The *Journal of Cotton Science 12:210–227 (2008)*
http://journal.cotton.org, © The Cotton Foundation 2008

ECONOMICS AND MARKETING

Economics of Management Zone Delineation in Cotton Precision Agriculture

M. Velandia*, R. M. Rejesus, K. Bronson, E. Segarra

ABSTRACT

The concept of precision agriculture is based on the ability to improve the management of production factors using site-specific information. The optimal configuration of management zones for more precise management of farm inputs is one of the most important components in precision farming. The objective of this study is to develop a management zone delineation procedure based on a spatial statistics approach and evaluate its economic impact for Texas cotton production. Using an optimization model that utilizes a yield response function estimated from field experiment data through spatial econometric methods, we evaluate the economics of the management zone delineation procedure. We found that applying variable N rates based on the management zones delineated would result in higher cotton yields and higher net returns, relative to a uniform rate application based on field information and a variable rate application based on landscape position. This is indicative of the potential economic value of using a spatial statistics approach to management zone delineation in cotton production.

Precision agriculture (PA) refers to the use of different site-specific technologies in production agriculture, such as Global Positioning Systems (GPS), computer-controlled variable rate application technologies (CVRT), and geo-referenced yield maps. These technologies are typically used to obtain information about yield and/or soil characteristics at different points in a field to establish more efficient management strategies that explicitly consider heterogeneity among the different locations within a field. PA is based on the premise that management of production factors can be improved (and profitability potentially enhanced) when producers take advantage of site-specific information and variable rate application technologies (Hurley, Oishi and Malzer, 2005).

For precision fertilizer application in particular, one would ideally want to collect dense data about soil attributes in order to variably apply this input based on the continuous variation of the attributes in the field. However, producers typically have sparse soil attribute data due to the prohibitive expense of collecting dense soil data. Note that soil attribute data are usually collected using chemically analyzed soil samples that are randomly chosen over the whole field. Although there has been progress in developing soil sensors (e.g. pH sensors) for collecting denser soil attribute data (Griffin et al., 2004), most of the soil characteristic data needed for more precise fertilizer application are still typically collected using soil samples (e.g. collecting soil nitrogen (N) data). Hence, the sparse data environment makes it appealing to simply identify and delineate discrete management zones that can be variably and more precisely managed. Management zones are defined as geographical areas that can be treated as homogenous, so that input application and decision making can be treated separately for each zone.

In light of the sparse data environment that producers are typically faced with, optimally configuring management zones for more precise management of farm inputs becomes a very important issue in precision farming. The profitability of precision technologies in this case will depend on the producer's ability to divide his/her field into appropriate management zones that are based on sparse soil characteristic data. If fields can be divided into appropriate management zones even in sparse data environments, then the use of variable rate input technology could be implemented more efficiently and returns from this precision application may be enhanced.

M. Velandia*, Department of Agricultural Economics, University of Tennessee, 314C Morgan Hall, Knoxville, TN 37996-4518; R. M. Rejesus, Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC 27695-8109; K. Bronson, Texas Agricultural Experiment Station, Texas A&M University, Lubbock, TX 79401, E. Segarra, Department of Agricultural and Applied Economics, Texas Tech University, Lubbock, TX 79409-2132
Corresponding author: mvelandia@utk.edu

A number of studies examine the use of yield and soil properties to visually establish management zones (Fridgen et al., 2000, Franzen et. al., 2000; Fleming et. al., 2000; Nolan et. al., 2000, Basnet, B., et. al., 2003, Diker et al. 2003).[1] But in recent years, there have been studies that explored more rigorous statistical methods to delineate management zones based on spatial precision agriculture data. These types of studies have usually relied on traditional clustering or fuzzy clustering techniques which look for the identification of naturally occurring clusters in the data through algorithms (Stafford et. al., 1998; Fridgen et. al., 2000, Ping et al., 2005). However, studies that use clustering techniques typically do not take into account the underlying spatial autocorrelation in the yield or soil data when delineating management zones (with the exception of Ping et al. (2005)).

While cluster analysis allows one to arrange the data into different classes according to similarity measures (taking into account a particular variable or several variables), taking spatial autocorrelation into account will make the management zone configuration based not only on the information about a particular characteristic but also their location in space. The 'First Law of Geography' suggests that everything is related to everything else but things that are near are going to have a higher level of similarity than things that are distant in space (Tobler, 1979). Hence, consideration of spatial autocorrelation in management zone delineation procedures may provide better insight into the spatial patterns of the field and more effectively suggest zones for use as management units. A management zone delineation procedure that takes spatial autocorrelation into account is a fairly new method that still needs to be further developed in a precision farming context. In this regard, advances in spatial statistics allow us to develop this type of delineation approach in order to address this gap in the literature.

Aside from developing a new management zone delineation procedure that considers spatial autocorrelation, another important issue that needs to be examined is the economic implications of using this type of management zone delineation procedure. Although there have been numerous studies that have looked at various economic aspects of precision agriculture[2], there have only been a few studies that explicitly addressed the economics of alternative management zone delineation procedures (See, for example, Thrikawala et. al. (1998); Basin, et al. (2003); and Dillon et al. (2003)). Thus, this paper not only develops a practical spatial method to delineate management zones in cotton production, but it also contributes to the emerging literature on the economics of management zone delineation procedures in precision agriculture.

The objective of this paper is to develop, and economically evaluate different management zone delineation procedures that can be used for precision nitrogen (N) fertilizer application in Texas cotton production. In particular, a univariate management zone delineation procedure based on a spatial statistics technique (called Exploratory Spatial Data Analysis (ESDA)) is developed and evaluated. The economic returns to a variable rate N application based on this delineation technique is then compared to a traditional uniform rate application and a variable rate application based on a management zone delineation procedure using landscape position. In order to appropriately assess the economics of these different N application procedures, spatial econometric techniques in yield response estimation were carefully applied, and its importance in this context was also shown. (See Anselin, Bongiovanni, and Lowenberg-DeBoer, 2004 for more discussion of this issue)[3].

The remainder of this article is organized as follows. The next section describes the field experimental design and the procedures used for collecting/standardizing the spatial data. The empirical methodology for developing the 'spatial' management zone delineation procedure and the economic analysis is discussed in the third section. Results and concluding comments are presented in the fourth and fifth sections, respectively.

[1] Some examples of yield and soil properties that have been used to visually establish management zones are (among others): variability of yield within landscape and elevation positions, soil color, texture and electrical conductivity.

[2] See Lambert and Lowenberg-DeBoer (2000) and Griffin et al. (2004) for an extensive review of the literature on the economics/profitability of precision agriculture technologies as applied to different crops.

[3] Note that Anselin, Bongiovanni, and Lowenberg-DeBoer (2004) discuss these issues in the context of management zones delineated using landscape position, which is why we chose to compare our delineation approach to a delineation procedure based on landscape position, as well as a blanket or uniform rate approach with no zone delineation.

DESCRIPTION OF THE DATA

Experimental Design of the Field. The data used in this article are based on an agronomic cotton experiment designed to study nitrogen (N) use for cotton production in the Southern High Plains of Texas (Lamesa, TX) during the 2002 growing season. The study area is located 60 miles south of Lubbock, TX and the total area is about 35 acres under 119 acres of a center pivot irrigation system. The experiment is a randomized complete block design with three replicates and each replicate was within a center pivot irrigation span. There were three N treatments – variable-rate N, blanket-rate N, and zero N. The fertilizer N rate for both variable and blanket N rate was calculated using the N supply requirement of 120 lb/acre for a yield goal of approximately 981 lb lint/acre (See Zhang et al. (1998) for the agronomic basis of this recommendation). The blanket rate of N fertilizer was based on the 0-2 foot soil NO_3 – N content of the blanket-N plots. Inverse distance interpolation of all the original soil sample values (0-2 foot NO_3 – N) was used to create variable-rate application maps. Irrigation levels were 63%, 74%, and 84% of the estimated evapo-transpiration replacement. The three N treatments were replicated under each irrigation level, (See Figure 1). [4]

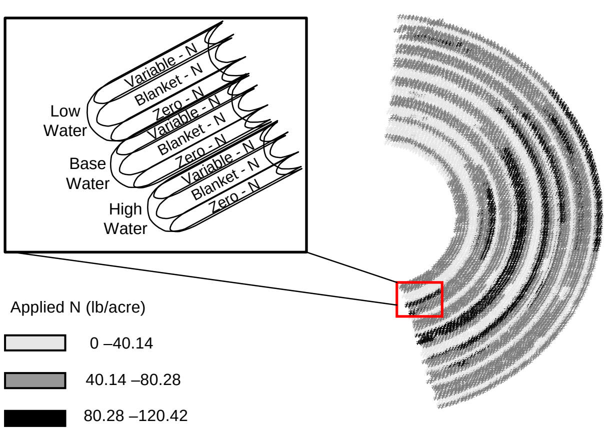

Figure 1. Design of Nitrogen Fertilizer Experiment, 2002, Lamesa, Texas

The site-specific soil characteristics data were originally collected as point data. In March 2002, soil samples were collected on 135 Differential Global Positioning System (DGPS) referenced points. The samples were then sent to a laboratory for chemical analysis. The results of this chemical analysis provided the necessary point data about the different soil characteristics in the field. For this study the information for extractable NO_3 – N at two foot depth was used as the basis for estimating the amount of N in the soil.

The yield data used in this study are from a dense data set collected from a yield monitor (5592 points). A John Deere 7445 four-row stripper harvester equipped with a Micro-Trak optical yield monitoring system (Micro- Trak, Eagle Lake, MN) was used to harvest seed cotton in the field. These seed cotton weights were adjusted and a single percentage turnout of lint content from a local commercial gin was used to calculate lbs of lint per acre as our measure of yield (Bronson, et al 2006). A dense applied N data set was also utilized in this study, which came from the precision N applicator used for the experiment. The dense applied N data also had the same number of points as with the dense yield data from the yield monitor (5592 points).

At this point it is important to note that the yield data used in this study were adjusted based on the water treatments/irrigation levels applied in the experiments (i.e. high (84%), medium (74%) and low (63%) water treatments). A coefficient was added (subtracted) for the low water treatment (for the high water treatment) that takes out the effect of water on yield. The coefficient used for this procedure is based on an analysis of covariance adjustment (see Bronson et al., 2006 for details). By using this adjustment procedure, the effect of N on yields per acre was better isolated.[5]

Interpolation Procedures. Since data are available at different spatial scales (e.g. some are dense point data, like the yield and applied N

[4] Although there are only three N-treatments, there is significant variability in the N-rates applied for the whole field, which allows this experimental data to be used for yield response estimation. First, there is variability in the variable rate treatment plots because of the different rates applied for each point in that particular treatment plot (V in Figure 1). Second, there is also variability across the blanket rate replicated plots because application rates differ depending on the average soil NO_3-N for each blanket rate treatment plot.

[5] Another approach to control for water treatment effect is to include water levels in the yield response estimation. However, we opted to do the adjustment described here because it allows for estimation of a parsimonious yield response function that facilitates the economic evaluation. We believe that this adjustment is an acceptable approach in the agronomic literature (See Bronson et al., 2006) to control for water treatments and we have found no study that clearly invalidates this approach.

data, while some are sparse point data, like the soil $NO_3 - N$), an interpolation method is used to create a consistent spatial data set for all pertinent variables. Interpolation is typically used when it is expensive to collect more data points on the soil characteristics of interest, which is the case for our experiment (e.g. collecting dense data on soil $NO_3 - N$).

The kriging interpolation method was used to convert the dense data and the sparse 135 DGPS data into consistent grids with a size of 28m x 28m. A 28 x 28 m design was chosen over other grid sizes (8 x 8 m, 16 x 16 m and 24 x 24 m) after checking by visual inspection that this grid size generated enough information between the DGPS referenced points and that the resulting grid pattern closely follows the actual patterns presented in the point data (see Figure 2).[6] For the case of the dense data sets (e.g. yield and applied N data), the interpolation procedure spatially averages the data using different weights for the points used (according to the spatial autocorrelation patterns) to convert it in to a regular grid layout. Therefore, the interpolation procedure for a dense data set is akin to spatially averaging the data by assigning different weights to different points. This interpolation and conversion approach allows us to produce a uniform spatial design of 215 grids and made it possible to create a continuous map with the same grid size for each variable of interest (i.e. yield, applied N, and $NO_3 - N$ values).

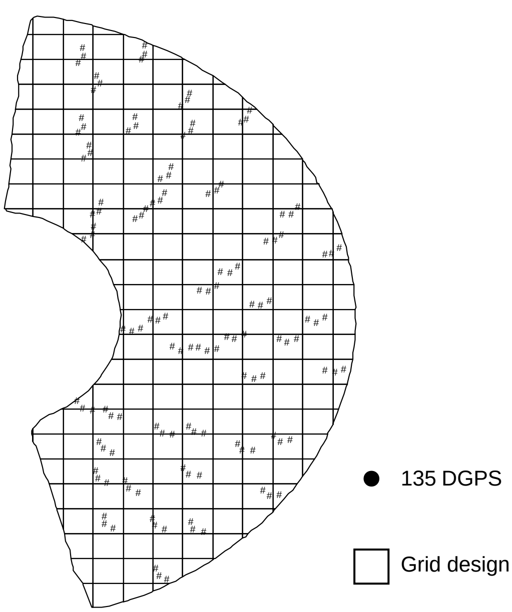

Figure 2. Sparse data vs. Grid Design, 2002, Lamesa, Texas

The kriging interpolation method was chosen over other interpolation methods because it is an approach that offers minimum error variance (Siska and Hung, 2001). In addition, kriging generates information from a linear combination of the actual information (which is similar to the other interpolation methods), but it also takes into account the spatial autocorrelation patterns inherent to the data. However, Anselin (2001) indicated that any interpolation method may introduce exaggerated spatial patterns that affect inference from the data. Further, the Purdue Site Specific Management Center (SSMC) expressed concerns with the common practice of using interpolated data to deal with differences in spatial resolution or scale because it can potentially lead to wrong decisions. As with Anselin (2001), the Purdue SSMC group points out that these interpolation procedures might introduce spatial patterns that do not exist in the original data (Erickson, 2005; Griffin, Brown, and, Lowenberg-DeBoer, 2005).

However, there does not seem to be any consensus in the literature as to whether or not the use of interpolated data is indeed not valid. For example, even in light of the potential problems with this approach, Anselin (2001) himself mentioned that kriging is still one method that may result in unbiased predictors, although this may only be true under certain assumptions (see p. 706 in Anselin, 2001).

Further note that the use of interpolated data in spatial economic analysis is an especially important issue when interpolating grid data from sparse point data (i.e. "interpolating up") because one is interpolating more grid data from less point data (i.e. in our case, generating 215 grid data from 135 soil $NO_3 - N$ point data).[7] Measurement errors and exaggerated spatial patterns are more likely

[6] Although we used a simple procedure for choosing the grid size, our visual approach was guided by recommendations from the agronomist in our project team so that the spatial pattern of the DGPS data remains consistent in the grid configuration. We believe that there is no single widely-accepted approach for choosing grid size in the agricultural economics literature (See Hertz, and Hibbard, 1993 and Bongiovanni and Lowenberg-DeBoer, 2000). In addition, given that this is not the focus of this paper, we opted for a more straightforward but agronomically consistent approach in choosing the grid size. Further analysis of the implications of using alternative grid sizes may be an interesting topic for future research.

[7] This issue is akin to the "zonation effect" of the Modified Areal Unit Problem (MAUP). See Armheim (1995).

in this type of scenario (Anselin, 2001). Spatially averaging or interpolating from a dense data to a sparser grid data (i.e. in our case, using 5592 dense yield/applied N points to generate 215 grids) is typically a less contentious issue since one is "interpolating down" and a similar approach has been used in a number of studies already (See, for example, Anselin, Bongiovanni and Lowenberg-DeBoer (2004)).

In light of the concerns regarding the use of interpolated data for economic analysis in precision agriculture, cross validation techniques and variogram analysis were used to assess how well the interpolated data coincide with the actual point data. These analytical approaches are the most common approaches used in the literature for validating interpolated data sets (See Isaaks, and Srivastava, 1989; Chiles and Delfiner, 1999; Barnes, 2004; Mueller, 2004; Panagopoulos et al., 2006). Results of the cross validation and variogram analyses suggest that, in general, the interpolated grid data based on the kriging approach accurately approximates the actual data[8]. In the interest of space, the detailed cross validation scatterplots and variogram graphs used in the analysis are not presented here but are available from the authors upon request.

EMPIRICAL METHODOLOGY AND ESTIMATION PROCEDURES

General Description of the ESDA Approach. A spatial statistics approach called Exploratory Spatial Data Analysis (ESDA) was used as the main procedure for establishing management zones. ESDA is a method that combines different techniques to visualize spatial distributions, identify patterns of different locations, and identify patterns of association between these locations. This method is based on the concept of spatial autocorrelation (or spatial dependence), which is the relationship between spatial units, and makes use of the concept of distance between locations. Positive spatial autocorrelation is the idea that locations or grids with similar values of a specific characteristic are near in space. This means that, in the presence of positive spatial autocorrelation, certain grids located close to each other share similar characteristics (Messner and Anselin, 2002, p. 10). On the other hand, patterns where neighbors of a grid have the opposite characteristics are defined as having negative spatial autocorrelation (Lambert, Lowenberg-Deboer, and Bongiovanni, 2004).

The step-by-step procedure for establishing the ESDA approach to management zone delineation can be briefly described as follows: (1) Define the 'neighborhood' structure of each grid; (2) Establish a 'weight matrix' for the 'neighborhood' structure defined; (3) Test for the presence of spatial autocorrelation; (4) Graphically visualize the spatial correlation structure (if step (3) indicates that there is spatial autocorrelation); and (5) Establish the management zones. The first step is to define the 'neighbors' of each grid. This allows one to assess if there are any spatial relationships between these grids, which can then serve as the basis for management zones. Since our data is arranged as a grid structure, typically either a "rook" structure (four neighbors to each cell, north, south, east and west) or a "queen" structure (eight neighbor to each cell) is used as the criterion for defining the neighborhood structure (Anselin, Bongiovanni, and Lowenberg-Deboer , 2004).[9]

Once the neighborhood structure was defined , the contiguity relations of each grid within a neighborhood must be formally characterized using a spatial weights matrix (Bivand, 1998). A spatial weights matrix (**W**) is an N x N (where N is the number of observations), positive definite matrix with elements w_{ij}, where w_{ij} correspond to a pair of observations at locations i and j. By convention the diagonal elements of the weight matrix are set to be zero, implying that each location is not a neighbor of itself. Non-zero elements (w_{ij} =1) means that locations i and j are neighbors. Typically, the spatial weights matrices are also row-standardized to facilitate comparison of spatial characteristics across rows.

The Moran's I statistic, which is a function of the spatial weights matrix, is then used to test for

[8] For example, the variogram analysis of the soil NO3 – N data shows a goodness-of-fit measure (R2) of 0.88, which indicates a fairly accurate representation of the data even if the soil NO3 – N in this study was interpolated based on sparse point data.

[9] As suggested by one reviewer, a more "universal" definition of a rook structure is when the neighbors of a grid (or a cell/polygon) only share a border of non-zero length. The queen structure, on the other hand, is when the neighbors of a cell can be defined as those that shares a border (including vertices). However, the definitions of the rook and queen structures are consistent with the square grid structure in our interpolated data.

the presence of spatial autocorrelation (Anselin, 1988). Specifically, the global Moran's I statistic is calculated as follows:

$$(1) \qquad I = \frac{n}{S_O}\frac{x'Wx}{x'x},$$

where n is the number of observations; $S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$; W is the spatial weight matrix; and x is the vector containing the variable of interest. The null hypothesis of the test is that there is no association between the value observed at a location and the values observed at the neighboring sites. The alternative is that the values of the neighboring sites are statistically similar.

Application of the ESDA Approach in Delineating Management Zones for Nitrogen. $NO_3 - N$ in the soil was used as our variable of interest in the Moran's I specification, given that nitrogen is one of the main limitations for cotton production in the Southern High Plains of Texas and this variable has historically been used to establish N application rates in the area. A common practice in the this area is to establish the N rate applications given the pre-plant soil $NO_3 - N$ tests at 0 to 2 foot depths (Bronson, et al. 2006). Given this common practice extractable soil $NO_3 - N$ information was used as the criteria to design the management zones under the ESDA approach. Also, this approach will allow for comparison of these results with the uniform rate N approach where the $NO_3 - N$ was used as a guide to decide the single N rate to apply uniformly across the field. This study makes a contribution to the literature in this regard because this study is the first (as far as we know) to economically evaluate a management zone delineation procedure based on a spatial autocorrelation statistic specifically calculated from soil nitrate levels.

Using soil nitrate as the variable of interest, the computed global Moran's I statistic, based on the "rook" and "queen" neighborhood structures and weights matrix defined above, are 0.796 and 0.795 respectively, both with p-value of <0.01. This indicates that null hypothesis is rejected and that there is spatial autocorrelation in the data, for both "rook", and "queen" neighborhood structures. Based on this result, a Moran scatterplot is created and three management zones based on this scatterplot is then determined (Figure 3).[10] Note that the design of the management zones does not vary with the choice of neighborhood structure. There are three management zones established based on this procedure. Management zone High - High (MZ H-H) represents high nitrate areas (i.e. grids with high nitrate levels have "neighbors" also with high nitrate levels). Management zone Low - Low (MZ L-L) represents low nitrate areas (i.e. grids with low nitrate levels have "neighbors" with low nitrate levels). Lastly, management zone High - Low - High (H-L-H) represents the area with a mix of high and low nitrate levels (i.e. grids with low nitrate levels have "neighbors" with high nitrate levels, and vice-versa). The delineated management zones could then be utilized to implement a variable rate nitrogen application program.

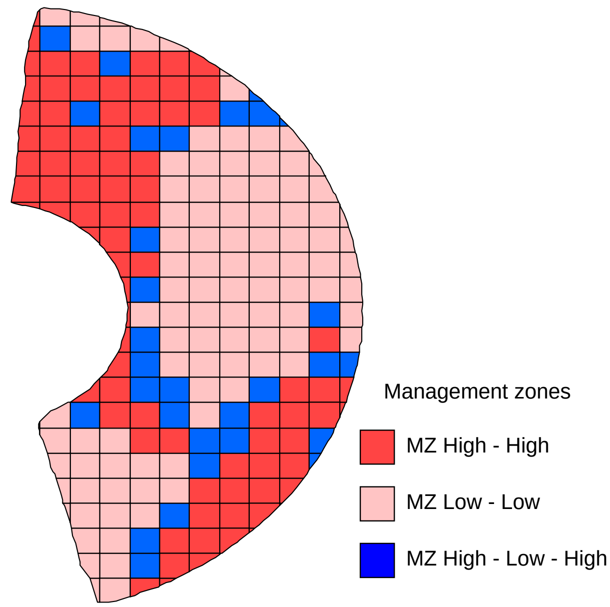

Figure 3. Delineated Management Zones based on NO_3 - N from the ESDA Procedure, 2002, Lamesa, Texas

Yield Response Estimation Methods: Spatial Econometric Approaches. To economically evaluate a variable rate fertilization program that is based on the management zone delineation procedure developed above, one must first properly estimate a yield response function for each management zone in the field. Initially

[10] For readers familiar with the Moran scatterplot, the management zones are based on the quadrants in the scatterplot. For example, one management zone can be defined for the grids in the upper right quadrant of the scatterplot because these are the grids with neighbors that have similar high values. As can be seen in the proceeding discussion, the management zone based on the cells in the upper right quadrant have neighbors with high N values (i.e. clustering of high N values) and we define this as the High-High management zone. Similar arguments can be used for establishing a management zone based on the grids in the lower left quadrant of the Moran scatterplot (i.e. clustering of grids with low N values that make this the Low-Low management zone) and the cells in the remaining quadrants (i.e. grids with clustering of opposite high-low or low-high N values). In the interest of space, the Moran scatterplot is not reported here, but is available from the authors upon request.

this yield response function can be estimated using standard ordinary least squares (OLS) regression with parameters varying based on the specified management zones. However, recent studies have indicated that spatial econometric methods may be more appropriate when estimating a yield response function with varying parameters for each management zone, especially when spatial precision agriculture data is utilized in the estimation procedure (See, among others, Anselin, Bongiovanni, and Lowenberg-Deboer, 2004; Lambert and Lowenberg-Deboer, 2000; Lambert, Lowenberg-Deboer, and Bongiovanni, 2004; Liu, Swinton, and Miller, 2006). In this article, spatial econometric methods are primarily used (or what some call, more specifically, as spatial process models) to estimate the yield response function (discussed more formally below). But OLS was also used to serve as a basis of comparison which allows one to see the potential inference errors when spatial econometric techniques are not utilized in the economic analysis.

Consistent with previous studies, a yield response function with a quadratic specification was estimated:

$$(2) \qquad Yield_i = \alpha_i + \beta_i N_i + \gamma_i N_i^2 + \varepsilon_i,$$

where $Yield_i$ is the cotton yield, N_i is the total N available to the plant (applied N + N available in the soil), α_i and β_i are parameters to be estimated, and i indexes the management zone. Note that for uniform rate N application techniques, it is assumed that the yield response is homogenous over the whole field and, therefore, equation (2) is estimated where the parameters α and β are assumed to be the same for the whole field (i.e. parameters are not varying). That is, equation (2) was estimated without taking into consideration the subscript i for all the parameters (i.e. or alternatively, it can be assumed that $i = 1$ in this case).

In contrast, for variable rate N application techniques, one assumes that there is heterogeneity in the yield response for each management zone i (where $i = 1, \dots n$) in the field. Hence, the yield response equation in (2) is estimated with varying parameters (α_i and β_i) for each management zone i. In this case, a spatial switching regression or a “spatial regimes” approach is utilized (See Anselin 1988) that jointly estimates the different parameters for the different zones defined in the previous sub-section. This model is represented in matrix form as:

$$(3) \begin{bmatrix} y_{H-H} \\ y_{L-L} \\ y_{H-L-H} \end{bmatrix} = \begin{bmatrix} x_{H-H} & 0 & 0 \\ 0 & x_{L-L} & 0 \\ 0 & 0 & x_{H-L-H} \end{bmatrix} \begin{bmatrix} \beta_{H-H} \\ \beta_{L-L} \\ \beta_{H-L-H} \end{bmatrix} + \begin{bmatrix} \mu_{H-H} \\ \mu_{L-L} \\ \mu_{H-L-H} \end{bmatrix}$$

where y, x, and μ are the yield, vector of covariates (N and N^2 in this case), and unobserved error, respectively, for each management zone – H-H, L-L, and H-L-H; and β’s are parameter vectors of the yield response function to be estimated for each management zone.

A check for the stability of the coefficients across management zones was performed by using a standard Chow test. The null hypothesis of this test is represented as follows: $\beta_{H-H} = \beta_{L-L} = {}_{H-L-H}$. If one fails to reject the null hypothesis, the response function for each management zone is the same; therefore, uniform rate approach estimation is appropriate. On the other hand, the rejection of this hypothesis demonstrates a need to divide the field into management zones because the response to N is different for each management zone.

From the estimated regression of the yield response function, the presence of spatial autocorrelation or spatial dependence in the residuals is then evaluated for both the spatial regimes (or variable rate) and the uniform rate response specification. If it is present, then appropriate spatial econometric techniques need to be implemented to account for the spatial autocorrelation in the residuals. Ignoring such autocorrelation will yield OLS estimates that are inefficient and will bias the standard errors, t-test statistics, measures of fit, and specification tests, rendering statistical inference unreliable (Anselin, 1988).

There are several spatial econometric techniques that can be used to incorporate spatial dependence in yield response function estimation -- i) a geostatistical approach, ii) a spatial regression approach, iii) a polynomial trend approach, and iv) a classical nearest neighbor approach (Lambert, Lowenberg-Deboer, and Bongiovanni, 2004). In this paper, the spatial regression approach is used to account for spatial autocorrelation in the yield response function.

Note that there is no clear consensus in the literature as to which approach is superior for modeling spatial dependence in yield response functions.[11]

---

[11] Lambert, Lowenberg-Deboer, and Bongiovanni (2004), for example, found that the parameter estimates from the different estimation procedures tend to be very similar. Hurley, Oishi, and Malzer (2005) had mixed results where they found that a spatial regression model works best in one of their fields, while a geostatistical approach is more appropriate in another field. Zimmerman and Harville (1991) and Lark (2000) further mention that controlling for spatial dependence in these models are very important, but there is no clear advantage as to which methodology to use to account for this dependence.

However, it was decided to use a spatial regression approach because it is more practical when the spatial layout of the data is a grid structure and the spatial autocorrelation can be characterized as discrete relationships between specific grids or polygons (see Lambert, Lowenberg-Deboer, and Bongiovanni, 2004).

Under the spatial regression approach, one way to incorporate spatial dependence is to include an additional independent variable called the spatial lagged dependent variable. This spatial lagged variable is simply the dependent variable weighted by the matrix containing the neighborhood structure. Formally, a spatial lag model is expressed as:

(4) $y = \rho Wy + X\beta + \varepsilon_i$,

where ρ is a spatial autoregressive coefficient, ε is a vector of error terms, W defines the weight matrix that defines the neighborhood structure, and y defines our dependent variable, which in our case is yield; X is the matrix containing all the independent variables of our model (applied N, and the square term of applied N), and β is the vector containing all the parameters associated with the independent variables. This procedure is called the *spatial lag model*, and it is mostly used when the main concern is to explicitly correct or to evaluate the existence and strength of the interaction between locations in the space.

Within the spatial regression approach, another way to incorporate spatial dependence in linear regression models is through the error structure. This method is called the *spatial error model* (SAR), and it is mostly used when the concern is correcting for the potentially biasing influence of spatial autocorrelation due to the use of spatial data (Anselin, 1999). The formal expression of the spatial error model can be written as:

(5) $y = X\beta + \mu$,

where: $\mu = \lambda W \mu + \xi$.

Note that the spatial structure of the data is represented in the specification of the error term (μ) through the parameter lambda (λ).

To determine which spatial regression model fits the data better, an empirical test based on the Lagrange Multiplier (LM) principle can be used (Anselin, 1988). Standard LM tests against a spatial error alternative and against a spatial lag alternative is used in this study. A robust version of these LM statistics is calculated as well because these robust statistics take spatial misspecifications into account. The robust version of the LM statistic for the SAR model is used to test for a spatial error process when the model specification contains a spatially lagged dependent variable. The robust version of the LM statistic for a spatial lagged model is used to test for a spatially lagged dependent variable in the presence of a spatial error process. Based on the decision rules outlined by Anselin and Florax (1995), a spatial error model is deemed more appropriate if the robust LM-error statistic is statistically significant while the robust LM-lag statistic is not.

An additional robust test used in this paper is the recently developed statistic by Kelejian and Robinson (1992). In contrast to the Moran's I and Lagrange Multiplier tests, this test does not require normality for the error terms. It is also applicable to both linear and nonlinear regressions and requires less information about the exact form of the spatial weights matrix. This is also a large sample test which means that it might not be of great power for small samples. If the LM and Robust LM tests are not conclusive, the alternative test of Kelejian and Robinson (1992) is used to find the appropriate way to account for spatial autocorrelation in the model.

The two important components that underlie the econometric results from the models above are the choice of neighborhood structure and the yield response estimation technique. The rook neighborhood structure is used as the basis for the spatial weights matrix in our delineation of the management zones and also in modeling the yield response function above. Standard OLS techniques and Kelejian and Prucha's (1999) Generalized Method of Moments (GM) approach are the techniques used to estimate the parameters of the yield response function in order to undertake the economic analysis (described in more detail below). Unlike Maximum Likelihood (ML) estimation, the GM estimation approach does not rely on a strong distributional assumption about the error term. Nevertheless, in order to check for the sensitivity of the economic results, the economic effect of using one alternative neighborhood structure and several different alternative estimation techniques were examined. Specifically, the sensitivity of our results to the use of a queen neighborhood structure and/or to the use of the following estimation procedures: Maximum Likelihood (ML), a two-step GMM approach (GM-Two step), an iterated GM approach that does not account for groupwise heteroskedasticity (GM-Iterated), and an iterated GM approach that does consider groupwise heteroskedasticity (GM-GH) were investigated. This sensitivity analysis allows for exploration of the robustness of the results to alternative neighborhood structure and estimation procedures.

The Economic Optimization Model and Net Return Evaluation. The economic model to assess the impact of the management zone delineation procedure is based on an optimization model for spatial profit (or net return) maximization.

The procedure used is consistent with previous studies that undertook economic (or profitability) analysis of precision technologies (See, among others, Lowenberg- Deboer and Boehlje, 1996; Bongiovanni and Lowenberg- Deboer, 1998; Anselin, Bongiovanni, and Lowenberg-Deboer, 2004; Bullock, Lowenberg-DeBoer, and Swinton, 2002). Specifically, the estimated parameters of the cotton yield response functions are used to formulate an economic optimization model that maximizes profit for a representative field. In this model, net returns were maximized over the following costs: fertilizer cost, cost of soil samples, and cost of retrofitting equipment appropriate for each case ($0 acre $^{-1}$ for uniform rate and $1.69 acre $^{-1}$ for variable rate, see Bronson et. al, 2006). Together with the estimated yield response parameters, information about cotton output price, N fertilizer price, and the fixed costs for cotton production in the Texas High Plains were collected and used in the economic optimization model described in detail below (Bronson et al. 2006).

More formally, the economic optimization model is:

$$\text{(6)} \quad \text{Max}\pi = \sum_{i=1}^{m} (A\omega_i[P_c(\alpha_i + \beta_i N_i + \gamma_i N_i^2) - r_N N_i - ø])$$

where:

π = Total net returns over N fertilizer and fixed cost ($)

A = total area (34.59 acres)

ω_i = Proportion of total land area allocated to management unit i (i.e. for the management zones based on the spatial statistics approach developed in this study, MZ H-H = 40%, MZ L-L = 47%, MZ H-L-H = 13%; for a uniform rate approach ω_i will simply be equal to one.)

i = Management unit (either the whole field or the management zones)

m = Total number of management units (m = 1 for uniform rate approach and m = 3 for a variable rate approach based on the management zones delineated using the procedure in this paper)

P_c = Price of cotton ($0.47 per lb, see Bronson et. al, 2006)

N_i = Qty. of N applied in management unit i (in lbs/acre); N is the choice variable

r_N = Price of N fertilizer applied ($0.35/lb, see Bronson et. al, 2006)

ø = Fixed cost (For the uniform rate, this cost includes the cost of regular N analysis of a soil sample ($0.3 acre $^{-1}$). For the variable rate, this cost includes cost of retrofitting the equipment for variable rate N application plus the cost of N analysis of the soil samples ($ 9.71 acre $^{-1}$), see Bronson et al. 2006).

Using the framework in equation (6), the net returns were computed from the following scenarios: (1) a uniform N rate application based on the recommended rate by Bronson et al. 2006, which was defined as the uniform rate agronomic (URA), (2) a uniform N rate application based on an economic optima (URE), (3) a variable rate N application based on the economic optima for each of the management zones established through the spatial statistics procedure above (VRN), and (4) a variable rate N application based on the economic optima for each of the management zones delineated based on landscape position (VRL).[12]

[12] Scenarios 1 and 2 implicitly assumethat the producer does not have information about the variability of the field and, therefore, could only apply uniformly. Scenarios 3 implicitly assumes that the producer utilizes the information about the management zones delineated using the approach in this paper to variable apply N. Scenario 4 implicitly assumes that the producer utilizes information about landscape position to variably apply N. As mentioned in the introduction, the VRL scenario (scenario 4) is included in the analysis so that we can compare: (1) the net returns from a variable rate application program based on the delineation procedure designed in this study, versus (2) the net returns from a variable rate application program that uses the more common approach of dividing the field into different management zones based on landscape position. The three zones in the VRL scenario are: the south-facing side slope (MZS), the north-facing side slope (MZN), and the bottom slope (MZB) (See Figure 4).

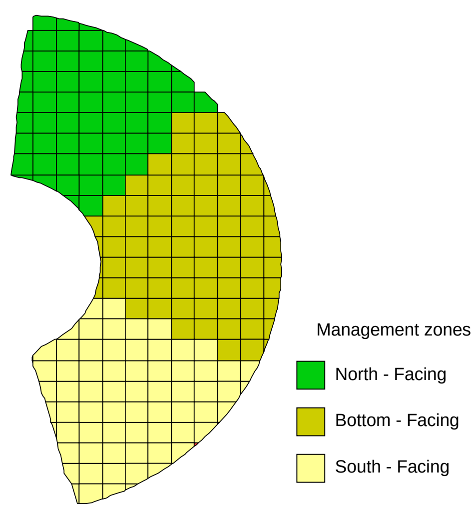

Figure 4. Delineated Management Zones based on landscape position, 2002, Lamesa, Texas

For the URA scenario, the net return was calculated for this scenario by plugging in the agronomic N recommendation of 52 lbs/acre (See Bronson et al., 2006 for the agronomic basis of this recommendation) to the profit maximization model (in equation (6)) where the estimated parameters of the uniform rate response function is utilized. For the URE scenario, the parameters of the uniform rate response function is again used to first implement the maximization algorithm in equation (6) in order to calculate the economically optimal uniform N rate application (N^*) for this scenario. Once the actual value of N^* is found, this value is then plugged back in to the objective function in equation (6) to get the net returns for the URE scenario.

For the VRN scenario, the yield response function estimated with management zone-specific parameters is used to first ascertain the economically optimal variable N rates for each zone (i.e. N^* for the H-H zone (N^*_{H-H}), N^* for the L-L zone (N^*_{L-L}), and N^* for the H-L-H zone (N^*_{H-L-H})). These optimal variable N rates (for each portion of the field) are then inserted in the objective function of equation (6) to calculate the net returns for the VRN scenario. Note that the net returns for VRN are calculated in (6) using the variable rate yield response parameters of the management zones delineated using the spatial statistics procedure developed in this paper. As with the VRN scenario, the optimal variable N rates for the VRL scenario needs to be found first (in this case, optimal rates for the zones based on the three landscape positions: the south-facing side slope (N^*_{South}), the north-facing side slope (N^*_{North}), and the bottom slope (N^*_{Bottom})). These optimal variable rates are then plugged into the objective function in equation (6), where the parameters of the yield response function based on the landscape position delineation is used. The net returns for all four scenarios are then compared to examine which application and zone delineation procedure generated the highest net returns.

RESULTS AND DISCUSSION

Yield Response Estimation Results and Model Diagnostics. The yield response estimation results for the OLS and SAR specifications of the different N application approaches (uniform, VRN, and VRL) are summarized in Tables 1, 2, and 3.

Table 1. Parameter estimates of the cotton yield response function for uniform rate case (URA and URE), 2002, Lamesa, Texas

Variables	OLS (Ordinary Least Squares)		SAR-GM-Iterated (Spatial Error Model)	
	Parameter estimate	P-value	Parameter estimate	P-value
Constant	987.35	0.0000	824.92	0.0000
N	0.23	0.9298	3.59	0.1207
N^2	-0.01	0.6169	-0.02	0.0414
Lambda[1]	NA	NA	0.35	0.0000

Measures of fit	OLS	SAR-GM-Iterated
Log Likelihood	-1489.17	NA
AIC[2]	2984.35	2982.81

Diagnostic tests	d.f.	Value	P-value	Value	P-value
Jarque – Bera (Normality)	2	81.00	0.0000		
Koenker-Basset test (heteroskedasticity)	2	2.03	0.3616	NA	NA
Lagrange multiplier(error)	1	26.72	0.0000	NA	NA
Robust LM(error)	1	3.41	0.0649	NA	NA
Kelejian-Robinson (error)	3	20.75	0.0001		
Lagrange multiplier (lag)	1	25.10	0.0000	NA	NA
Robust LM (lag)	1	1.79	0.1808	NA	NA

[1] **Lambda is the estimated spatial autoregressive coefficient that characterizes the spatial structure of the unobserved error terms in the spatial error model.**

[2] **Akaike's Information Criteria (AIC) is calculated using Hurvich and Tsai's (1989) correction factor to reduce bias: AIC= $2n \ln(\sigma) + n \ln(2\Pi) + n((n+k)/(n-2-k))$.**

These results were estimated using the SpaceStat® software. Starting with the estimates for OLS in Table 1, the coefficients of both N and N^2 have the expected signs but they are not significant (except for the constant term). In Table 2, the OLS estimates for the coefficients of both N and N^2 for management zones H-H and L-L also have the expected signs, but the H-L-H management zone does not. Only four of the coefficients estimated by OLS are significant in Table 2.

In Table 3, the estimated response function is presented for the case of management zones delineated based on landscape position (south facing (MZS), north facing (MZN), and bottom (MZB)). Using OLS, all the coefficients follow an a priori expectation except for the coefficients associated with the north facing landscape position. Only four of the coefficients estimated by OLS are significant in Table 3. According to the Chow test for structural stability, we fail to reject the null hypothesis in this case, which means that the OLS coefficients are stable over all the management zones in Tables 2 and 3. These results suggest that a uniform single response function will be appropriate for this cotton production function. But note that this stability becomes invalid with OLS if spatial autocorrelation in the residuals is present (Anselin 1988).

Table 2. Parameter estimates of the cotton yield response function for the management zones delineated using the spatial statistics approach (Spatial regimes), 2002, Lamesa, Texas

Variables	OLS (Ordinary Least Squares)		SAR-GM-GHET (Spatial Error Model- Groupwise Heteroskedasticty)	
	Parameter estimate	P-value	Parameter estimate	P-value
MZ H - H	449.92	0.3485	432.89	0.3696
N x MZHH	8.71	0.2324	9.82	0.1753
N^2 x MZHH	-0.04	0.1840	-0.04	0.1125
MZ L - L	668.03	0.0027	483.32	0.0030
N x MZLL	9.35	0.0778	12.56	0.0011
N^2 x MZLL	-0.06	0.0318	-0.08	0.0004
MZ H – L - H	1491.17	0.0123	1303.69	0.0092
N x MZHLH	-8.43	0.4553	-5.32	0.5736
N^2 x MZHLH	0.03	0.5912	0.02	0.7200
Lambda[1]	NA	NA	0.37	0.0000

Measures of fit	OLS	SAR-GM-GHET
Log Likelihood	-1483.72	NA
AIC[2]	2985.44	2973.785

Diagnostic tests	d.f.	Value	P-value	Value	P-value
Jarque-Bera Normality	2	61.16	0.0000	NA	NA
Koenker-Basset Test	2	5.99	0.0500	NA	NA
Lagrange multiplier(error)	1	25.09	0.0000	NA	NA
Robust LM(error)	1	0.14	0.7075	NA	NA
Kelejian-Robinson (error)	9	33.81	0.0001	NA	NA
Lagrange multiplier (lag)	1	25.29	0.0000	NA	NA
Robust LM (lag)	1	0.34	0.5595	NA	NA
Chow, Chow-Wald Test	6, 206	1.79	0.1032	12.26	0.0564

[1] **Lambda is the estimated spatial autoregressive coefficient that characterizes the spatial structure of the unobserved error terms in the spatial error model.**

[2] **Akaike's Information Criteria (AIC) is calculated using Hurvich and Tsai's (1989) correction factor to reduce bias: AIC= $2n \ln(\sigma) + n \ln(2\Pi) + n((n + k)/(n - 2 - k))$.**

Table 3. Parameter estimates of the cotton yield response function for the management zones delineated using the landscape position approach, 2002, Lamesa, Texas

Variables	OLS (Ordinary Least Squares)		SAR-GM-Iterated (Spatial Error Model)	
	Parameter estimate	P-value	Parameter estimate	P-value
MZS	901.70	0.0111	865.90	0.0084
N x MZS	1.09	0.8485	2.88	0.5808
N^2 x MZS	-0.01	0.7254	-0.02	0.3528
MZN	1216.27	0.0000	853.40	0.0000
N x MZN	-5.51	0.2499	2.07	0.6109
N^2 x MZN	0.02	0.2957	-0.01	0.5658
MZB	721.01	0.0036	729.43	0.0013
N x MZB	7.45	0.1392	7.03	0.1336
N^2 x MZB	-0.05	0.0683	-0.04	0.0710
Lambda[1]	NA	NA	0.36	0.0000

Measures of fit	OLS	SAR-GM-Iterated
Log Likelihood	-1486.27	NA
AIC[2]	2990.54	2982.35

Diagnostic tests	d.f.	Value	P-value	Value	P-value
Jarque-Bera Normality	2	88.12	0.0000	NA	NA
Koenker-Basset Test	2	0.97	0.6166	NA	NA
Lagrange multiplier(error)	1	20.70	0.0000	NA	NA
Robust LM(error)	1	0.03	0.8689	NA	NA
Kelejian-Robinson (error)	9	20.19	0.0168	NA	NA
Lagrange multiplier (lag)	1	21.35	0.0000	NA	NA
Robust LM (lag)	1	0.68	0.4105	NA	NA
Chow, Chow-Wald Test	6,206	0.94	0.4671	2.15	0.9053

[1] **Lambda is the estimated spatial autoregressive coefficient that characterizes the spatial structure of the unobserved error terms in the spatial error model.**

[2] **Akaike's Information Criteria (AIC) is calculated using Hurvich and Tsai's (1989) correction factor to reduce bias: AIC= $2n \ln(\sigma) + n \ln(2\Pi) + n((n + k)/(n - 2 - k))$.**

Diagnostics for spatial autocorrelation in all of the OLS models indicated that this problem is present and that a spatial error model is the proper specification to correct for spatial autocorrelation. The standard LM and robust-LM tests in Table 1 clearly show that the spatial error model is the proper specification for the uniform rate (Anselin and Florax, 1995). However, for Tables 2 and 3 the traditional LM and robust LM diagnostics statistics for spatial dependence are not conclusive. Both the robust LM-error and the robust LM-lag statistics are not significant. Given these results the more recent Kelejian-Robinson test is used to decide the proper specification needed to correct the spatial dependence in these models (Kelejian and Robinson, 1992). In this case, the Kelejian and Robinson test points to the spatial error specification as the appropriate model to correct for spatial autocorrelation in this case. In addition, standard diagnostics for groupwise heteroskedasticity (across zones) suggests the presence of this misspecification for the case of spatial variable rate response function in Table 2 (See the Koenker and Bassett test). For this particular case, both spatial autocorrelation as well as heteroskedasticity are accounted for in the SAR model in Table 2.

After accounting for spatial autocorrelation in the uniform rate model (Table 1) the coefficients estimates for the constant, N and N^2 coefficients change relative to the OLS estimates and they are all significant (at the 5% level of significance for N and marginally significant for N^2 at 12% level). The fit of the model improves when the spatial error structure is modeled, as indicated by a decrease in the Akaike Information Criteria (AIC) from 2984.35 to 2982.81. The improvement of the model was also to be expected because of the highly significant spatial error coefficient (i.e. the lambda (λ) coefficient in equation 5).

Both spatial autocorrelation and groupwise heteroskedasticity are accounted for in the spatial error model in Table 2. The coefficients in this case vary slightly relative to the OLS estimates, but the significance levels for the individual coefficients improve. Under the spatial error model all the coefficients are significant (at a lower significance levels that varies from 1% to 10%) except for the constant term and N coefficient for the H-H management zone and the N and N^2 coefficients associated with MZ H-L-H (where there is no pattern of spatial association given the Moran's I results). After correcting for spatial autocorrelation and groupwise heteroskedasticity, the Chow test presented in Table 2 shows evidence of structural instability. This suggests that allowing for different yield response functions for each management zone is the proper functional specification. Again, the fit of the model improves when the spatial error structure is modeled as indicated by a decrease in AIC from 2985.44 to 2973.785.

After incorporating spatial autocorrelation in the VRL response function presented in Table 3 (management zones based on landscape position), all the coefficients are consistent with our expectations, but there is no improvement in the significance levels. The Chow test (for both OLS and spatial error models) shows stability of the estimated coefficients among all the landscape management zones. This result indicates that a single yield response function for the whole field may be appropriate and, consequently, that a uniform rate approach may be preferred over the variable rate approach based on landscape position. Nevertheless, these VRL parameter estimates are used in the optimization models in order to compare the net returns from a variable rate N application using landscape position versus a variable rate N application based on the spatial statistics approach to management zone delineation.

Estimated Net Returns, Yields, and N Levels. A comparison of the returns from different N rates is given in Table 4. Based on the estimated response function(s) and the optimization model described in the previous section, point estimates of the yield, the optimal N application levels, and the corresponding net returns for each of the different application techniques considered – URA, URE, VRN, and VRL were calculated. The URA was used to represent the N rates currently practiced in Southern High Plains of Texas. Each of these application scenarios was examined by using a yield response function estimated both by OLS and by using the spatial error model (SAR) estimated by Generalized Moments (GM). This allows the potential magnitude of inference or recommendation errors that could be committed to be seen, when spatial autocorrelation is not properly accounted for in the yield response estimation.

In general, the OLS technique tends to underestimate the difference between the net benefits from the spatial variable rate application (VRN) and the uniform rate application procedures (see Table 4a). On the other hand, the difference between the estimates using VRL approach versus the uniform rate approaches tend to be overestimated when OLS is

Table 4a. Estimated net returns under uniform rate and variable rate application methods and different estimation procedures, 2002, Lamesa, Texas

	OLS	SAR-GM-Iterated/SAR-GM-GHET	Difference (OLS-SAR)
	Net Returns ($/acre)		
Uniform rate, agronomic optimum (URA, 52 lb/ac)	443.69	428.68	15.01
Uniform rate, economic optimum (URE, 70.16 lb/ac)	463.75	431.44	32.31
Variable rate, spatial statistics approach (VRN)	458.20	446.88	11.32
Variable rate, landscape position (VRL)	484.70	433.86	50.84
Differences across application techniques			
URE vs. URA (URE – URA)	20.07	2.75	17.32
VRN vs. URA (VRN – URA)	14.52	18.20	-3.68
VRN vs. URE (VRN – URE)	-5.55	15.44	-20.99
VRL vs. URA (VRL – URA)	41.01	5.18	35.83
VRL vs. URE (VRL – URE)	20.95	2.42	18.53
VRN vs. VRL (VRN – VRL)	-26.50	13.02	-39.52

used. These results reinforce findings in previous studies that incorrectly used OLS to estimate yield response functions (i.e. OLS may provide misleading findings). Hence, using the more appropriate yield response function that takes spatial autocorrelation into account, it is found that the VRN approach tends to have a higher net return relative to the URA ($18.20/acre) and URE scenarios ($15.44/acre), respectively. Thus, the economic optimization results show that the returns to the VRN approach can more than cover the variable fertilizer cost, as well as the other fixed costs considered in this analysis. The net returns from a VRL approach also tend to be higher relative to the uniform rate approaches. Furthermore, the VRN approach based on the spatial statistics approach to management zone delineation is also shown to be more profitable than VRL ($13.02/acre). Note that this result is consistent with the a priori expectation since VRN is a more precise approach and because the field has more variability than what is reflected in the landscape position.

The estimated yields and N levels for the different N fertilizer application scenarios considered are presented in Tables 4a and 4b. When spatial autocorrelation is considered in the yield response estimation, the VRN scenario based on the spatial approach to management zone delineation tend to have higher yields relative to all the other application techniques. The average N application levels used in the VRN approach also tend to be higher relative to the other application procedures (when spatial autocorrelation is accounted for). Note,

Table 4b. Estimated yields under uniform rate and variable rate application methods and different estimation procedures, 2002, Lamesa, Texas

	OLS	SAR-GM-Iterated /SAR-GM-GHET	Difference (OLS-SAR)
	--- Yield (lb/acre) ---		
Uniform rate, agronomic optimum (URA, 52 lb/ac)	983.37	951.45	31.92
Uniform rate, economic optimum (URE, 70.16 lb/ac)	987.35	969.32	18.03
Variable rate, spatial statistics approach (VRN)	1055.47	1033.46	22.01
Variable rate, landscape position (VRL)	1059.18	972.96	86.22
Differences across application techniques			
URE vs. URA (URE – URA)	3.98	17.87	-13.89
VRN vs. URA (VRN – URA)	72.09	82.01	-9.92
VRN vs. URE (VRN – URE)	68.12	64.14	3.98
VRL vs. URA (VRL – URA)	75.81	21.51	54.30
VRL vs. URE (VRL – URE)	71.83	3.64	68.19
VRN vs. VRL (VRN – VRL)	-3.71	60.50	-64.21

Table 4c. Estimated N levels under uniform rate and variable rate application methods and different estimation procedures, 2002, Lamesa, Texas

	OLS	SAR-GM-Iterated/SAR-GM-GHET	Difference (OLS-SAR)
	--- N application (lb/acre) ---		
Uniform rate, agronomic optimum (URA, 52 lb/ac)	52.00	52.00	0.00
Uniform rate, economic optimum (URE, 70.16 lb/ac)	0.00	68.13	-68.13
Variable rate, spatial statistics approach (VRN)	75.63	78.43	-2.80
Variable rate, landscape position (VRL)	31.79	62.97	-31.18
Differences across application techniques			
URE vs. URA (URE – URA)	-52.00	16.13	-68.13
VRN vs. URA (VRN – URA)	23.63	26.43	-2.80
VRN vs. URE (VRN – URE)	75.63	10.30	65.33
VRL vs. URA (VRL – URA)	-20.21	10.97	-31.18
VRL vs. URE (VRL – URE)	31.79	-5.16	36.95
VRN vs. VRL (VRN – VRL)	43.84	15.45	28.39

however, that the VRN scenario tends to more efficiently utilize N because it applies less N in zones with high soil nitrate levels and more N in zones with low soil nitrate levels (results available from the authors upon request). Therefore, even if N application is higher (on average) for the VRN relative to URA, URE, and VRL, the more efficient use of the N fertilizer may still possibly reduce nitrate run-off in the soil and, consequently, reduce non-point source pollution. This is an empirical question beyond the scope of this study.

Sensitivity Analysis. As a sensitivity analysis, the effect of using an alternative neighborhood structure and alternative estimation techniques on the difference in the net returns across application scenarios (Table 5) was examined. In general, it was found that changing the estimation technique does not significantly affect the signs and the magnitudes of the difference in net returns. The economic inference that the VRN approach has higher returns relative to the other approaches still holds. Using the queen structure as the neighborhood specification, however, produced major changes in the magnitudes of the difference in net returns. Note that the economic inference of higher returns relative to the uniform rate approaches still hold in this case, but the inference of the VRN providing higher returns relative to the VRL approach do not hold anymore. Hence, it seems that these results are robust to the estimation method but may be sensitive to the choice of neighborhood structure. This result indicates that further investigation of the effects of neighborhood choice may be warranted, but this is beyond the scope of this study. Nevertheless, a robust result was found in that the VRN approach tends to produce higher net returns relative to the uniform rate approaches (URA and URE) regardless of neighborhood structure and estimation method.

Table 5. Sensitivity of the differences in net returns under alternative neighborhood structure and estimation method assumptions, 2002, Lamesa, Texas

Neighborhood structure[1]	Difference in net returns ($/acre) across application techniques[2]					
Estimation Method[3]	URE-URA	VRN-URA	VRN-URE	VRL-URA	VRL-URE	VRN-VRL
Rook Structure						
OLS	20.07	14.52	-5.55	41.01	20.95	-26.50
SEM(ML)	3.25	22.98	19.73	5.92	2.51	16.59
SEM-GH (ML)	3.41	18.14	14.73	7.07	3.82	12.81
SEM (GM-Two step)	2.04	18.09	16.05	5.63	3.59	13.04
SEM (GM-Iterated)	2.75	18.15	15.39	5.17	2.41	13.57
SEM (GM-GH)	2.63	19.20	16.57	6.19	3.56	13.60
Queen Structure						
OLS	20.07	14.52	-5.55	41.01	20.95	-26.50
SEM(ML)	0.12	7.60	7.48	0.11	-0.18	-2.38
SEM-GH (ML)	0.29	3.92	3.63	0.83	0.71	-5.86
SEM (GM-Two step)	0.25	6.32	6.07	3.84	3.58	-6.93
SEM (GM-Iterated)	0.06	5.47	5.41	0.54	0.48	-4.48
SEM (GM-GH)	0.01	6.09	6.08	1.29	1.28	-4.62

Note:

[1] **The neighborhood structures considered are rook and queen. Note that these structures are assumed both in the delineation of the management zones for the spatial approach and in specifying the error structure in the SEM model;**

[2] **The alternative estimation methods considered (aside from the traditional OLS and SEM (ML)) are: SEM groupwise heteroskedasticty estimation (SEM-GH),SEM using two stage Generalized Moments (GM-Two step), SEM using iterated Generalized Moments (GM-Iterated), and SEM using Generalized Moments considering groupwise hetroskedasticty (GM-GH).**

[3] **Application techniques are: uniform rate based on agronomic optimum (URA), uniform rate based on economic optimum (URE), variable rate based on the global spatial approach (VRN), and variable rate based on landscape position (VRL).**

CONCLUSIONS AND RECOMMENDATIONS

Based on an ESDA approach that utilizes a spatial autocorrelation statistic, a procedure was developed for delineating management zones using precision agriculture data from cotton in the Texas high plains. The ESDA approach to management zone delineation is a statistical method that could serve as a guide for producers to recognize relevant spatial patterns in their field and to manage it more effectively. An optimization model evaluates the economic impact of more precisely applying N fertilizer based on these management zones versus (a) the more conventional method of using a uniform rate for the whole field, and (b) an alternative variable rate application procedure based on landscape position. The main input of this optimization model is a cotton yield response function estimated using spatial econometric techniques that accounts for spatial dependence in the residuals

The results of this analysis reinforce observations in past studies that incorrectly estimating yield response functions without correcting for spatial dependence may lead to misleading inferences about the economic impact of variable rate technologies. In this study, an approach that uses standard OLS regression to estimate the yield response function tends to underestimate the net benefits of variable rate application in cotton production. But when a yield response function is used that accounts for spatial autocorrelation, it is found that applying variable N rates based on the management zones delineated (using the spatial statistics approach) provides higher net returns relative to conventional uniform rate application irrespective of neighborhood structure and estimation method. The variable rate approach using the delineated management zones also produce higher net returns as compared to a variable rate application based on landscape position, but this result does not seem to be invariant to the choice of neighborhood structure assumed in the analysis. As mentioned above, further analysis that examines sensitivity of economic inferences under different neighborhood structures may be a worthwhile undertaking in the future. Recent advances in spatial econometric theory may provide future directions on how to select the appropriate neighborhood structure to impose for a particular data set (Holloway et al., 2006).

While the presented VRT profitability results are interesting, there are a couple of caveats to keep in mind when interpreting these results. First, the analysis done in this paper is “ex post”. This means that the ex post economic calculations are used to understand the implications of the estimated response on profitability, but they do not necessarily represent rules for an ex ante economic analysis. The second caveat is that the results from this study pertain only to a single year. If the response function is stable from one year to another, then one can take the one year analysis as a representative approach for the subsequent years. However, if variability is high from year to year, the results are only representative for a given state of nature observed at certain point in time (Anselin, Bongiovanni, and Lowenberg-DeBoer, 2004). Hence, a multi-year analysis would be an interesting extension of this study.

Another area for future applied work includes the development and refinement of a user-friendly management zone decision tool based on the spatial statistics approach described above. The challenge is developing an interactive interface that is understandable to producers and is easy to use. Producers that have site-specific data for their fields may then be able to utilize this tool to delineate management zones for more precise management of their farm. Collaboration with Extension faculty is needed for the successful development of this tool and for the dissemination of the tool to producers.

REFERENCES:

Anselin, L. *Spatial Econometrics: Methods and Models*. Dordrecht, Netherlands: Kluwer Academic, 1988.

Anselin, L. “Spatial Econometrics.” Working paper, Bruton Center, School of Social Sciences, University of Texas at Dallas, Richardson, TX, 1999.

Anselin L. “Spatial Effects in Econometric Practice in Environmental and Resource Economics” *American Journal of Agricultural Economics* 83 (August 2001): 705-710.

Anselin L., Bongiovanni R. and Lowenberg-DeBoer J. “A spatial econometric approach to the economics of site-specific nitrogen management in corn production.” *American Journal of Agricultural Economics* 86 (August 2004): 675-687.

Anselin, L and Florax, R. *New Directions in Spatial Econometrics*. Berlin: Springer-Verlag, 1995.

Armhein C 1995 Searching for the elusive aggregation effect: Evidence from statistical simulations. *Environment & Planning A*, Jan95, Vol. 27 Issue 1, p105

Basnet, B., Kelly, R., Jensen, T., Strong, W., Apan, A. and Butler, D. "Delineation of Management Zones using Multiple Crop Yield Data." Paper presented at the International Soil Tillage Research Organization Conference, 16th Triennial Conference, Brisbane, Australia, July 13 -18, 2003.

Barnes, R. "Variogarm Tutorial." Golden Software, Inc. 2004.

Bivand, R. "A review of spatial statistical techniques for location studies." Internet site: http://www.nhh.no/geo/gib/gib1998/gib98-3/lund.html (Accessed October 12, 2003).

Bullock, D. S., Lowenberg-Deboer, J. Swinton, S. "Adding value to spatially managed inputs by understanding site-specific yield response." *Agricultural Economics* 27(November 2002): 233- 245.

Bongiovanni, Rodolfo, and J. Lowenberg-DeBoer, "Economics of Variable Rate Lime in Indiana," *Precision Agriculture* 22: 1 (2000), p. 55-70.

Bronson, K., Keeling, W., Booker, J.D., Chua, T., Wheeler, T., Boman, R., and Lascano, R. **"**Site-Specific irrigation and Nitrogen Management for Cotton Production in the Southern High Plains" *Agronomy Journal* 98(2006): 212–219.

Chiles, J.P., and Delfiner, P. "Geostatistics: Modeling Spatial Uncertainty" John Wiley and Sons, New York, 695 pp. 1999

Diker, K., Bucheleiter, G. W., Farahani, H.J., Heermann, D.F., Brodahl, M.K. "Frecuency analysis of yield for delineating management zones." Paper presented at the 6th Intrenational Conference on Precision Agriculture and other precision Management, Minneapolis, MN, July 14 – 17, 2002.

Dillon, C., Mueller, T., Shearer, S. "An economic optimization model for management zone configuration" Paper presented at the 4th European Conference on Precision Agriculture, Berlin, Germany, June 15 -19, 2003.

Erickson, B. *Step by Step yield monitor data analysis*. Site Specific Management Center (SSMC) Newsletter, Purdue University, August 2005.

Fleming, K.L., Westfall, D.G., and Bausch, W.C. "Evaluating Management Zone Technology and Grid Soil Sampling for Variable Rate Nitrogen Application." Paper presented at the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, July 16-19, 2000.

Franzen, D.W. , Halvorson, A.D., and Hofman, V.L. "Management Zones for Soil N and P Levels in The Northern Great Plains." Paper presented at the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, July 16-19, 2000.

Fridgen, J.J., Kitchen, N.R, and Sudduth, K.A. "Variability of Soil and Landscape Attributes Within Sub – Field Management zones." Paper presented at the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, July 16-19, 2000.

Griffin, T., Brown, J.P., and Lowenberg-DeBoer, J. 2005. Yield Monitor Data Analysis: Data Acquisition, Management, and Analysis Protocol. Available on line at http://www.purdue.edu/ssmc

Griffin, T., Lowenberg-DeBoer, J., Lambert, D.M., Peone, J., Payne, T.,and Daberkow, S.G. "Adoption, profitability, and making better use of precision farming data." Staff paper, Department of Agriculture Economics, Purdue University, June , 2004.

Holloway, G., D. Lacombe, and J. LeSage. "Spatial Econometric for Bio-Economic and Land-Use Modeling." Invited Paper presented at the 2006 IAAE Meetings, Gold Coast, Australia (August 12-18, 2006).

Hurley, T., Oishi, K., G., Malzer, G. "Estimating the potential value of variable rate nitrogen applications: A comparison of spatial econometric and geostatistical models." *Journal of Agricultural and Resources Economics* 30 (2) (2005): 231-249.

Hurvich, C.C., Tsai C."Regression and time series model selection in small samples." *Biometrika* 76(2) (1989): 297-307

Isaaks, E.H. and Srivastava, R.M. *Applied Geostatistics.* New York: Oxford University Press, Inc., 1989.

Kelejian, H., and Robinson, D.P., "Spatial autocorrelation: a new computationally simple test with an application to per capita county policy expenditures." *Regional Science and Urban Economics* 22 (1992): 317-331.

Kelejian, H., and Prucha, I.R., "A Generalized Moments Estimator for the Autorregresive Parameter in a Spatial Model." *Int. Econ. Rev.*40 (1999): 509-533.

Lambert, D.M., and Lowenberg-DeBoer, J. "Precision Agriculture Profitability Review" Working paper, Site Specific Management Center, School of Agriculture, Purdue University. 2000.

Liu, Y., Swinton, S., and Miller, N. "Is Site Specific Yield Response Consistent over time? Does it Pay?" *American Journal of Agricultural Economics* 88 (May 2006): 471-483.

Lowenberg-DeBoer, J. and M. Boehlje. "Revolution, Evaluation or Deadend: Economic Perspectives on Precision Agriculture." Paper presented at the 3rd International Conference on Precision Agriculture. Minneapolis, MN, June 23-26, 1996.

Lambert, D.M., Lowenberg-DeBoer, J., and Bongiovanni, R. "A Comparison of Four Spatial Regression Models for Yield Monitor Data: A case study from Argentina." *Precision Agriculture* (5) (2004): 579-600

Messner, S. and Anselin, L. "Spatial analyses of homicide with aerial data." Working paper, University of Illinois at Urbana-Champaign, 2002.

Mueller, T.G., Pusuluri, N.B., Mathias, K.K., Cornelius, P.R., Barnhisel, R.I., and Shearer, S.A. "Map Quality for Ordinary Kriging and Inverse Distance Weighted Interpolation" Soil Sci Soc Am J 68 (2004): 2042-2047.

Nolan, S.C., Goddard, T.W., Lohstraeter, G., and Coen, G.M. "Assessing Management Units on Rolling Topography." Paper presented at the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, July 16-19, 2000.

Panagopoulos, T., Jesus, J., Antunes, M.D.C., and Beltrão, J. "Analysis of spatial interpolation for optimizing management of a salinized field cultivated with lettuce." *European Journal of Agronomy* 24 (January 2006):1-10.

Ping, J.L., Green C.J., Bronson K.F., Zatman R.E., and Dobermann A. "Delineating potential management zones for cotton based on yields and soil properties." *Soil Sci Soc Am J* 170(May 2005): 371 – 385.

Siska, P. and Hung, K. "Assessment of Krigging Accuracy in the GIS Environment" Presented at The 21st Annual ESRI International Conference, San Diego, CA, 2001.

Stafford, J., Lark R.M., and Bolam H.C. "Using Yield Maps to Regionalize Fields into potencial Management Units." Paper presented at the 4th International Conference on Precision Agriculture, St. Paul, Minnesota, July 19-22, 1998.

Thrikawala, S., A. Weersink, and G. Kachanoski. "Management Unit Size and Efficiency Gains from Nitrogen Fertilizer Application." *Agricultural Systems*.56(April 1998):513-531.

Tobler, W. "Cellular geography." *Philosophy in Geography*. Gale, S., Olsson, G., eds., Reidel, Dordrecht, 1979.

Zhang, H., B. Raun, J. Hattey, G. Johnson, and N. Basta. 1998. *OSU soil test interpretations*. Publ. F-2225. Oklahoma Coop. Ext. Serv., Oklahoma State Univ., Stillwater.