Staff Paper P04-2 March 2004

Staff Paper Series

Estimating Site-Specific Nitrogen Crop Response Functions: A Conceptual Framework and Geostatistical Model

By

Terrance M. Hurley

Gary L. Malzer

Bernard Kilian

DEPARTMENT OF APPLIED ECONOMICS
COLLEGE OF AGRICULTURAL, FOOD, AND ENVIRONMENTAL SCIENCES
UNIVERSITY OF MINNESOTA

Staff Paper P04-2 March 2004

Estimating Site-Specific Nitrogen Crop Response Functions: A Conceptual Framework and Geostatistical Model

Terrance M. Hurley*
Department of Applied Economics
University of Minnesota

Gary L. Malzer
Department of Soil, Water, and Climate
University of Minnesota

Bernard Kilian
Instituto Centroamericano de Administración de Empresas (INCAE)
Latin American Center for Competitiveness and Sustainable Development
Alejuela, Costa Rica

* Corresponding author–thurley@apec.umn.edu

Copies of this publication are available at http://agecon.lib.umn.edu/. Information on other titles in this series may be obtained from: Waite Library, University of Minnesota, Department of Applied Economics, 232 Classroom Office Building, 1994 Buford Avenue, St. Paul, MN 55108, U.S.A.

Estimating Site-Specific Nitrogen Crop Response Functions:

A Conceptual Framework and Geostatistical Model

Abstract

Confirming the precision agriculture hypothesis for variable rate nitrogen applications (VRA) is challenging. To confront this challenge, researchers have used increasingly sophisticated statistical models to estimate and compare site-specific crop response functions (SSCRFs). While progress has been made, it has been hampered by the lack of a conceptual framework to guide the development of appropriate statistical models. This paper provides such a framework and demonstrates its utility by developing a heteroscedastic, fixed and random effects, geostatistical model to test if VRA can increase nitrogen returns. The novelty of the model is the inclusion of site, spatial, treatment, and treatment strip heteroscedasticity and correlation. Applied to data collected in 1995 from two corn nitrogen response experiments in South Central Minnesota, results demonstrate the importance of including site, spatial, treatment, and treatment strip effects in the estimation of SSCRFs. Results also indicate a significant potential for VRA to increase nitrogen returns and that these potential returns increase as the area of the management unit decreases. At one location, there was greater than a 95% chance that VRA could have increased profitability if the cost of implementing VRA was less than 14.5 \$ ha^{-1}. At the other location, if implementation costs were less than 48.3 \$ ha^{-1}, there was greater than a 95% chance of increased profitability.

Abbreviations and Notation

FGLS	Feasible Generalized Least Squares
LRS	Likelihood Ratio Statistic
ML	Maximum Likelihood
OLS	Ordinary Least Squares
PA	Precision Agriculture
SSCRF	Site-Specific Crop Response Function
UMN	University of Minnesota
URA	Uniform Rate Nitrogen Application
VRA	Variable Rate Nitrogen Application
*	Significant for $p < 0.05$
**	Significant for $p < 0.01$
***	Significant for $p < 0.001$
x	Units of Variable/Managed Input
x^*	Optimal Units of Variable/Managed Input
z	Units of Fixed/Unmanaged Input
$y = f(x, z)$	Units of Crop Yield as a Function of Variable and Fixed Inputs
p_y and p_x	Price Per Unit of Crop Yield and Variable Input
k_x and k_z	Indexes for Units of Variable and Fixed Input
$\beta_{k_x k_z} = \frac{\partial^{k_x+k_z} f(x,z)}{\partial x^{k_x} \partial z^{k_z}}$	k_x and k_z Order Cross Partial Derivative of Crop Yield With Respect to the Variable and Fixed Input
N	Number of Observations
i	Observation Index
e_i	Approximation and Measurement Error
K_x and K_z	Integer Constants

α_{k_x}	k_xth Estimable Mean Parameter
Z_{k_z}	k_zth Unobserved Real Constant
ξ_i	Unobserved Error
R and r_i	Number of Field Sites/Partitions and Site Assigned to the ith Observation
$\alpha_{k_x r}$	k_xth Estimable Mean Parameter for Site r
$Z_{k_z r}$	k_zth Unobserved Real Constant for Site r
E[·]	Expectation Operator
d_{ij}	Spatial Distance Between Observations i and j
C_0	Estimable Spatial Nugget Semi-Variogram Parameter
C_1	Estimable Semi-Variogram Distance Correlation Parameter
a	Estimable Semi-Variogram Range/Shape Parameter
$g(d_{ij}, a)$	Semi-Variogram Distance Function
C_x	Estimable Variable Input Correlation Parameter
x_{ij}	Indicator Variable for Observations With the Same Variable Input
C_s	Estimable Treatment Strip Correlation Parameter
s_{ij}	Indicator Variable Equal for Observations from the Same Strip
$\gamma(d_{ij})$	Semi-Variogram
σ_{rs}	Covariance for Site r and Treatment Strip s
σ_r and σ_s	Estimable Covariance Parameters for Site r and Treatment Strip s
π	Returns to Nitrogen Above the Cost of Nitrogen

The precision agriculture (PA) hypothesis asserts that varying management activities between or within fields can benefit farmers or the environment. A necessary condition for PA is that the productivity of management activities must vary between or within fields due to factors typically not managed by a farmer. Validating this necessary condition is challenging due to the inherent difficulties of collecting, analyzing, and interpreting appropriate data.

One approach that has emerged to test the PA hypothesis for variable rate nitrogen applications (VRA) is the estimation and comparison of site-specific crop response functions (SSCRFs) using multiple regression analysis (e.g. Davis et al., 1996; Malzer et al., 1996; Bongiovanni and Lowenberg-DeBoer, 2000 and 2001; Lambert et al., 2002; Hurley et al. 2002a,b; and Mamo et al., 2003). Early applications relied on ordinary least squares (OLS), which does not account for heteroscedastic or correlated errors. While OLS estimates may remain unbiased even with heteroscedasticity and correlation, they are typically not efficient and can convey a false sense of precision (Schabenbeger and Pierce, 2002). Having confirmed the presence of spatial correlation, recent applications have used more sophisticated statistical models to address this problem. Still, the conceptual foundations used to justify these models are seldom explicit, making it difficult to judge the merit of the method.

The purpose of this paper is to provide a conceptual framework to illuminate how SSCRFs can be used to test the PA hypothesis. The framework is useful because it identifies an appropriate hypothesis and explains recent evidence of site and treatment dependent heteroscedasticity and spatial correlation (Hernandez and Mulla, 2002; Hurley et al., 2002a; and Lambert et al., 2002). The framework is used to guide the development of a heteroscedastic, fixed and random effects, geostatistical model for estimating SSCRFs and testing the PA hypothesis using field data from a common experimental design. The novelty of the model is the inclusion of site, spatial, treatment, and treatment strip dependent heteroscedasticity and correlation. The model is applied to 1995 field data to demonstrate the importance of the conceptual results, test the PA hypothesis, and estimate the potential value of PA.

Conceptual Framework

Implications of the PA Hypothesis

The precision agriculture (PA) hypothesis asserts that farmers or the environment can benefit from varying management within or between fields. To understand this hypothesis from a farmer's perspective (analogous arguments exist for the environment), suppose crop yield y (e.g. corn kg ha^{-1}) depends on two types of inputs. The first, denoted by x, are variable inputs or a farmer's managed inputs (e.g. nitrogen). The second, denoted by z, are fixed inputs or a farmer's unmanaged inputs (e.g. soil type, rainfall, and topography). The general relationship between yield, and variable and fixed inputs is described as $y = f(x, z)$, which is assumed continuously differentiable in x and z. For convenience, y, x, and z are treated as scalars.

If a farmer's objective is to manage the variable input to optimize the net return, the classic rule from economic theory says to choose x^* such that $p_y \frac{\partial f(x^*, z)}{\partial x} = p_x$ where p_y and p_x are the price per unit of crop yield and variable input. In economic parlance, the rule states that an input's value of marginal product should equal its marginal cost. The optimal amount of variable input depends on the crop price, variable input price, and most importantly for PA, amount of fixed input. How the optimal amount of variable input depends on the amount of fixed input is found using the implicit function theorem:

$\frac{\partial x^*}{\partial z} = -p_y \frac{\partial^2 f(x^*, z)}{\partial x \partial z} \frac{\partial x^2}{\partial^2 f(x^*, z)}$. Note that the optimal amount of variable input does not change with the amount of fixed input if $\frac{\partial^2 f(x^*, z)}{\partial x \partial z} = 0$, which means there is no interaction between the variable and fixed input. For example, if soil organic matter does not influence crop response to nitrogen, there is no value to varying nitrogen applications in response to variation in soil organic matter.

Testing the PA Hypothesis with an Observable Fixed Input

Observational and experimental field data provide an opportunity to test the PA hypothesis, but the development of appropriate statistical models has proven challenging. To understand why, consider the Taylor series expansion,

(1) $$f(x,z) = \beta_{00} + \sum_{k_x=1}^{\infty} \beta_{k_x 0} x^{k_x} + \sum_{k_z=1}^{\infty} \beta_{0 k_z} z^{k_z} + \sum_{k_x=1}^{\infty} \sum_{k_z=1}^{\infty} \beta_{k_x k_z} x^{k_x} z^{k_z}$$

= Constant + Main Effect of x + Main Effect of z + Interaction Between x and z

and

(2) $$\frac{\partial^2 f(x,z)}{\partial x \partial z} = \sum_{k_x=1}^{\infty} \sum_{k_z=1}^{\infty} \beta_{k_x k_z} k_x k_z x^{k_x - 1} z^{k_z - 1}$$

where $\beta_{k_x k_z} = \frac{\partial^{k_x + k_z} f(0,0)}{k_x! k_z! \partial x^{k_x} \partial z^{k_z}}$ for all k_x and k_z are real constants that indicate how variable and fixed inputs combine to influence yield. Equation (1) is a general decomposition of yield into the familiar constant, main, and interaction effects. Equation (2) suggests the null hypothesis $\beta_{k_x k_z} = 0$ for all $k_x > 0$ and $k_z > 0$, which implies PA cannot be used to the benefit of a farmer or the environment because there is no interaction effect.

Consider a set of data collected from a controlled field experiment: (y_i, x_i, z_i) for $i = 1, 2, \ldots, N$. An individual data point consists of y_i, an observed yield; x_i, an observed variable input; and z_i, an observed fixed input. To test the PA hypothesis with this data, the constant β coefficients in equation (1) must be estimated, a task that is generally not feasible.

The first obstacle is the dimension of the problem. Since the true relationship between yield and inputs is seldom (if ever) known, some approximation is necessary. Additionally, there is the potential for measurement error. Both problems are universal and the common solution (explicit or implicit) is to truncate the expansion in equation (1) and add an error:

(3) $$y_i = \beta_{00} + \sum_{k_x=1}^{K_x} \beta_{k_x 0} x_i^{k_x} + \sum_{k_z=1}^{K_z} \beta_{0 k_z} z_i^{k_z} + \sum_{k_x=1}^{K_x} \sum_{k_z=1}^{K_z} \beta_{k_x k_z} x_i^{k_x} z_i^{k_z} + e_i$$

= Constant + Main Effect of x + Main Effect of z + Interaction Between x and z + Error

where K_x and K_z are integers and e_i includes the approximation error due to truncation and measurement error in yield and inputs. Equation (3) is a generalized linear regression model, so the parameters for the constant, main, and interaction effects can be estimated using a variety of techniques. For example, if it is

reasonable to assume e_i is independently and identically distributed with zero mean for $i = 1, 2, \ldots, N$, ordinary least squares (OLS) is appropriate. If the variance of error differs between observations (i.e. there is heteroscedasticity) or errors are correlated (e.g. spatially), feasible generalized least squares (FGLS) or maximum likelihood (ML) with a heteroscedastic and correlated covariance matrix is appropriate. Depending on the method, the PA hypothesis can be tested using the F or likelihood ratio statistic (LRS).

Testing the PA Hypothesis with an Unobservable Fixed Input

Another obstacle more specific to PA is that z_i is often unobserved. A researcher or farmer may suspect some fixed input interacts with the variable input, but not know which fixed input is important. Confirming the PA hypothesis without knowledge of important fixed inputs is useful because it indicates whether searching for such inputs is worth an effort. If the PA hypothesis cannot be confirmed generally or the value of discovering which fixed inputs are important is small, it makes sense to devote research effort elsewhere.

When z_i is unobserved, it can be treated as another source of error. Equation (3) becomes

$$y_i = \alpha_0 + \sum_{k_x=1}^{K_x} \alpha_{k_x} x_i^{k_x} + \xi_i \tag{4}$$

= Constant + Main Effect of x + Error

where $\alpha_{k_x} = \beta_{k_x 0} + \sum_{k_z=1}^{K_z} \beta_{k_x k_z} Z_{k_z}$ for $k_x = 0,\ldots,K_x$, and Z_{k_z} are real constants and

$\xi_i = \sum_{k_z=1}^{K_z} \left(\beta_{0k_z} + \sum_{k_x=1}^{K_x} \beta_{k_x k_z} x_i^{k_x} \right) \left(z_i^{k_z} - Z_{k_z} \right) + e_i$ is the regression error. Under the traditional assumption that the expected value of the regression error is zero, Z_{k_z} is an average of $z_i^{k_z}$. Two important implications emerge from equation (4). First, the parameters associated with the constant and main effect of the variable input depend on the interaction between the variable and fixed inputs. Second, there is another source of error attributable to the unobserved fixed input that is dependent on the variable input and interactions between the variable and fixed input.

Note that in a perfectly controlled experiment, the value of the fixed input is constant for all observations: $z_i^{k_z} - Z_{k_z} = 0$ for all i and k_z. Therefore, the only source of error is related to approximation and measurement. Unfortunately, most field experiments are not perfectly controlled, so error attributable to variation in the unobserved fixed input can be important.

Testing the null hypothesis for PA using equation (4) is complicated by the fact that the interaction parameters of interest are inextricably embedded in the parameter estimates for the main effect of the variable input and in the error. This complication highlights the utility of estimating SSCRFs to test the PA hypothesis. Suppose the data is partitioned by dividing the field into R distinct sites such that $r_i \in \{1,\ldots,R\}$ denotes the ith observation's assigned site. Separate α parameters can be estimated for each site by rewriting equation (4) as

$$(5) \qquad y_i = \alpha_{0r_i} + \sum_{k_x=1}^{K_x} \alpha_{k_x r_i} x_i^{k_x} + \xi_i$$

= Constant for Site r_i + Main Effect of x for Site r_i + Error

where $\alpha_{k_x r_i} = \beta_{k_x 0} + \sum_{k_z=1}^{K_z} \beta_{k_x k_z} Z_{k_z r_i}$ for $k_x = 0,\ldots,K_x$, and $Z_{k_z r_i}$ are real constants; and

$\xi_i = \sum_{k_z=1}^{K_z} \left(\beta_{0k_z} + \sum_{k_x=1}^{K_x} \beta_{k_x k_z} x_i^{k_x} \right) \left(z_i^{k_z} - Z_{k_z r_i} \right) + e_i$ is the regression error. Under the assumption that the expected value of the regression error is zero, $Z_{k_z r_i}$ is an average of $z_i^{k_z}$ for all observations falling in site r_i.

Equation (5) shows that the parameters for the main effect of the variable input at a site can be decomposed into a main effect of the variable input ($\beta_{k_x 0}$) that does not vary by site and an interaction $\left(\sum_{k_z=1}^{K_z} \beta_{k_x k_z} Z_{k_z r_i} \right)$ that does vary by site. Therefore, if the PA hypothesis is true and $Z_{k_z r}$ vary by site, the parameters for the main effect of the variable input will vary by site. This implies that if the null

hypothesis ($\beta_{k_x k_z} = 0$ for $k_x = 1,\ldots K_x$ and $k_z = 1,\ldots K_z$) is correct, $\alpha_{k_x r_i} = \alpha_{k_x r_j}$ for $k_x = 1,\ldots K_x$, and all r_i and r_j.

SSCRFs allow the PA hypothesis to be tested by comparing parameter estimates for the main effect of the variable input in equation (5) across sites— parameters for which efficient and unbiased estimates can usually be obtained even in the presence of heteroscedastic and correlated errors. It is important to note that this test does not imply the equality of site constants (α_{0r_i} for all r_i). When there is no interaction between the variable and fixed input, check plot yields (yields with no variable input) can vary across sites, even though crop response to the variable input does not. Equation (5) shows this is possible because the main effect of the fixed input is absorbed into the site constants.

Using equation (5) to test the PA hypothesis is still not trivial because of the covariance

(6) $\mathrm{E}[\xi_i \xi_j] =$

$$\sum_{k_z=1}^{K_z}\sum_{k_z'=1}^{K_z}\left(\beta_{0k_z}+\sum_{k_x=1}^{K_x}\beta_{k_xk_z}x_i^{k_x}\right)\left(\beta_{0k_z'}+\sum_{k_x=1}^{K_x}\beta_{k_xk_z'}x_j^{k_x}\right)\mathrm{E}\left[\left(z_i^{k_z}-Z_{k_zr_i}\right)\left(z_j^{k_z'}-Z_{k_z'r_j}\right)\right]$$

$$+\sum_{k_z'=1}^{K_z}\left(\beta_{0k_z'}+\sum_{k_x=1}^{K_x}\beta_{k_xk_z'}x_j^{k_x}\right)\mathrm{E}\left[\left(z_j^{k_z'}-Z_{k_z'r_j}\right)e_i\right]$$

$$+\sum_{k_z=1}^{K_z}\left(\beta_{0k_z}+\sum_{k_x=1}^{K_x}\beta_{k_xk_z}x_i^{k_x}\right)\mathrm{E}\left[\left(z_i^{k_z}-Z_{k_zr_i}\right)e_j\right]+\mathrm{E}[e_ie_j].$$

Equation (6) provides an explanation for three phenomena reported in the literature. The first and most common is spatial correlation where regression errors tend to be more correlated for observations that are closer in distance to each other. If fixed inputs are spatially correlated, $z_i^{k_z} - Z_{k_z r_i}$ and $z_j^{k_z'} - Z_{k_z' r_j}$ will be spatially correlated. Hernandez and Mulla (2002) also reports semi-variogram estimates that vary by treatment, a result explained by the dependence of equation (6) on the variable input, $x_i^{k_x}$ and $x_j^{k_x}$, when the PA hypothesis is true. Hurley et al. (2002a) and Lambert et al. (2002) report site-specific

heteroscedasticity, a result consistent with the dependence of equation (6) on $Z_{k_z r_i}$ and $Z_{k_z r_j}$. Each of these phenomena implies estimates from equation (5) using OLS will be inefficient.

A variety of methods have been proposed to deal with the estimation problems posed by these phenomena. Spatial econometric and geostatistical models have been estimated to address problems arising from spatial correlation. Hernandez and Mulla (2002) estimate treatment specific semi-variograms to deal with treatment dependent spatial correlation. Hurley et al. (2002a) and Lambert et al. (2002) incorporate site dependent heteroscedasticity using OLS, spatial econometric, and geostatistical models. None of these models or others we are aware of address site, spatial, and treatment dependent heteroscedasticity and correlation jointly.

The practical relevance of these problems is now explored using data from a common experimental design. The experiment was constructed to test within field variation in corn response to nitrogen. After discussing the design details, a new statistical model is proposed using insights gleaned from the experimental design and conceptual framework.

Materials and Methods

Experimental

Data were collected in 1995 from two production fields near Hanska and Morgan (Brown and Redwood Co. in South Central Minnesota). These sites are located on a higher elevation of glacial till lowland plain that comprises the majority of the Counties. Most soils at these locations belong to the Clarion-Nicollet-Webster association or similar soil series/ associations. The area is nearly level to gently sloping, and the soils range from poorly to moderately well drained. All soils were mollisols, ranging from fine-loamy, mixed, mesic Typic Haplaquolls (the Webster clay loam) to fine-loamy, mixed, mesic typic Hapludolls (the Clarion loam). The climate is interior continental with cold winters and moderately hot summers with occasional cool periods. Total annual precipitation ranges from 635 to 711 mm, which is normally adequate for corn, since 80% falls during the growing season. The 1994 crop was soybean and no manure applications had occurred in the last five growing seasons.

Each 4.5 ha experimental field plots was 164 by 274 m. Within this area six 27.3 by 274 m replications of six 4.6 by 274 m treatment strips were established in a randomized complete block design (for an illustration of an analogous design see Mamo et al., 2003). The strips in each replication included nitrogen rates of 0, 67, 101, 134, 168, and 202 kg ha^{-1} applied as anhydrous ammonia. Treatments were applied on November 4, 1994 using a radar controlled variable rate applicator to ensure a constant application rate within each strip.

Corn (cv. Pioneer 3531) was planted during the first week of May in 0.76 m rows at approximately 76,500 seeds ha^{-1}. Grain yield was determined by harvesting the center two rows (six row strips) with a Massey Ferguson® plot combine equipped with a ground distance monitor and a computerized Harvestmaster® weigh cell. Each of the 36 strips was divided into 17 4.6 by 15.2 m harvest segments. Approximately 8 m was discarded from the end of each strip to eliminate border effects. No headlands were harvested. The experiment produced 612 yield observations at each location. Sub-samples of grain were collected to adjust yields to reflect 15.5 percent moisture. Dikici (2000) reports more details and a descriptive summary of the data.

Empirical

Estimating equation (5) with these data provide an opportunity to test the PA hypothesis for VRA. One feature of these data is that they provide observations for each of the six treatments in 102 15.2 by 27.6 m sites at each location. Therefore, equation (5) can be used to estimate up to 102 SSCRFs with a full complement of treatments. Another feature is that treatments were randomly assigned across, but not within, strips. This lack of randomization within strips may introduce additional correlation.

The conceptual framework and experimental design suggest that estimation of equation (5) using OLS is not efficient. OLS estimates of the standard errors for the α parameters can be either upward or downward biased. The conceptual framework shows the covariance of regression errors will exhibit site and treatment spatial dependencies. Lack of randomization within strips suggests the covariance of regression errors may also exhibit strip dependencies. Therefore, estimates of equation (5) should incorporate an error structure that permits strip as well as site and treatment spatial dependencies.

The proposed model is based on the geostatistical framework. First, let $K_x = 2$, so equation (5) becomes

$$y_i = \alpha_{0r_i} + \alpha_{1r_i} x_i + \alpha_{2r_i} x_i^{\,2} + \xi_i \tag{7}$$

where y_i is corn yield (t ha^{-1}) and x_i is applied nitrogen (kg ha^{-1}) for the ith observation. The covariance of ξ_i and ξ_j is assumed to be

$$\mathrm{E}\left[\left(\xi_i \xi_j\right)\right] = \begin{cases} \sigma_{r_i s_i}^{\,2}, & \text{for } i = j \\ \sigma_{r_i s_i} \sigma_{r_j s_j} \left[C_1\left(1 - g\left(d_{ij}, a\right)\right) + C_x x_{ij} + C_s s_{ij}\right], & \text{for } i \neq j \end{cases} \tag{8}$$

where $\sigma_{r_i s_i}^{\,2} > 0$ and $\sigma_{r_j s_j}^{\,2} > 0$ are the site and strip specific variances for observations i and j; d_{ij} is the distance in meters between observations i and j; x_{ij} is an indicator variable equal to 1.0 if observations i and j had the same treatment applied and 0.0 otherwise; s_{ij} is an indicator variable equal to 1.0 if observations i and j came from the same strip and 0.0 otherwise; $C_1 \geq 0$, $C_s \geq 0$, and $C_x \geq 0$ are spatial, strip, and treatment correlation parameters that assume positive correlation; $1 \geq g(d_{ij}, a) \geq 0$ is a permissible semi-variogram distance function (e.g. see McBratney and Webster, 1986); and a is a range or shape parameter for the semi-variogram distance function.

Dividing equation (8) by $\sigma_{r_i s_i}$ and $\sigma_{r_j s_j}$ results in the correlation coefficient. When $i \neq j$, this correlation coefficient is comprised of three elements: spatial correlation ($C_1(1 - g(d_{ij},a))$, strip correlation ($C_s s_{ij}$), and treatment correlation ($C_x x_{ij}$). Since the correlation coefficient must always lie between 1.0 and -1.0, $1.0 \geq C_1 + C_s + C_x \geq 0.0$ assuming spatial, strip, and treatment correlation are positive to ensure the covariance matrix satisfies the necessary regularity conditions (i.e. is positive definite).

The classical geostatistical approach decomposes variation in the dependent variable into a trend, local variance (nugget), and distance effect. Equations (7) and (8) accomplish a similar decomposition, but add heteroscedasticity, strip effects, and treatment effects. The trend is captured by $\alpha_{0r_i} + \alpha_{1r_i} x_i + \alpha_{2r_i} x_i^{\,2}$, which is site specific. The semi-variogram is

$$(9)\qquad \gamma(d_{ij}) = \begin{cases} 0, & \text{for } d_{ij} = 0 \\ \sigma_{r_i s_i}\sigma_{r_j s_j}\left(C_0 + C_1 g(d_{ij}, a) + C_x(1 - x_{ij}) + C_s(1 - s_{ij})\right), & \text{for } d_{ij} > 0 \end{cases}$$

where $\sigma_{r_i s_i}\sigma_{r_j s_j} C_0 = \sigma_{r_i s_i}\sigma_{r_j s_j}(1 - C_1 - C_s - C_x)$ can be interpreted as the nugget and $\sigma_{r_i s_i}\sigma_{r_j s_j}$ as the sill. Equation (9) shows precisely how the standard geostatistical model is modified by heteroscedasticity, and strip and treatment correlation.

Estimation

Equations (7) and (8) can be estimated using a variety of methods after choosing how to divide the field and a distance function for spatial correlation (Schabenberger and Pierce, 2002). The method employed uses FGLS for the α parameters. Estimates of the covariance parameters (C_1, C_s, C_x, a, and σ^2_{rs} for all $r \in R$ and $s \in \{1,\ldots,36\}$) are obtained using maximum likelihood (ML) after substituting the FGLS estimator for the α parameters. The α parameters are substituted or profiled in this manner because the FGLS estimator for α is the ML estimator given the covariance parameters. The procedure also substantially speeds computation.

The data can be divided into 102 sites with a full complement of treatments, but with only a single observation per strip in each of these sites, it is not possible to identify strip correlation. Therefore, fewer sites are necessary given these data. To illustrate the benefit of estimating equations (7) and (8) for smaller management units, two site partitions are explored. The first divides each location into six contiguous sites of about 0.75 ha; four sites with 108 observations and two sites with 90 (Figure 1). The second divides each location into 48 contiguous sites of about 0.094 ha; six sites with 18 observations and 42 sites with 12.

While these partitions are not the only possible choices and may not be the best choices for each field, refuting the null hypothesis with either is sufficient and can still be used to estimate a lower bound for the potential value of VRA. Note that it is possible to use the model to determine the best way to divide the field by comparing model fit for alternative partitions. This process is time consuming however and beyond the scope of the current paper.

With these two partitions, the variance parameters for every possible site and strip combination (σ^2_{rs} for all $r \in \{1,\ldots,R\}$ and $s \in \{1,\ldots,36\}$) cannot be identified without additional simplifying assumptions. The identification problem is analogous to trying to use an independent variable in a multiple regression analysis that is a linear combination of other independent variables. To identify the model, the site and strip variances were assumed to be multiplicatively separable (i.e. $\sigma_{rs}^2 = \sigma_r^2\sigma_s^2$) and σ_s^2 is set to 1 for $s \in \{1, 19\}$ for six sites and $s \in \{1, 7, 13, 19, 25, 31\}$ for 48 sites. Additively separable variances (i.e. $\sigma_{rs}^2 = \sigma_r^2 + \sigma_s^2$) were also explored, but did not fit the data as well.

There are a variety of possible distance functions. However, the computational intensity of the model restricts the practicality of comparing lots of functions. Since the primary purpose of the paper is to explore the value of incorporating site, treatment, and strip dependencies into a model with spatially correlated errors, attention is focused on a single distance function. Comparing the fit of a standard geostatistical model at both locations based on the maximized log-likelihood using the exponential, Gaussian, and spherical distance functions suggested the Gaussian model fit best. Therefore, the full model with site, treatment, and strip spatial dependencies was estimated with the Gaussian function:

$$g\left(d_{ij}, a\right) = 1 - e^{-\left(\frac{d_{ij}}{a}\right)^2}.$$

Hypotheses

Eight models based on equations (7) and (8) were estimated for each location to test a variety of hypotheses. Table 1 summarizes, while detailing the applicable model restrictions. Model 1 used six sites in an ML analogy to OLS. Model 2 used six sites in a standard geostatistical model. Model 3 used six sites, while adding site and treatment heteroscedasticity and correlation to model 2. Model 4 used six sites, while adding strip heteroscedasticity and correlation to model 3. Model 5 is similar to model 4 except it assumed no interaction between nitrogen and fixed inputs. Models 6, 7, and 8 used 48 sites, but were otherwise identical to models 1, 4, and 5.

Hypothesis testing was accomplished using the likelihood ratio statistic (LRS) since the models are nested. The LRS is twice the difference in the maximized likelihood of the unrestricted and restricted

model. The statistic is asymptotically distributed χ^2 with the degrees of freedom equal to the number of parameter restrictions.

The benefit of incorporating site, spatial, treatment, and strip dependent heteroscedasticity and correlation was evaluated by comparing model 1 and 2, 2 and 3, 3 and 4, and 6 and 7. The comparison of 1 and 2 evaluates the importance of spatial correlation. The comparison of 2 and 3 evaluates the importance of conditioning the variance and spatial correlation on the site and treatment. The comparison of 3 and 4 evaluates the importance of also conditioning on strips. These three comparisons are all based on six sites. The comparison of 6 and 7, evaluates the importance of incorporating site, spatial, treatment, and strip effects with smaller management units (0.094 vs. 0.75 ha).

Comparing models 4 and 5, and 7 and 8 tests the PA hypothesis. The comparison between 4 and 5 evaluates whether there were significant differences in crop response to nitrogen between the six sites in the first partition. The comparison between 7 and 8 evaluates whether there were significant differences in crop response to nitrogen between the 48 sites in the second partition. If there is a significant difference in crop response to nitrogen within a field, VRA can potentially improve nitrogen returns.

Finally, comparing 4 and 7 evaluates variation in crop response functions within the six sites in the first partition. The test determines if dividing a field into smaller management units significantly improves explanatory power.

Potential Value of VRA

The potential value of the increased nitrogen return from VRA was calculated using coefficient estimates for the α parameters in equations (7) and (8). The estimated nitrogen return above fertilizer costs was defined as $\pi = \left(\sum_{i=1}^{612} \frac{p_y\left(\alpha_{0r_i} + \alpha_{1r_i} x_i + \alpha_{2r_i} {x_i}^2\right) - p_x x_i}{612} \right)$. The optimal VRA was calculated by choosing x_i for $i = 1,..,612$ to maximize π. Alternatively, an optimal uniform rate (URA) was calculated by choosing $x = x_i$ for $i = 1,..,612$ to maximize π. These optimal rates were constrained between 0 and 202 kg ha^{-1} to avoid predicting yields outside the range of available data. Nitrogen returns for the optimal VRA and URA were compared to the University of Minnesota (UMN) recommendation (140 kg ha^{-1} for both

Hanska and Morgan) to determine the potential value of VRA within and between fields assuming the price of corn and nitrogen were 98.21 \$ t^{-1} (2.50 \$ bu^{-1}) and 0.374 \$ kg^{-1} (0.17 \$ lbs^{-1}).

Let π^{VRA}, π^{URA}, and π^{UMN} be the estimated nitrogen return for the optimal VRA, optimal URA, and UMN rate. The potential return to switching to the optimal VRA from the UMN rate was calculated as $\pi^{VRA} - \pi^{UMN}$, which represents the potential value of varying nitrogen applications within a field using VRA. This potential value is exclusive of the cost of implementing a VRA strategy (e.g. the cost of information acquisition and variable rate application equipment or services). The standard deviation and 90 percent confidence interval were calculated using a Taylor series expansion (see Caselle and Berger, 1990, pp. 328-331) and assuming normality.

The potential value to switching to the optimal VRA from the UMN rate was decomposed as $\pi^{VRA} - \pi^{UMN} = \pi^{VRA} - \pi^{URA} + \pi^{URA} - \pi^{UMN}$. The potential value of VRA due to switching to the optimal URA from the UMN rate or of getting the right average rate for a field is $\pi^{URA} - \pi^{UMN}$. The potential value of VRA due to switching to the optimal VRA from the optimal URA or to varying the right average rate optimally within a field is $\pi^{VRA} - \pi^{URA}$.

Results

Hypotheses Tests

The regression errors from the SSCRF estimates exhibited significant site, spatial, treatment, and strip dependent heteroscedasticity and correlation. Table 2 reports the maximized log-likelihood for each model, and the LRS and degrees of freedom for each model comparison. Model 1 was rejected in favor of 2 at both locations confirming spatial correlation. Model 2 was rejected in favor of 3 supporting the implications of the conceptual framework. Model 3 was rejected in favor of 4 indicating significant strip dependent heteroscedasticity and correlation. Model 1 was rejected in favor of 4 and 6 was rejected in favor of 7, so dividing fields into smaller management units did not change the importance of site, spatial, treatment, and strip dependent heteroscedasticity and correlation.

There was significant within field variation in corn response to nitrogen, so there was the potential for VRA to improve nitrogen returns. Model 5 was rejected in favor of 4 indicating that nitrogen crop

response varied significantly between the six sites in the first partition at both locations. Model 8 was rejected in favor of 7, indicating that nitrogen crop response varied significantly between the 48 sites in the second partition at both locations. Model 4 was rejected in favor of 7, which means SSCRFs differed significantly within the six sites of the first partition at both locations. Nitrogen returns could be improved by varying nitrogen applications across smaller management units.

Error Structure

Table 3 reports the correlation parameters along with the shape parameter (*a*) and the average standard deviation for selected models.

Spatial and treatment strip correlation were substantial, but treatment correlation was not. Spatial correlation is reduced but not eliminated by estimating SSCRFs for smaller management units, which implies that estimating SSCRFs for smaller management units captures more within field variation in unmanaged inputs. The average standard deviation of error is also reduced when SSCRFs were estimated for smaller management units. Spatial correlation explained between 51 and 63%, while strip correlation explained between 15 and 18% of the semi-variance sill depending on the model and location. The magnitude of the spatial and treatment strip correlation was similar for Hanska and Morgan in models 4 and 7. Comparing model 4 and 7 shows the proportion of the semi-variance sill explained spatially and the shape parameter are lower with 48 rather than six sites. Both factors imply correlation diminishes faster with distance, when smaller management units are used to estimate SSCRFs.

Potential Value of VRA

Figure 2 reports estimates of the potential value of VRA and the decomposition of this value into the effect of switching to the optimal URA from the UMN rate, and to the optimal VRA from the optimal URA. While the results of Table 1 show that model 7 is the best fitting model, results for other models are also reported in order to demonstrate the practical importance of using a model that incorporates site, treatment, and strip as well as spatial effects.

Model 7 indicates that the potential value of switching to the optimal VRA from the UMN rate was 27.54 \$ ha^{-1} and 65.87 \$ ha^{-1} for Hanska and Morgan, with a 95% chance this value exceeded 14.48 \$ ha^{-1} and 48.31 \$ ha^{-1}. For Hanska and Morgan, 10 and 69% of this value could have been achieved by

applying a better uniform rate, while the balance would have required varying applications optimally between the 48 sites within each field. The UMN rate was close to the optimal rate on average for Hanska, but not Morgan. Therefore, most of the VRA benefit for Hanska would have come from varying the application rate optimally within the field, while for Morgan most of the benefit would come from increasing the average application rate for the whole field.

The estimated potential value of VRA increases as the size of the management unit decreases, but the precision of the estimate (width of the confidence interval) may increase or decrease. Estimating SSCRFs with 48 instead of six sites (model 7 vs. 4) increased the estimated potential value of VRA by 133 and 88% for Hanska and Morgan. For Hanska, smaller management units increased the width of the confidence interval for the estimate by about 2%, while for Morgan it decreased it by about 30%.

More of the spatial variability in corn yields and corn response to nitrogen was captured by estimating more SSCRFs for smaller sites within the field. This allows nitrogen application rates to be better tailored to within field variability and increases the potential nitrogen return. It also reduced the error in the estimated SSCRFs, which tended to reduce the width of the confidence intervals making the estimate more precise. However, estimating more SSCRFs increased the number of estimated parameters reducing the model's degrees of freedom, which tended to increase the width of the confidence intervals making the estimate less precise. This result reflects the classic tradeoff between degrees of freedom and error reduction that comes from increasing the number of estimated parameters. For Hanska, the loss of degrees of freedom dominates, so the confidence interval got wider and the estimate became less precise with smaller management units. For Morgan, the reduction in error dominated, so the confidence interval shrank and the estimate became more precise with smaller management units.

Comparing model 4 to 1-3 and 7 to 6 provides insight into the practical importance of using a model with site, treatment, and strip as well as spatial effects. Two features of this comparison are of particular interest.

First, for Hanska, models 1-4 produced similar estimates of the potential value of VRA. Models 6 and 7 also produced similar estimates. These results are consistent with the findings of Lambert et al. (2002). However, for Morgan, the estimate for model 4 is notably lower than the estimates for 1-3 and the

estimate for model 7 is notably lower than for 6. These results are contrary to the findings of Lambert et al. (2002).

The notable reduction in the value of VRA for Morgan using models 4 and 7 can be explained by the increased precision of the estimates of the quadratic parameters in equation (7). Figures 3 and 4 explain why by reporting and illustrating the estimated SSCRFs for models 1 and 4. All the parameter estimates are significant ($p < 0.05$) for model 1 and 4 at Hanska and both models produced similar crop response functions for each site. Both models indicated the response functions were concave (a positive linear and negative quadratic parameter) implying limited nitrogen returns at Hanska. For Morgan, both models produced significant estimates for the constant and linear parameters, but not for the quadratic parameters. Only model 4 produced significant estimates for all quadratic parameters. For sites 1-4, model 1 produced larger linear estimates, but smaller insignificant quadratic estimates implying linear response functions or unlimited nitrogen returns. Model 4 produced smaller linear estimates, but larger significant quadratic estimates implying concave response functions or limited nitrogen returns. The unlimited nitrogen returns implied by model 1 for sites 1-4 result in larger predicted yield increases and a higher estimated value for the optimal VRA.

Replacing model 1 with either 2 or 3 did not qualitatively change results reported in Figures 3 and 4. Ignoring the significant strip effects found in models 4 and 7 resulted in less precise estimates of the quadratic parameters for Morgan. With imprecise quadratic estimates, returns to nitrogen appeared unlimited and the estimate of the potential value of VRA was biased upward. For Hanska, relatively precise estimates were obtained without strip effects because there was a wide enough range of nitrogen treatments employed in the experimental design to clearly delineate when nitrogen returns became limited (e.g. there were many observations where higher treatment rates were associated with lower yields). For Morgan, precise estimates were not obtained without the inclusion of strip effects because there was not a wide enough range of treatments employed to clearly delineate when nitrogen returns became limited (e.g. there were few observations where higher treatment rates were associated with lower yields).

Second, the confidence intervals for model 4 were wider than for 1, 2, and 3 as were the confidence intervals for model 7 when compared to 6. While these results seem to suggest OLS produced

more precise estimates for the potential value of VRA, this is an erroneous conclusion. OLS confidence intervals are reliable only if it is reasonable to assume errors are homoscedastic and uncorrelated. Table 1 rejected these assumptions, so the OLS confidence intervals are unreliable and even worse convey a false sense of precision. For example, with 48 sites OLS can lead to the false conclusion that there was greater than a 95% chance that the potential value of VRA exceeds 15 \$ ha^{-1} for Hanska (model 6 vs. 7 in Figure 2).

Lambert et al. (2002) finds that including spatial correlation improved the precision of the estimated value of VRA. Comparing models 1 and 2 supports this conclusion. However, also including site, treatment, and strip effects reverses this conclusion. Therefore, accounting for spatial correlation without considering site, treatment, and strip effects resulted in even narrower confidence intervals that exacerbate the false sense of precision obtained from OLS.

Figures 5 and 6 report more detailed spatial results for the best fitting model (Model 7). The figures highlight the degree of within field variability at both locations. For Hanska and Morgan, estimated check strip yields ranged from 2.7 to 8.2 t ha^{-1} and 3.6 to 9.8 t ha^{-1} with an average of 6.2 and 6.3 t ha^{-1}. The optimal nitrogen rates ranged from 97 to 202 kg ha^{-1} for Hanska with an average of 154 kg ha^{-1}. These rates correspond to yields ranging from 6.5 to 11.2 t ha^{-1} with an average of 9.4 t ha^{-1}. For Morgan, the optimal application rates ranged from 109 to 202 kg ha^{-1} with an average of 184 kg ha^{-1}. Corresponding yields ranged from 8.3 to 12.9 t ha^{-1} with an average of 10.7 t ha^{-1}. The increase in return when compared to the UMN rate ranged from 0.0 to 176.1 \$ ha^{-1} for Hanska and 0.0 to 274.2 \$ ha^{-1} for Morgan. The standard deviation of this increased return ranged from 0.3 to 76.3 \$ ha^{-1} for Hanska and 0.2 to 78.5 \$ ha^{-1} for Morgan.

The results are concluded by discussing the sensitivity of the potential value of VRA to the price of corn and nitrogen for model 7. Holding the price of nitrogen constant at 0.374 \$ kg^{-1} (0.17 \$ lbs^{-1}) and letting the price of corn increase from 78.57 \$ t^{-1} (2.00 \$ bu^{-1}) to 117.86 \$ t^{-1} (3.00 \$ bu^{-1}), the potential value of VRA compared to the UMN recommended rate increases linearly from 20.75 to 34.54 \$ ha^{-1} for Hanska and from 49.50 to 82.33 \$ ha^{-1} for Morgan, the percentage of this value attributable to using the

optimal uniform rate increases from 5 to 13 for Hanska and 64 to 72 for Morgan. Therefore, the importance of getting the right average rate for a field increases with an increase in the price of corn.

Holding the price of corn constant at 98.21 \$ t^{-1} (2.50 \$ bu^{-1}) and letting the price of nitrogen increase from 0.15 \$ kg^{-1} (0.07 \$ lbs^{-1}) to 0.59 \$ kg^{-1} (0.27 \$ lbs^{-1}), the potential value of VRA compared to the UMN recommended rate decreases linearly from 32.52 to 24.27 \$ ha^{-1} for Hanska and from 75.75 to 57.09 \$ ha^{-1} for Morgan, the percentage of this value attributable to using the optimal uniform rate decreases from 23 to 1 for Hanska and 78 to 56 for Morgan. Therefore, the importance of varying the right average rate optimally within a field increases with an increase in the price of nitrogen.

Summary and Conclusions

Confirming the PA hypothesis for VRA has proven challenging. To confront this challenge, researchers are using increasingly sophisticated statistical models to estimate and compare SSCRFs. While progress has been made, it has been hampered by the lack of a clear conceptual framework to guide and motivate the development of appropriate statistical models. The purpose of this paper was to propose such a framework. The framework was used to identify a testable hypothesis and develop a statistical model to evaluate that hypothesis. The model was then applied to 1995 data from two fields in South Central Minnesota.

Effort to improve models for testing the PA hypothesis has focused on spatial correlation. Recently however, problems with site-specific and treatment dependent heteroscedasticity and correlation have been identified. Our conceptual framework shows why this is not a surprise and our results show this is not the end of the story for data from a common experimental design. We also find important strip heteroscedasticity and correlation. Failing to account for strip effects resulted in estimates of the potential value of VRA that were too high and confidence intervals that convey a false sense of precision because they were too narrow.

The conceptual and empirical models we developed are most applicable to a single year of data. Though, the models could be extended to multiple years. An important consideration for a multi-year extension of the model is the need to differentiate between unmanaged inputs that are temporally stable (e.g. topography and soil type) and those that are not (e.g. rainfall and temperature) (Bullock et al., 2002).

Within the context of our conceptual model, one could include two rather than one vector of unmanaged inputs: one that varies with time and one that does not. Empirically, additional parameters would have to be estimated for the fixed effect of time invariant unmanaged inputs and random effect of time variant unmanaged inputs.

The conceptual model points to the importance of site and treatment dependent heteroscedasticity and spatial correlation. These results are generally applicable to any field experiment where soils, rainfall, and other important agronomic factors other than the treatment may vary substantially across the experimental plot. The important strip effects found in our analysis are specific to complete randomized block design experiments that divide treatment strips within a block into multiple observations. Experimental designs that randomize more completely eliminate this complication.

Due to the computational intensity of the model and scope of our objectives, we did not systematically explore a wide variety of assumptions regarding the structure of spatial correlation and heteroscedasticity. Specifically, we focused on a multiplicative form of heteroscedasticity and Guassian spatial correlation. Alternatively, one could explore other forms of heteroscedasticity. With increasing computer power and new experiments with more observations per site, estimating the most general form of heteroscedasticity in our empirical model may soon be practical. There are also a wide variety of both isotropic and anisotropic models of spatial correlation that could be explored in future work.

The range of treatments employed in our experiments was well suited for the Hanska location, but not for the Morgan location, which is why we see a greater divergence between the estimated models using the Morgan data. It is also why we had to constrain our estimates of the optimal nitrogen rates for many of the sites at Morgan; therefore, the estimated potential value of VRA is likely downward biased.

Finally, our analysis of the potential value of VRA does not include implementation costs. These costs will vary depending on how this potential is tapped (e.g. the information used to guide applications and size of management units). A farmer who uses soil nitrate testing to tap this potential may have lower implementation costs than a farmer who runs controlled field experiments; however, controlled field experiments may provide better information. While demonstrating the potential of VRA under varied field conditions is important, more effort could be devoted to finding better ways to tap this potential.

References

Bongiovanni, R. and J. Lowenberg-DeBoer. 2000. Nitrogen Management In Corn Using Site-Specific Crop Response Estimates From a Spatial Regression Model. *In* P.C. Robert et al. (eds.) Precision Agriculture: Proc. Int. Conf. 5th, Minneapolis, MN. 16-19 July 2000. ASA, CSSA, and SSSA, Madison, WI.

Bongiovanni, R. and J. Lowenberg-DeBoer. 2001. Precision Agriculture: Economics of Nitrogen Management in Corn Using Site-Specific Crop Response Estimates From a Spatial Regression Model. Selected Paper: AAEA Annual Meeting, Chicago, IL, August 6, 2001. http://agecon.lib.umn.edu/.

Bullock, D.S., J. Lowenberg-Deboer, and S.M. Swinton. 2002. Adding Value to Spatially Managed Inputs by Understanding Site-Specific Yield Response. Agric. Econ. 27:233-245.

Casella, G., and R. L. Berger. 1990. Statistical Inference, Duxbury Press, Belmont, CA.

Davis, J.G., G.L. Malzer, P.J. Copeland, J.A. Lamb, P.C. Robert, and T.W. Bruulsema. 1996. Using Yield Variability to Characterize Spatial Crop Response to Applied N. *In* P.C. Robert et al. (eds.) Precision Agriculture: Proc. Int. Conf. 3rd, Minneapolis, MN. ASA, CSSA, and SSSA, Madison, WI, pp. 513-519.

Dikici, H. 2000. Seasonal Nitrogen Availability: A Site Specific Approach. Ph.D Thesis, University of Minnesota, St. Paul, MN.

Hernandez, J.A. and D.J. Mulla. 2002. Comparing Statistical Analysis for Landscape Scale Experimental Designs. *In* P.C. Robert et al. (eds.) Precision Agriculture: Proc. 6th Int. Conf., Minneapolis, MN. 14-17 July 2002. ASA, CSSA, and SSSA, Madison, WI.

Hurley, T.M., G. Malzer, and B. Killian. 2002a. A test of within field variation of corn response to nitrogen in central Minnesota. *In* A. Werner and A. Jarfe (eds.) Precision Agriculture: Herausforderung an Integrative Forschung, Entwicklung und Anwendung in der Praxis. Bonn, Germany 16-17 March 2002. Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL), Bartningstraße 49, 64289 Darmstadt. pp. 413-421.

Hurley, T.M., G. Malzer, and B. Killian. 2002b. A Test Of Within Field Variation For Corn Response To Nitrogen In Central Minnesota. *In* P.C. Robert et al. (eds.) Precision Agriculture: Proc. 6th Int. Conf., Minneapolis, MN. 14-17 July 2002. ASA, CSSA, and SSSA, Madison, WI.

Lambert, D., R. Bongiovanni, and J. Lowenberg-DeBoer. 2002. Spatial Regression, an Alternative Statistical Analysis for Landscape Scale On-Farm Trials: Case Study of Soil Density Trials in Central Illinois. *In* P.C. Robert et al. (eds.) Precision Agriculture: Proc. 6th Int. Conf., Minneapolis, MN. 14-17 July 2002. ASA, CSSA, and SSSA, Madison, WI.

Malzer, G.L., P.J. Copeland, J.G. Davis, J.A. Lamb, P.C. Robert, and T.W. Bruulsema. 1996. Spatial Variability of Profitability in Site-Specific N Management. *In* P.C. Robert et al. (eds.) Precision Agriculture: Proc. Int. Conf. 3rd, Minneapolis, MN. ASA, CSSA, and SSSA, Madison, WI, pp. 967-975.

Mamo, M., G.L. Malzer, D.J. Mulla, D.R. Huggins, and J. Strock. 2003. Spatial and Temporal Variation in Economically Optimum Nitrogen Rate for Corn. Agron. J. 95:958-964.

McBratney, A.B. and R. Webster. 1986. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. J. Soil Sci. 37:617-639.

Schabenberger, O. and F.J. Pierce. 2002. Contemporary Statistical Models for the Plant and Soil Sciences. CRC Press, Boca Raton, FL.

Figure 1: Partition of experimental plot into 6 and 48 sites.

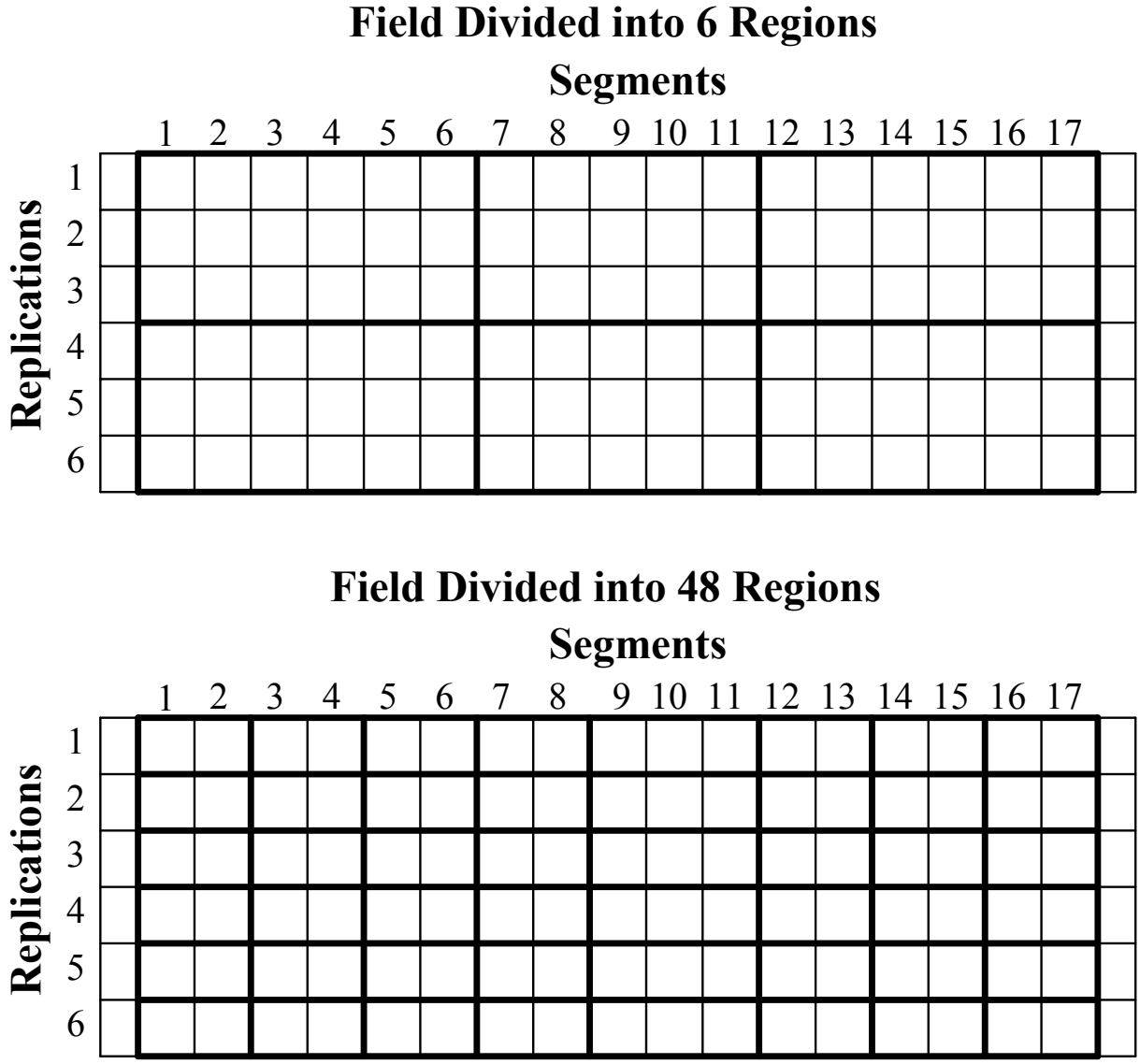


Figure 2: Estimates of the potential value of precision agriculture (exclusive of implementation costs) for switching from the University recommendation to the optimal uniform nitrogen application rate, the optimal uniform to the optimal variable rate, and the University recommendation to the optimal variable rate.

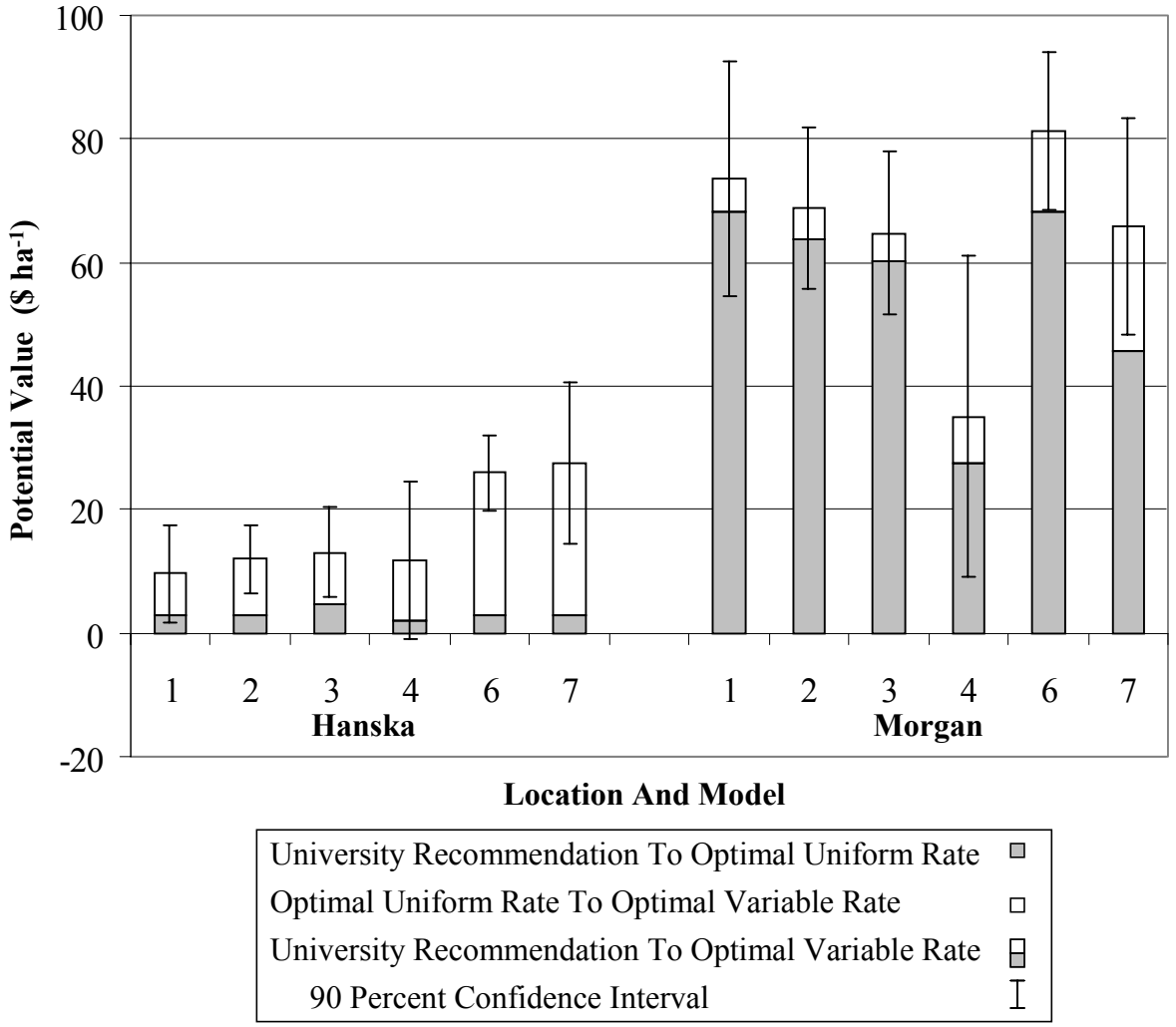
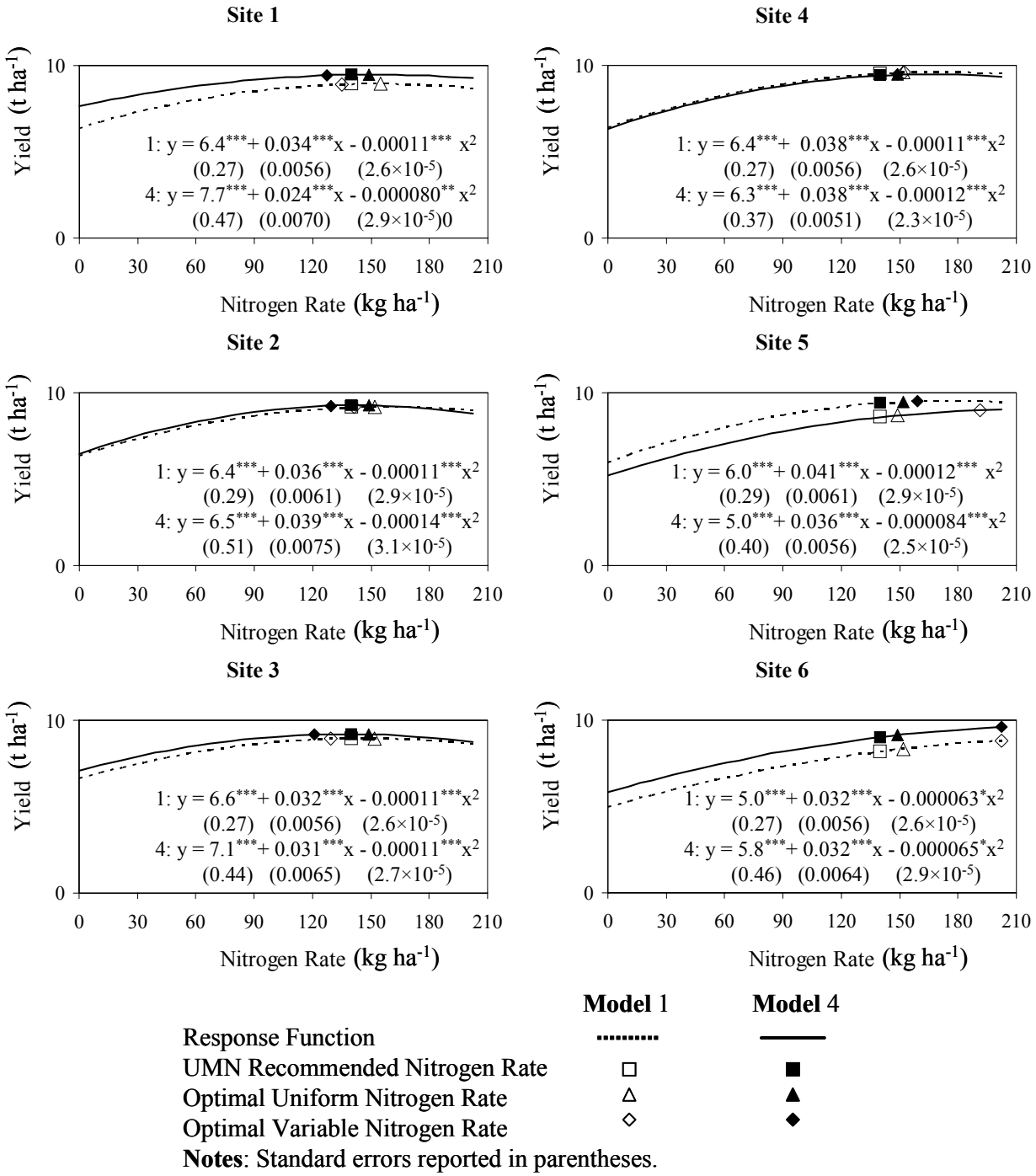
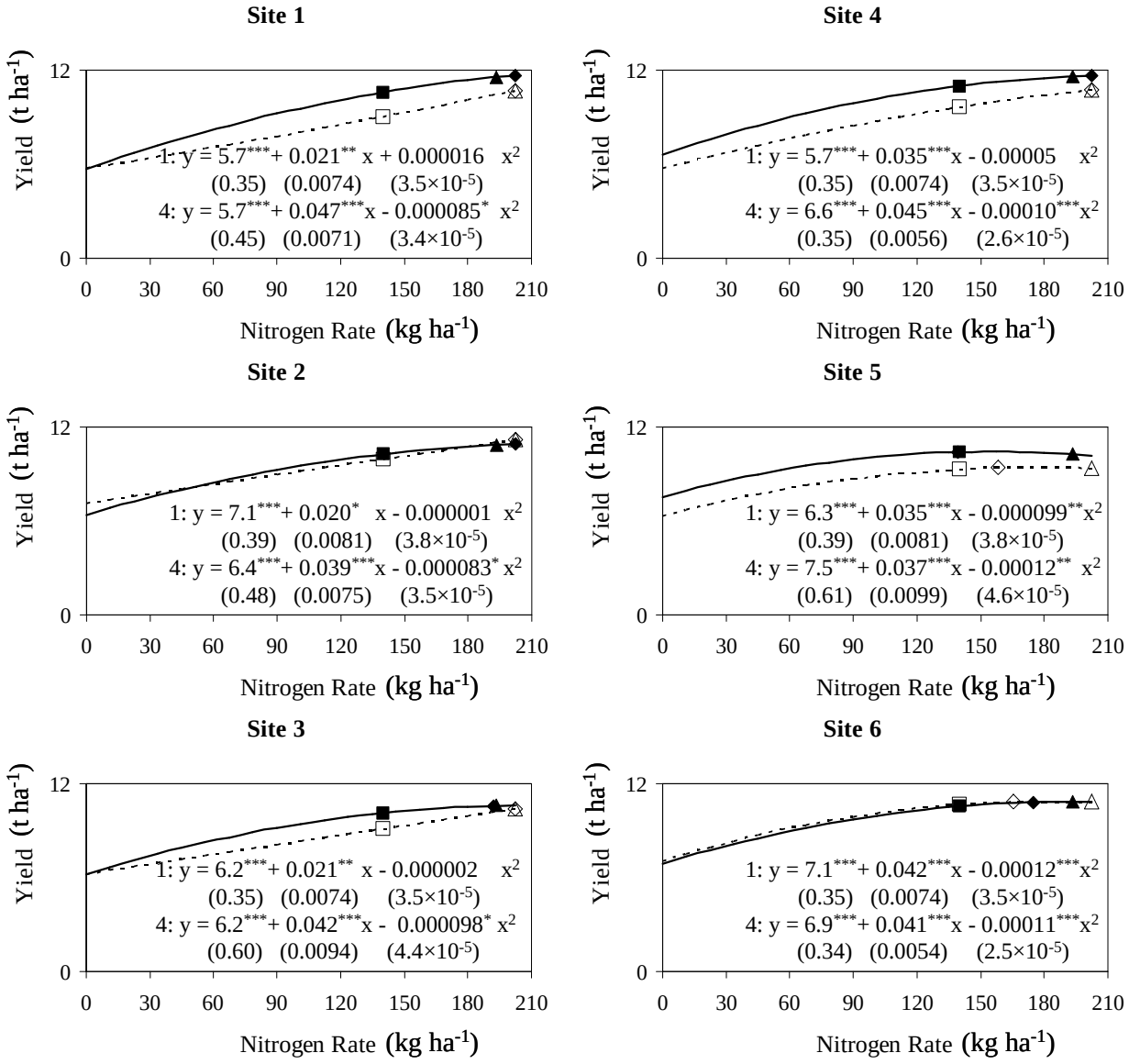


Figure 3: Hanska crop response function estimates, and University of Minnesota (UMN) recommended, optimal uniform, and optimal variable nitrogen rates for models 1 and 4.



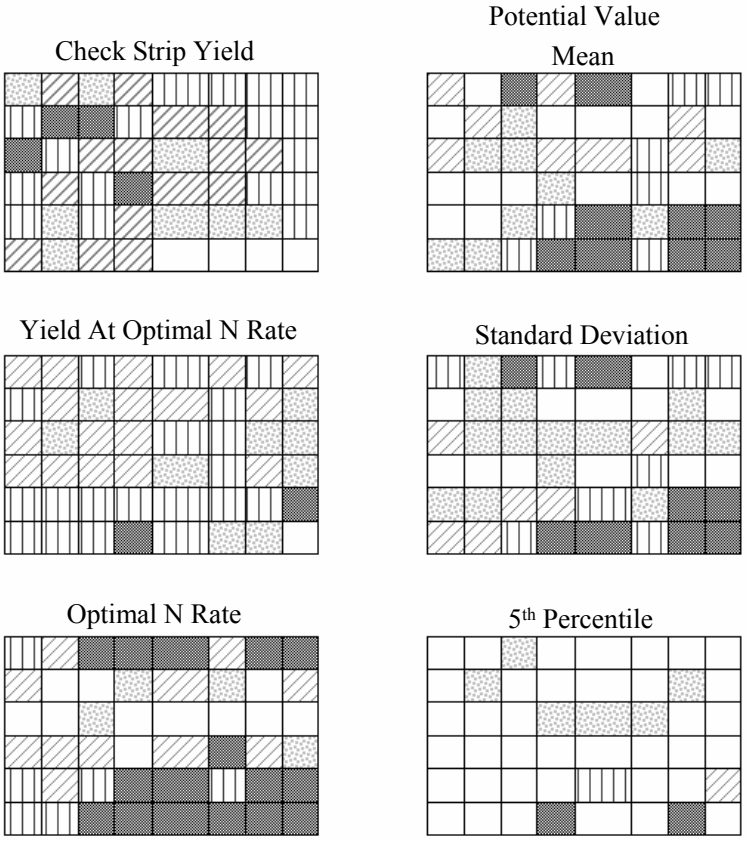
Notes: Standard errors reported in parentheses.

Figure 4: Morgan crop response function estimates, and University of Minnesota (UMN) recommended, optimal uniform, and optimal variable nitrogen rates for models 1 and 4.



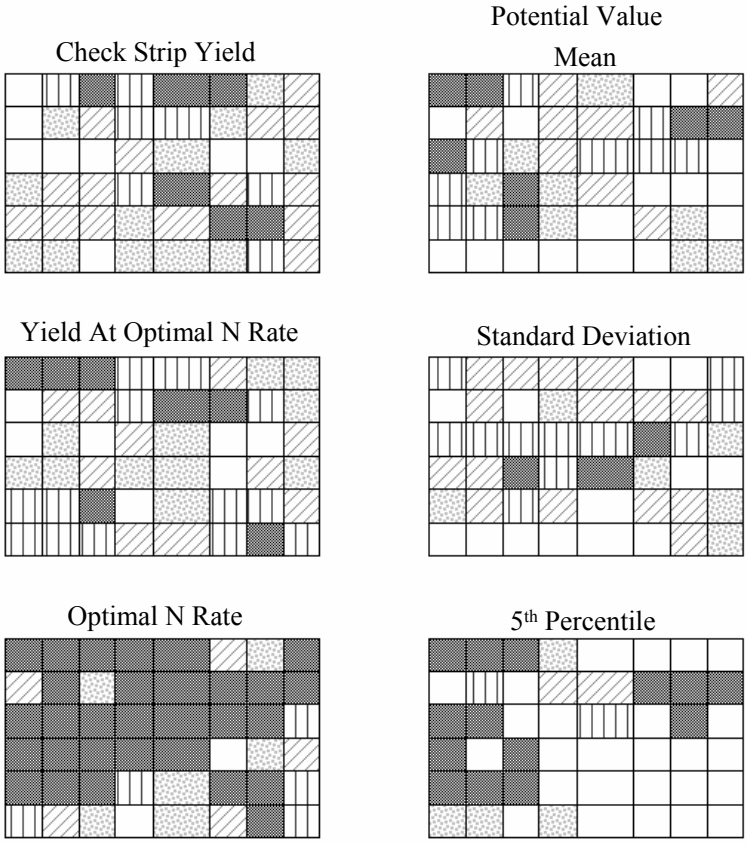
Notes: Standard errors reported in parentheses.

Figure 5: Check strip yield, yield at optimal nitrogen rate, optimal nitrogen rate, and potential value (exclusive of implementation costs) of switching to the optimal variable rate from the University of Minnesota recommended rated by site for Model 7 at Hanska.



	Check Strip Yield *t ha^{-1}*	Yield At Optimal N Rate *t ha^{-1}*	Optimal N Rate *kg ha^{-1}*	Increased Return: Mean *$ ha^{-1}*	Increased Return: Standard Deviation *$ ha^{-1}*	Increased Return: 5th Percentile *$ ha^{-1}*
	2.7 - 4.3	6.5 – 6.5	97 – 115	0.0 – 2.5	0.3 – 2.8	< 0
	4.3 - 5.7	6.5 – 9.0	115 – 127	2.5 – 6.7	2.8 – 7.9	0 – 10
	5.7 - 6.4	9.0 – 9.5	127 – 144	6.7 – 15.9	7.9 – 23.5	10 – 20
	6.4 - 7.4	9.5 – 10.2	144 – 170	15.9 – 48.3	23.5 – 26.0	20 – 30
	7.4 - 8.2	10.2 – 11.2	170 – 202	48.3 – 176.1	46.0 – 76.3	> 30
Average	6.2	9.4	154			

Figure 6: Check strip yield, yield at optimal nitrogen rate, optimal nitrogen rate, and potential value (exclusive of implementation costs) of switching to the optimal variable rate from the University of Minnesota recommended rated by site for Model 7 at Morgan.



	Check Strip Yield	Yield At Optimal N Rate	Optimal N Rate	Increased Return Mean	Increased Return Standard Deviation	Increased Return 5th Percentile
	t ha^{-1}	*t ha^{-1}*	*kg ha^{-1}*	*$ ha^{-1}*	*$ ha^{-1}*	*$ ha^{-1}*
	3.6 - 4.7	8.3 – 9.5	109 – 124	0.0 – 16.7	0.2 – 10.7	< 0
	4.7 - 5.7	9.5 – 10.3	124 – 147	16.7 – 40.1	10.7 – 24.9	0 – 10
	5.7 - 6.8	10.3 – 10.9	147 – 161	40.1 – 76.4	24.9 – 44.0	10 – 20
	6.8 - 8.0	10.9 – 11.8	161 – 186	76.4 – 151.7	44.0 – 59.3	20 – 30
	8.0 - 9.8	11.8 – 12.9	186 – 202	151.7 – 274.2	59.3 – 78.5	> 30
Average	6.3	10.7	184			

Table 1: Description of different models with restrictions reported in parentheses.

		Assumptions		
Model	**Sites**	**Site Specific Nitrogen Response**	**Correlation**	**Heteroscedasticity**
1	6	Yes	None ($C_1 = C_x = C_s = 0$)	None ($\sigma_{r_i s_i}^2 = \sigma^2$ for all r_i and s_i)
2	6	Yes	Guassian Spatial ($C_x = C_s = 0$)	None ($\sigma_{r_i s_i}^2 = \sigma^2$ for all r_i and s_i)
3	6	Yes	Guassian Spatial and Treatment ($C_s = 0$)	Site and Treatment ($\sigma_{r_i s_i}^2 = \sigma_{r_i x_i}^2$ for all r_i and s_i)
4	6	Yes	Guassian Spatial, Treatment, and Strip	Site, Treatment, and Strip
5	6	No ($\alpha_{k_x r_i} = \alpha_{k_x r_j}$ for all $k_x > 0$, r_i, and r_j)	Guassian Spatial, Treatment, and Strip	Site, Treatment, and Strip
6	48	Yes	None ($C_1 = C_x = C_s = 0$)	None ($\sigma_{r_i s_i}^2 = \sigma^2$ for all r_i and s_i)
7	48	Yes	Guassian Spatial, Treatment, and Strip	Site, Treatment, and Strip
8	48	No ($\alpha_{k_x r_i} = \alpha_{k_x r_j}$ for all $k_x > 0$, r_i, and r_j)	Guassian Spatial, Treatment, and Strip	Site, Treatment, and Strip

Table 2: Maximized log-likelihood and model comparisons.

Model	Location	
	Hanska	**Morgan**
	Maximized Log-Likelihood (*Estimated Parameters*)	
1	-2653.71 (19)	-2825.80 (19)
2	-2493.95 (21)	-2672.83 (21)
3	-2432.59 (32)	-2639.95 (32)
4	-2322.91 (58)	-2527.79 (58)
5	-2336.55 (48)	-2548.26 (48)
6	-2408.05 (145)	-2590.80 (145)
7	-2145.04 (226)	-2340.72 (226)
8	-2240.18 (132)	-2478.64 (132)
Model Comparisons (Restricted vs. Unrestricted)		
	Likelihood Ratio Statistic (*Degrees of Freedom*)	
1 vs. 2	319.5^{***} (2)	305.9^{***} (2)
2 vs. 3	122.7^{***} (11)	65.8^{***} (11)
3 vs. 4	219.3^{***} (26)	224.3^{***} (26)
1 vs. 4	661.6^{***} (39)	596.0^{***} (39)
6 vs. 7	526.0^{***} (81)	500.2^{***} (81)
5 vs. 4	27.3^{**} (10)	40.9^{***} (10)
8 vs. 7	190.3^{***} (94)	275.8^{***} (94)
4 vs. 7	355.7^{***} (168)	374.1^{***} (168)

Table 3: Semi-variance parameters and average standard deviation.

	Model					
Parameter	**1**	**2**	**3**	**4**	**6**	**7**
Hanska						
Nugget ($1 - C_1 - C_x - C_s$)	1.00	0.37	0.40	0.26	1.00	0.33
Spatial (C_1)	-	0.63	0.60	0.58	-	0.52
Treatment (C_x)	-	-	0.00	0.00	-	0.00
Strip (C_s)	-	-	-	0.16	-	0.15
Range (a m)	-	27.2	25.5	27.1	-	21.3
Average Standard Deviation†	1.16	1.25	1.20	1.23	0.78	0.89
Morgan						
Nugget ($1 - C_1 - C_x - C_s$)	1.00	0.45	0.41	0.28	1.00	0.31
Spatial (C_1)	-	0.55	0.59	0.57	-	0.51
Treatment (C_x)	-	-	0.00	0.00	-	0.00
Strip (C_s)	-	-	-	0.15	-	0.18
Range (a m)	-	26.3	30.3	30.4	-	16.1
Average Standard Deviation†	1.54	1.55	1.62	1.68	1.05	1.21

†Calculated as $\sqrt{\frac{\sum_{i=1}^{N} {\sigma_{r_i s_i}}^2}{N}}$ t ha^{-1} where N is the number of observations and ${\sigma_{r_i s_i}}^2$ is the estimated standard deviation for observation i by site and treatment strip.