Journal of Agricultural and Resource Economics 30(2):231–249

Estimating the Potential Value of Variable Rate Nitrogen Applications: A Comparison of Spatial Econometric and Geostatistical Models

Terrance M. Hurley, Kikuo Oishi, and Gary L. Malzer

Site-specific crop response functions (SSCRFs) are useful for estimating the value of variable rate nitrogen applications (VRA), but appropriate statistical models are necessary. Problems estimating SSCRFs using experimental field data include region, spatial, treatment, and strip dependent heteroskedasticity and correlation. We develop a spatial autoregressive error (SARE) model for dealing with these problems and compare results with previous analysis based on a geostatistical (GEO) model. VRA value estimates for the two models differ notably for 1995 data from Southern Minnesota. Furthermore, findings show that the results of a comparison of model performance are location specific.

Key words: geostatistics, precision agriculture, site-specific crop response functions, spatial autoregressive error, variable rate nitrogen application

Introduction

Precision agriculture (PA) uses site- (or region-) specific information to improve management. The PA hypothesis asserts that varying production inputs between or within fields can benefit farmers or the environment. Crop response to production inputs can vary between or within fields due to variation in conditions such as soil type, organic matter, topography, and drainage. If field conditions interact with production inputs, varying these inputs in response to site-specific information on varied field conditions can increase input productivity.

Recent attempts to test the PA hypothesis for variable rate nitrogen applications (VRA) on corn have used multiple regression analysis to estimate site-specific crop response functions (SSCRFs) based on data from randomized complete block design field experiments (e.g., Davis et al., 1996; Malzer et al., 1996; Bongiovanni and Lowenberg-DeBoer, 2000, 2001; Lambert, Bongiovanni, and Lowenberg-DeBoer, 2002; and Hurley, Malzer, and Kilian, 2002a,b, 2004). Early work relied on ordinary least squares (OLS), the simplest form of multiple regression analysis. OLS assumes regression errors are homoskedastic and uncorrelated—an assumption soundly rejected by later work. Thus, OLS produces inefficient parameter estimates and may over- or understate statistical significance (Schabenberger and Pierce, 2002).

Terrance M. Hurley is associate professor and Kikuo Oishi is research assistant, both in the Department of Applied Economics, University of Minnesota. Gary L. Malzer is professor, Department of Soil, Water, and Climate, University of Minnesota. The authors gratefully acknowledge the financial research assistance support provided by the Minnesota Corn Research and Promotion Council.

Review coordinated by David Aadland.

The first problem identified with OLS was spatial correlation. More recently, problems with region, treatment, and strip dependent heteroskedasticity and correlation have been identified. Lambert, Bongiovanni, and Lowenberg-DeBoer (2002), hereafter referred to as LBL, use a geostatistical model (Cressie, 1993) to deal with spatial correlation. They also use a spatial econometric model (Anselin, 1988) to deal with spatial correlation and region heteroskedasticity. A comparison of the geostatistical and spatial econometric models favored the latter for field data from Argentina. However, the geostatistical model did not address region heteroskedasticity. Neither model examined treatment and strip heteroskedasticity and correlation.

Hurley, Malzer, and Kilian (2004), denoted here as HMK, present a conceptual framework to explain the presence of region, spatial, and treatment dependent heteroskedasticity and correlation in SSCRF estimates. Their paper also argues that the presence of strip heteroskedasticity and correlation is to be expected in common randomized complete block design field experiments because treatments are randomized between, but not within strips. These insights are used to motivate a geostatistical model with heteroskedasticity, and fixed and random effects. The authors confirm the presence of region, spatial, treatment, and strip dependent heteroskedasticity and correlation using Minnesota field data, and then estimate the potential value of VRA. However, given the findings reported by LBL, it is natural to wonder whether the geostatistical framework used by HMK is a good choice.

The purpose of this study is to develop and present a spatial econometric model for estimating SSCRFs that incorporates region, treatment, and strip dependent heteroskedasticity and correlation. The comparison of the geostatistical and spatial econometric models is then revisited by comparing results from this new spatial econometric model to the geostatistical model reported by HMK.

Methods

Conceptual Framework

Following HMK, define crop yield, y, as depending on two types of inputs: $y = f(x, z)$, where x represents variable inputs and z represents fixed inputs. Variable inputs can be thought of as a farmer's managed inputs (e.g., nitrogen and pesticides). Fixed inputs can be thought of as inputs that influence yield, but are not actively managed by a farmer (e.g., soil type, rainfall, and topography). For expositional brevity, assume x and z are scalars.

The precision agriculture (PA) hypothesis implies farmers or the environment can benefit from varying the amount of variable input between or within fields in response to between- or within-field variation in the availability of a fixed input. For example, to maximize the net return to a variable input, a farmer should set the value of marginal product equal to its marginal cost; ignoring well-known caveats,

$$p_y\left(\frac{\partial f(x^*, z)}{\partial x}\right) = p_x,$$

where p_y and p_x are the price of crop yield and the variable input, and x^* is the optimal level of variable input. The implicit function theorem implies this optimal level of variable input varies in response to the fixed input when variable and fixed inputs interact:

$$\frac{\partial x^*}{\partial z} = -\frac{\partial^2 f(x^*, z)}{\partial x \partial z} \frac{\partial x^2}{\partial^2 f(x^*, z)} \neq 0$$

when

$$\frac{\partial^2 f(x^*, z)}{\partial x \partial z} \neq 0.$$

For instance, if soil type varies within a field, a farmer should adjust inputs like nitrogen within the field only if there is some type of interaction between soil type and nitrogen.

If $f(x, z)$ is continuously differentiable in x and z, a Taylor series expansion yields:

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

and

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

where

$$\beta_{k_x k_z} = \frac{\partial^{k_x + k_z} f(0,0)}{k_x! k_z! \partial x^{k_x} \partial z^{k_z}} \; \forall \; k_x \text{ and } k_z$$

are real constants that indicate how variable and fixed inputs combine to influence yield, and $f(0, 0)$ is yield with no variable or fixed inputs. Equation (2) suggests the null hypothesis $\beta_{k_x k_z} = 0 \; \forall \; k_x > 0$ and $k_z > 0$, which implies PA cannot be used to the benefit of a farmer or the environment because there is no interaction or synergy between the variable and fixed inputs.

HMK show how to test for this interaction by estimating SSCRFs using data from a randomized complete block design experiment. To review their argument, consider a set of data (y_i, x_i, z_i, e_i) for $i = 1, 2, ..., N$ collected from such an experiment. An individual data point consists of y_i, an observed yield; x_i, an observed variable input; z_i, an unobserved fixed input; and e_i, an error due to, for example, imperfect measurement and approximation. Furthermore, suppose these data are pooled for R distinct regions within a field such that $r_i \in \{1, ..., R\}$ represents the region assigned to the ith observation. Rewrite equation (1) as:

$$y_i = \alpha_{0 r_i} + \alpha_{1 r_i} x_i + \alpha_{2 r_i} x_i^2 + \xi_i, \tag{3}$$

where

$$\xi_i = \sum_{k_z=1}^{\infty} \left(\beta_{0 k_z} + \beta_{1 k_z} x_i + \beta_{2 k_z} x_i^2 \right) \left(z_i^{k_z} - Z_{k_z r_i} \right) + e_i$$

is the regression error; $Z_{k_z r_i} \; \forall \; k_z$ and r_i are unobserved constants; and

$$\alpha_{k_x r_i} = \beta_{k_x 0} + \sum_{k_z=1}^{\infty} \beta_{k_x k_z} Z_{k_z r_i}$$

are estimable parameters. Under the assumption that the expected value of the regression error is zero, $Z_{k_z r_i}$ reflects the expectation of $z_i^{k_z}$ given $i \in r_i$. Now if the PA null hypothesis is true and $\beta_{k_x k_z} = 0 \; \forall \; k_x \in \{1, 2\}$ and $k_z > 0$, then $\alpha_{k_x r_i} = \alpha_{k_x r_j} \; \forall \; k_x \in \{1, 2\}, r_i$, and

r_j. Therefore, if the linear or quadratic effect of a variable input on yield differs between regions, it should be possible for a farmer or the environment to benefit from varying the level of that input between regions.

It is important to note that the unobserved fixed input has two distinct effects in equation (3), a fixed effect and a random effect. Each α parameter consists of two components. One component, $\beta_{k_x 0}$, is identical for all regions in the field. The other component,

$$\sum_{k_z=1}^{\infty} \beta_{k_x k_z} Z_{k_z r_i},$$

is a region-specific constant, also known as a fixed effect in statistical vernacular. The error also consists of two components. One component, e_i, does not depend on the amount of fixed input. The other component,

$$\sum_{k_z=1}^{\infty} \left(\beta_{0k_z} + \beta_{1k_z} x_i + \beta_{2k_z} x_i^2\right)\left(z_i^{k_z} - Z_{k_z r_i}\right),$$

depends on the unobserved amount of fixed input and how it varies from the regional expectation, $z_i^{k_z} - Z_{k_z r_i}$, which is known as a random effect in statistical vernacular. Note that this random effect is also influenced by the amount of variable input.

The challenge to testing the PA hypothesis by estimating equation (3) comes from the structure of the error covariance:

$$\begin{aligned}(4)\qquad \mathrm{E}(\xi_i \xi_j) = {} & \sum_{k_z=1}^{\infty}\sum_{k_z'=1}^{\infty} \varepsilon(x_i, k_z)\varepsilon(x_j, k_z')\mathrm{E}\left[\left(z_i^{k_z} - Z_{k_z r_i}\right)\left(z_j^{k_z'} - Z_{k_z' r_j}\right)\right] \\ & + \sum_{k_z'=1}^{\infty} \varepsilon(x_j, k_z')\mathrm{E}\left[\left(z_j^{k_z'} - Z_{k_z' r_j}\right)e_i\right] \\ & + \sum_{k_z=1}^{\infty} \varepsilon(x_i, k_z)\mathrm{E}\left[\left(z_i^{k_z} - Z_{k_z r_i}\right)e_j\right] + \mathrm{E}[e_i e_j],\end{aligned}$$

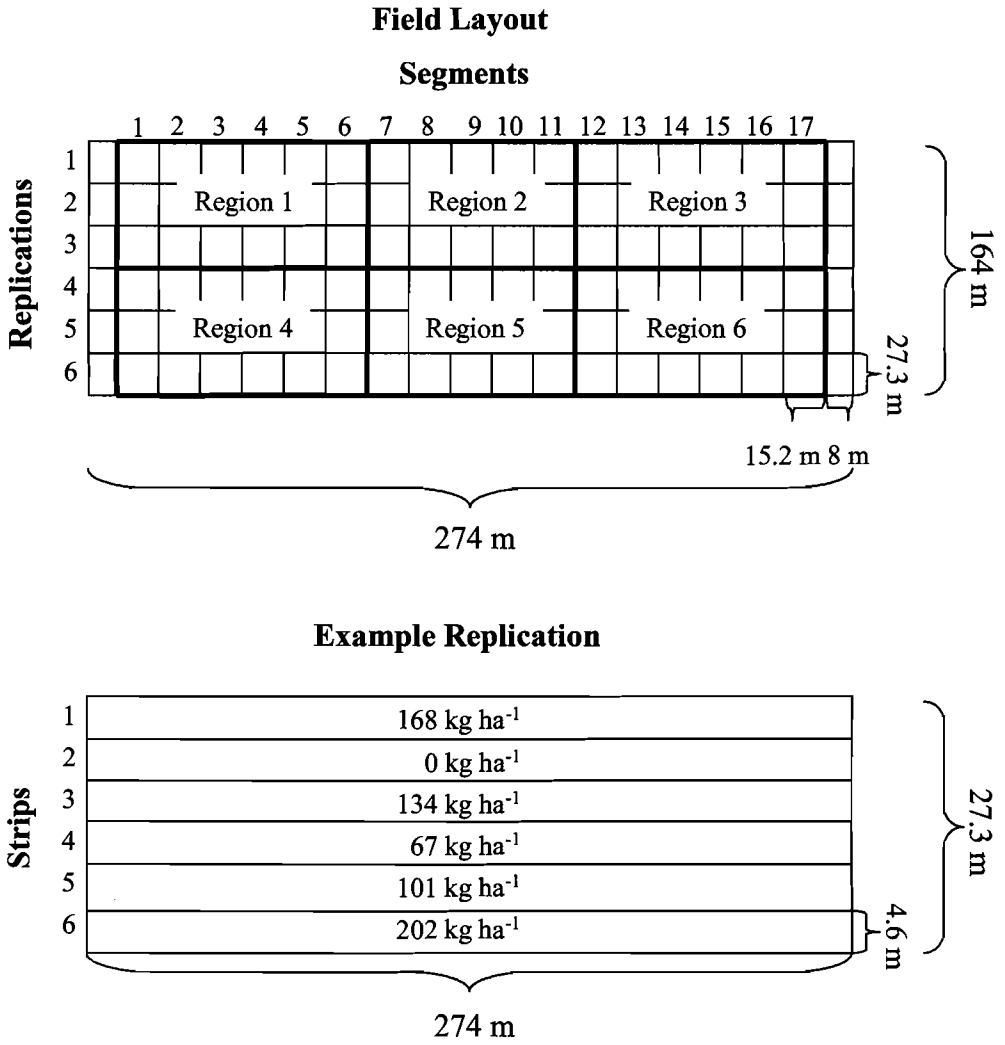
where $\varepsilon(x, k) = \beta_{0k} + \beta_{2k} x + \beta_{2k} x^2$. Spatial correlation in equation (4) can result from the random effect due to unobserved fixed inputs,

$$\mathrm{E}\left[\left(z_i^{k_z} - Z_{k_z r_i}\right)\left(z_j^{k_z'} - Z_{k_z' r_j}\right)\right],$$

when these fixed inputs are spatially correlated. Region and treatment dependent heteroskedasticity and correlation are likely because x_i and x_j vary by treatment, and $Z_{k_z r_i}$ and $Z_{k_z r_j}$ vary by region.

Experimental Methods

The data used to estimate SSCRFs and test the PA hypothesis were collected in 1995 from two production field experiments in the towns of Hanska (Brown County) and Morgan (Redwood County), Minnesota. Each location was 164 m wide and 274 m long (4.5 ha). Within this area, six replications of six treatments were established in a randomized complete block design (figure 1). The six replications ran the length of the field. The six randomized treatments within each replication also ran the length of the



Note: For each replication, treatments are randomly assigned to a strip. The example of treatment assignments above is taken from the first replication.

Figure 1. Illustration of complete randomized block experimental design

field in 36 (6 replications × 6 treatments) 4.6 m wide strips that included nitrogen rates of 0, 67, 101, 134, 168, and 202 kg/ha applied as anhydrous ammonia. Treatments were applied on November 4, 1994, using a radar-controlled variable rate applicator to compensate for variations in applicator speed and ensure a constant application rate within each strip.

Corn (*Zea mays* L.; cv. Pioneer® 3531) was planted during the first week of May in 0.76 m rows at approximately 76,500 seeds/ha. 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 seventeen 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 per field. Subsamples of grain were collected from each area of yield measurement to determine moisture and adjust yields to reflect 15.5% moisture. Dikici (2000) reports additional details and a descriptive summary of the data.

This common type of randomized complete block experimental design raises another issue of concern not captured in equation (4). While nitrogen treatments were randomized across the width of the experimental plot, they were not randomized down the length. Treatments were randomized between, but not within strips. This lack of randomization in one direction can result in the type of strip heteroskedasticity and correlation identified in HMK.

Combined, the conceptual framework and experimental design suggest regression errors may exhibit region, spatial, treatment, and strip heteroskedasticity and correlation. Therefore, OLS will be inefficient. To obtain more efficient parameter estimates, a model with a richer error structure is necessary. While HMK showed how to extend the traditional GEO model to include region, treatment, and strip effects, LBL demonstrated that SARE type models tended to outperform most other classes of models including the GEO class. Hence, we focus our attention on (*a*) developing a SARE model that incorporates region, treatment, and strip effects; (*b*) demonstrating that this new SARE model outperforms previously used SARE models; and (*c*) comparing our new SARE model to the GEO model reported in HMK.

GEO Framework

HMK assume the covariance of the regression errors is:

$$\text{(5)}\qquad \mathrm{E}\left[(\xi_i \xi_j)\right] = \begin{cases} \sigma^2_{r_i s_i} & \text{for } i = j, \\ \sigma_{r_i s_i}\sigma_{r_j s_j}\left[C_1\left(1 - g_1(d_{ij}, a)\right) + C_x x_{ij} + C_s s_{ij}\right] & \text{for } i \neq j, \end{cases}$$

where $\sigma^2_{r_i s_i} > 0$ and $\sigma^2_{r_j s_j} > 0$ are the region- 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 x is the same for observations i and j, 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; $C_1 \geq 0$, $C_x \geq 0$, and $C_s \geq 0$ are spatial, treatment, and strip correlation parameters that assume positive correlation; $1 \geq g_1(d_{ij}, a) \geq 0$ is a permissible semi-variogram distance function (e.g., see McBratney and Webster, 1986); and a is a shape or range parameter for the semi-variogram distance function.

Dividing equation (5) 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: (*a*) spatial correlation ($C_1(1-g_1(d_{ij}, a))$, (*b*) treatment correlation ($C_x x_{ij}$), and (*c*) strip correlation ($C_s s_{ij}$). Since the correlation coefficient must always lie between 1.0 and –1.0, $1.0 \geq C_1 + C_x + C_s \geq 0.0$ assuming spatial, treatment, and strip correlation are positive, which helps ensure the covariance matrix satisfies the necessary regularity conditions (i.e., positive definite).

The classical geostatistical approach decomposes variation in the dependent variable into a trend, local variance (nugget), and distance effect. Equations (3) and (5) accomplish a similar decomposition, but add region, treatment, and strip effects. The trend is captured by $\alpha_{0r_i} + \alpha_{1r_i} x_i + \alpha_{2r_i} x_i^2$, which is region specific. The semi-variogram is:

$$\gamma_{ij} = \begin{cases} 0 & \text{for } d_{ij} = 0, \\ \sigma_{r_i s_i} \sigma_{r_j s_j} \left(C_0 + C_1 g(d_{ij}, a) + C_x(1 - x_{ij}) + C_s(1 - s_{ij}) \right) & \text{for } d_{ij} > 0, \end{cases} \tag{6}$$

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 for the *i*th and *j*th observations. Equation (6) shows explicitly how the standard geostatistical model was modified to include region, treatment, and strip effects.

SARE Framework

An alternative method for capturing spatial relationships is a spatial autoregressive error (SARE) model. Redefine equation (1) using matrix notation as:

$$\mathbf{y} = \mathbf{XA} + \xi, \tag{7}$$

where $\mathbf{y}$ is an $N \times 1$ vector of observed yields, $\mathbf{X}$ is an $N \times R(K_x + 1)$ matrix of region-specific variable inputs, $\mathbf{A}$ is an $R(K_x + 1) \times 1$ vector of regression coefficients, and ξ is an $N \times 1$ vector of regression errors. Write ξ in the heteroskedastic autoregressive form:

$$\Psi^{-0.5}\xi = \left[\lambda_1 \mathbf{W}_1 + \lambda_x \mathbf{W}_x + \lambda_s \mathbf{W}_s \right] \Psi^{-0.5}\xi + \varepsilon, \tag{8}$$

where Ψ is an $N \times N$ diagonal matrix with the *i*th diagonal element equal to a region and strip heteroskedasticity weight $\psi^2_{r_i s_i}$; $\mathbf{W}_1$, $\mathbf{W}_x$, and $\mathbf{W}_s$ are $N \times N$ exogenous weighting matrices; λ_1, λ_x, and λ_s are estimable spatial, treatment, and strip dependence parameters; and ε is an $N \times 1$ vector of independent standard normal errors such that $\mathrm{E}[\varepsilon] = \mathbf{0}$ and $\mathrm{E}[\varepsilon\varepsilon'] = \mathbf{I}$ (Anselin, 1988).[1] The covariance matrix is expressed as:

$$\mathrm{E}[\xi\xi'] = \Psi^{0.5} \Omega \Psi^{0.5}, \tag{9}$$

where

$$\Omega = \left[\mathbf{I} - \lambda_1 \mathbf{W}_1 - \lambda_x \mathbf{W}_x - \lambda_s \mathbf{W}_s \right]^{-1} \left[\mathbf{I} - \lambda_1 \mathbf{W}_1' - \lambda_x \mathbf{W}_x' - \lambda_s \mathbf{W}_s' \right]^{-1}.$$

---

[1] We also considered the spatial autoregressive lag model: $\mathbf{y} = \lambda_1 \mathbf{W}_1 \mathbf{y} + \mathbf{XA} + \varepsilon$. The maximized value of the log likelihood for this model is –2,503.71 for Hanska and –2,674.35 for Morgan. Both of these values are lower than the maximized log likelihoods for model 2 in table 1, which is the spatial autoregressive counterpart with the same number of parameters. These results are consistent with Lambert, Bongiovanni, and Lowenberg-DeBoer (2002), and suggest that the spatial autoregressive framework fits the data better than the spatial lag framework—which is why our efforts here focus on improving the traditional spatial autoregressive framework.

Equation (9) captures spatial correlation by defining the weighting matrix $\mathbf{W}_1$ in terms of spatial contiguity, so that λ_1 measures the intensity of spatial dependence. For example, a positive weight can be assigned for all columns in the ith row that correspond to observations immediately adjacent and diagonal to the ith observation (i.e., queen contiguity). Treatment correlation can be captured by defining the weighting matrix $\mathbf{W}_x$ in terms of treatment contiguity, so that λ_x measures the intensity of treatment dependence. For example, a positive weight was assigned for all columns in the ith row corresponding to observations that had the same treatment level as the ith observation. Strip correlation can be captured by defining the weighting matrix $\mathbf{W}_s$ in terms of strip contiguity, so that λ_s measures the intensity of strip dependence. For example, a positive weight was assigned for all columns in the ith row that correspond to observations from the same strip as the ith observation.

GEO versus SARE

Before describing the estimation procedure, a discussion of the similarities and differences between the GEO and SARE frameworks is warranted. Both frameworks control for large-scale spatial trends using regional fixed effects (i.e., different α coefficients are estimated for different regions within a field). Both models deal explicitly with heteroskedastic region, treatment, and strip random effects. Where the two models differ is how they treat these random effects. For the geostatistical model, $E[\xi\xi'] = \Sigma^{0.5}\mathbf{P}\Sigma^{0.5}$, where Σ is a diagonal matrix of variances and $\mathbf{P}$ is the correlation matrix. Recall that the SARE model has a similar decomposition, $E[\xi\xi'] = \Psi^{0.5}\Omega\Psi^{0.5}$. In general, however, Ψ will not equal the diagonal variance matrix and Ω will not equal the correlation matrix. Therefore, while the covariance matrix for the GEO model is multiplicatively separable in the variance and correlation parameters, the covariance of the SARE model is not. In the SARE model, λ_1, λ_x, and λ_s specify to some extent the degree of spatial, treatment, and strip heteroskedasticity, as well as the degree of spatial, treatment, and strip correlation.[2] Alternatively, in the GEO model, C_1, C_x, C_s, and a only specify the degree of spatial, treatment, and strip correlation, and not heteroskedasticity.

Estimation

Specifying SSCRFs as quadratic functions, which is consistent with both HMK and LBL, we estimate the SARE model defined by equations (3) and (9) with our data. Estimation can be accomplished in a variety of ways (Schabenberger and Pierce, 2002). The method used here profiles the α parameters by substituting the feasible generalized least squares (FGLS) estimator, and estimates the variance and covariance parameters (λ_1, λ_x, λ_s, and $\psi_{rs}^2 \; \forall \, r \in \{1, ..., R\}$ and $s \in \{1, ..., 36\}$) using maximum likelihood (ML). The likelihood function was optimized using the unconstrained optimization routine provided in Matlab's® optimization toolbox. Standard errors for the α parameters are calculated using FGLS and the maximum-likelihood estimates for the variance and covariance parameters. This estimation procedure is identical to HMK, so differences in the results should not be attributable to a difference in the estimation technique.

---

[2] Wall (2004) discusses and illustrates the inherent difficulties of interpreting dependence parameters in the SARE framework in relation to the correlation structure of the errors.

In an effort to maintain comparability, we employ as many of the same assumptions as possible. Because we are using the same data as HMK, we likewise use the same six regions of the field (see figure 1). These six regions pool observations from replications 1–3 and 4–6 and segments 1–5, 7–11, and 12–17. We also define $\psi^2_{rs} = \psi^2_r \psi^2_s$, and set $\psi^2_s = 1$ for $s \in \{1, 19\}$ for identification. LBL specify spatial correlation using queen contiguity with uniform weights that sum to one across rows, a procedure also adopted here. The weighting matrices for treatment and strip dependence likewise have uniform weights that sum to one across rows. HMK assume $C_1 \geq 0$, $C_x \geq 0$, and $C_s \geq 0$, implying positive correlation. We place different restrictions on λ_1, λ_x, and λ_s because they can imply more than just correlation. Instead, we restrict $|\lambda_1 + \lambda_x + \lambda_s| < 1$ and $\text{sign}(\lambda_1) = \text{sign}(\lambda_x) = \text{sign}(\lambda_s)$, which imply $\mathbf{I} - \lambda_1 \mathbf{W}_1 - \lambda_x \mathbf{W}_x - \lambda_s \mathbf{W}_s = \mathbf{I} - \lambda^o \mathbf{W}^o$ such that $|\lambda^o| < 1$ and $\mathbf{W}^o$ is a row randomized matrix.[3]

Under these conditions, Horn and Johnson (1993) show that the covariance matrix must be positive definite and that the implied correlation between observations must be smaller for observations which are further apart in terms of spatial, treatment, and strip contiguity. Note that $\lambda^o = \lambda_1 + \lambda_x + \lambda_s$ can now be interpreted as the magnitude and direction of autocorrelation, while λ_1/λ^o, λ_x/λ^o, and λ_s/λ^o can be interpreted as the proportion of this autocorrelation attributable to spatial, treatment, and strip dependence. It is also important to note these restrictions do not ensure the same positive correlations assumed by HMK.

Hypothesis Tests

The importance of incorporating region, treatment, and strip effects in the estimation of SSCRFs using a SARE model is evaluated by estimating a series of five nested models and comparing these models based on the likelihood-ratio statistic (LRS). The likelihood-ratio statistic is defined as $LRS = 2(\log L_U - \log L_R)$, where $\log L_U$ and $\log L_R$ are the maximized log likelihoods for the unrestricted and restricted models. The LRS is asymptotically distributed χ^2 with the degrees of freedom equal to the number of parameter restrictions.

Model 1 assumes $\psi^2_{r_i s_i} = \sigma^2 \; \forall \, i$, $\lambda_1 = 0$, $\lambda_s = 0$, and $\lambda_x = 0$, which is the ML analogy to OLS. Model 2 assumes $\psi^2_{r_i s_i} = \psi^2 \; \forall \, i$, $\lambda_s = 0$, and $\lambda_x = 0$, which is the same specification as the standard SARE model used by LBL and analogous to HMK's second model. Model 3 assumes $\psi^2_{r_i s_i} = \psi^2_{r_i} \; \forall \, i$, $\lambda_s = 0$, and $\lambda_x = 0$, which is analogous to the group-wise heteroskedastic SARE model in LBL, but distinct from any model reported in HMK. Model 4 assumes $\psi^2_{r_i s_i} = \psi^2_{r_i x_i} \; \forall \, i$ and $\lambda_s = 0$, which incorporates treatment heteroskedasticity and correlation into model 3 and is analogous to HMK's third model, but distinct from any model reported in LBL. Model 5 is unrestricted; it incorporates strip heteroskedasticity and correlation into model 4 and is analogous to HMK's fourth model, but again is distinct from any model reported in LBL.

---

[3] These restrictions were imposed by defining

$$\lambda_1 + \lambda_x + \lambda_s = \frac{2e^{\tau}}{1 + e^{\tau}} - 1, \quad \frac{\lambda_1}{\lambda_1 + \lambda_x + \lambda_s} = \frac{e^{\tau_1}}{1 + e^{\tau_1} + e^{\tau_x}},$$

$$\frac{\lambda_x}{\lambda_1 + \lambda_x + \lambda_s} = \frac{e^{\tau_x}}{1 + e^{\tau_1} + e^{\tau_x}}, \quad \text{and} \quad \frac{\lambda_s}{\lambda_1 + \lambda_x + \lambda_s} = \frac{1}{1 + e^{\tau_1} + e^{\tau_x}},$$

and maximizing the log likelihood function over τ, τ_1, and τ_x instead of λ_1, λ_x, and λ_s.

The PA hypothesis for variable rate nitrogen applications (VRA) is tested by estimating model 6, which restricts model 5 by assuming $\alpha_{k_x r_i} = \alpha_{k_x r_j} \forall\, k_x \in \{1, 2\}$, i, and j. The hypothesis is tested using the likelihood-ratio statistic, since model 6 is nested in model 5.

The SARE and analogous GEO models reported in HMK are not nested. Therefore, the likelihood-ratio statistic cannot be used for comparison. Following LBL, we compare these models using the Akaike information criterion (AIC), taken as $AIC = -2\log L + 2K$, where $\log L$ is the maximized log likelihood and K is the number of parameter estimates for the model. The AIC rewards models with a higher maximized log likelihood, but penalizes ones with more parameters. Therefore, it trades off model fit for model parsimony. A lower AIC is preferable.

Model Validation

A comprehensive validation of our model and estimation procedure using methods such as Monte Carlo simulation is computationally impractical. Still, exploring how the model and estimation procedures perform on simulated data is useful. To accomplish this simulation for the preferred SARE model, the estimated parameters were used to construct the covariance matrix for the error. We then transformed a vector of 612 independent standard normal variates using a Cholesky decomposition and added these transformed variates to the yields predicted from our quadratic yield functions. The models were then re-estimated using these simulated yield data.

Potential Value of VRA

Estimates of the potential value of increased nitrogen returns from VRA are calculated using the α coefficient estimates from equations (4) and (5). The estimated nitrogen return above fertilizer costs is defined as:

$$\pi = \left(\sum_{i=1}^{612} \frac{p_y\left(\alpha_{0r_i} + \alpha_{1r_i} x_i + \alpha_{2r_i} x_i^2\right) - p_x x_i}{612} \right),$$

where p_y and p_x are the price of corn and nitrogen, respectively. The optimal VRA is calculated by choosing x_i for $i = 1, \ldots, 612$ to maximize π. Alternatively, an optimal uniform rate (URA) is calculated by choosing $x = x_i$ for $i = 1, \ldots, 612$ to maximize π. These optimal rates are constrained between 0 and 202 kg/ha to avoid predicting yields outside the range of the available data. Nitrogen returns for the optimal VRA and URA are compared to the University of Minnesota (UMN) recommendation (135 kg/ha for Hanska and 140 kg/ha for Morgan) to determine the potential value of VRA within and between fields assuming the price of corn and nitrogen equal \$98.21/t and \$0.374/kg.

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 is calculated as $\pi^{VRA} - \pi^{UMN}$, and 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 is calculated using a Taylor series expansion (see Casella and Berger, 1990, pp. 328–331).

Table 1. Log Likelihood and Model Comparisons for Spatial Autoregressive Error (SARE) Model

	Location			
	Hanska		Morgan	
Model	Maximized Log Likelihood	Estimated Parameters	Maximized Log Likelihood	Estimated Parameters
1 [a]	-2,653.71	19	-2,825.80	19
2 [a,c]	-2,489.99	20	-2,648.54	20
3 [a,c,d]	-2,470.05	25	-2,628.28	25
4 [a,c,d,e]	-2,431.70	31	-2,619.36	31
5 [a,c,d,e,f]	-2,336.43	61	-2,524.69	61
6 [b,c,d,e,f]	-2,367.70	51	-2,564.40	51
Model Comparisons (Restricted vs. Unrestricted)	Likelihood-Ratio Statistic	Degrees of Freedom	Likelihood-Ratio Statistic	Degrees of Freedom
1 vs. 2	327.44***	1	354.51***	1
2 vs. 3	39.88***	5	40.53***	5
3 vs. 4	76.68***	6	17.83***	6
4 vs. 5	190.55***	30	189.34***	30
6 vs. 5	62.53***	10	79.43***	10

Note: Double and triple asterisks (*) denote statistical significance at $p < 0.05$ and $p < 0.01$, respectively.
[a] Assumes within-field variation in crop response to nitrogen.
[b] Assumes no within-field variation in crop response to nitrogen.
[c] Assumes spatial correlation is present.
[d] Assumes region-specific heteroskedasticity is present.
[e] Assumes treatment heteroskedasticity and correlation are present.
[f] Assumes strip heteroskedasticity and correlation are present.

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

Results

Hypothesis Tests

We begin by comparing the alternative versions of the SARE models to test the importance of incorporating region, spatial, treatment, and strip dependent heteroskedasticity and correlation. The PA hypothesis for VRA is also tested. Table 1 reports the maximized log likelihood and number of estimated parameters for each location and model and the likelihood-ratio statistic and degrees of freedom for each model comparison.

Regression errors for the SSCRF estimates exhibit region, spatial, treatment, and strip dependent heteroskedasticity and correlation. Model 1 is rejected in favor of model 2 for both locations, which indicates errors are spatially dependent. Model 2 is rejected in favor of model 3, indicating, in agreement with LBL, that there is significant region dependent heteroskedasticity. Model 3 is rejected in favor of model 4, which shows there

is also treatment dependent heteroskedasticity and correlation. Model 4 is rejected in favor of model 5, so strip dependent heteroskedasticity and correlation are important. These results are consistent with those reported by HMK.

There is the potential for precision agriculture to improve nitrogen returns. Model 6 is rejected in favor of model 5 for both the Hanska and Morgan locations, indicating crop response to nitrogen varied significantly among the six regions within each field (see table 1). This result is consistent with the findings of both LBL and HMK.

Model Fit and Validation

Table 2 reports the regional response coefficients and *t*-statistics for the best fitting SARE model. It reports estimates for the spatial, treatment, and strip dependence parameters and the likelihood-ratio statistic for the individual significance of each; the average estimated standard deviation; and the maximized log likelihood. Table 2 also reports this same information obtained from estimating each model with data simulated assuming the actual estimates are true.[4]

The response function coefficient estimates for Hanska and Morgan all have the expected sign. They are all significant at 5% with the exception of the quadratic coefficients for Hanska regions 1 and 6. Individually, the spatial dependence parameter is relatively large and significant for both Hanska and Morgan. The strip dependence parameter is smaller, but still significant for Hanska. It is not significant for Morgan. The treatment dependence parameter is not significant for either Hanska or Morgan. Since the treatment and strip parameters are individually insignificant for Morgan, we also tested and rejected the hypothesis that these parameters are jointly insignificant ($\chi^2_{[2]} = 9.88$). Further comparisons of the maximized log likelihoods for the restricted and unrestricted models indicated that including the treatment dependence parameter is preferable to including the strip dependence parameter. However, including one or the other is preferable to not including both.

The actual and simulated parameter estimates are of the same sign and magnitude. In terms of individual hypothesis tests for the response coefficients, there are two differences between the actual and simulated data for a 5% level of significance. For the actual data, the quadratic coefficients for regions 1 and 6 at Hanska are statistically insignificant, which is not the case with the simulated data. Individual hypothesis tests for the spatial, treatment, and strip dependence parameters are consistent across the actual and simulated data.

GEO versus SARE

We now consider a comparison of the SARE and GEO frameworks. Table 3 reports the Akaike information criterion (AIC) for selected SARE models. It also reports the AIC for analogous models from HMK. For Hanska, the GEO model 5 has a lower AIC than the SARE model 5. The reverse is true for Morgan. When strip heteroskedasticity and correlation are not accounted for, the SARE models have lower AICs than the GEO models for both locations. These results are consistent with LBL, who found a group-wise heteroskedastic SARE model (model 3) performed better than the standard SARE and

[4] More rigorous results based on Monte Carlo methods were not possible due to the computational intensity of the models.

Table 2. Crop Response Estimates (*t*-Statistics), Covariance Parameters ($\chi^2_{[1]}$), and Maximized Log Likelihood for Spatial Autoregressive Error Model 5 with Actual and Simulated Data (corn *t*/ha for nitrogen kg/ha)

	Hanska		Morgan	
Parameter / Region	Actual	Simulated	Actual	Simulated
Constant				
1	9.43***	8.27***	6.07***	6.14***
	(22.14)	(25.88)	(20.03)	(21.72)
2	8.42***	7.92***	6.97***	6.79***
	(20.55)	(22.13)	(25.36)	(22.31)
3	9.13***	8.11***	6.35***	6.35***
	(25.38)	(29.68)	(19.30)	(20.64)
4	4.90***	5.15***	7.71***	7.83***
	(15.39)	(19.68)	(32.02)	(37.54)
5	3.97***	3.75***	8.95***	9.32***
	(11.14)	(12.25)	(26.48)	(33.69)
6	4.50***	4.01***	7.85***	7.87***
	(11.86)	(13.20)	(37.57)	(38.48)
Linear				
1	0.009**	0.018***	0.046***	0.044***
	(1.70)	(3.91)	(10.41)	(10.99)
2	0.025***	0.027***	0.041***	0.051***
	(4.58)	(4.83)	(9.71)	(11.10)
3	0.009**	0.019***	0.052***	0.056***
	(2.03)	(4.60)	(10.93)	(12.95)
4	0.036***	0.030***	0.050***	0.054***
	(12.71)	(12.29)	(14.15)	(15.75)
5	0.031***	0.034***	0.040***	0.037***
	(8.92)	(10.84)	(7.61)	(7.92)
6	0.025***	0.028***	0.041**	0.040**
	(7.38)	(9.97)	(13.41)	(11.95)
Quadratic				
1	-0.000021	-0.000047**	-0.000076***	-0.000074***
	(0.90)	(2.26)	(3.69)	(4.06)
2	-0.000082***	-0.000079***	-0.000103***	-0.000150***
	(3.48)	(3.23)	(5.15)	(7.17)
3	-0.000021	-0.000052***	-0.000156***	-0.000180***
	(1.08)	(2.90)	(6.97)	(9.02)
4	-0.000090***	-0.000061***	-0.000128***	-0.000145***
	(6.89)	(5.33)	(8.33)	(10.11)
5	-0.000047***	-0.000052***	-0.000145***	-0.000130***
	(2.91)	(3.48)	(6.28)	(6.58)
6	-0.000017	-0.000022**	-0.000110***	-0.000095***
	(1.07)	(1.65)	(8.31)	(6.76)
Spatial (λ_1)	0.83***	0.85***	0.87***	0.87***
	(188.64)	(320.54)	(268.26)	(318.28)
Treatment (λ_x)	0.00	0.00	0.06	0.06
	(0.00)	(0.00)	(2.97)	(0.88)
Strip (λ_s)	0.14***	0.13***	0.05	0.01
	(22.10)	(19.49)	(1.13)	(0.02)
Average Standard Deviation [a]	2.12	2.03	2.24	1.54
Maximized Log Likelihood	-2,336.43	-2,280.62	-2,524.69	-2,491.83

Note: Double and triple asterisks (*) denote statistical significance at $p < 0.05$ and $p < 0.01$, respectively.

[a] Calculated as follows, where $\sigma_{r_i s_i}$ is the standard deviation, and r_i and s_i are the site and treatment strip for the ith observation:

$$\sum_{i=1}^{N} \sigma_{r_i s_i} / N \ t/\text{ha}.$$

Table 3. Spatial Autoregressive Error (SARE) and Geostatistical (GEO) Model Comparisons Using the Akaike Information Criterion (AIC)

	Hanska		Morgan	
Model	SARE	GEO[a]	SARE	GEO[a]
2	5,020.0	5,030.0	5,337.1	5,387.6
3	4,990.1	—	5,306.6	—
4	4,925.4	4,929.2	5,300.7	5,344.0
5	4,794.9	4,769.8	5,171.4	5,179.6

[a] AIC calculated based on values reported in Hurley, Malzer, and Kilian (2004). Hurley, Malzer, and Kilian's second, third, and fourth models correspond to the SARE models 2, 4, and 5.

GEO models (model 2). The results of their comparison between the standard SARE and GEO models were dependent on the type of distance function used for the GEO estimates. Our results indicate that the standard SARE model outperforms the standard GEO model estimated with a Guassian distance function when region, treatment, and strip effects are not taken into account.

Variation in Corn Response to Nitrogen

Figures 2 and 3 illustrate the differences in our estimates of corn response to nitrogen along with the differences in the optimal nitrogen rates and yields between the SARE and GEO models. For Hanska (figure 2), the SARE model predicts higher overall yields for half the field (regions 1–3) and lower overall yields for half (regions 4–6). In general, the response functions appear less concave for the SARE model. For regions 1–5, the optimal nitrogen rate is higher for the SARE model, and for region 6 it is the same. For Morgan (figure 3), the SARE model consistently predicts higher overall yields for all regions. For regions 2–5, the response functions also appear to be more concave. For regions 2–6, the optimal nitrogen rate is lower with the SARE model. For region 1, it is the same.

Comparing the optimal VRA to the UMN rate, the SARE model produces a larger estimate of the potential return to VRA for Hanska, but a smaller estimate for Morgan (see table 4). The SARE model produces lower estimates for the standard deviation, which results in tighter confidence intervals. The point estimate obtained from the GEO model falls outside the confidence interval estimated for the SARE model for both locations. Furthermore, the estimated confidence intervals from each model do not even overlap for Hanska.

Comparing the optimal VRA to the URA rate, the SARE model produces a larger estimate for both locations. The SARE model produces lower estimates of the standard deviation and tighter confidence intervals. The point estimate obtained from the GEO model falls outside the SARE model's confidence interval for Morgan, but not for Hanska.

Comparing the optimal URA to the UMN rate, the SARE model produces a larger estimate for Hanska, but a smaller estimate for Morgan. The SARE model produces a higher estimate of the standard deviation and wider confidence intervals for Hanska, but a lower standard deviation and tighter confidence intervals for Morgan. However, for both locations the SARE model produces a smaller coefficient of variation. The point estimate obtained from the GEO model again falls outside the SARE model's confidence interval for both locations. There is also no overlap in the models' confidence intervals for Hanska.

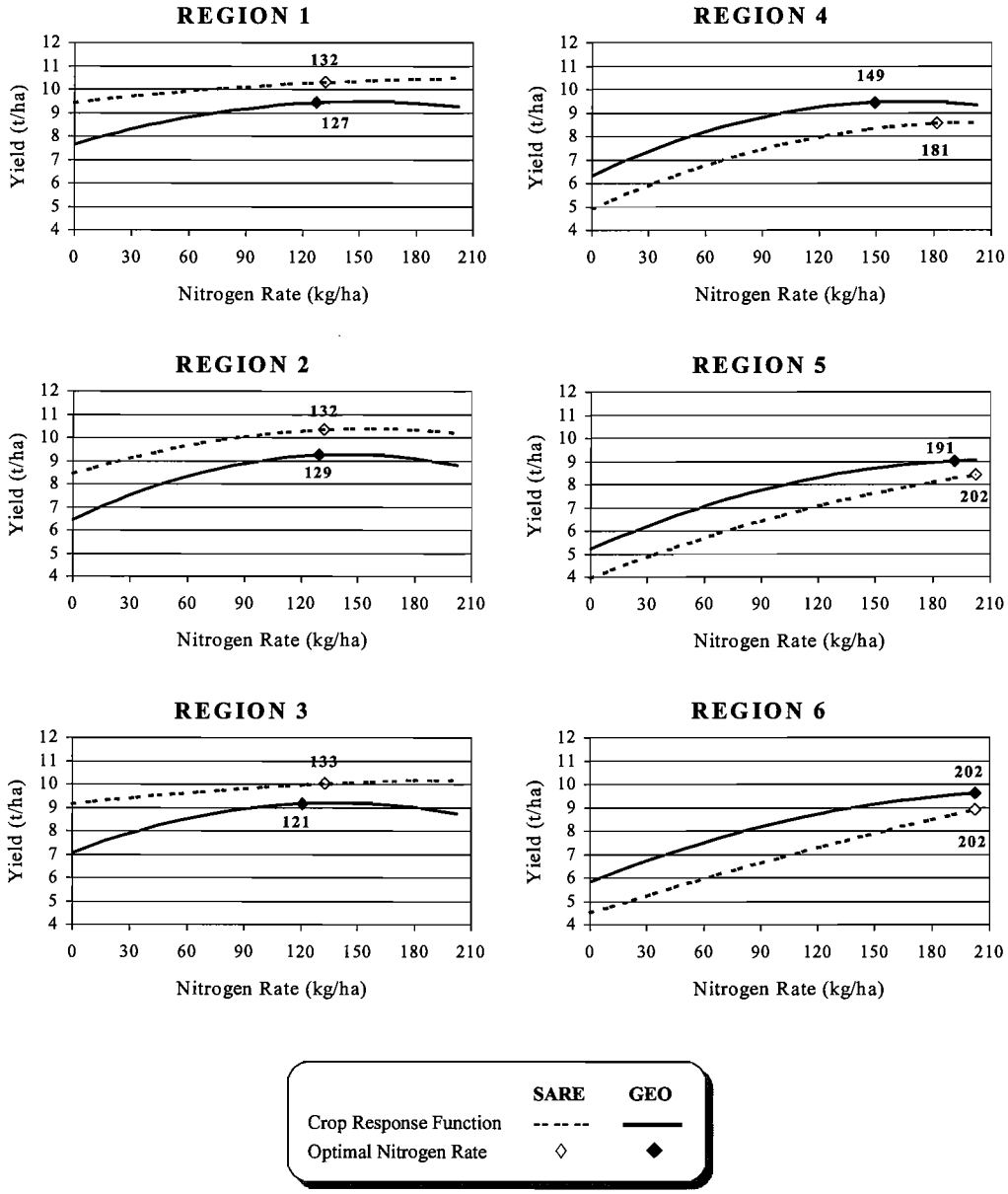


Figure 2. Estimated crop response functions and optimal nitrogen rates for Hanska (model 5)

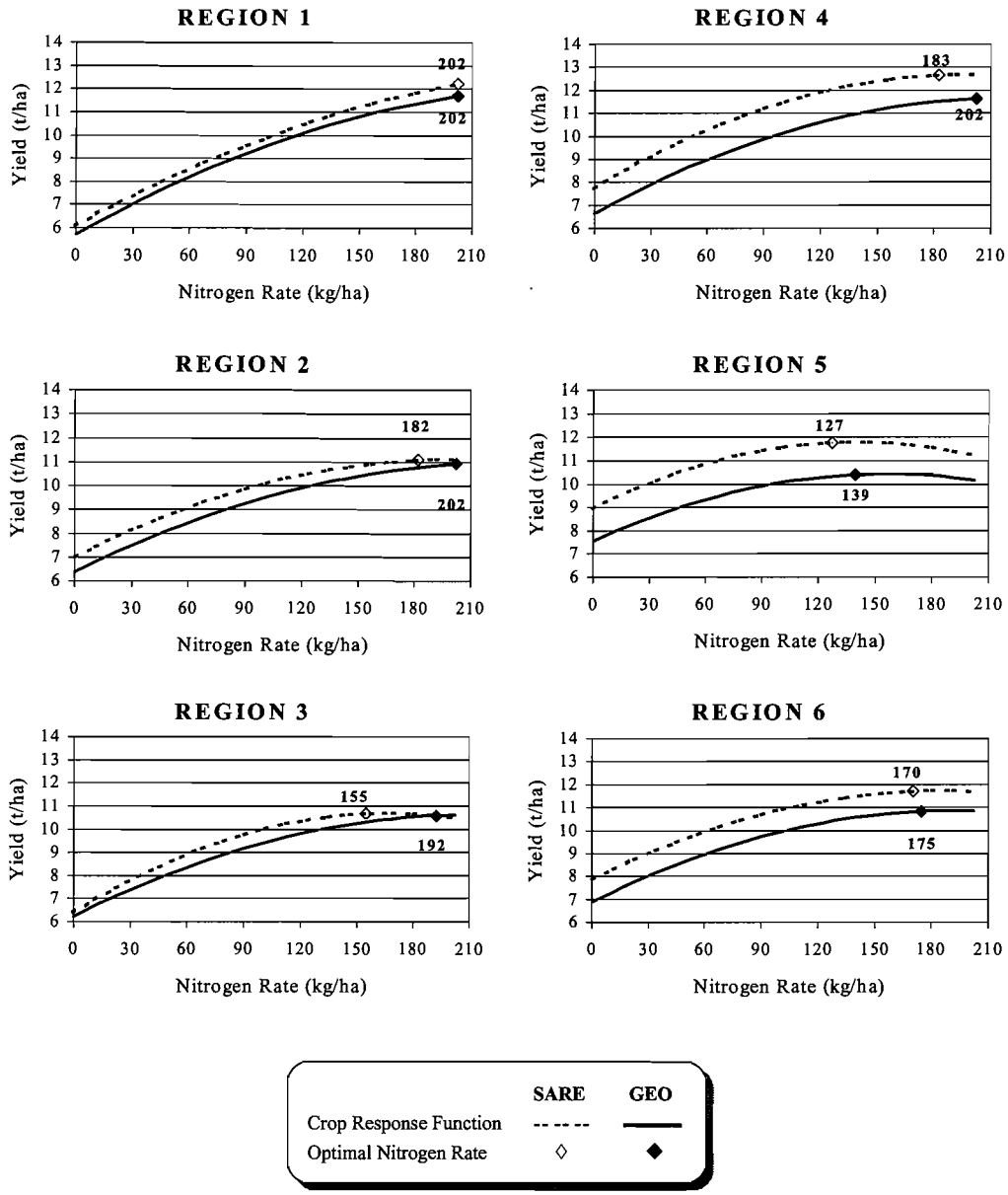


Figure 3. Estimated crop response functions and optimal nitrogen rates for Morgan (model 5)

Table 4. Difference in Returns to Nitrogen, $/ha (5th percentile, 95th percentile)

	Hanska		Morgan	
Comparison	SARE	GEO[a]	SARE	GEO[a]
VRA[b] vs. UMN[c]	33.32 (25.03, 41.61)	11.83 (-1.00, 24.65)	26.06 (19.25, 32.88)	35.03 (8.96, 61.10)
VRA vs. URA[d]	10.04 (0.48, 19.60)	9.85 (-0.89, 20.59)	12.35 (7.99, 16.71)	7.64 (-1.88, 17.16)
URA vs. UMN	23.28 (10.16, 36.41)	1.98 (-1.70, 5.65)	13.71 (8.18, 19.24)	27.39 (-0.44, 55.22)

Note: Percentiles assume estimates are normally distributed.
[a]Calculated based on Hurley, Malzer, and Kilian's (2004) fourth model, which is the geostatistical analogue to the SARE model 5.
[b]Optimal variable rate nitrogen application.
[c]University of Minnesota recommended uniform rate nitrogen application.
[d]Optimal uniform rate nitrogen application.

Whipker and Akridge (2003) report the results of a survey of agronomy dealerships throughout the United States. The results of their survey provide some insight into whether these estimates of the potential value of precision agriculture are high enough to justify any increased implementation costs (e.g., cost of acquiring and interpreting site-specific information and the cost of carrying out variable rate applications). The results of the survey indicated that the average cost of global positioning satellite (GPS) soil sampling was $15.30/ha. Average yield monitor data analysis cost was $3.14/ha. Average geographic information system (GIS) mapping cost was $8.80/ha. The average cost of agronomic recommendations based on GPS or GIS was $6.42/ha. The average cost of GPS controller driven, single nutrient, variable rate fertilizer application was $13.12/ha. Combined, these costs equal $46.78/ha, which exceeds the estimated potential value of variable rate nitrogen applications for both Hanska and Morgan regardless of whether or not we use the SARE or GEO estimates. However, with the GEO model, $46.78/ha falls inside Morgan's 90% confidence interval for switching from the UMN recommendation to the URA or VRA. Therefore, while both models offer the same conclusion for Hanska, they give somewhat different conclusions for Morgan.

Conclusions

Testing the precision agriculture (PA) hypothesis for variable rate nitrogen applications (VRA) has proven challenging due to the difficulties of collecting, analyzing, and interpreting appropriate data. Site-specific crop response functions (SSCRFs) provide a useful tool for evaluating the PA hypothesis, but appropriate statistical models are necessary. Problems that can hinder the estimation of SSCRFs using data from randomized complete block design field experiments include region, spatial, treatment, and strip heteroskedasticity and correlation. The purpose of this paper was to propose a spatial autoregressive error (SARE) model for dealing with these problems and to compare the results of this model with the geostatistical (GEO) model proposed by Hurley, Malzer, and Kilian (2004).

Results using the spatial autoregressive model confirm previous results based on a geostatistical model. There is significant region, spatial, treatment, and strip

heteroskedasticity and correlation, and potential for VRA to improve nitrogen returns. A comparison of the SARE and GEO models based on the Akaike information criterion favored the GEO model for one experimental location and the SARE model for the other. The SARE model, however, consistently produced lower estimates for the coefficient of variation. The SARE model also tends to be computationally less intensive and more practical to estimate because it can be accomplished with fewer matrix inversions. Alternatively, covariance parameter estimates for the GEO model have a more direct interpretation in terms of heteroskedasticity and correlation.

Lambert, Bongiovanni, and Lowenberg-DeBoer (2002) show comparisons of fit between SARE and GEO models may be sensitive to the distance function used for the GEO semi-variogram. The results are also likely to be sensitive to the specification of the spatial weight matrix employed by the SARE model. However, they conclude that the estimates of the potential value of VRA, at least for their Argentina field, were not very sensitive to whether a SARE or GEO model was used to estimate SSCRFs. Therefore, from a practical standpoint, the choice of models was not so critical. Our results suggest the choice of model could be critical. The two models produced estimates for the optimal nitrogen rates and potential value of VRA that were notably different.

It is important to understand that there are a variety of differences between the Lambert, Bongiovanni, and Lowenberg-DeBoer (2002) and the Hurley, Malzer, and Kilian (2004) data sets. For example, the experiments were from different locations and different years. Treatment rates differed. One used a yield monitor to measure yields, while the other employed a plot combine. One used topographical and other information to divide the field into regions, while the other used a rectangular grid. Still, with all these differences, we find the same qualitative results as Lambert, Bongiovanni, and Lowenberg-DeBoer for the same types of comparisons using the Hurley, Malzer, and Kilian data—both SARE and GEO models do better than OLS, and a SARE model with group (region)-wise heteroskedasticity performs better than SARE or GEO models without. We also find the same type of results with the SARE model as those reported by Hurley, Malzer, and Kilian with the GEO model—important spatial, region, treatment, and strip effects. Lambert, Bongiovanni, and Lowenberg-DeBoer reported important spatial and region effects. While they did not report any treatment or strip effects, neither did they rule them out. Given the theoretical underpinnings of the spatial, region, and treatment effects, they are likely to be important in any SSCRF estimation. The strip effect is likely attributable to an idiosyncratic feature of the experimental design—the lack of within-strip randomization. Therefore, strip effects are unlikely to be of concern when estimating SSCRFs with data from experiments that randomize treatments within as well as between strips.

As a final note, we pursued the SARE framework instead of the spatial lag (SL) framework because our initial comparison of the two favored the SARE framework. However, this initial comparison did not include the region, treatment, and strip heteroskedasticity and correlation that were found to be important embellishments to the SARE and GEO models. Future work might try including these factors in an SL model to see if these initial rankings are robust. It should be noted that within the SL framework, it is possible to handle region, treatment, and strip effects using either the GEO or the SARE strategy.

[*Received July 2003; final revision received May 2005.*]

References

Anselin, L. *Spatial Econometrics: Methods and Models.* London: Kluwer Academic Publishers, 1988.

Bongiovanni, R., and J. Lowenberg-DeBoer. "Nitrogen Management in Corn Using Site-Specific Crop Response Estimates from a Spatial Regression Model." In *Proceedings of the 5th International Conference on Precision Agriculture*, eds., P. C. Robert, R. H. Rust, and W. E. Larson. Madison, WI: American Society of Agronomy/Crop Science Society of America/Soil Science Society of America, 2000.

———. "Precision Agriculture: Economics of Nitrogen Management in Corn Using Site-Specific Crop Response Estimates from a Spatial Regression Model." Paper presented at the annual meetings of the American Agricultural Economics Association, Chicago, IL, 6 August 2001.

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

Cressie, N. A. C. *Statistics for Spatial Data.* New York: John Wiley and Sons, 1993.

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 *Precision Agriculture: Proceedings of the 3rd International Conference*, eds., P. C. Robert, R. H. Rust, and W. E. Larson, pp. 513–519. Madison, WI: American Society of Agronomy/Crop Science Society of America/Soil Science Society of America, 1996.

Dikici, H. "Seasonal Nitrogen Availability: A Site-Specific Approach." Dept. of Soil, Water, and Climate, University of Minnesota, Minneapolis, 2000.

Horn, R., and C. Johnson. *Matrix Analysis.* New York: Cambridge University Press, 1993.

Hurley, T. M., G. Malzer, and B. Kilian. "A Test of Within-Field Variation of Corn Response to Nitrogen in Central Minnesota." In *Precision Agriculture: Herausforderung an Integrative Forschung, Entwicklung und Anwendung in der Praxis*, eds., A. Werner and A. Jarfe, pp. 413–421. Darmstadt, Germany: Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL), 2002a.

———. "A Test of Within-Field Variation for Corn Response to Nitrogen in Central Minnesota." Paper presented at the 6th International Congress on Precision Agriculture and Other Precision Resource Management, Minneapolis, MN, 15 July 2002b.

———. "Estimating Site-Specific Nitrogen Crop Response Functions: A Conceptual Framework and Geostatistical Model." *Agronomy J.* 96(2004):1331–1343.

Lambert, D., R. Bongiovanni, and J. Lowenberg-DeBoer. "Spatial Regression, an Alternative Statistical Analysis for Landscape Scale On-Farm Trials: Case Study of Soil Density Trials in Central Illinois." Paper presented at the 6th International Congress on Precision Agriculture and Other Precision Resource Management, Minneapolis, MN, 15 July 2002.

Malzer, G. L., P. J. Copeland, J. G. Davis, J. A. Lamb, P. C. Robert, and T. W. Bruulsema. "Spatial Variability of Profitability in Site-Specific N Management." In *Precision Agriculture: Proceedings of the 3rd International Conference*, eds., P. C. Robert, R. H. Rust, and W. E. Larson, pp. 967–975. Madison, WI: American Society of Agronomy/Crop Science Society of America/Soil Science Society of America, 1996.

McBratney, A. B., and R. Webster. "Choosing Functions for Semi-variograms of Soil Properties and Fitting Them to Sampling Estimates." *J. Soil Sci.* 37(1986):617–639.

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

Wall, M. "A Close Look at the Spatial Correlation Structure Implied by the CAR and SAR Models." *J. Statistical Planning and Inference* 121(2004):311–324.

Whipker, L. D., and J. T. Akridge. "2003 Precision Agricultural Services Dealership Survey Results." Staff Paper No. 3-10, Center for Food and Agricultural Business, Purdue University, 2003.