also considered.

Table B-3. Simulated farm-level corn yields and field-level yields

Year	Farm-level (bu/ac)	F1B (bu/ac)	F1C (bu/ac)	F2 (bu/ac)	F3B (bu/ac)	F4 (bu/ac)	F5 (bu/ac)	F7 (bu/ac)	F8B (bu/ac)
1922	181	170	176	121	174	192	170	199	195
1923	169	159	164	130	162	179	159	186	183
1924	112	105	109	97	108	119	105	123	121
1925	117	110	113	102	112	124	110	129	126
1926	145	136	141	112	139	154	136	160	157
1927	186	175	180	125	179	197	175	205	201
1928	112	105	109	97	108	119	105	123	121
1929	110	103	107	96	106	117	103	121	119
1930	171	161	166	115	164	181	161	188	185
1931	143	134	139	110	137	152	134	157	154
1932	133	125	129	102	128	141	125	146	144
1933	141	133	137	109	135	149	133	155	152
1934	135	127	131	104	130	143	127	149	146
1935	169	159	164	130	162	179	159	186	183
1936	130	122	126	100	125	138	122	143	140
1937	174	164	169	117	167	184	164	191	188
1938	175	165	170	117	168	186	165	193	189
1939	116	109	113	101	111	123	109	128	125
1940	113	106	110	98	108	120	106	124	122
1941	137	129	133	105	132	145	129	151	148
1942	173	163	168	116	166	183	163	190	187
1943	138	130	134	106	132	146	130	152	149
1943	140	132	136	108	134	148	132	154	151
1945	144	135	140	111	138	153	135	158	156
1946	183	172	178	123	176	194	172	201	198
1947	113	106	110	98	108	120	106	124	122
1948	117	110	113	102	112	124	110	129	126
1949	143	134	139	110	137	152	134	157	154
1950	174	164	169	117	167	184	164	191	188
1951	147	138	143	113	141	156	138	162	159
1952	141	133	137	109	135	149	133	155	152
1953	127	119	123	110	122	135	119	140	137
1954	108	102	105	94	104	114	102	119	117
1955	170	160	165	114	163	180	160	187	184
1956	110	103	107	96	106	117	103	121	119
1957	119	112	115	104	114	126	112	131	129
1958	145	136	141	112	139	154	136	160	157
1959	176	165	171	118	169	187	165	194	190
1960	144	135	140	111	138	153	135	158	156
1961	107	101	104	93	103	113	101	118	116
1962	185	174	179	124	178	196	174	204	200
1963	116	109	113	101	111	123	109	128	125
1964	155	146	150	119	149	164	146	171	167
1965	185	174	179	124	178	196	174	204	200

(to be continued)

Table B-3. Simulated farm-level corn yields and field-level yields (*continued*)

Year	Farm-level (bu/ac)	F1B (bu/ac)	F1C (bu/ac)	F2 (bu/ac)	F3B (bu/ac)	F4 (bu/ac)	F5 (bu/ac)	F7 (bu/ac)	F8B (bu/ac)
1966	111	104	108	97	107	118	104	122	120
1967	113	106	110	98	108	120	106	124	122
1968	141	133	137	109	135	149	133	155	152
1969	150	141	146	116	144	159	141	165	162
1970	140	132	136	108	134	148	132	154	151
1971	174	164	169	117	167	184	164	191	188
1972	109	102	106	95	105	116	102	120	118
1973	111	104	108	97	107	118	104	122	120
1974	142	133	138	109	136	151	133	156	153
1975	176	165	171	118	169	187	165	194	190
1976	166	156	161	128	159	176	156	183	179
1977	137	129	133	105	132	145	129	151	148
1978	107	101	104	93	103	113	101	118	116
1979	171	161	166	115	164	181	161	188	185
1980	105	99	102	91	101	111	99	116	113
1981	107	101	104	93	103	113	101	118	116
1982	170	160	165	114	163	180	160	187	184
1983	119	112	115	104	114	126	112	131	129
1984	141	133	137	109	135	149	133	155	152
1985	155	146	150	119	149	164	146	171	167
1986	133	125	129	102	128	141	125	146	144
1987	122	115	118	106	117	129	115	134	132
1988	136	128	132	105	131	144	128	150	147
1989	151	142	146	116	145	160	142	166	163
1990	169	159	164	130	162	179	159	186	183
1991	141	133	137	109	135	149	133	155	152
1992	186	183	175	126	186	196	180	205	201
1993	107	101	104	81	103	113	98	119	131
1994	125	104	106	109	111	123	118	138	135
1995	178	167	173	118	171	189	177	188	182
1996	170	160	191	114	162	179	160	187	184
1997	107	101	104	105	103	113	87	125	111
1998	126	125	119	110	124	145	118	139	136
Mean	142	133	138	109	136	150	133	156	153
Std.dev.	26	24	26	10	25	27	25	28	27

Table B-4. Rescaling standards for corn yield for each field [a]

Field	Pattern	Scaling standard
Farm-level[b]	No more than two high yields (>160) in consecutive years. After two high consecutive yields, yield may drop to very low level (around 110 bu/ac). Once yields drop to very low level, it takes one or more years to raise the yields.	For first one or two simulated yield above 145, multiply by 1.15. If the third or fourth years are still over 140, then multiply them by 0.75.[b]
F1B	The yields are 6% lower on average (1992, 1994, 1996, and 1998) than the farm-level.	The simulated farm-level yield sequence is multiplied by 0.94.
F1C	The yields are 3% lower on average (1992, 1994, 1996, and 1998) than the farm-level.	The simulated farm-level yield sequence is multiplied by 0.97.
F2	In bad year (1993 and 1997), yields are 13% lower than farm-level. In other year (1995), yields are 32% lower than farm-level (good year for farm-level). By average, it is 22% lower than the farm-level.	The simulated yields are multiplied by 0.78.
F3B	The yields are 4% lower on average (1992, 1994, 1996, and 1998) than the farm-level.	The simulated farm-level yield sequence is multiplied by 0.96.
F4	The yields are 6% higher on average (1992, 1994, 1996, and 1998) than the farm-level.	The simulated farm-level yield sequence is multiplied by 1.06.
F5	The yields are 6% lower on average (1992, 1993, 1995, and 1997) than the farm-level.	The simulated farm-level yield sequence is multiplied by 0.94.
F7	The yields are 10% higher on average (1993, 1995, and 1997) than the farm-level.	The simulated farm-level yield sequence is multiplied by 1.10.
F8B	The yields are 8% higher on average (1993, 1995, and 1997) than the farm-level.	The simulated farm-level yield sequence is multiplied by 1.08.

a. In determining scaling standards, pattern comparison for wheat in the years when wheat were planted is also considered.

b. Since the EPIC simulation did not reflect the large temporal farm-level yield variation, the simulated farm-level sequence need to be rescaled to reflect the very high yield level as well as the sudden drop from high yields to low yield.

Appendix C. Within-field yield variability by functional zones

Table C-1. Within-field yield variation by functional zones for case fields 1995-1999[a]

Field	Zone #	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv
		---- 1995 crop ----			---- 1996 crop ----			---- 1997 crop ----			---- 1998 crop ----			---- 1999 crop ----		
F1B	1	35			54	11	20	29	6	22	27	9	32	18	5	30
	2	51	4	8	78	4	5	48	4	9	58	7	12	32	3	10
	3	62	2	4	89	3	4	60	3	4	79	6	7	42	3	6
	4	71	2	3	101	3	3	68	2	3	97	3	3	51	3	5
	5	80	2	2	111	3	3	75	2	2	104	2	2	59	2	3
	6	85	1	2	120	2	2	82	2	3	111	2	1	66	2	3
	7	91	2	2	129	3	2	90	2	2	117	1	1	75	2	3
	8	97	2	2	143	3	2	97	3	3	119	1	0	83	2	2
	9				151	2	1				121					
	10				156	1	1				125	1	1			
	11				162	2	1				131	2	1			
	12				168	2	1				137	2	1			
	13				175	2	1				143	2	1			
	14				182	2	1				151	2	2			
	15				190	3	1				160	3	2			
F1C	1	46			111			25	8	34	26	17	65	26	11	41
	2	60	2	4	120	1	1	46	3	7	63	7	11	51	4	8
	3	70	3	4	132	4	3	57	3	5	86	4	5	65	3	4
	4	79	2	3	141	2	2	64	2	3	100	1	1	73	2	3
	5	84	1	1	148	2	1	69	2	2	107	1	1	82	3	3
	6	88	1	1	160	3	2	74	1	2	113	0	0	93	3	3
	7	91	1	1	169	2	1	79	2	2	116	1	0	102	3	3
	8	96	1	2	176	1	1	85	3	3	119	1	1	109	2	2
	9				182	2	1				122	1	1			
	10				186	0	0				126	1	1			
	11				188						133	2	1			
	12				192	1	1				140	2	2			
	13				197	2	1				148	3	2			
	14				204	2	1				155	1	1			
	15				213	4	2				173					

a. Number of functional zones for wheat is 8 and for corn is 15. When only one SMU is in certain functional zone, the standard deviation and CV are not estimated.

Table C-1. Within-field yield variation by functional zones for case fields 1995-1999[a] (continued)

Field	Zone #	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv
		---- 1995 crop ---			---- 1996 crop ----			---- 1997 crop ----			---- 1998 crop ----			---- 1999 crop ----		
F3B	1	40			68	3	4	35	5	13	53	7	13	57	3	5
	2	63	3	6	83			50	3	7	78	4	6	68	2	3
	3	72	3	4	88	1	2	61	3	6	91	3	3	75	2	2
	4	82	2	3	102			71	3	4	100	2	2	82	2	2
	5	88	2	2	109	3	3	81	2	3	107	2	2	89	2	2
	6	94	1	1	119	3	3	90	3	3	115	1	1	95	2	2
	7	98	1	1	129	3	2	101	3	3	120	1	1	101	2	2
	8	102	2	2	138	2	2	110	3	3	122	0	0	107	2	2
	9				149	2	1				123	0	0			
	10				156	2	1				125	1	0			
	11				160	1	1				127	1	1			
	12				165	1	1				132	2	1			
	13				171	2	1				138	2	1			
	14				178	2	1				144	2	1			
	15				183	2	1				152	3	2			
F4	1	38	7	18	56	14	24	15	5	33	42	10	23	19	7	37
	2	53	2	3	75	5	6	42	7	16	73	5	7	49	5	10
	3	62	2	4	87	4	4	60	4	7	88	3	4	62	4	6
	4	71	3	5	100	4	4	73	3	4	100	3	3	73	3	4
	5	84	4	4	111	1	1	82	3	3	111	3	3	82	2	3
	6	95	2	3	120	3	2	92	3	3	119	3	2	90	3	3
	7	103	2	2	135	2	1	98	2	2	127	2	2	99	2	2
	8	111	3	3	142	3	2	104	2	2	136	2	1	105	2	2
	9				151	2	2				141	1	1			
	10				161	3	2				145	1	1			
	11				171	2	1				149	1	1			
	12				181	3	2				153	2	1			
	13				189	3	1				160	2	1			
	14				202	3	2				167	2	1			
	15				214	8	4				175	3	1			

a. Number of functional zones for wheat is 8 and for corn is 15. When only one SMU is in certain functional zone, the standard deviation and CV are not estimated.

Table C-1. Within-field yield variation by functional zones for case fields 1995-1999[a] (continued)

Field	Zone #	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv
		---- 1995 crop ---			---- 1996 crop ----			---- 1997 crop ----			---- 1998 crop ----			---- 1999 crop ----		
F2	1	31	8	25	10	5	47	42	16	38	14	5	38	40	10	25
	2	51	6	12	26	4	13	70	6	9	28	2	7	63	4	6
	3	76	5	7	35	1	4	87	3	4	36	2	5	73	3	4
	4	92	3	3	41	2	4	94	2	2	40	1	3	82	1	1
	5	100	2	2	47	1	3	99			44	1	3	85	1	1
	6	107	2	2	52	1	2	101			49	2	3	87		
	7	113			58	2	4	103	1	1	56	2	4	88		
	8	115			66	3	5	107	1	1	64	3	4	89		
	9	122	3	2				112	2	2				92	1	1
	10	133	3	2				120	3	2				96	1	1
	11	142	2	2				129	2	2				101	2	2
	12	150	2	2				139	3	2				112	3	2
	13	159	2	1				147	2	1				119	2	2
	14	167	2	1										126	2	2
F5	1	50	13	27	17	8	45	12	5	45	34	7	19	52		
	2	77	1	1	37	5	13	33	6	17	51	3	6	81	4	5
	3	91	6	6	51	2	4	52	4	7	61	2	4	90	3	3
	4	103	1	0	58	2	3	63	3	4	66	1	2	99	3	3
	5	111	3	3	63	1	2	76	2	3	69	1	1	110	2	2
	6	122	4	3	66	1	2	81	1	2	73	1	2	116	1	1
	7	131	3	2	71	1	2	84	1	1	76	2	2	120	1	1
	8	141	3	2	76	2	3	88	1	1	88	1	1	122		
	9	156	3	2				91	1	1				123	0	0
	10	166	2	1				97	2	2				125	1	1
	11	175	2	1				105	2	2				129	1	1
	12	182	2	1				112	2	2				135	2	2
	13	189	2	1				120	3	3				143	2	1
	14	196	2	1				131	2	1				149	2	1
	15	203	2	1				137	1	1				154	2	2

a. Number of functional zones for wheat is 8 and for corn is 15. When only one SMU is in certain functional zone, the standard deviation and CV are not estimated.

Table C-1. Within-field yield variation by functional zones for case fields 1995-1999[a] (continued)

field	Zone #	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv	mean	std	cv
		---- 1995 crop ---			---- 1996 crop ----			---- 1997 crop ----			---- 1998 crop ----			---- 1999 crop ----		
F7	1	63	4	6	16	7	45	28	14	50	36	9	23	16	11	68
	2	102	2	2	42	5	12	61	7	11	53	3	6	56	7	12
	3	122			55	3	5	79	5	6	62	2	4	72	4	5
	4	128	2	2	61	1	2	98	3	3	69	2	3	85	3	4
	5	140	3	2	65	1	1	109	2	2	76	2	2	93	2	2
	6	152	3	2	68	1	1	116	2	1	81	2	2	98	1	1
	7	163	2	1	72	1	2	120	1	1	88	2	2	102	1	1
	8	171	2	1	77	2	2	122	0	0	95	2	2	105	1	1
	9	177	2	1				123	0	0				108	1	1
	10	183	1	1				125	1	0				114	2	2
	11	187	1	1				126	0	0				121	2	1
	12	192	1	1				129	1	1				128	2	2
	13	197	2	1				135	2	1				138	3	2
	14	206	4	2				141	2	2				149	3	2
	15	221	5	2				149	3	2				159	3	2
F8B	1	39	20	51	32	9	27	27	9	33	31	8	25	32	10	30
	2	105			50	3	6	48	5	11	49	4	8	75	7	9
	3	110			58	2	3	65	5	8	57	1	3	91	2	3
	4	122	1	1	62	1	1	87	5	6	61	1	2	100	2	2
	5	135	1	1	65	1	1	101	3	3	66	1	2	106	2	2
	6	141	3	2	68	1	1	108	1	1	70	1	1	112	1	1
	7	154	2	1	72	1	2	109			74	1	2	117	1	0
	8	164	2	1	76	2	2	110			80	2	3	118	0	0
	9	171	2	1				111	0	0				120	1	1
	10	176	1	1				114	1	1				124	2	1
	11	181	1	1				117	1	1				131	2	2
	12	185	2	1				123	2	2				141	2	2
	13	191	2	1				129	2	2				149	2	2
	14	198	2	1				138	2	2				158	3	2
	15	206	3	1				144	2	1						

a. Number of functional zones for wheat is 8 and for corn is 15. When only one SMU is in certain functional zone, the standard deviation and CV are not estimated.

Appendix D. SAS programs used in ex post, ex ante, and sensitivity analysis

Part I (original model)

```
/****************************************************************************
* Program name = fldtransition.sas
*
* Abstract: This program estimates field-level transition matrix and test
*           for the existence of Markov chain. Several steps involved:
*           1. Define state of nature for each year;
*             1.1. Count frequency of states for each fields;
*           2. Define transition for each year;
*             2.2. Count frequency of each transition;
*             2.3. Calculate transition matrix for each field;
*           3. Test for the existence of first-order Markov chain;
*           4. Determines field-level optimal yields for 1999.
*              4.1. If 1st-order Markov chain exist, use 1998 yield and the
*                   corresponding transition probs to determine;
*              4.2. If 1st-order Markov chain does not exist, assign equal
*                   prob to each year to determine.
*           5. Determine the expected field-level yield for 1999.
*              5.1. If temporal Markov chain does not exist, then long-term
*                   averages will be used;
*              5.2. If temporal Markov chain exists, then by transition matrix
*                   and 1998 yield, the field-level can be calculated for 1999.
* Input: Field-level wheat yield ('c:\data\thesis\richmond\wtfld_new.txt'),
*        Field-level corn yield  ('c:\data\thesis\richmond\cnfld_new.txt'),
****************************************************************************/

options ps=3000;

libname bigtrans "c:\data\thesis\richmond\sasdata";

   ***** Wheat and corn field level yield from 1923 to 1998 and 1922 to 1998, resp. ;
   %macro fldtrans(cropfld,cropname);
         data bigtrans.&cropname;
              infile &cropfld dlm=',';
              input year f1-f8;  *order: F1B, F1C, F2, F3B, F4, F5, F7, F8B;
         run;
         proc sort; by year;run;
   %mend;
   %fldtrans('c:\data\thesis\richmond\wtfld_new.txt',WHEATFLD);
   %fldtrans('c:\data\thesis\richmond\cnfld_new.txt',CORNFLD);

%macro means(cropfile,   /*name of the crop yield file---input*/
             crop,       /*indicator of the crop */
             states,     /*indicate wheat or corn state of nature for 8 fields*/
             trans,      /*transition matrix*/
             fac         /*std factor used to calculate normal range*/
             );
  proc univariate data=bigtrans.&cropfile noprint;
      var f1-f8;
      output out=two mean=m1-m8 std=std1-std8;
  run;
  data &states;
     set two;
     %do i=1 %to 8;
         call symput('b'||trim(left(&i)),m&i-std&i * &fac);
         call symput('u'||trim(left(&i)),m&i+std&i * &fac);
```

```
   %end;

   if upcase(&crop)='W' then do;   ***specific for F2 wheat;
         call symput('b3',m3-std3*0.5);
         call symput('u3',m3+std3*0.7);
      end;
run;

data &cropfile;
   set bigtrans.&cropfile;
   %do i=1 %to 8;
       s1&i = (f&i<&&b&i); **for count of state;
       s3&i = (f&i>&&u&i);
       s2&i = 1 - s1&i - s3&i;
       ss&i = s1&i + 2*s2&i + 3*s3&i; **state of year;
   %end;
   tranyr=(_n_>1); *from 2nd year on;
   %do i=1 %to 3;
       %do j=1 %to 8;
           s1&i&j = (s&i&j=1 and lag1(s1&j)=1); *count transition;
           s2&i&j = (s&i&j=1 and lag1(s2&j)=1);
           s3&i&j = (s&i&j=1 and lag1(s3&j)=1);
       %end;
   %end;

run;

%let vars=s111-s118 s121-s128 s131-s138
          s211-s218 s221-s228 s231-s238
          s311-s318 s321-s328 s331-s338
          s11 -s18  s21 -s28  s31 -s38  tranyr;

proc univariate data=&cropfile noprint;
   output out=states sum=&vars;
   var &vars;
run;

data F1 f2 f3 f4 f5 f6 f7 f8;
   set states;
   %do i=1 %to 8; *calculate transistion prob matrx and output to indiv. Datasets;
       fld='f'||trim(left(&i));
       n11=s11&i;n21=s21&i;n31=s31&i;
       n12=s12&i;n22=s22&i;n32=s32&i;
       n13=s13&i;n23=s23&i;n33=s33&i;
       s1=s1&i; s2=s2&i; s3=s3&i;

       t11=s11&i/(s11&i+s12&i+s13&i);
       t12=s12&i/(s11&i+s12&i+s13&i);
       t13=1-t11-t12;

       t21=s21&i/(s21&i+s22&i+s23&i);
       t22=s22&i/(s21&i+s22&i+s23&i);
       t23=1-t21-t22;

       t31=s31&i/(s31&i+s32&i+s33&i);
       t32=s32&i/(s31&i+s32&i+s33&i);
       t33=1-t31-t32;
       output f&i;
   %end;
run;

data &trans; *transition matrices put together;
   set F1 f2 f3 f4 f5 f6 f7 f8;
   keep fld tranyr s1-s3 t11-t13 t21-t23 t31-t33 n11-n13 n21-n23 n31-n33;
run;

data testfreq;
   set &trans;
   keep fld n11-n13 n21-n23 n31-n33;
run;
```

```
proc transpose data=testfreq out=testfreq;
   by fld;
run;

data testfreq; *prepare for test on the existence of Markov chain;
  set testfreq;
  if substr(trim(left(_name_)),1,2)='N1' then start='1';
  if substr(trim(left(_name_)),1,2)='N2' then start='2';
  if substr(trim(left(_name_)),1,2)='N3' then start='3';
  if substr(trim(left(_name_)),3,1)='1' then finish='1';
  if substr(trim(left(_name_)),3,1)='2' then finish='2';
  if substr(trim(left(_name_)),3,1)='3' then finish='3';
run;
proc sort data=testfreq;by fld;run;

proc freq data=testfreq; *Fisher’s exact test for existence of MC;
   tables start*finish /exact;
   weight col1;
   by fld;
run;  *Results show that for wheat, MC does not exist;

 data stat1998;
   set &cropfile;
   if year=1998;
   %do i=1 %to 8;
       s98_&i=(s1&i=1)+2*(s2&i=1)+3*(s3&i=1); /*states of nature in 1998*/
   %end;
 run;

 proc iml;  /******now to calculate the optimal yield for 1999****/
    use &cropfile;
    read all var{year} into year;
    %do i=1 %to 8;
        read all var{f&i} into f&i;
        read all var{ss&i} into ss&i;
    %end;
    ylds=f1||f2||f3||f4||f5||f6||f7||f8;
    state=ss1||ss2||ss3||ss4||ss5||ss6||ss7||ss8;
    t=nrow(ylds);
    close &cropfile;

    staterec=year||state;
    varnames={year s1 s2 s3 s4 s5 s6 s7 s8};
    create bigtrans.&states from staterec [colname=varnames]; *states of nature;
    append from staterec;

    use &trans;
    %do i=1 %to 3;
        read all var{s&i} into s&i;
        read all var{t1&i} into t1&i;
        read all var{t2&i} into t2&i;
        read all var{t3&i} into t3&i;
    %end;
    close &trans;

    nprice=0.25;     /*N price per pound*/
    haul=0.15;       /*hauling cost per bushel*/

    If &crop='W' then
       do;
         cprice=2.87;       /*wheat price*/
         Nbushel=1.25;      /*N needed per bushel*/
         penalty=0.8;       /*yield loss per pound N underapplied*/
      end;
    else
      do;
         cprice=2.10; /*wheat price*/
         Nbushel=1;   /*N needed per bushel*/
         penalty=1;   /*yield loss per pound N underapplied*/
      end;
```

```
    %do i=1 %to 8;
        optimal&i=J(t,4,0); /*yield level, N level, minimum cost, etc.*/
        if &crop='W' then
           do;
              prob1=1/t; prob2=1/t; prob3=1/t;
           end;
        else
           do;
              prob1=t21[&i,1]/s1[&i,1];
              prob2=t22[&i,1]/s2[&i,1];
              prob3=t23[&i,1]/s3[&i,1];
           end;

        yield=0; *store the expected yield for the field in 1999;
        n1=0;n2=0;n3=0;
        do j=1 to t;
           if &crop='W' then nrate=ylds[j,&i]*Nbushel;    /*N applied for jth level*/
           if &crop='C' then nrate=ylds[j,&i]*Nbushel-20; /*N applied for jth level*/
           cost=0; *store the expected cost given Nrate;

           do k=1 to t;
              if &crop='W' then
                 nneed=ylds[k,&i]*Nbushel;    /* N needed for kth year for wheat*/
              else
                 nneed=ylds[k,&i]*Nbushel-20; /* N needed for kth year for corn */

              /* below 3 if's are for Markov chain optimal yields */
              if state[k,&i]=1 then prob=prob1;
              if state[k,&i]=2 then prob=prob2;
              if state[k,&i]=3 then prob=prob3;

              cost=cost+prob*(nprice*max(nrate-nneed,0)
                    -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
           end;
           optimal&i[j,1]=&i;          *mark of the field;
           optimal&i[j,2]=ylds[j,&i];  *yield;
           optimal&i[j,3]=nrate;       *N applied for this yield;
           optimal&i[j,4]=cost;        *expected cost for this N rate;

           if state[j,&i]=1 then n1=n1+ylds[j,&i];
           if state[j,&i]=2 then n2=n2+ylds[j,&i];
           if state[j,&i]=3 then n3=n3+ylds[j,&i];
        end;

        yield=prob1*n1 + prob2*n2 + prob3*n3;
        print "Expected yield for field &i in 1998 is:" yield "total year is: " t;
        varnames={fieldno yield N cost};
        create cost&i from optimal&i [colname=varnames];
        append from optimal&i;
        close cost&i;
        sort cost&i by cost;
    %end;
quit;

/* optimal field-level yields */
%do i=1 %to 8;
    data cost&i;
         set cost&i;
         if _n_=1;  *smallest cost;
    run;
%end;

data cost;
     set cost1 cost2 cost3 cost4 cost5 cost6 cost7 cost8;
run;

proc print data=cost;run;

data bigtrans.&trans;  *permanent file for transition matrices;
   set &trans;
run;
```

```
%mend;

%means(wheatfld,'W',wfldstat,wtrans,0.7);
%means(cornfld,'C',cfldstat,ctrans,0.8);



/* * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * *
 * prog name = soilunif.sas
 *
 * Calculate soil-zone uniform optimal Yields and expected costs
 *
 * output: onscreen display
 * * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * */

libname big 'c:\data\thesis\richmond\sasdata';

      /******* This macro generate .sd2 from .dbf files ******
         %macro dbf2sd2(filein,fileout);

             proc import datafile=&filein out=big.&fileout dbms=dbf replace;
                  format soil 3. count 3.;
             run;
         %mend;
         %dbf2sd2('c:\data\thesis\sas\data\g30f1b.dbf', g30f1b);
         %dbf2sd2('c:\data\thesis\sas\data\g30f1c.dbf', g30f1c);
         %dbf2sd2('c:\data\thesis\sas\data\g30f3b.dbf', g30f3b);
         %dbf2sd2('c:\data\thesis\sas\data\g30f4.dbf',  g30f4);
         %dbf2sd2('c:\data\thesis\sas\data\g30f2.dbf',  g30f2);
         %dbf2sd2('c:\data\thesis\sas\data\g30f5.dbf',  g30f5);
         %dbf2sd2('c:\data\thesis\sas\data\g30f7.dbf',  g30f7);
         %dbf2sd2('c:\data\thesis\sas\data\g30f8b.dbf', g30f8b);
         endsas;
      ********************/
data facs;
   input crop $ soil fld fac; /*Rescaling factors based on Table 3.1 and 3.6 */
   cards;
         W     148    1      0.36145
         W     151    1      0.96386
         W     168    1      0.36145
         W     171    1      1.08434
         W     172    1      1.08434
         W     148    2      0.37975
         W     171    2      1.13924
         W     172    2      1.13924
         W     151    3      0.93023
         W     168    3      0.34884
         W     171    3      1.04651
         W     172    3      1.04651
         W     148    4      0.35294
         W     171    4      1.05882
         W     172    4      1.05882
         C     168    1      0.41667
         C     171    1      1.02564
         C     172    1      1.02564
         C     148    2      0.43333
         C     151    2      0.86667
         C     171    2      1.06667
         C     172    2      1.06667
         C     151    3      0.82803
         C     161    3      0.76433
         C     171    3      1.01911
         C     172    3      1.01911
         C     148    4      0.44828
         C     151    4      0.89655
         C     171    4      1.10345
         C     172    4      1.10345
      ;
run;
```

```
%macro soilunif (fldsoil,   /*Field file containing soil number and cell counts*/
                 crop,      /*Indicate what crop it is */
                 fldyld,    /*field-level yield */
                    ord,    /*ith field for given crop. Eg. 3rd fld for wht is F3B*/
                 trans,     /*transition matrix */
                 states     /*states of nature for each field */
                 )
                 ;
   proc sort data=big.&fldsoil out=&fldsoil (keep=soil count);
      by soil;
   run;
   proc univariate data=&fldsoil noprint;
      output out=one sum=area;
      var count;
      by soil;
   run;
   proc sort data=one;by soil;run;

   data two;
      set facs;
      if crop=&crop and fld=&ord+0;
      keep soil fac;
   run;
   proc sort data=two;by soil;run;

   data &fldsoil;
      merge one (in=fld) two;
      by soil;
      area=area*9*0.00025;
      count=1;  /* to generate a macro variable of number of soil zones */
      if fld;
   run;
           /***generate the macro number of soil zones in the field ***/
               proc univariate data=&fldsoil noprint;
                    output out=numbersl sum=aaaa;
                    var count;
               run;
               data numbersl;
                    set numbersl;
                    call symput('msoil',aaaa);
               run;

   /*proc print data=&fldsoil;title1 "field is: &fldsoil" ; run;*/

   data fldyld;
      set big.&fldyld;
      if &crop='W' then do;
          if &ord+0=1 then yield=f1;
          if &ord+0=2 then yield=f2;
          if &ord+0=3 then yield=f4;
          if &ord+0=4 then yield=f5;
          end;
      else do;
          if &ord+0=1 then yield=f3;
          if &ord+0=2 then yield=f6;
          if &ord+0=3 then yield=f7;
          if &ord+0=4 then yield=f8;
          end;
      keep year yield;
   run;
   /*proc print data=fldyld;run;*/

   data trans;
      set big.&trans;
      if &crop='W' then do;
         if _n_=1 or _n_=2 or _n_=4 or _n_=5;
         end;
      else do;
         if _n_=3 or _n_=6 or _n_=7 or _n_=8;
         end;
   run;
```

```
data trans;
   set trans;
   if _n_=&ord+0;
run;

data state;
    set big.&states;
    if &crop='W' then do;
       stat1=s1;stat2=s2;stat3=s4;stat4=s5;
       end;
    else do;
       stat1=s3;stat2=s6;stat3=s7;stat4=s8;
       end;
    stat=stat&ord;
    keep year stat;
run;

proc iml;
   use &fldsoil;
   read all var{area soil fac};
   close &fldsoil;
   nsoil=nrow(soil);
   fldarea=0;
   do i=1 to nsoil;
       fldarea=fldarea+area[i,1];
   end;

   use fldyld;
   read all var{yield} into orgyld;
   close fldyld;
   t=nrow(orgyld);

   /*generate a yield sequence for each soil in the field*/
   soilyld=J(t,nsoil,0);
   do i=1 to t;
      do j=1 to nsoil;
         soilyld[i,j]=round(orgyld[i,1]*fac[j,1]);
      end;
   end;
            /*Further, generate a combined yield sequece: stack together*/
            bigyld=soilyld[,1];
            do j=1 to nsoil;
               bigyld=bigyld//soilyld[,j];
            end;
   use trans;
   read all var{s1 s2 s3 t11 t12 t13 t21 t22 t23 t31 t32 t33};
   close trans;

   use state;
   read all var{year stat};
   close state;

  /**** Now adopt from fldtransition.sas to calculated optimal yield and costs ****/
   nprice=0.25;        /*N price per pound*/
   haul=0.15;          /*hauling cost per bushel*/
   If &crop='W' then do;
         cprice=2.87;  /*wheat price*/
         Nbushel=1.25; /*N needed per bushel*/
         penalty=0.8;  /*yield loss per pound N underapplied*/
         prob1=1/t;    /* used to calculated optimal yield and cost*/
         prob2=1/t;
         prob3=1/t; /*above is valid because no Markov chain detected for wheat*/
      end;
   else do;
         cprice=2.10; /*wheat price*/
         Nbushel=1;   /*N needed per bushel*/
         penalty=1;   /*yield loss per pound N underapplied*/
         prob1=t21[1,1]/s1[1,1];
         prob2=t22[1,1]/s2[1,1];
         prob3=t23[1,1]/s3[1,1];
      end;  /*above is valid because states in 1998 for corn are all normal*/
```

```
  bigopt=J(t*nsoil,3,0);        /*target yield will be evaluated for t*nsoil values*/
  n1=J(1,nsoil,0); n2=n1; n3=n1;/*In order to calculate fld-level exp yld */

  do i=1 to t*nsoil;
      if &crop='W' then nrate=bigyld[i,1]*Nbushel; /*N rate applied for wheat*/
      else nrate=bigyld[i,1]*Nbushel-20;           /*N rate applied for corn*/

        %let szon=1;
        %do %while (&szon<&msoil+1);
            cost&szon=0;    /*Each zone has a exp cost.
                              Fld-level cost is weighted average of them*/
            expyld&szon=0;  /*Exp Yield for each soil zone */
            %let szon=%eval(&szon+1);
        %end;

      do j=1 to t;
         if &crop='W' then do;
               %let szon=1;
               %do %while (&szon<&msoil+1);
                   nneed&szon=soilyld[j,&szon]*Nbushel; /* N need for wht on zone */
                   %let szon=%eval(&szon+1);
               %end;
            end;
         else do;
                 %let szon=1;
                 %do %while (&szon<&msoil+1);
                     nneed&szon=soilyld[j,&szon]*Nbushel-20; /*N need for crn on zone*/
                     %let szon=%eval(&szon+1);
                 %end;
              end;

                /* below 3 if's are for Markov chain optimal yields */
                   if stat[j,1]=1 then prob=prob1;
                   if stat[j,1]=2 then prob=prob2;
                   if stat[j,1]=3 then prob=prob3;
         %let szon=1;
         %do %while (&szon<&msoil+1);
             cost&szon=cost&szon+prob*(nprice*max(nrate-nneed&szon,0)
                      -((cprice-haul)*penalty-Nprice)*min(nrate-nneed&szon,0));
             %let szon=%eval(&szon+1);
         %end;

      end;

      %let szon=1;cost=0;
      %do %while (&szon<&msoil+1);
          cost=cost + cost&szon*area[&szon+0,1]/fldarea; /*expected cost given Nrate*/
          %let szon=%eval(&szon+1);
      %end;

      bigopt[i,1]=bigyld[i,1]; /*Actual (simulated) crop yield*/
      bigopt[i,2]=nrate;       /*N applied for this yield over the whole field*/
      bigopt[i,3]=cost;        /*Expected cost for this N rate */

      if mod(i,t)=0 then j=t;
      else j=mod(i,t);
           /**In order to calculated field-level expected yield**/
              do szon=1 to nsoil;
                 if stat[j,1]=1 then n1[1,szon]=n1[1,szon] + bigyld[i,1];
                 if stat[j,1]=2 then n2[1,szon]=n2[1,szon] + bigyld[i,1];
                 if stat[j,1]=3 then n3[1,szon]=n3[1,szon] + bigyld[i,1];
              end;

      %let szon=1;
      %do %while (&szon<&msoil+1);
          expyld&szon=expyld&szon + prob1*n1[1,&szon]
                     + prob2*n2[1,&szon] + prob3*n3[1,&szon];
          %let szon=%eval(&szon+1);
      %end;
end;
```

```
    expyld=0;    /*expected field-level yield*/
    %let szon=1;
    %do %while (&szon<&msoil+1);
        expyld=expyld + expyld&szon*area[&szon,1]/fldarea;
        %let szon=%eval(&szon+1);
    %end;

    varnames={Yield Nrate expcost};
    create okay from bigopt [colname=varnames];
    append from bigopt;
    close okay;
    sort okay by expcost;

  quit;

  /* optimal field-level yields and expected cost*/
  data okay; set okay;if _n_=1;run;
  proc print data=okay;
       title1 "the field is ======= &fldsoil =========";
  run;

%mend;

%soilunif(g30f1b,'W',wheatfld,1,wtrans,wfldstat);
%soilunif(g30f1c,'W',wheatfld,2,wtrans,wfldstat);
%soilunif(g30f3b,'W',wheatfld,3,wtrans,wfldstat);
%soilunif(g30f4,'W',wheatfld,4,wtrans,wfldstat);

%soilunif(g30f2,'C',cornfld,1,ctrans,cfldstat);
%soilunif(g30f5,'C',cornfld,2,ctrans,cfldstat);
%soilunif(g30f7,'C',cornfld,3,ctrans,cfldstat);
%soilunif(g30f8b,'C',cornfld,4,ctrans,cfldstat);
```

```
/*******************************************************************************************
* prog name = cluster_spatial.sas
*
* Abstract: this program groups SMUs in each field into 5 clusters according to obs yld
*           in 1995 and 1997. Calculates the area of each cluster, mean yield (each yr),
*           and test for the significance of cluster difference.
*
* input:  big.g30f1b.sd2 etc;
*         big.whtseed.sd2 etc;
*
* output: Table 3-15 and revised G30F1b.sd2 etc. (G30f1b.sd2 should have cluster numbers)
*
* Next: SMUtransition.sas
*******************************************************************************************/

libname big 'c:\data\thesis\richmond\sasdata\';

       /*** Test to see if field yield in 1999 is diff among soil zones ***
        %macro try (fld,crop95, crop97, crop99);
             data one;
                 set big.&fld;
                 diff=&crop97 - &crop95;
             run;

             proc glm
                 data=one;
                 class soil;
                 model &crop99 = soil;
             run;
             proc glm
                 data=one;
                 class cluster;
                 model diff = cluster;
             run;
```

```
        %mend;

        %try(g30f1b,w95, w97,w99);
        %try(g30f1c,w95, w97,w99);
        %try(g30f3b,w95, w97,w99);
        %try(g30f4, w95, w97,w99);
        %try(g30f2, c95, c97,c99);
        %try(g30f5, c95, c97,c99);
        %try(g30f7, c95, c97,c99);
        %try(g30f8b,c95, c97,c99);
        endsas
        *********/

%macro clust (orgfld,   /*original field yield file     */
              fld,      /*name of the field (f1b etc.)  */
              fldname,  /*indicator of field name ('F1B')*/
              crop,     /*indicator of 1995,97, 99 crop  */
              crop95,   /*variable name for 1995 yield in original file*/
              crop97,   /*variable name for 1997 yield in original file*/
              seed      /*seed file for clustering*/
              )
              ;

    proc sort
         data = big.&orgfld
         out = &fld;
         by &crop95 &crop97;
    run;

    proc fastclus data=&fld
                  seed=big.&seed
                  out=&fld
                  mean=clus
                  radius=8
                  maxc=5
                  maxiter=140
                  noprint;
         var &crop95 &crop97;
    run;

    proc sort
         data=&fld (drop=distance);
         by cluster;
         run;

 /*   proc sort
         data=clus (rename=(&crop95=mclus95 &crop97=mclus97) keep=cluster &crop95 &crop97);
         by cluster;
         run;

    data &fld;
       merge &fld (in=org) clus;
       by cluster;
       if org;
    run;
 */

    proc glm
         data=&fld;
         class cluster;
         model &crop95 = cluster;
    run;

    proc glm
         data=&fld;
         class cluster;
         model &crop97 = cluster;
    run;
```

```
    data big.&orgfld;
       set &fld;
    run;

%mend;
%clust(g30f1b, f1b, 'F1B','W',w95,w97,whtseed);
%clust(g30f1c, f1c, 'F1C','W',w95,w97,whtseed);
%clust(g30f3b, f3b, 'F3B','W',w95,w97,whtseed);
%clust(g30f4 , f4 , 'F4' ,'W',w95,w97,whtseed);
%clust(g30f2 , f2 , 'F2' ,'C',c95,c97,crnseed);
%clust(g30f5 , f5 , 'F5' ,'C',c95,c97,crnseed);
%clust(g30f7 , f7 , 'F7' ,'C',c95,c97,crnseed);
%clust(g30f8b, f8b, 'F8B','C',c95,c97,crnseed);



/****************************************************************************************
* Program name = SMUtransition.sas
*
* Abstract: This program estimates within-field transition matrix (spatial transition).
*           Several steps are involved:
*           1. Mark states of nature for each SMU within each functional zone;
*           2. Test for existence of Markov Chain;
*           3. Output transition matrices and states of nature;
*
* input: smu_resc.sd2 (output of soilmean.sas and soilrescale.sas)
*        Table 3-15.
*
* output: &smutrans.sd2, &smufld.sd2. e.g. smutrf1b.sd2, smuf1b.sd2, Permanent file.
*         (To be used by another two programs to calculate SMU optimal yields)
*****************************************************************************************/

option ps=3000;

*first step: rescale yield of each SMU to the level of 1999 field-level;

/* y1999: Expected field-level yield for 1999
   y1997: 1997 yield need to be adjusted by this factor
   y1995: 1995 yield need to be adjusted by this factor*/
data fldlevel;
  input crop $ fld $ y1995 y1997 y1999; /*Field level rescaling. From Table 3-  */
  ratio5=y1999/y1995;
  ratio7=y1999/y1997;
  datalines;
   W F1B  85  74  61
   W F1C  86  68  64
   W F3B  93  88  72
   W F4  101  92  74
   C F2  118 105 118
   C F5  177  87 143
   C F7  188 125 167
   C F8B 182 111 164
   ;
run;

libname big 'c:\data\thesis\richmond\sasdata';

%macro grd30(fldfile,  /*  location of the file                  */
             fldname,  /*  name of the field                     */
             crop,     /*  name of the crop                      */
             crop95,   /*  original variable name for 1995 yield */
             crop97,   /*  original variable name for 1997 yield */
             crop99,   /*  original variable name for 1999 yield */
             smutrans, /*  SMU transition matrix for the field   */
             smufld    /*  SMU states and yields for the field   */
             );

   /* rescale the orginal yield data for 1995 and 1997 in order to estimate transition*/
      proc sort data=fldlevel
                     (keep= fld ratio5 ratio7);
```

```
        by fld;
   run;

   data field;
        set big.&fldfile;
        length fld $ 8;
        fld=&fldname;
   run;

   proc sort data=field;
        by fld;

   run;

   data field;
        merge field (in=org) fldlevel;
        by fld;
        if org;
        y1995=round(ratio5*&crop95,1); /*95 SMU yld adjust to level of predicted 1999*/
        y1997=round(ratio7*&crop97,1); /*97 SMU yld adjust to level of predicted 1999*/
        y1999=&crop99;
   keep x y fld cluster soil count y1995 y1997 y1999 &crop95 &crop97;

/* Get mean yield and std for each soil in order
   to determine the state of nature for each SMU */
proc sort
   data=field;
   by cluster;
run;

proc univariate data=field noprint;
      var y1995 y1997;
      by cluster;
      output out=meanstd mean=m95 m97 std=var95 var97;
      title1 'Field = ' &fldname;
run;

data field;
   merge field (in=fldin) meanstd (in=clus);
      by cluster;
      if fldin and clus;
run;

/* Determine state of nature for each SMU in 1995 and 1997 */
data field;
   set field;
   m95=round(m95,1);
   m97=round(m97,1);

   if &crop='W' then
      do;
        v95=round(var95*0.7,1);
        v97=round(var97*0.7,1);
     end;
   else
      do;
         v95=round(var95*0.8,1);
         v97=round(var97*0.8,1);
      end;
       s95=(y1995 ge m95+v95)*3 + (y1995 lt m95+v95)*(y1995 ge m95-v95)*2
                                + (y1995 lt m95-v95)*1;
       s97=(y1997 ge m97+v97)*3 + (y1997 lt m97+v97)*(y1997 ge m97-v97)*2
                                + (y1997 lt m97-v97)*1;
run;

/* Testing if the transiion exists */
proc freq
   data=field;
   output out=out_p exact; /*this out is to get p-value (var name = p_exact2)*/
      tables s95*s97 /exact out=trantest outpct; /*percents for transition*/
run;
```

```
proc print
   data=trantest;
   var s95 s97 pct_row;
run;

      /* get ready to deal with the situation where no Markov exists **/
           data out_p;
                set out_p;
                call symput('pvalue',p_exact2);
           run;
           data stat9597;
              set field;
              s95_1 = (s95=1);s95_2 = (s95=2);s95_3 = (s95=3);
              s97_1 = (s97=1);s97_2 = (s97=2);s97_3 = (s97=3);
           run;
           proc univariate data=stat9597 noprint;
                output out=stat9597 sum=s95_1-s95_3 s97_1 - s97_3;
                var s95_1-s95_3 s97_1 - s97_3;
           run;
           data stat9597;
                set stat9597;
                call symput("s951", s95_1);
                call symput("s952", s95_2);
                call symput("s953", s95_3);
                call symput("s971", s97_1);
                call symput("s972", s97_2);
                call symput("s973", s97_3);
           run;

proc sort data=trantest; by s95 s97; run;

data trantest;
  retain fr0m to1 to2 to3;
     format to1 to2 to3 5.3;
  set trantest;
  by s95 s97;

  if s95=1 then fr0m="FROM 1";
  if s95=2 then fr0m="FROM 2";
  if s95=3 then fr0m="FROM 3";

  if &pvalue <= 0.05 then do;
        if s97=1 then to1=pct_row*0.01;
        if s97=2 then to2=pct_row*0.01;
        if s97=3 then to3=pct_row*0.01;
     end;
  else do;
          to1=(&s951 + &s971) / ((&s951 + &s952 + &s953)*2);
          to2=(&s952 + &s972) / ((&s951 + &s952 + &s953)*2);
          to3=1-to1-to2;
       end;

  if last.s95 then
     do;
        output;
        to1=0;to2=0;to3=0;
    end;
     keep fr0m to1 to2 to3;
run;

data big.&smutrans;
   set trantest;
   if to1=. then to1=0;
   if to2=. then to2=0;
   if to3=. then to3=0;
run;

proc sort
    data=field
    out=big.&smufld (drop= var95 var97 v95 v97);
```

```
       by cluster s95 s97;
   run;


%mend;

%grd30(g30f1b, 'F1B', 'W', W95, W97,w99,smutrf1b,smuf1b);
%grd30(g30f1c, 'F1C', 'W', W95, W97,w99,smutrf1c,smuf1c);
%grd30(g30f3b, 'F3B', 'W', W95, W97,w99,smutrf3b,smuf3b);
%grd30(g30f4,  'F4',  'W', W95, W97,w99,smutrf4, smuf4);
%grd30(g30f2,  'F2',  'C',  C95, C97,c99,smutrf2,smuf2);
%grd30(g30f5,  'F5',  'C',  C95, C97,c99,smutrf5,smuf5);
%grd30(g30f7,  'F7',  'C',  C95, C97,c99,smutrf7,smuf7);
%grd30(g30f8b, 'F8B', 'C',  C95, C97,c99,smutrf8b,smuf8b);







/* * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * *
 * program name = smu_FV.sas
 *
 * Purpose: Calculate SMU variable rate OPTIMAL Yields and expected costs.
 *          Same as for functional zone variable strategy.
 *          So this program is for generic SMU (100 counts).next program will
 *          assign to each real SMU a pair of N rate and optimal yield.
 * NOTE:    functional zones are based on clustering.
 *
 * Input:   smutrf1b.sd2 etc and smuf1b.sd2 etc.Generated by SMUtransition.sas
 *
 * output:  smu_var0.sd2
 *
 * date:    3JUN2001
 * * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * */

libname big 'c:\data\thesis\richmond\sasdata';

%macro smuvar (crop,    /*indicate crop is wheat or corn                */
               fldname, /*field name                                    */
               f,       /*file name for the field (used only once)     */
               trans,   /*transition matrix                            */
               field,   /*field with count,cluster,y1995,y1997, and s97*/
               clusno   /*the cluste                                    */
               )
               ;
   data y95;
      set big.&field;
      stat=s97;
      yield=y1995;
      area1=0;area2=0;areat3=0; /*used to count areas of stat in 1997*/
      if cluster=&clusno;
   run;

   data y97;
      set big.&field;
      stat=s97;
      yield=y1997;
      if stat=1 then area1=count*0.00025*9; else area1=0; /*count areas of stat in 1997*/
      if stat=2 then area2=count*0.00025*9; else area2=0; /*count areas of stat in 1997*/
      if stat=3 then area3=count*0.00025*9; else area3=0; /*count areas of stat in 1997*/
      if cluster=&clusno;
   run;

   data obsyld;
      set y95 y97;  /*Observed yields. Used as candidates for zone target*/
      if stat=1 then s1=1;
      if stat=2 then s2=1; else s2=0;
      if stat=3 then s3=1; else s3=0;
   run;
```

```
      /*obtain # of each states observed*/
        proc univariate data=obsyld noprint;
             output out=one sum=s1-s3 area1-area3;
             var s1-s3 area1-area3;
        run;

        data one;
           set one;
           call symput('s1',s1);
           call symput('s2',s2);
           call symput('s3',s3);
           call symput('area1',area1);
           call symput('area2',area2);
           call symput('area3',area3);
        run;


proc iml;

   use obsyld;
   read all var{yield count stat};
   close obsyld;
   t=nrow(yield); /* # of yields to try */

   use big.&trans;
   read all var{to1 to2 to3};
   close big.&trans;

  /**** Now adopt from SOILUNIF.sas to calculated optimal yield and costs ****/

   nprice=0.25;        /*N price per pound*/
   haul=0.15;          /*hauling cost per bushel*/
   If &crop='W' then do;
         cprice=2.87;  /*wheat price*/
         Nbushel=1.25; /*N needed per bushel*/
         penalty=0.8;  /*yield loss per pound N underapplied*/
      end;
   else do;
         cprice=2.10; /*wheat price*/
         Nbushel=1;   /*N needed per bushel*/
         penalty=1;   /*yield loss per pound N underapplied*/
      end;

   do k=1 to 3;

      prob1=to1[k,1]/&s1;  /*suppose SMU is state i in 1997 */
      prob2=to2[k,1]/&s2;
      prob3=to3[k,1]/&s3;

      bigopt=J(t,3,0);     /*target yield will be evaluated for t*nsoil values*/
      expyld=J(t,1,1);     /*coloum for expected yield */
      area=J(t,1,1);       /*column for area for each SMU-type*/
      n1=0; n2=0; n3=0;    /*In order to calculate exp yld */

      do i=1 to t;
         if &crop='W' then nrate=yield[i,1]*Nbushel; /*N rate applied for wheat*/
         else nrate=yield[i,1]*Nbushel-20;           /*N rate applied for corn*/

         cost=0;
         do j=1 to t;
            if &crop='W' then do;
                   nneed=yield[j,1]*Nbushel;    /* N needed for wheat on zone szon*/
               end;
            else do;
                   nneed=yield[j,1]*Nbushel-20; /* N needed for corn on zone szon */
               end;

            /* below 3 if's are for Markov chain optimal yields */
                if stat[j,1]=1 then prob=prob1;
                if stat[j,1]=2 then prob=prob2;
                if stat[j,1]=3 then prob=prob3;
```

```
              cost=cost+prob*(nprice*max(nrate-nneed,0)
                       -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
           end;

           bigopt[i,1]=yield[i,1];  /*Actual (observed) crop yield for a SMU*/
           bigopt[i,2]=nrate;       /*N applied for this yield over a SMU*/
           bigopt[i,3]=cost;        /*Expected cost for this N rate */

           if stat[i,1]=1 then n1=n1 + yield[i,1];
           if stat[i,1]=2 then n2=n2 + yield[i,1];
           if stat[i,1]=3 then n3=n3 + yield[i,1];
        end;
        expyld=(prob1*n1 + prob2*n2 + prob3*n3)*expyld;

        if k=1 then area=area*&area1;
        if k=2 then area=area*&area2;
        if k=3 then area=area*&area3;

        bigopt=area||expyld||bigopt;

        epp=expyld[1,1];
        print &fldname &clusno epp k;

        varnames={area expyld optYld Nrate expcost};
        if k=1 then do;
           create okay1 from bigopt [colname=varnames];
           append from bigopt;
           close okay1;
           sort okay1 by expcost;
        end;
        if k=2 then do;
           create okay2 from bigopt [colname=varnames];
           append from bigopt;
           close okay2;
           sort okay2 by expcost;
        end;
        if k=3 then do;
           create okay3 from bigopt [colname=varnames];
           append from bigopt;
           close okay3;
           sort okay3 by expcost;
        end;

     end;

  quit;

  /* optimal SMU yields and expected costs*/

  %do k=1 %to 3;
      data okay&k;
        length fld $ 8;
        format clusno 1. s97 1.;
        set okay&k;
        if _n_=1;
        fld=&fldname;
        clusno=&clusno;
        s97=&k+0; /*state of nature in 1997*/
        run;
  %end;
  data f&clusno;
     set okay1 okay2 okay3;
  run;
%mend;

%smuvar('W','F1B',f1b,smutrf1b,smuf1b,1); /*soil zones are:148 151 168 171 172*/
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,2);
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,3);
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,4);
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,5);
data FF1b;
```

```
  set f1 f2 f3 f4 f5;
run;

%smuvar('W','F1C',f1c,smutrf1c,smuf1c,1); /*soil zones are:148 171 172*/
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,2);
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,3);
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,4);
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,5);

data FF1c;
  set f1 f2 f3 f4 f5;
run;

%smuvar('W','F3B',f3b,smutrf3b,smuf3b,1); /*soil zones are:151 168 171 172*/
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,2);
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,3);
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,4);
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,5);
data FF3b;
  set f1 f2 f3 f4 f5;
run;

%smuvar('W','F4',f4,smutrf4,smuf4,1); /*soil zones are:148 171 172*/
%smuvar('W','F4',f4,smutrf4,smuf4,2);
%smuvar('W','F4',f4,smutrf4,smuf4,3);
%smuvar('W','F4',f4,smutrf4,smuf4,4);
%smuvar('W','F4',f4,smutrf4,smuf4,5);

data FF4;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F2',f2,smutrf2,smuf2,1); /*soil zones are:168 171 172*/
%smuvar('C','F2',f2,smutrf2,smuf2,2);
%smuvar('C','F2',f2,smutrf2,smuf2,3);
%smuvar('C','F2',f2,smutrf2,smuf2,4);
%smuvar('C','F2',f2,smutrf2,smuf2,5);
data FF2;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F5',f5,smutrf5,smuf5,1); /*soil zones are:148 151 171 172*/
%smuvar('C','F5',f5,smutrf5,smuf5,2);
%smuvar('C','F5',f5,smutrf5,smuf5,3);
%smuvar('C','F5',f5,smutrf5,smuf5,4);
%smuvar('C','F5',f5,smutrf5,smuf5,5);
data FF5;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F7',f7,smutrf7,smuf7,1); /*soil zones are:151 161 171 172*/
%smuvar('C','F7',f7,smutrf7,smuf7,2);
%smuvar('C','F7',f7,smutrf7,smuf7,3);
%smuvar('C','F7',f7,smutrf7,smuf7,4);
%smuvar('C','F7',f7,smutrf7,smuf7,5);
data FF7;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F8B',f8b,smutrf8B,smuf8B,1); /*soil zones are:148 151 171 172*/
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,2);
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,3);
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,4);
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,5);
data FF8b;
  set f1 f2 f3 f4 f5;
run;

data big.smu_var0;
  set ff1b ff1c ff3b ff4 ff2 ff5 ff7 ff8b;
run;
```

```
/***** NEXT: ---- determine NRATE for each SMU in 1999 *******
   *****
   ***** THIS IS DEFERRED TO PROGRAM SMU_NRATE1999.SAS
   ******************************************************/




/* * * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * ********
 * program name = smu_nrate1999.sas
 *
 * Abstract: 1.Assign for each SMU the N rate in 1999 under variable rate application.
 *           2.Determine field-level N rate in 1999 under functional uniform application.
 *             2a. Also optimal yields (field-level).
 *             2b. Also expected costs (field-level).
 *
 *input: c:\data\thesis\richmond\sasdata\smu_var0.sd2
 *       (generic SMU: fld cluster, state, area, Nrate for 1999)
 *      c:\data\thesis\richmond\sasdata\smuf1b.sd2, ..., smuf8b.sd2
 *       (actual SMU: count,cluster,fld,s97 etc).
 *
 *NOTE: here functional zones are based on clusters.
 *
 *output: Screen output
 * * * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * *********/

libname big 'c:\data\thesis\richmond\sasdata';

/****** 1: assign N rate in 1999 to each SMU under variable rate application ******/

data smu_var0 (rename=(clusno=cluster nrate=n99_FV));
   set big.smu_var0;
run;

proc sort data=smu_var0;
     by fld cluster s97;
run;
        /*******calculated weighted avg of exp and opt yld and cost for each field
                with SMU variable information. For Table 3-22.
               data one; set smu_var0;
                    expyld=area*expyld;
                    optyld=area*optyld;
                    expcost=area*expcost;
               run;

               proc sort data=one; by fld;run;

               proc univariate data=one noprint;
                    output out=one sum=expyld optyld expcost totarea;
                    var expyld optyld expcost area;
                    by fld;
               run;

               data one; set one;
                    expyld=expyld/totarea;
                    optyld=optyld/totarea;
                    expcost=expcost/totarea;
               run;
               proc print data=one; run;
         *********************/

%macro smuvar (fldname,  /*field name                                  */
               field,    /*the field data (name,count,cluster,yield etc) */
               crop,     /*crop indicator                              */
               trans,    /*within-field transition matrix file         */
               fldfunc   /*the functional zone information for a field*/
               );
   proc sort data=big.&field out=&field;
      by fld cluster s97;
   run;
```

```
data &field;
   merge &field (in=onebyone) smu_var0;
   by fld cluster s97;
   fzone=put(cluster,3.)||'_'||put(s97,1.); *generate mark for functional zones;
   if onebyone;
   drop area expyld optyld expcost;
run;

/*proc print data=&field;run;*/

data &trans; *the transition matrix for the all SMU in a field;
   set big.&trans;
   if _n_=1 then do;
            call symput('p11',to1);
            call symput('p12',to2);
            call symput('p13',to3);
            end;
   if _n_=2 then do;
            call symput('p21',to1);
            call symput('p22',to2);
            call symput('p23',to3);
            end;
   if _n_=3 then do;
            call symput('p31',to1);
            call symput('p32',to2);
            call symput('p33',to3);
            end;
run;

/* now can treat each functional zone the same way as a soil zone. So can adopt
   the program from SOILUNIF.SAS as the following */

data &fldfunc;
   set smu_var0 (where=(fld=&fldname));
   fzone=put(cluster,1.)||'_'||put(s97,1.); *generate mark for functional zones;
   count=1;
   keep fzone area count;
run;

        /***generate the macro number of func zones in the field ***/

            proc univariate data=&fldfunc noprint;
                 output out=numbersl sum=aaaa;
                 var count;
            run;
            data numbersl;
                 set numbersl;
                 call symput('nfunc',aaaa);
            run;

/** dataset of candidate target yield for the field**/
data y95;
   set &field;
   state=s95;
   yield=y1995;
run;

data y97;
   set &field;
   state=s97;
   yield=y1997;
run;

data obsyld;
   set y95 y97;  /*Observed yields. Used as candidates for zone target*/
run;
/*proc print data=obsyld;run;     */

     /*order # for each sol zone in order to single out a yld seq for each func zone*/
       proc sort data=&field out=one;
```

```
            by cluster;
       run;
       proc univariate data=one noprint;
            output out=one sum=sno;
            var count;  /*actually count is not used at all later*/
            by cluster;
       run;
       data one;
            set one;
            sno=_n_;
        run;
       proc sort data=one (keep=cluster sno);
            by cluster;
       run;
          /*assign to the obsyld the soil zone number*/
          proc sort data=obsyld;
               by cluster;
          run;
          data obsyld;
             merge one obsyld (in=obs);
             by cluster;
          run;
             /*now generate yld seq for each func zone*/
               %let i=1;
               %do %while (&i < 6);
                   data seq&i;
                      set obsyld;
                      if sno=&i+0;
                      syld=yield;
                   run;
                   /*proc print data=seq&i;run;*/
                   %let i=%eval(&i+1);
               %end;

proc iml;

   use obsyld;
   read all var{yield} into bigyld;
   close obsyld;
   t=nrow(bigyld);

   use &fldfunc;
   read all var{area fzone};
   close &fldfunc;
   fldarea=0;
   do i=1 to &nfunc;
       fldarea=fldarea+area[i,1];
   end;

   %let i=1;
   %do %while (&i < 6);
       use seq&i;
       read all var{syld} into syld&i;  /*yield seq for a cluster*/
       read all var{state} into stat&i; /*state of nature this obs belong*/
       nsmu_s&i=nrow(syld&i);
       %let i=%eval(&i+1);
   %end;

   nprice=0.25;        /*N price per pound*/
   haul=0.15;          /*hauling cost per bushel*/
   If &crop='W' then do;
         cprice=2.87;  /*wheat price*/
         Nbushel=1.25; /*N needed per bushel*/
         penalty=0.8;  /*yield loss per pound N underapplied*/
      end;
   else do;
         cprice=2.10; /*corn price*/
         Nbushel=1;   /*N needed per bushel*/
         penalty=1;   /*yield loss per pound N underapplied*/
      end;
   bigopt=J(t,3,0);        /*target yield will be evaluated for all candidate targets*/
```

```
/**** Now begin to calculate expected costs ****/

 do i=1 to t;
     if &crop='W' then nrate=bigyld[i,1]*Nbushel; /*N rate applied for wheat*/
     else nrate=bigyld[i,1]*Nbushel-20;           /*N rate applied for corn*/

     funcount=1; /*count the # of functional zone as recorded in &fldfunc dataset*/
     %let szon=1;
     %do %while (&szon < 6);     /*for each zone, calculate cost1 to cost3*/
         cost1&szon=0;    /* subzone exp cost (previous state = 1)         */
         cost2&szon=0;    /* subzone exp cost (previous state = 2)         */
         cost3&szon=0;    /* subzone exp cost (previous state = 3)         */
         n1=0;n2=0;n3=0;  /* number of each state of nature observed in subzone*/

         do j=1 to nsmu_s&szon;                /*# of observed yields for this row  */
             if &crop='W' then do;
                nneed=syld&szon[j,1]*Nbushel;   /* N need for wheat on this subzone*/
                end;
             else do;
                nneed=syld&szon[j,1]*Nbushel-20;/* N need for corn on this subzone */
                end;

             /*calculate cost to each SMU in the zone if given prev stat of nat*/
             if stat&szon[j,1]=1 then do;
                n1=n1+1;
                cost1&szon=cost1&szon + &p11*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                cost2&szon=cost2&szon + &p21*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                cost3&szon=cost3&szon + &p31*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                end;
             if stat&szon[j,1]=2 then do;
                n2=n2+1;
                cost1&szon=cost1&szon + &p12*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                cost2&szon=cost2&szon + &p22*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                cost3&szon=cost3&szon + &p32*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                end;
             if stat&szon[j,1]=3 then do;
                n3=n3+1;
                cost1&szon=cost1&szon + &p13*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                cost2&szon=cost2&szon + &p23*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                cost3&szon=cost3&szon + &p33*(nprice*max(nrate-nneed,0)
                          -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                end;
         end;

         /*calculate zone-level cost*/
           if n1=0 then n1=1;
           if n2=0 then n2=1;
           if n3=0 then n3=1;

           cost&szon= cost1&szon*area[funcount,1]/n1
                    + cost2&szon*area[funcount+1,1]/n2
                    + cost3&szon*area[funcount+2,1]/n3;
           funcount=funcount+3;
           *print cost&szon;
         %let szon=%eval(&szon+1);
     %end;

     /*calculate field-level cost*/
      cost=0;
      %let szon=1;
      %do %while (&szon < 6);
```

```
               cost=cost+cost&szon;
               %let szon=%eval(&szon+1);
           %end;
           cost=cost/fldarea;
           *print cost1 cost2 cost3 cost4 cost5 cost nrate fldarea;
           *print fldarea;
           bigopt[i,1]=bigyld[i,1]; /*Actual (simulated) crop yield*/
           bigopt[i,2]=nrate;       /*N applied for this yield over the whole field*/
           bigopt[i,3]=cost;        /*Expected cost for this N rate */
      end;

      varnames={Yield Nrate expcost};
      create okay from bigopt [colname=varnames];
      append from bigopt;
      close okay;
      sort okay by expcost;

  quit;

  /* optimal field-level yields and expected cost*/
  data okay; set okay;if _n_=1;run;
  proc print data=okay;
       title1 "FIELD = &fldname";
  run;

%mend;

%smuvar('F1B',smuf1b,'W', smutrf1b,funcf1b);
%smuvar('F1C',smuf1c,'W', smutrf1c,funcf1c);
%smuvar('F3B',smuf3b,'W', smutrf3b,funcf3b);
%smuvar('F4', smuf4, 'W', smutrf4, funcf4);
%smuvar('F2', smuf2, 'C', smutrf2, funcf2);
%smuvar('F5', smuf5, 'C', smutrf5, funcf5);
%smuvar('F7', smuf7, 'C', smutrf7, funcf7);
%smuvar('F8B',smuf8b,'C', smutrf8b,funcf8b);



/**********************************************************************************
 * program name = exp_fu_T3_24.sas
 *
 * Generate Table 3-24: field level 1999 expected and optimal yield unider functional
 * zone variable strategy.
 *
 * NOTE: functional zones are based on cluter analyses.
 *
 * Input: big.smu_var0.sd2
 *
 * output: screen output.
 **********************************************************************************/
libname big 'c:\data\thesis\richmond\sasdata';

proc sort
     data=big.smu_var0
     out=one;
     by fld;
run;

data one;
    set one;
    by fld;
    eyld=expyld*area;
    eopt=optyld*area;
    ecos=expcost*area;
run;

proc univariate data=one noprint;
   output out=one sum=area eyld eopt ecos;
   var area eyld eopt ecos;
   by fld;
run;
```

```
data one;
   set one;
   eyld=eyld/area;eopt=eopt/area;ecos=ecos/area;
run;

proc print data=one;var fld area eyld eopt ecos;run;endsas;




/***************************************************************************************
*progname=combine.sas;
*
*Abstract: Combine all information together for each field based on 30x30 grids.
*          1. Field name,soil name,cluster #, grid counts, and actual yields 1995-1999.
*          2. State of nature in 1997
*          3. Determined N rate for all strategies (ex post or ex ante)
*
*Input: big.smu_var0.sd2;   ---- functional zone optimal yields
*       big.smuf1b.sd2 etc; ---- state of nature in 1997
*       big.g30g1b.sd2 etc; ---- original yields in 30x30 grids
*       big.n_notran.sd2    ---- N rate for soil zone uniform w/o transition considered.
*       other data generated by other programs and reported in Chapter 3.
*
*output: Combined dataset for each field to evaluate inf value and vra value
*
*date: 2Jun2001
***************************************************************************************/
options nodate;
libname big 'c:\data\thesis\richmond\sasdata';
libname com 'c:\data\thesis\richmond\sasdata\finally';

/*The following are from big.n_notran.sd2. The conv and soil unif w/o trans */
/*Also reported in Table 3-18*/
data notr_sc;
   notr_suw=113;notr_suc=140; notr_cuw=113;notr_cuc=140;
   input fld $;
   cards;
   F1B
   F1C
   F2
   F3B
   F4
   F5
   F7
   F8B
   ;
run;
proc sort data=notr_sc;by fld;run;

/*data from Table 3-19 and 3-22 (2Jun01 version) and smu_nrate1999.sas output (for table 3-
25)*/
data tran_sc; /*soil zone uniform and conventional w trans considered*/
   input fld $ ntr_cu ntr_su ntr_fu;
   cards;
   F1B   93  95  94
   F1C   96 103  93
   F3B  110 109 108
   F4   113 118 101
   F2   117 121 139
   F5   150 156 162
   F7   174 178 167
   F8B  170 186 181
   ;
run;
proc sort data=tran_sc;by fld;run;

data unif;
   merge notr_sc tran_sc;
   by fld;
```

```
run;

/*The following comes from Table 3-20 and Table 3-21.
  The N rate for soil zone variable application */
data ntr_sv;
  input fld $ soil ntr_sv;
  cards;
   F1B  148 34
   F1B  151 90
   F1B  168 34
   F1B  171 101
   F1B  172 101
   F1C  148 36
   F1C  171 109
   F1C  172 109
   F3B  151 102
   F3B  168 38
   F3B  171 115
   F3B  172 115
   F4   148 40
   F4   171 120
   F4   172 120
   F2   168 37
   F2   171 121
   F2   172 121
   F5   148 54
   F5   151 127
   F5   171 161
   F5   172 161
   F7   151 141
   F7   161 128
   F7   171 178
   F7   172 178
   F8B  148 65
   F8B  151 150
   F8B  171 190
   F8B  172 190
    ;
run;

/*now assign state of nature in 1997 to original field data*/
%macro bigmer (fldname,
               fldfile, /*field file containing state of nature in 1997*/
               original,/*original yield data                         */
               outfile, /*output combined file                        */
               );

           proc sort data=ntr_sv; by fld;run;
           data middle; /*combine notran and tran*/
                merge unif ntr_sv (in=va);
                by fld;
                if va and fld=&fldname;
           run;
           proc sort data=middle;by fld soil;run;

        /* the following is for functional zone variable */
        data ntr_fv;
           set big.smu_var0;
           cluster=clusno;
           ntr_fv=nrate;
           if fld=&fldname;
           keep fld cluster s97 ntr_fv;
        run;
        proc sort data=ntr_fv;by fld cluster s97;run;

        proc sort
             data=big.&fldfile
             out=&fldfile (keep=fld cluster soil x y s97);
             by fld cluster s97;
        run;
        data ntr_fv;
```

```
             merge ntr_fv &fldfile (in=fd);
             by fld cluster s97;
             if fd;
        run;

        proc sort
             data=ntr_fv;
             by fld soil;
        run;

        data middle;
             merge middle ntr_fv (in=fv);
             by fld soil;
             if fv;
        run;

    proc sort
         data=middle;
         by x y;
    run;

    proc sort
         data=big.&original
         out=&original;
         by x y;
    run;

    data com.&outfile;
       merge middle &original (in=org);
       by x y;
       if org;
   run;

%mend;
%bigmer('F1B',smuf1b,g30F1b,comf1b);
%bigmer('F1C',smuf1c,g30F1c,comf1c);
%bigmer('F3B',smuf3b,g30F3b,comf3b);
%bigmer('F4' ,smuf4 ,g30F4 ,comf4 );
%bigmer('F2' ,smuf2 ,g30F2 ,comf2 );
%bigmer('F5' ,smuf5 ,g30F5 ,comf5 );
%bigmer('F7' ,smuf7 ,g30F7 ,comf7 );
%bigmer('F8B',smuf8b,g30F8b,comf8b);



/***************************************************************************************
*progname=evaluate.sas;
*
*  Abstract: Evaluate all information values and values of viriable application in ex
*  ante analyses
*
*Input: final.comf1b.sd2 etc.
*       information cost and cost of VRA (from Chapter 3).
*
*output: Results for ex ante analyese. Output file is: final.allout.sd2
***************************************************************************************/
options nodate;
libname final 'c:\data\thesis\richmond\sasdata\finally';

%macro evaluate(fldname,     /*name of the field                   */
                comfld,     /*field file with combined information*/
                crop99,      /*the crop planted in 1999            */
                cropyld,     /*yield of crop in 1999 for each SMU  */
                outfile      /*the output file for all evaluation  */
                )
                ;

     data &comfld;
        set final.&comfld;
        area=count*9*0.00025;
        yield=&cropyld*area;
```

```
    keep fld area  yield &cropyld ntr_cu ntr_su ntr_fu ntr_sv ntr_fv;
 run;
 proc sort data=&comfld;by fld; run;


/*The the calculation*/
 data &comfld;
    set &comfld;

    /*first, N applied*/
    napp_cu = ntr_cu * area;
    napp_su = ntr_su * area;
    napp_sv = ntr_sv * area;
    napp_fu = ntr_fu * area;
    napp_fv = ntr_fv * area;

    /*second, yld penalty*/
    if &crop99 = 'W' then do;
            ypen_cu = -min(ntr_cu - &cropyld * 1.25, 0) * 0.8 * area;
            ypen_su = -min(ntr_su - &cropyld * 1.25, 0) * 0.8 * area;
            ypen_sv = -min(ntr_sv - &cropyld * 1.25, 0) * 0.8 * area;
            ypen_fu = -min(ntr_fu - &cropyld * 1.25, 0) * 0.8 * area;
            ypen_fv = -min(ntr_fv - &cropyld * 1.25, 0) * 0.8 * area;
         end;
    else do;
            ypen_cu = -min(ntr_cu - &cropyld + 20, 0) * area;
            ypen_su = -min(ntr_su - &cropyld + 20, 0) * area;
            ypen_sv = -min(ntr_sv - &cropyld + 20, 0) * area;
            ypen_fu = -min(ntr_fu - &cropyld + 20, 0) * area;
            ypen_fv = -min(ntr_fv - &cropyld + 20, 0) * area;
         end;

    /*third, net N */
    if &crop99='W' then do;
            netn_cu = max(ntr_cu - &cropyld*1.25 + 2.8, 0) *area;
            netn_su = max(ntr_su - &cropyld*1.25 + 2.8, 0) *area;
            netn_sv = max(ntr_sv - &cropyld*1.25 + 2.8, 0) *area;
            netn_fu = max(ntr_fu - &cropyld*1.25 + 2.8, 0) *area;
            netn_fv = max(ntr_fv - &cropyld*1.25 + 2.8, 0) *area;
         end;
    else do;
            netn_cu = max(ntr_cu - (&cropyld-20)*0.9 + 1.3, 0) *area;
            netn_su = max(ntr_su - (&cropyld-20)*0.9 + 1.3, 0) *area;
            netn_sv = max(ntr_sv - (&cropyld-20)*0.9 + 1.3, 0) *area;
            netn_fu = max(ntr_fu - (&cropyld-20)*0.9 + 1.3, 0) *area;
            netn_fv = max(ntr_fv - (&cropyld-20)*0.9 + 1.3, 0) *area;
         end;

    /*cost of inaccurate N appl*/
   if &crop99='W' then do;
            cost_cu = (0.25*max(ntr_cu - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                       -0.25)*min(ntr_cu - &cropyld*1.25,0))*area;
            cost_su = (0.25*max(ntr_su - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                       -0.25)*min(ntr_su - &cropyld*1.25,0))*area;
            cost_sv = (0.25*max(ntr_sv - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                       -0.25)*min(ntr_sv - &cropyld*1.25,0))*area;
            cost_fu = (0.25*max(ntr_fu - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                       -0.25)*min(ntr_fu - &cropyld*1.25,0))*area;
            cost_fv = (0.25*max(ntr_fv - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                       -0.25)*min(ntr_fv - &cropyld*1.25,0))*area;
        end;
   else do;
            cost_cu = (0.25*max(ntr_cu - &cropyld + 20, 0)-(2.10-0.15-0.25)
                       *min(ntr_cu - &cropyld + 20,0))*area;
            cost_su = (0.25*max(ntr_su - &cropyld + 20, 0)-(2.10-0.15-0.25)
                       *min(ntr_su - &cropyld + 20,0))*area;
            cost_sv = (0.25*max(ntr_sv - &cropyld + 20, 0)-(2.10-0.15-0.25)
                       *min(ntr_sv - &cropyld + 20,0))*area;
            cost_fu = (0.25*max(ntr_fu - &cropyld + 20, 0)-(2.10-0.15-0.25)
                       *min(ntr_fu - &cropyld + 20,0))*area;
            cost_fv = (0.25*max(ntr_fv - &cropyld + 20, 0)-(2.10-0.15-0.25)
```

```
                        *min(ntr_fv - &cropyld + 20,0))*area;
              end;
       run;

       proc sort data=&comfld;by fld;run; /*only to keep fld name in following PROC*/

       proc univariate data=&comfld noprint;
            output out=&outfile
                   sum=area napp_cu napp_su napp_sv napp_fu napp_fv
                            ypen_cu ypen_su ypen_sv ypen_fu ypen_fv
                            netn_cu netn_su netn_sv netn_fu netn_fv
                            cost_cu cost_su cost_sv cost_fu cost_fv
                       yield;
            var area napp_cu napp_su napp_sv napp_fu napp_fv
                     ypen_cu ypen_su ypen_sv ypen_fu ypen_fv
                     netn_cu netn_su netn_sv netn_fu netn_fv
                     cost_cu cost_su cost_sv cost_fu cost_fv
                yield;
            by fld;
       run;

       /*calculated to the per acre base */
       data &outfile;
            set &outfile;

            yield = yield /area;

            napp_cu = napp_cu / area;
            napp_su = napp_su / area;
            napp_sv = napp_sv / area;
            napp_fu = napp_fu / area;
            napp_fv = napp_fv / area;

            ypen_cu = ypen_cu / area;
            ypen_su = ypen_su / area;
            ypen_sv = ypen_sv / area;
            ypen_fu = ypen_fu / area;
            ypen_fv = ypen_fv / area;

            netn_cu = netn_cu / area;
            netn_su = netn_su / area;
            netn_sv = netn_sv / area;
            netn_fu = netn_fu / area;
            netn_fv = netn_fv / area;

            cost_cu = cost_cu / area;
            cost_su = cost_su / area;
            cost_sv = cost_sv / area + 2.09;
            cost_fu = cost_fu / area + 2.61;
            cost_fv = cost_fv / area + 5.22;

            /*now information value and VRA value*/
            info_s = cost_cu - cost_su;
            info_f = cost_cu - cost_fu;

            var_s = cost_su - cost_sv;
            var_f = cost_fu - cost_fv;

           /*finally, net N reduction */
            nless_su = netn_cu - netn_su;
            nless_sv = netn_cu - netn_sv;
            nless_fu = netn_cu - netn_fu;
            nless_fv = netn_cu - netn_fv;

            crop=&crop99;
       run;

  %mend;
  %evaluate('F1B', comf1b, 'W', W99, outf1b);
  %evaluate('F1C', comf1c, 'W', W99, outf1c);
  %evaluate('F3B', comf3b, 'W', W99, outf3b);
```

```
%evaluate('F4' , comf4 , 'W', W99, outf4);
%evaluate('F2' , comf2 , 'C', C99, outf2);
%evaluate('F5' , comf5 , 'C', C99, outf5);
%evaluate('F7' , comf7 , 'C', C99, outf7);
%evaluate('F8B', comf8b, 'C', C99, outf8b);

data final.allout;
    set outf1b outf1c outf3b outf4 outf2 outf5 outf7 outf8b;
run;
```

Part II (sensitivity analysis)

```
/*sensitivity: assume expected yields in 1999 are correct ***/

/* * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * * *  * * * * *
 * prog name = soilunif.sas
 *
 * Calculate soil-zone variable application rates.
 *
 * output: onscreen display
 * * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * * *  * * * * */
data facs;
   input crop $ soil fld fac real99; /*Rescaling factors */
   format ntr_su 3.;
   if crop='W' then ntr_su = real99*fac*1.25;
   else ntr_su=real99*fac - 20;

   cards;
         W     148    1      0.36145 44
         W     151    1      0.96386 44
         W     168    1      0.36145 44
         W     171    1      1.08434 44
         W     172    1      1.08434 44
         W     148    2      0.37975 92
         W     171    2      1.13924 92
         W     172    2      1.13924 92
         W     151    3      0.93023 94
         W     168    3      0.34884 94
         W     171    3      1.04651 94
         W     172    3      1.04651 94
         W     148    4      0.35294 87
         W     171    4      1.05882 87
         W     172    4      1.05882 87
         C     168    1      0.41667 92
         C     171    1      1.02564 92
         C     172    1      1.02564 92
         C     148    2      0.43333 125
         C     151    2      0.86667 125
         C     171    2      1.06667 125
         C     172    2      1.06667 125
         C     151    3      0.82803 106
         C     161    3      0.76433 106
         C     171    3      1.01911 106
         C     172    3      1.01911 106
         C     148    4      0.44828 119
         C     151    4      0.89655 119
         C     171    4      1.10345 119
         C     172    4      1.10345 119
      ;
run;
```

```
/*** sensitivity analysis: field-level for 1999 are real ****/
/* * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * *
 * prog name = sen_su.sas
 *
 * Calculate soil-zone uniform optimal Yields and expected costs
 *
 * output: onscreen display
 * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * */
libname big 'c:\data\thesis\richmond\sasdata';
libname sen 'c:\data\thesis\richmond\sensitivity';

      /******* This macro generate .sd2 from .dbf files ******
         %macro dbf2sd2(filein,fileout);

             proc import datafile=&filein out=big.&fileout dbms=dbf replace;
                  format soil 3. count 3.;
             run;
         %mend;
         %dbf2sd2('c:\data\thesis\sas\data\g30f1b.dbf', g30f1b);
         %dbf2sd2('c:\data\thesis\sas\data\g30f1c.dbf', g30f1c);
         %dbf2sd2('c:\data\thesis\sas\data\g30f3b.dbf', g30f3b);
         %dbf2sd2('c:\data\thesis\sas\data\g30f4.dbf',  g30f4);
         %dbf2sd2('c:\data\thesis\sas\data\g30f2.dbf',  g30f2);
         %dbf2sd2('c:\data\thesis\sas\data\g30f5.dbf',  g30f5);
         %dbf2sd2('c:\data\thesis\sas\data\g30f7.dbf',  g30f7);
         %dbf2sd2('c:\data\thesis\sas\data\g30f8b.dbf', g30f8b);
         endsas;
      ********************/
data facs;
   input crop $ soil fld fac exp99 real99; /*Rescaling factors:from soilmean.sas based on
Table 3.1 and 3.6 */
   cards;
         W     148    1      0.36145   61  44
         W     151    1      0.96386   61  44
         W     168    1      0.36145   61  44
         W     171    1      1.08434   61  44
         W     172    1      1.08434   61  44
         W     148    2      0.37975   64  92
         W     171    2      1.13924   64  92
         W     172    2      1.13924   64  92
         W     151    3      0.93023   72  94
         W     168    3      0.34884   72  94
         W     171    3      1.04651   72  94
         W     172    3      1.04651   72  94
         W     148    4      0.35294   74  87
         W     171    4      1.05882   74  87
         W     172    4      1.05882   74  87
         C     168    1      0.41667  118  92
         C     171    1      1.02564  118  92
         C     172    1      1.02564  118  92
         C     148    2      0.43333  143 125
         C     151    2      0.86667  143 125
         C     171    2      1.06667  143 125
         C     172    2      1.06667  143 125
         C     151    3      0.82803  167 106
         C     161    3      0.76433  167 106
         C     171    3      1.01911  167 106
         C     172    3      1.01911  167 106
         C     148    4      0.44828  164 119
         C     151    4      0.89655  164 119
         C     171    4      1.10345  164 119
         C     172    4      1.10345  164 119
      ;
run;

data facs;
   set facs;
     fac=fac*real99/exp99;
run;

%macro soilunif (fldsoil,   /*Field file containing soil number and cell counts*/
```

```
              crop,      /*Indicate what crop it is */
              fldyld,    /*field-level yield */
                 ord,    /*ith field for given crop. Eg. 3rd fld for wht is F3B*/
              trans,     /*transition matrix */
              states     /*states of nature for each field */
              )
              ;
proc sort data=big.&fldsoil out=&fldsoil (keep=soil count);
   by soil;
run;
proc univariate data=&fldsoil noprint;
   output out=one sum=area;
   var count;
   by soil;
run;
proc sort data=one;by soil;run;

data two;
   set facs;
   if crop=&crop and fld=&ord+0;
   keep soil fac;
run;
proc sort data=two;by soil;run;

data &fldsoil;
   merge one (in=fld) two;
   by soil;
   area=area*9*0.00025;
   count=1;  /* to generate a macro variable of number of soil zones */
   if fld;
run;
        /***generate the macro number of soil zones in the field ***/
            proc univariate data=&fldsoil noprint;
                 output out=numbersl sum=aaaa;
                 var count;
            run;
            data numbersl;
                 set numbersl;
                 call symput('msoil',aaaa);
            run;

data fldyld;
   set big.&fldyld;
   if &crop='W' then do;
       if &ord+0=1 then yield=f1;
       if &ord+0=2 then yield=f2;
       if &ord+0=3 then yield=f4;
       if &ord+0=4 then yield=f5;
       end;
   else do;
       if &ord+0=1 then yield=f3;
       if &ord+0=2 then yield=f6;
       if &ord+0=3 then yield=f7;
       if &ord+0=4 then yield=f8;
       end;
   keep year yield;
run;
/*proc print data=fldyld;run;*/

data trans;
   set big.&trans;
   if &crop='W' then do;
      if _n_=1 or _n_=2 or _n_=4 or _n_=5;
      end;
   else do;
      if _n_=3 or _n_=6 or _n_=7 or _n_=8;
      end;
run;
data trans;
   set trans;
   if _n_=&ord+0;
```

```
run;

data state;
    set big.&states;
    if &crop='W' then do;
       stat1=s1;stat2=s2;stat3=s4;stat4=s5;
       end;
    else do;
       stat1=s3;stat2=s6;stat3=s7;stat4=s8;
       end;
    stat=stat&ord;
    keep year stat;
run;

proc iml;

   use &fldsoil;
   read all var{area soil fac};
   close &fldsoil;
   nsoil=nrow(soil);
   fldarea=0;
   do i=1 to nsoil;
       fldarea=fldarea+area[i,1];
   end;

   use fldyld;
   read all var{yield} into orgyld;
   close fldyld;
   t=nrow(orgyld);

   /*generate a yield sequence for each soil in the field*/
   soilyld=J(t,nsoil,0);
   do i=1 to t;
      do j=1 to nsoil;
         soilyld[i,j]=round(orgyld[i,1]*fac[j,1]);
      end;
   end;
          /*Further, generate a combined yield sequece: stack together*/
            bigyld=soilyld[,1];
            do j=1 to nsoil;
               bigyld=bigyld//soilyld[,j];
            end;
   use trans;
   read all var{s1 s2 s3 t11 t12 t13 t21 t22 t23 t31 t32 t33};
   close trans;

   use state;
   read all var{year stat};
   close state;

  /**** Now adopt from fldtransition.sas to calculated optimal yield and costs ****/

   nprice=0.25;        /*N price per pound*/
   haul=0.15;          /*hauling cost per bushel*/
   If &crop='W' then do;
         cprice=2.87;  /*wheat price*/
         Nbushel=1.25; /*N needed per bushel*/
         penalty=0.8;  /*yield loss per pound N underapplied*/
         prob1=1/t;    /* used to calculated optimal yield and cost*/
         prob2=1/t;
         prob3=1/t; /*above is valid because no Markov chain detected for wheat*/
      end;
   else do;
         cprice=2.10; /*wheat price*/
         Nbushel=1;   /*N needed per bushel*/
         penalty=1;   /*yield loss per pound N underapplied*/
         prob1=t21[1,1]/s1[1,1];
         prob2=t22[1,1]/s2[1,1];
         prob3=t23[1,1]/s3[1,1];
      end;  /*above is valid because states in 1998 for corn are all normal*/
```

```
  bigopt=J(t*nsoil,3,0);        /*target yield will be evaluated for t*nsoil values*/
  n1=J(1,nsoil,0); n2=n1; n3=n1;/*In order to calculate fld-level exp yld */

  do i=1 to t*nsoil;
      if &crop='W' then nrate=bigyld[i,1]*Nbushel; /*N rate applied for wheat*/
      else nrate=bigyld[i,1]*Nbushel-20;           /*N rate applied for corn*/

        %let szon=1;
        %do %while (&szon<&msoil+1);
            cost&szon=0;    /*Each zone has a exp cost.
                              Fld-level is weighted average of them*/
            expyld&szon=0;  /*Exp Yield for each soil zone */
            %let szon=%eval(&szon+1);
        %end;

      do j=1 to t;
         if &crop='W' then do;
               %let szon=1;
               %do %while (&szon<&msoil+1);
                   nneed&szon=soilyld[j,&szon]*Nbushel;   /* N for wheat on zone szon*/
                   %let szon=%eval(&szon+1);
               %end;
            end;
         else do;
                 %let szon=1;
                 %do %while (&szon<&msoil+1);
                     nneed&szon=soilyld[j,&szon]*Nbushel-20;/*N for corn on zone*/
                     %let szon=%eval(&szon+1);
                 %end;
              end;

                /* below 3 if's are for Markov chain optimal yields */
                   if stat[j,1]=1 then prob=prob1;
                   if stat[j,1]=2 then prob=prob2;
                   if stat[j,1]=3 then prob=prob3;
         %let szon=1;
         %do %while (&szon<&msoil+1);
             cost&szon=cost&szon+prob*(nprice*max(nrate-nneed&szon,0)
                      -((cprice-haul)*penalty-Nprice)*min(nrate-nneed&szon,0));
             %let szon=%eval(&szon+1);
         %end;

      end;

      %let szon=1;cost=0;
      %do %while (&szon<&msoil+1);
          cost=cost + cost&szon*area[&szon+0,1]/fldarea; /*expected cost given Nrate*/
          %let szon=%eval(&szon+1);
      %end;

      bigopt[i,1]=bigyld[i,1]; /*Actual (simulated) crop yield*/
      bigopt[i,2]=nrate;       /*N applied for this yield over the whole field*/
      bigopt[i,3]=cost;        /*Expected cost for this N rate */

      if mod(i,t)=0 then j=t;
      else j=mod(i,t);
           /**In order to calculated field-level expected yield**/
              do szon=1 to nsoil;
                 if stat[j,1]=1 then n1[1,szon]=n1[1,szon] + bigyld[i,1];
                 if stat[j,1]=2 then n2[1,szon]=n2[1,szon] + bigyld[i,1];
                 if stat[j,1]=3 then n3[1,szon]=n3[1,szon] + bigyld[i,1];
              end;

      %let szon=1;
      %do %while (&szon<&msoil+1);
          expyld&szon=expyld&szon + prob1*n1[1,&szon]
                                  + prob2*n2[1,&szon] + prob3*n3[1,&szon];
          %let szon=%eval(&szon+1);
      %end;
end;
```

```
    expyld=0;    /*expected field-level yield*/
    %let szon=1;
    %do %while (&szon<&msoil+1);
        expyld=expyld + expyld&szon*area[&szon,1]/fldarea;
        %let szon=%eval(&szon+1);
    %end;
    *print "Exp yld (wgtd avg) in 1999:" expyld "   total year:" t;

    varnames={Yield Nrate expcost};
    create okay from bigopt [colname=varnames];
    append from bigopt;
    close okay;
    sort okay by expcost;

  quit;

  /* optimal field-level yields and expected cost*/
  data okay; set okay;if _n_=1;run;
  proc print data=okay;
       title1 "the field is ======= &fldsoil =========";
  run;

%mend;

%soilunif(g30f1b,'W',wheatfld,1,wtrans,wfldstat);
%soilunif(g30f1c,'W',wheatfld,2,wtrans,wfldstat);
%soilunif(g30f3b,'W',wheatfld,3,wtrans,wfldstat);
%soilunif(g30f4,'W',wheatfld,4,wtrans,wfldstat);

%soilunif(g30f2,'C',cornfld,1,ctrans,cfldstat);
%soilunif(g30f5,'C',cornfld,2,ctrans,cfldstat);
%soilunif(g30f7,'C',cornfld,3,ctrans,cfldstat);
%soilunif(g30f8b,'C',cornfld,4,ctrans,cfldstat);
```

```
* sensitivity: assume expected yields of 1999 are right *****/

/* * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * * *  * * * * *
 * program name = smu_FV.sas
 *
 * Purpose: Calculate SMU variable rate OPTIMAL Yields and expected costs.
 *          Same as for functional zone variable strategy.
 *          So this program is for generic SMU (100 counts).next program will
 *          assign to each real SMU a pair of N rate and optimal yield.
 * NOTE:    functional zones are based on clustering.
 *
 * Input:   smutrf1b.sd2 etc and smuf1b.sd2 etc.Generated by SMUtransition.sas
 *
 * output:  smu_var0.sd2
 * * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * * *  * * * * */

libname big 'c:\data\thesis\richmond\sasdata';

%macro smuvar (crop,    /*indicate crop is wheat or corn               */
               fldname, /*field name                                   */
               f,       /*file name for the field (used only once)     */
               trans,   /*transition matrix                            */
               field,   /*field with count,cluster,y1995,y1997, and s97*/
               clusno   /*the cluste                                   */
               )
               ;
   data y95;
      set big.&field;
```

```
      stat=s97;
      yield=y1995;
      area1=0;area2=0;areat3=0; /*used to count areas of stat in 1997*/
      if cluster=&clusno;
   run;

   data y97;
      set big.&field;
      stat=s97;
      yield=y1997;
      if stat=1 then area1=count*0.00025*9; else area1=0; /*used to count areas of stat in
1997*/
      if stat=2 then area2=count*0.00025*9; else area2=0; /*used to count areas of stat in
1997*/
      if stat=3 then area3=count*0.00025*9; else area3=0; /*used to count areas of stat in
1997*/
      if cluster=&clusno;
   run;

   data obsyld;
      set y95 y97;  /*Observed yields. Used as candidates for zone target*/
      if stat=1 then s1=1;
      if stat=2 then s2=1; else s2=0;
      if stat=3 then s3=1; else s3=0;
   run;
         /*obtain # of each states observed*/
           proc univariate data=obsyld noprint;
                output out=one sum=s1-s3 area1-area3;
                var s1-s3 area1-area3;
           run;

           data one;
              set one;
              call symput('s1',s1);
              call symput('s2',s2);
              call symput('s3',s3);
              call symput('area1',area1);
              call symput('area2',area2);
              call symput('area3',area3);
           run;

/*not needed for variable SMU scenario
   proc sort
        data=obsyld
        out=&field (drop=yield) nodupkey; * SMU in the soil zone;
        by cellno;
   run;

   *proc print data=&field;run;

           * obtain the total counts for the soil zone and put in a macro variable;
              proc univariate data=&field noprint;
                   output out=one sum=totcon;
                   var count;
              run;
              data one;
                   set one;
                   call symput("area", totcon);
              run;

      use &field;
      read all var{count stat};
      close obsyld;
      y=nrow(stat); * # of SMU to evaluate;
*********/

   proc iml;

      use obsyld;
      read all var{yield count stat};
      close obsyld;
```

```
 t=nrow(yield); /* # of yields to try */

 use big.&trans;
 read all var{to1 to2 to3};
 close big.&trans;

/**** Now adopt from SOILUNIF.sas to calculated optimal yield and costs ****/

 nprice=0.25;        /*N price per pound*/
 haul=0.15;          /*hauling cost per bushel*/
 If &crop='W' then do;
       cprice=2.87;  /*wheat price*/
       Nbushel=1.25; /*N needed per bushel*/
       penalty=0.8;  /*yield loss per pound N underapplied*/
    end;
 else do;
       cprice=2.10; /*wheat price*/
       Nbushel=1;   /*N needed per bushel*/
       penalty=1;   /*yield loss per pound N underapplied*/
    end;

 do k=1 to 3;

    prob1=to1[k,1]/&s1;  /*suppose SMU is state i in 1997 */
    prob2=to2[k,1]/&s2;
    prob3=to3[k,1]/&s3;

    bigopt=J(t,3,0);     /*target yield will be evaluated for t*nsoil values*/
    expyld=J(t,1,1);     /*coloum for expected yield */
    area=J(t,1,1);       /*column for area for each SMU-type*/
    n1=0; n2=0; n3=0;    /*In order to calculate exp yld */

    do i=1 to t;
       if &crop='W' then nrate=yield[i,1]*Nbushel; /*N rate applied for wheat*/
       else nrate=yield[i,1]*Nbushel-20;           /*N rate applied for corn*/

       cost=0;
       do j=1 to t;
          if &crop='W' then do;
                 nneed=yield[j,1]*Nbushel;    /* N needed for wheat on zone szon*/
             end;
          else do;
                 nneed=yield[j,1]*Nbushel-20; /* N needed for corn on zone szon */
             end;

          /* below 3 if's are for Markov chain optimal yields */
              if stat[j,1]=1 then prob=prob1;
              if stat[j,1]=2 then prob=prob2;
              if stat[j,1]=3 then prob=prob3;
          cost=cost+prob*(nprice*max(nrate-nneed,0)
                   -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
       end;

       bigopt[i,1]=yield[i,1];  /*Actual (observed) crop yield for a SMU*/
       bigopt[i,2]=nrate;       /*N applied for this yield over a SMU*/
       bigopt[i,3]=cost;        /*Expected cost for this N rate */

       if stat[i,1]=1 then n1=n1 + yield[i,1];
       if stat[i,1]=2 then n2=n2 + yield[i,1];
       if stat[i,1]=3 then n3=n3 + yield[i,1];
    end;
    expyld=(prob1*n1 + prob2*n2 + prob3*n3)*expyld;

    if k=1 then area=area*&area1;
    if k=2 then area=area*&area2;
    if k=3 then area=area*&area3;

    bigopt=area||expyld||bigopt;

    epp=expyld[1,1];
    print &fldname &clusno epp k;
```

```
         varnames={area expyld optYld Nrate expcost};
         if k=1 then do;
            create okay1 from bigopt [colname=varnames];
            append from bigopt;
            close okay1;
            sort okay1 by expcost;
         end;
         if k=2 then do;
            create okay2 from bigopt [colname=varnames];
            append from bigopt;
            close okay2;
            sort okay2 by expcost;
         end;
         if k=3 then do;
            create okay3 from bigopt [colname=varnames];
            append from bigopt;
            close okay3;
            sort okay3 by expcost;
         end;

      end;

  quit;

  /* optimal SMU yields and expected costs*/

  %do k=1 %to 3;
      data okay&k;
        length fld $ 8;
        format clusno 1. s97 1.;
        set okay&k;
        if _n_=1;
        fld=&fldname;
        clusno=&clusno;
        s97=&k+0; /*state of nature in 1997*/
        run;
  %end;
  data f&clusno;
     set okay1 okay2 okay3;
  run;
%mend;

%smuvar('W','F1B',f1b,smutrf1b,smuf1b,1); /*soil zones are:148 151 168 171 172*/
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,2);
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,3);
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,4);
%smuvar('W','F1B',f1b,smutrf1b,smuf1b,5);
data FF1b;
  set f1 f2 f3 f4 f5;
run;

%smuvar('W','F1C',f1c,smutrf1c,smuf1c,1); /*soil zones are:148 171 172*/
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,2);
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,3);
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,4);
%smuvar('W','F1C',f1c,smutrf1c,smuf1c,5);

data FF1c;
  set f1 f2 f3 f4 f5;
run;

%smuvar('W','F3B',f3b,smutrf3b,smuf3b,1); /*soil zones are:151 168 171 172*/
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,2);
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,3);
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,4);
%smuvar('W','F3B',f3b,smutrf3b,smuf3b,5);
data FF3b;
  set f1 f2 f3 f4 f5;
run;
```

```
%smuvar('W','F4',f4,smutrf4,smuf4,1); /*soil zones are:148 171 172*/
%smuvar('W','F4',f4,smutrf4,smuf4,2);
%smuvar('W','F4',f4,smutrf4,smuf4,3);
%smuvar('W','F4',f4,smutrf4,smuf4,4);
%smuvar('W','F4',f4,smutrf4,smuf4,5);

data FF4;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F2',f2,smutrf2,smuf2,1); /*soil zones are:168 171 172*/
%smuvar('C','F2',f2,smutrf2,smuf2,2);
%smuvar('C','F2',f2,smutrf2,smuf2,3);
%smuvar('C','F2',f2,smutrf2,smuf2,4);
%smuvar('C','F2',f2,smutrf2,smuf2,5);
data FF2;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F5',f5,smutrf5,smuf5,1); /*soil zones are:148 151 171 172*/
%smuvar('C','F5',f5,smutrf5,smuf5,2);
%smuvar('C','F5',f5,smutrf5,smuf5,3);
%smuvar('C','F5',f5,smutrf5,smuf5,4);
%smuvar('C','F5',f5,smutrf5,smuf5,5);
data FF5;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F7',f7,smutrf7,smuf7,1); /*soil zones are:151 161 171 172*/
%smuvar('C','F7',f7,smutrf7,smuf7,2);
%smuvar('C','F7',f7,smutrf7,smuf7,3);
%smuvar('C','F7',f7,smutrf7,smuf7,4);
%smuvar('C','F7',f7,smutrf7,smuf7,5);
data FF7;
  set f1 f2 f3 f4 f5;
run;

%smuvar('C','F8B',f8b,smutrf8B,smuf8B,1); /*soil zones are:148 151 171 172*/
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,2);
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,3);
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,4);
%smuvar('C','F8B',f8b,smutrf8B,smuf8B,5);
data FF8b;
  set f1 f2 f3 f4 f5;
run;

data big.smu_var0;
  set ff1b ff1c ff3b ff4 ff2 ff5 ff7 ff8b;
run;





/* sensitivity: expected yields in 1999 are correct */
/* * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * *  * * * * ********
 * program name = smu_nrate1999.sas
 *
 * Abstract: 1.Assign for each SMU the N rate in 1999 under variable rate application.
 *           2.Determine field-level N rate in 1999 under functional uniform application.
 *             2a. Also optimal yields (field-level).
 *             2b. Also expected costs (field-level).
 *
 *input: c:\data\thesis\richmond\sasdata\smu_var0.sd2
 *       (generic SMU: fld cluster, state, area, Nrate for 1999)
 *      c:\data\thesis\richmond\sasdata\smuf1b.sd2, ..., smuf8b.sd2
 *       (actual SMU: count,cluster,fld,s97 etc).
 *
 *NOTE: here functional zones are based on clusters.
 *
```

```
 *output: Screen output
 * * * * * * * * * * * * * * * * * * * *  * * * * *  * * * * * * * * *  * * * * *********/

libname big 'c:\data\thesis\richmond\sasdata';

/****** First step: assign N rate in 1999 to each SMU under variable rate application
******/

data smu_var0 (rename=(clusno=cluster nrate=n99_FV));
   set big.smu_var0;
run;

proc sort data=smu_var0;
     by fld cluster s97;
run;
        /*******calculated weighted avg of exp and opt yld and cost for each field
                with SMU variable information. For Table 3-22.
               data one; set smu_var0;
                    expyld=area*expyld;
                    optyld=area*optyld;
                    expcost=area*expcost;
               run;

               proc sort data=one; by fld;run;

               proc univariate data=one noprint;
                    output out=one sum=expyld optyld expcost totarea;
                    var expyld optyld expcost area;
                    by fld;
               run;

               data one; set one;
                    expyld=expyld/totarea;
                    optyld=optyld/totarea;
                    expcost=expcost/totarea;
               run;
               proc print data=one; run;
         *********************/



%macro smuvar (fldname,  /*field name                                 */
               field,    /*the field data (name,count,cluster,yield etc) */
               crop,     /*crop indicator                             */
               trans,    /*within-field transition matrix file        */
               fldfunc   /*the functional zone information for a field*/
               );
   proc sort data=big.&field out=&field;
      by fld cluster s97;
   run;

   data &field;
      merge &field (in=onebyone) smu_var0;
      by fld cluster s97;
      fzone=put(cluster,3.)||'_'||put(s97,1.); *generate mark for functional zones;
      if onebyone;
      drop area expyld optyld expcost;
   run;

   /*proc print data=&field;run;*/

   data &trans; *the transition matrix for the all SMU in a field;
      set big.&trans;
      if _n_=1 then do;
               call symput('p11',to1);
               call symput('p12',to2);
               call symput('p13',to3);
               end;
      if _n_=2 then do;
               call symput('p21',to1);
               call symput('p22',to2);
```

```
            call symput('p23',to3);
            end;
   if _n_=3 then do;
            call symput('p31',to1);
            call symput('p32',to2);
            call symput('p33',to3);
            end;
run;

/* now can treat each functional zone the same way as a soil zone. So can adopt
   the program from SOILUNIF.SAS as the following */

data &fldfunc;
   set smu_var0 (where=(fld=&fldname));
   fzone=put(cluster,1.)||'_'||put(s97,1.); *generate mark for functional zones;
   count=1;
   keep fzone area count;
run;

        /***generate the macro number of func zones in the field ***/

            proc univariate data=&fldfunc noprint;
                 output out=numbersl sum=aaaa;
                 var count;
            run;
            data numbersl;
                 set numbersl;
                 call symput('nfunc',aaaa);
            run;

/** dataset of candidate target yield for the field**/
data y95;
   set &field;
   state=s95;
   yield=y1995;
run;

data y97;
   set &field;
   state=s97;
   yield=y1997;
run;

data obsyld;
   set y95 y97;  /*Observed yields. Used as candidates for zone target*/
run;
/*proc print data=obsyld;run;     */

     /*add a order # for each sol zone */
       proc sort data=&field out=one;
            by cluster;
       run;
       proc univariate data=one noprint;
            output out=one sum=sno;
            var count;  /*actually count is not used at all later*/
            by cluster;
       run;
       data one;
            set one;
            sno=_n_;
        run;
       proc sort data=one (keep=cluster sno);
            by cluster;
       run;
          /*assign to the obsyld the soil zone number*/
          proc sort data=obsyld;
               by cluster;
          run;
          data obsyld;
             merge one obsyld (in=obs);
             by cluster;
```

```
          run;
             /*now generate yld seq for each func zone*/
               %let i=1;
               %do %while (&i < 6);
                   data seq&i;
                      set obsyld;
                      if sno=&i+0;
                      syld=yield;
                   run;
                   /*proc print data=seq&i;run;*/
                   %let i=%eval(&i+1);
               %end;

proc iml;

   use obsyld;
   read all var{yield} into bigyld;
   close obsyld;
   t=nrow(bigyld);

   use &fldfunc;
   read all var{area fzone};
   close &fldfunc;
   fldarea=0;
   do i=1 to &nfunc;
       fldarea=fldarea+area[i,1];
   end;

   %let i=1;
   %do %while (&i < 6);
       use seq&i;
       read all var{syld} into syld&i;  /*yield seq for a cluster*/
       read all var{state} into stat&i; /*state of nature this obs belong*/
       nsmu_s&i=nrow(syld&i);
       %let i=%eval(&i+1);
   %end;

   nprice=0.25;        /*N price per pound*/
   haul=0.15;          /*hauling cost per bushel*/
   If &crop='W' then do;
         cprice=2.87;  /*wheat price*/
         Nbushel=1.25; /*N needed per bushel*/
         penalty=0.8;  /*yield loss per pound N underapplied*/
      end;
   else do;
         cprice=2.10; /*corn price*/
         Nbushel=1;   /*N needed per bushel*/
         penalty=1;   /*yield loss per pound N underapplied*/
      end;
   bigopt=J(t,3,0);        /*target yield will be evaluated for all candidate targets*/


  /**** Now begin to calculate expected costs ****/

   do i=1 to t;
       if &crop='W' then nrate=bigyld[i,1]*Nbushel; /*N rate applied for wheat*/
       else nrate=bigyld[i,1]*Nbushel-20;           /*N rate applied for corn*/

       funcount=1; /*count the # of functional zone as recorded in &fldfunc dataset*/
       %let szon=1;
       %do %while (&szon < 6);     /*for each zone, calculate cost1 to cost3*/
           cost1&szon=0;    /* subzone exp cost (previous state = 1)         */
           cost2&szon=0;    /* subzone exp cost (previous state = 2)         */
           cost3&szon=0;    /* subzone exp cost (previous state = 3)         */
           n1=0;n2=0;n3=0;  /* number of each state of nature observed in subzone*/

           do j=1 to nsmu_s&szon;          /*# of observed yields for this row  */
               if &crop='W' then do;
                  nneed=syld&szon[j,1]*Nbushel;   /* N for wheat on this subzone*/
                  end;
               else do;
```

```
                   nneed=syld&szon[j,1]*Nbushel-20;/* N for corn on this subzone */
                   end;

                /*cost to each SMU in the zone if prev stat of nat is 1, 2, or 3 resp.*/
                if stat&szon[j,1]=1 then do;
                   n1=n1+1;
                   cost1&szon=cost1&szon + &p11*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   cost2&szon=cost2&szon + &p21*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   cost3&szon=cost3&szon + &p31*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   end;
                if stat&szon[j,1]=2 then do;
                   n2=n2+1;
                   cost1&szon=cost1&szon + &p12*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   cost2&szon=cost2&szon + &p22*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   cost3&szon=cost3&szon + &p32*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   end;
                if stat&szon[j,1]=3 then do;
                   n3=n3+1;
                   cost1&szon=cost1&szon + &p13*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   cost2&szon=cost2&szon + &p23*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   cost3&szon=cost3&szon + &p33*(nprice*max(nrate-nneed,0)
                             -((cprice-haul)*penalty-Nprice)*min(nrate-nneed,0));
                   end;
            end;

            /*calculate zone-level cost*/
              if n1=0 then n1=1;
              if n2=0 then n2=1;
              if n3=0 then n3=1;

              cost&szon= cost1&szon*area[funcount,1]/n1
                       + cost2&szon*area[funcount+1,1]/n2
                       + cost3&szon*area[funcount+2,1]/n3;
              funcount=funcount+3;
              *print cost&szon;
            %let szon=%eval(&szon+1);
        %end;

        /*calculate field-level cost*/
         cost=0;
         %let szon=1;
         %do %while (&szon < 6);
             cost=cost+cost&szon;
             %let szon=%eval(&szon+1);
         %end;
         cost=cost/fldarea;
         *print cost1 cost2 cost3 cost4 cost5 cost nrate fldarea;
         *print fldarea;
         bigopt[i,1]=bigyld[i,1]; /*Actual (simulated) crop yield*/
         bigopt[i,2]=nrate;       /*N applied for this yield over the whole field*/
         bigopt[i,3]=cost;        /*Expected cost for this N rate */
    end;

    varnames={Yield Nrate expcost};
    create okay from bigopt [colname=varnames];
    append from bigopt;
    close okay;
    sort okay by expcost;

quit;

/* optimal field-level yields and expected cost*/
data okay; set okay;if _n_=1;run;
```

```
  proc print data=okay;
       title1 "FIELD = &fldname";
  run;

%mend;

%smuvar('F1B',smuf1b,'W', smutrf1b,funcf1b);
%smuvar('F1C',smuf1c,'W', smutrf1c,funcf1c);
%smuvar('F3B',smuf3b,'W', smutrf3b,funcf3b);
%smuvar('F4', smuf4, 'W', smutrf4, funcf4);
%smuvar('F2', smuf2, 'C', smutrf2, funcf2);
%smuvar('F5', smuf5, 'C', smutrf5, funcf5);
%smuvar('F7', smuf7, 'C', smutrf7, funcf7);
%smuvar('F8B',smuf8b,'C', smutrf8b,funcf8b);




/*sensitivity: suppose expected yield are exact**/
/***************************************************************************************
* Program name = SMUtransition.sas
*
* Author: Wei Peng
*
* Abstract: This program estimates within-field transition matrix (spatial transition).
*
*           Several steps are involved:
*           1. Mark states of nature for each SMU within each functional zone;
*           2. Test for existence of Markov Chain;
*           3. Output transition matrices and states of nature;
*
* input: smu_resc.sd2 (output of soilmean.sas and soilrescale.sas)
*
* output: &smutrans.sd2, &smufld.sd2. e.g. smutrf1b.sd2, smuf1b.sd2, Permanent file.
*         (To be used by another two programs to calculate SMU optimal yields)
****************************************************************************************/

option ps=3000;
libname big 'c:\data\thesis\richmond\sasdata';

*first step: rescale yield of each SMU to the level of 1999 field-level;

/* y1999: Expected field-level yield for 1999
   y1997: 1997 yield need to be adjusted by this factor
   y1995: 1995 yield need to be adjusted by this factor*/

   /***** calculate the field-level yield for 1995, 1997, and 1999 ********/
          %macro avgyld (fldfile,
                         fldname,
                         cropname,
                         crop95,
                         crop97,
                         crop99
                         )
                         ;
                 data &fldfile;
                      set big.&fldfile;
                      &crop95=&crop95*count;
                      &crop97=&crop97*count;
                      &crop99=&crop99*count;
                      keep count &crop95 &crop97 &crop99;
                 run;

                 proc univariate data=&fldfile noprint;
                      output out=&fldfile sum=count &crop95 &crop97 &crop99;
                      var count &crop95 &crop97 &crop99;
                 run;

                 data &fldfile;
                      set &fldfile;
                      y1995=&crop95/count;
```

```
                      y1997=&crop97/count;
                      y1999=&crop99/count;
                      fld=&fldname;
                      crop = &cropname;
                      drop count &crop95 &crop97 &crop99;
                run;
           %mend;
           %avgyld(g30f1b, 'F1B', 'W', W95, W97, W99);
           %avgyld(g30f1c, 'F1C', 'W', W95, W97, W99);
           %avgyld(g30f3b, 'F3B', 'W', W95, W97, W99);
           %avgyld(g30f4 , 'F4' , 'W', W95, W97, W99);
           %avgyld(g30f2 , 'F2' , 'C', C95, C97, c99);
           %avgyld(g30f5 , 'F5' , 'C', C95, C97, c99);
           %avgyld(g30f7 , 'F7' , 'C', C95, C97, c99);
           %avgyld(g30f8b, 'F8B', 'C', C95, C97, c99);

           data fldlevel;
                set g30f1b g30f1c g30f3b g30f4 g30f2 g30f5 g30f7 g30f8b;
                ratio5=y1999/y1995;
                ratio7=y1999/y1997;
           run;

     /*** proc print data=fldlevel;run;endsas;

     OBS     Y1995      Y1997      Y1999     FLD    CROP     RATIO5     RATIO7

      1      85.037     74.164     44.457    F1B     W      0.52280    0.59944
      2      86.008     68.114     92.281    F1C     W      1.07293    1.35479
      3      92.512     87.975     94.095    F3B     W      1.01711    1.06956
      4     100.931     91.629     86.761    F4      W      0.85960    0.94687
      5     117.750    104.595     92.163    F2      C      0.78270    0.88114
      6     177.006     86.974    124.745    F5      C      0.70475    1.43428
      7     188.335    124.793    105.907    F7      C      0.56233    0.84866
      8     182.247    111.085    119.365    F8B     C      0.65497    1.07454

     ******/


%macro grd30(fldfile,  /*  location of the file                  */
             fldname,  /*  name of the field                     */
             crop,     /*  name of the crop                      */
             crop95,   /*  original variable name for 1995 yield */
             crop97,   /*  original variable name for 1997 yield */
             crop99,   /*  original variable name for 1999 yield */
             smutrans, /*  SMU transition matrix for the field   */
             smufld    /*  SMU states and yields for the field   */
             );

   /* rescale the orginal yield data for 1995 and 1997 in order to estimate transition*/
      proc sort data=fldlevel
                     (keep= fld ratio5 ratio7);
           by fld;
      run;

      data field;
           set big.&fldfile;
           length fld $ 8;
           fld=&fldname;
      run;

      proc sort data=field;
           by fld;

      run;

      data field;
           merge field (in=org) fldlevel;
           by fld;
           if org;
           y1995=ratio5*&crop95; /*SMU yld in 1995 adjusted to real 1999*/
           y1997=ratio7*&crop97; /*SMU yld in 1997 adjusted to real 1999*/
```

```
           y1999=&crop99;
      keep x y fld cluster soil count y1995 y1997 y1999 &crop95 &crop97;

   /* Get mean yield and std for each soil in order
      to determine the state of nature for each SMU */
   proc sort
      data=field;
      by cluster;
   run;

   proc univariate data=field noprint;
          var y1995 y1997;
          by cluster;
          output out=meanstd mean=m95 m97 std=var95 var97;
          title1 'Field = ' &fldname;
   run;

   data field;
      merge field (in=fldin) meanstd (in=clus);
          by cluster;
          if fldin and clus;
   run;

   /* Determine state of nature for each SMU in 1995 and 1997 */
   data field;
      set field;

      if &crop='W' then
         do;
           v95=var95*0.7;
           v97=var97*0.7;
        end;
      else
         do;
            v95=var95*0.8;
            v97=var97*0.8;
         end;
          s95=(y1995 ge m95+v95)*3 + (y1995 lt m95+v95)*(y1995 ge m95-v95)*2
                                   + (y1995 lt m95-v95)*1;
          s97=(y1997 ge m97+v97)*3 + (y1997 lt m97+v97)*(y1997 ge m97-v97)*2
                                   + (y1997 lt m97-v97)*1;
   run;

   /* Testing if the transiion exists */
   proc freq
      data=field;
      output out=out_p exact; /*this out is to get p-value (var name = p_exact2)*/
          tables s95*s97 /exact out=trantest outpct; /*this out is to get percents for
transition*/
   run;

   proc print
      data=trantest;
      var s95 s97 pct_row;
   run;

         /* get ready to deal with the situation where no Markov exists **/
              data out_p;
                   set out_p;
                   call symput('pvalue',p_exact2);
              run;
              data stat9597;
                 set field;
                 s95_1 = (s95=1);s95_2 = (s95=2);s95_3 = (s95=3);
                 s97_1 = (s97=1);s97_2 = (s97=2);s97_3 = (s97=3);
              run;
              proc univariate data=stat9597 noprint;
                   output out=stat9597 sum=s95_1-s95_3 s97_1 - s97_3;
                   var s95_1-s95_3 s97_1 - s97_3;
              run;
              data stat9597;
```

```
                   set stat9597;
                   call symput("s951", s95_1);
                   call symput("s952", s95_2);
                   call symput("s953", s95_3);
                   call symput("s971", s97_1);
                   call symput("s972", s97_2);
                   call symput("s973", s97_3);
              run;

   proc sort data=trantest; by s95 s97; run;

   data trantest;
     retain fr0m to1 to2 to3;
        format to1 to2 to3 5.3;
     set trantest;
     by s95 s97;

     if s95=1 then fr0m="FROM 1";
     if s95=2 then fr0m="FROM 2";
     if s95=3 then fr0m="FROM 3";

     if &pvalue <= 0.05 then do;
           if s97=1 then to1=pct_row*0.01;
           if s97=2 then to2=pct_row*0.01;
           if s97=3 then to3=pct_row*0.01;
        end;
     else do;
             to1=(&s951 + &s971) / ((&s951 + &s952 + &s953)*2);
             to2=(&s952 + &s972) / ((&s951 + &s952 + &s953)*2);
             to3=1-to1-to2;
          end;

     if last.s95 then
        do;
           output;
           to1=0;to2=0;to3=0;
       end;
        keep fr0m to1 to2 to3;
   run;

   data big.&smutrans;
      set trantest;
      if to1=. then to1=0;
      if to2=. then to2=0;
      if to3=. then to3=0;
   run;

   proc sort
       data=field
       out=big.&smufld (drop= var95 var97 v95 v97);
       by cluster s95 s97;
   run;


%mend;

%grd30(g30f1b, 'F1B', 'W', W95, W97,w99,smutrf1b,smuf1b);
%grd30(g30f1c, 'F1C', 'W', W95, W97,w99,smutrf1c,smuf1c);
%grd30(g30f3b, 'F3B', 'W', W95, W97,w99,smutrf3b,smuf3b);
%grd30(g30f4,  'F4',  'W', W95, W97,w99,smutrf4, smuf4);
%grd30(g30f2,  'F2',  'C',  C95, C97,c99,smutrf2,smuf2);
%grd30(g30f5,  'F5',  'C',  C95, C97,c99,smutrf5,smuf5);
%grd30(g30f7,  'F7',  'C',  C95, C97,c99,smutrf7,smuf7);
%grd30(g30f8b, 'F8B', 'C',  C95, C97,c99,smutrf8b,smuf8b);
```

```
/* sensitivity: expected yield in 1999 are correct  **/
/*************************************************************************************
*progname=combine.sas;
*
*Abstract: Combine all information together for each field based on 30x30 grids.
*          1. Field name,soil name,cluster #, grid counts, and actual yields 1995-1999.
*          2. State of nature in 1997
*          3. Determined N rate for all strategies (ex post or ex ante)
*
*Input: big.smu_var0.sd2;   ---- functional zone optimal yields
*       big.smuf1b.sd2 etc; ---- state of nature in 1997
*       big.g30g1b.sd2 etc; ---- original yields in 30x30 grids
*       big.n_notran.sd2    ---- N rate for soil zone uniform w/o transition considered.
*       other data generated by other programs and reported in Chapter 3.
*
*output: Combined dataset for each field to evaluate inf value and vra value
*
* NOTE: for soil strategies, only soil 171 172 rates prevail for given yields.
***************************************************************************************/

options nodate;

libname big 'c:\data\thesis\richmond\sasdata';
libname com 'c:\data\thesis\richmond\sasdata\finally';


/*The following are from big.n_notran.sd2. The conv and soil unif w/o trans */
/*Also reported in Table 3-18*/
data notr_sc;
   notr_suw=113;notr_suc=140; notr_cuw=113;notr_cuc=140;
   input fld $;
   cards;
   F1B
   F1C
   F2
   F3B
   F4
   F5
   F7
   F8B
    ;
run;
proc sort data=notr_sc;by fld;run;

/*data from Table 3-19 and 3-22 (2Jun01 version) and smu_nrate1999.sas output (for table 3-
25)*/
data tran_sc; /*soil zone uniform and conventional w trans considered*/
   input fld $ ntr_cu ntr_su ntr_fu;
   cards;
   F1B   56  60  68.1865
   F1C  115 131 134.786
   F3B  118 123 140.380
   F4   108 115 119.542
   F2    72  74 104.449
   F5   105 113 139.205
   F7    86  88  98.813
   F8B   99 111 126.137
    ;
run;


proc sort data=tran_sc;by fld;run;

data unif;
   merge notr_sc tran_sc;
   by fld;
run;

proc print data=unif;run;

/*The following comes from Table 3-20 and Table 3-21.
```

```
  The N rate for soil zone variable application */
data ntr_sv;
  input fld $ soil ntr_sv;
  cards;
   F1B  148 20
   F1B  151 53
   F1B  168 20
   F1B  171 60
   F1B  172 60
   F1C  148 44
   F1C  171 131
   F1C  172 131
   F3B  151 109
   F3B  168 41
   F3B  171 123
   F3B  172 123
   F4   148 38
   F4   171 115
   F4   172 115
   F2   168 18
   F2   171 74
   F2   172 74
   F5   148 34
   F5   151 88
   F5   171 113
   F5   172 113
   F7   151 68
   F7   161 61
   F7   171 88
   F7   172 88
   F8B  148 33
   F8B  151 87
   F8B  171 111
   F8B  172 111
    ;
run;



/*now assign state of nature in 1997 to original field data*/
%macro bigmer (fldname,
               fldfile, /*field file containing state of nature in 1997*/
               original,/*original yield data                         */
               outfile, /*output combined file                        */
               );

           proc sort data=ntr_sv; by fld;run;
           data middle; /*combine notran and tran*/
                merge unif ntr_sv (in=va);
                by fld;
                if va and fld=&fldname;
           run;
           proc sort data=middle;by fld soil;run;

        /* the following is for functional zone variable */
        data ntr_fv;
           set big.smu_var0;
           cluster=clusno;
           ntr_fv=nrate;
           if fld=&fldname;
           keep fld cluster s97 ntr_fv;
        run;
        proc sort data=ntr_fv;by fld cluster s97;run;

        proc sort
             data=big.&fldfile
             out=&fldfile (keep=fld cluster soil x y s97);
             by fld cluster s97;
        run;
        data ntr_fv;
```

```
             merge ntr_fv &fldfile (in=fd);
             by fld cluster s97;
             if fd;
        run;

        proc sort
             data=ntr_fv;
             by fld soil;
        run;

        data middle;
             merge middle ntr_fv (in=fv);
             by fld soil;
             if fv;
        run;
proc print data=middle;run;

    proc sort
         data=middle;
         by x y;
    run;

    proc sort
         data=big.&original
         out=&original;
         by x y;
    run;

    data com.&outfile;
       merge middle &original (in=org);
       by x y;
       if org;
   run;

%mend;
%bigmer('F1B',smuf1b,g30F1b,comf1b);
%bigmer('F1C',smuf1c,g30F1c,comf1c);
%bigmer('F3B',smuf3b,g30F3b,comf3b);
%bigmer('F4' ,smuf4 ,g30F4 ,comf4 );
%bigmer('F2' ,smuf2 ,g30F2 ,comf2 );
%bigmer('F5' ,smuf5 ,g30F5 ,comf5 );
%bigmer('F7' ,smuf7 ,g30F7 ,comf7 );
%bigmer('F8B',smuf8b,g30F8b,comf8b);




/*snesitivity: all expected yields are correct

/******************************************************************************************
*progname=evaluate.sas;
*
*Abstract: Evaluate all information values and values of viriable application in ex ante
*          analyses
*
*
*Input: final.comf1b.sd2 etc.
*       information cost and cost of VRA (from Chapter 3).
*
*
*output: Results for ex ante analyese. Output file is: final.allout.sd2
*
*date: 3Jun2001
***************************************************************************************/
options nodate;
libname final 'c:\data\thesis\richmond\sasdata\finally';

%macro evaluate(fldname,     /*name of the field                   */
                comfld,      /*field file with combined information*/
```

```
           crop99,      /*the crop planted in 1999            */
           cropyld,     /*yield of crop in 1999 for each SMU  */
           outfile,     /*the output file for all evaluation  */
           infsu,       /*force su apply same as cu           */
           infsv,       /*force sv apply same as cu          */
           inffu,       /*force fu apply same as cu           */
           inffv        /*force fv apply same as cu           */
           )
           ;

data &comfld;
   set final.&comfld;
   area=count*9*0.00025;
   yield=&cropyld*area;

                  /*second sensit analysis: variable application same as uniform   **/
                    ntr_su=ntr_su*(&infsu+0);
                    ntr_sv=ntr_sv*(&infsv+0);
                    ntr_fu=ntr_fu*(&inffu+0);
                    ntr_fv=ntr_fv*(&inffv+0);
   keep fld area  yield &cropyld ntr_cu ntr_su ntr_fu ntr_sv ntr_fv;
run;
proc sort data=&comfld;by fld; run;


/*The the calculation*/
data &comfld;
   set &comfld;

   /*first, N applied*/
   napp_cu = ntr_cu * area;
   napp_su = ntr_su * area;
   napp_sv = ntr_sv * area;
   napp_fu = ntr_fu * area;
   napp_fv = ntr_fv * area;

   /*second, yld penalty*/
   if &crop99 = 'W' then do;
           ypen_cu = -min(ntr_cu - &cropyld * 1.25, 0) * 0.8 * area;
           ypen_su = -min(ntr_su - &cropyld * 1.25, 0) * 0.8 * area;
           ypen_sv = -min(ntr_sv - &cropyld * 1.25, 0) * 0.8 * area;
           ypen_fu = -min(ntr_fu - &cropyld * 1.25, 0) * 0.8 * area;
           ypen_fv = -min(ntr_fv - &cropyld * 1.25, 0) * 0.8 * area;
        end;
   else do;
           ypen_cu = -min(ntr_cu - &cropyld + 20, 0) * area;
           ypen_su = -min(ntr_su - &cropyld + 20, 0) * area;
           ypen_sv = -min(ntr_sv - &cropyld + 20, 0) * area;
           ypen_fu = -min(ntr_fu - &cropyld + 20, 0) * area;
           ypen_fv = -min(ntr_fv - &cropyld + 20, 0) * area;
        end;

   /*third, net N */
   if &crop99='W' then do;
           netn_cu = max(ntr_cu - &cropyld*1.25 + 2.8, 0) *area;
           netn_su = max(ntr_su - &cropyld*1.25 + 2.8, 0) *area;
           netn_sv = max(ntr_sv - &cropyld*1.25 + 2.8, 0) *area;
           netn_fu = max(ntr_fu - &cropyld*1.25 + 2.8, 0) *area;
           netn_fv = max(ntr_fv - &cropyld*1.25 + 2.8, 0) *area;
        end;
   else do;
           netn_cu = max(ntr_cu - (&cropyld-20)*0.9 + 1.3, 0) *area;
           netn_su = max(ntr_su - (&cropyld-20)*0.9 + 1.3, 0) *area;
           netn_sv = max(ntr_sv - (&cropyld-20)*0.9 + 1.3, 0) *area;
           netn_fu = max(ntr_fu - (&cropyld-20)*0.9 + 1.3, 0) *area;
           netn_fv = max(ntr_fv - (&cropyld-20)*0.9 + 1.3, 0) *area;
        end;

   /*cost of inaccurate N appl*/
  if &crop99='W' then do;
           cost_cu = (0.25*max(ntr_cu - &cropyld*1.25, 0)-((2.87-0.15)*0.8
```

```
                     -0.25)*min(ntr_cu - &cropyld*1.25,0))*area;
           cost_su = (0.25*max(ntr_su - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                     -0.25)*min(ntr_su - &cropyld*1.25,0))*area;
           cost_sv = (0.25*max(ntr_sv - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                     -0.25)*min(ntr_sv - &cropyld*1.25,0))*area;
           cost_fu = (0.25*max(ntr_fu - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                     -0.25)*min(ntr_fu - &cropyld*1.25,0))*area;
           cost_fv = (0.25*max(ntr_fv - &cropyld*1.25, 0)-((2.87-0.15)*0.8
                     -0.25)*min(ntr_fv - &cropyld*1.25,0))*area;
       end;
  else do;
           cost_cu = (0.25*max(ntr_cu - &cropyld + 20, 0)-(2.10-0.15-0.25)
                     *min(ntr_cu - &cropyld + 20,0))*area;
           cost_su = (0.25*max(ntr_su - &cropyld + 20, 0)-(2.10-0.15-0.25)
                     *min(ntr_su - &cropyld + 20,0))*area;
           cost_sv = (0.25*max(ntr_sv - &cropyld + 20, 0)-(2.10-0.15-0.25)
                     *min(ntr_sv - &cropyld + 20,0))*area;
           cost_fu = (0.25*max(ntr_fu - &cropyld + 20, 0)-(2.10-0.15-0.25)
                     *min(ntr_fu - &cropyld + 20,0))*area;
           cost_fv = (0.25*max(ntr_fv - &cropyld + 20, 0)-(2.10-0.15-0.25)
                     *min(ntr_fv - &cropyld + 20,0))*area;
       end;
run;

proc sort data=&comfld;by fld;run; /*keep field name in the following PROC*/

proc univariate data=&comfld noprint;
     output out=&outfile
            sum=area napp_cu napp_su napp_sv napp_fu napp_fv
                     ypen_cu ypen_su ypen_sv ypen_fu ypen_fv
                     netn_cu netn_su netn_sv netn_fu netn_fv
                     cost_cu cost_su cost_sv cost_fu cost_fv
                yield;
     var area napp_cu napp_su napp_sv napp_fu napp_fv
              ypen_cu ypen_su ypen_sv ypen_fu ypen_fv
              netn_cu netn_su netn_sv netn_fu netn_fv
              cost_cu cost_su cost_sv cost_fu cost_fv
         yield;
     by fld;
run;

/*calculated to the per acre base */
data &outfile;
     set &outfile;

     yield = yield /area;

     napp_cu = napp_cu / area;
     napp_su = napp_su / area;
     napp_sv = napp_sv / area;
     napp_fu = napp_fu / area;
     napp_fv = napp_fv / area;

     ypen_cu = ypen_cu / area;
     ypen_su = ypen_su / area;
     ypen_sv = ypen_sv / area;
     ypen_fu = ypen_fu / area;
     ypen_fv = ypen_fv / area;

     netn_cu = netn_cu / area;
     netn_su = netn_su / area;
     netn_sv = netn_sv / area;
     netn_fu = netn_fu / area;
     netn_fv = netn_fv / area;

     cost_cu = cost_cu / area;
     cost_su = cost_su / area;
     cost_sv = cost_sv / area + 2.09;
     cost_fu = cost_fu / area + 2.61;
     cost_fv = cost_fv / area + 5.22;
```

```
          /*now information value and VRA value*/
          info_s = cost_cu - cost_su;
          info_f = cost_cu - cost_fu;

          var_s = cost_su - cost_sv;
          var_f = cost_fu - cost_fv;

         /*finally, net N reduction */
          nless_su = netn_cu - netn_su;
          nless_sv = netn_cu - netn_sv;
          nless_fu = netn_cu - netn_fu;
          nless_fv = netn_cu - netn_fv;

          crop=&crop99;
     run;

proc print data=&outfile;run;

%mend;

%evaluate('F1B', comf1b, 'W', W99, outf1b, 0.933,1.017,0.821,0.894);
%evaluate('F1C', comf1c, 'W', W99, outf1c, 0.878,1.001,0.853,0.908);
%evaluate('F3B', comf3b, 'W', W99, outf3b, 0.959,0.999,0.841,0.894);
%evaluate('F4' , comf4 , 'W', W99, outf4 , 0.939,0.992,0.903,0.927);
%evaluate('F2' , comf2 , 'C', C99, outf2 , 0.973,0.999,0.689,0.841);
%evaluate('F5' , comf5 , 'C', C99, outf5 , 0.929,1.000,0.754,0.806);
%evaluate('F7' , comf7 , 'C', C99, outf7 , 0.977,0.997,0.870,0.909);
%evaluate('F8B', comf8b, 'C', C99, outf8b, 0.892,1.001,0.785,0.837);

data final.s_allout;
    set outf1b outf1c outf3b outf4 outf2 outf5 outf7 outf8b;
run;
```

Vita

Wei Peng was born in China. He came to the United States on November 28, 1994 in his pursuit of liberty and an understanding of the market economy. In his seven years' stay at Virginia Tech, he completed an MS in Agricultural Economics, an MS in Statistics, and finally a PhD in Economics. He is satisfied with the worldly wisdom he has obtained from studying various economic theories over the years. Now he works as a biostatistician in the pharmaceutical industry, never tired of explaining basic economics to his naïve colleagues and friends.

The most valuable gain of his adventure in the United States is his encounter with Jesus Christ, his Lord and Savior. And now he lives with his wife Mary, son Rick, and daughter Nancy at Charlottesville, Virginia. He also takes classes in statistics at University of Virginia and is hoping that in the far future, he can get another PhD in biostatistics. That is all we can tell of him at this moment.