Modeling local and global spatial correlation in field-scale experiments

by **Griffin, T.W., Fitzgerald, G.J., Lowenberg-DeBoer, J. and Barnes, E.M.**

DOI: https://doi.org/10.1002/agj2.20266

Griffin, T.W., Fitzgerald, G.J., Lowenberg-DeBoer, J. and Barnes, E.M. Modeling local and global spatial correlation in field-scale experiments. *Agronomy Journal*.
22 April 2020

Running title: local spatial spillovers for landscape-scale experiments

Core ideas:

1. Accounting for local spatial correlation important to estimation
2. Cross regressive variables may explicitly model treatment edge effects
3. Alternative experimental designs may be acceptable given proper modeling

Modeling Local and Global Spatial Correlation in Field-Scale Experiments

Abbreviations:
GIS, geographic information system;
GM, general moments;
GMM, general method of moments;
GNSS, global navigation satellite system;
LM, Lagrange multiplier;
LSC, local spatial correlation;
ML, maximum likelihood;
OLS, ordinary least squares;
SMA, spatial moving average process model;
SEM, spatial error process model

Author list and affiliations:

Terry W. Griffin[1], **Glenn J. Fitzgerald**[2], **James Lowenberg-DeBoer**[3], **Edward M. Barnes**[4]

[1] Department of Agricultural Economics, Kansas State University, Manhattan, Kansas, USA
[2] Senior Research Scientist, Agriculture Victoria, Horsham, Victoria, AUS
[3] Harper Adams University, Newport, Shropshire, UK
[4] Cotton Incorporated, Cary, North Carolina, USA

ABSTRACT

Precision agriculture has renewed the interest of farmers and researchers to conduct on-farm planned comparisons and researchers with respect to field-scale research. Cotton yield monitor data collected on-the-go from planned field-scale on-farm experiments can be used to make improved decisions if analyzed appropriately. When farmers and researchers compare treatments implemented at larger block designs, treatment edge effects and spatial externalities need to be considered so that results are not biased. Spatial analysis methods are compared for field-scale research using site-specific data, paying due attention to local and global patterns of spatial correlation. Local spatial spillovers are explicitly modeled by spatial statistical techniques that

led to improved farm management decisions in combination with the limited replication strip trial data farmers currently collect.

INTRODUCTION

Precision agriculture has renewed the interest of farmers and researchers in conducting on-farm planned comparisons with respect to field-scale research. Yield monitor data can be collected as on-the-go and planned on-farm experiments, and can be implemented, harvested, and analyzed without interfering with field operations if precision technologies are appropriately used. Farmers conducting their own field-scale research often use limited-replication large block experimental designs (Cook et al., 2018; Marchant et al., 2019; Piepho et al., 2011), which are not well founded in classical statistics. Specifically, lack of homogeneity in soils and other factors as well as the typically limited number of replications, are at odds with classical statistical inference principles, potentially interfering with experimental yield results. While farmers are not as focused on statistical inference as researchers, they care intensely about the reliability of results. Are the yield differences observed the result of random variation, or repeatable mechanisms? The vision is that if statistical techniques can be developed to use yield monitor and other precision agriculture data to assess the reliability of inputs and agronomic management, those statistical techniques could be incorporated into easy to use decision support tools.

When farmers aim for a comparison of input products, application rates or agronomic management, larger treatment blocks are often substantially easier to implement than small plots. Large block designs are particularly important for some inputs specific to cotton (*Gossypium hirsutum L.*), such as tillage equipment or midseason insecticides, growth regulators, and defoliants applied with aerial applicators. Spatial analysis techniques have been evaluated at field scales in collaboration with farmers (Griffin et al., 2008). If information that is more reliable can be gleaned from the limited replication data that farmers are already collecting with cotton harvesters equipped with yield monitors, better farm management decisions can be implemented. These methods have been shown to be beneficial especially for cotton because of large scale input application practices. It is therefore important that a thorough methodology is used, capable of capturing and controlling heterogeneity and dependence across large treatment blocks. Spatial regression models have been shown to be capable of providing such a framework (Anselin et al., 2004; Liu et al., 2005; Liu et al. 2014, Liu et al., 2015; Griffin and Lowenberg-DeBoer, 2019).

One key problem with precision agriculture data, and particularly yield monitor data, is that the data are inherently spatially autocorrelated. This spatial correlation can have a rather wide or ‘global’ range, due to similarities in soil composition or hydrological characteristics over a substantial spatial range in the field. Within an experimental context, the experiment itself introduces local patterns of spatial correlation as well, due to nearby locations being subjected to the same treatment (Lambert et al., 2004; Hurley et al., 2004, Liu et al., 2015). Tests for spatial autocorrelation have power over other types of spatial effects such as spatial heterogeneity, therefore spatial diagnostics may detect spatial structure imposed by more elaborate experimental designs. The objective of this study was to determine whether appropriate spatial data analysis techniques modeling local and global spatial autocorrelation patterns in the data can contribute to better farm management decisions based on large-block field-scale on-farm data farmers currently collect with yield monitors. Specifically the concept of modeling local spatial spillovers induced by treatment edge effects was evaluated against aspatial and spatial regression models.

To investigate how spatial analysis methods can be applied to on-farm planned comparison research, this study uses spatial regression models to analyze a large block field-scale tillage experiment. In the example, yield monitor and georeferenced soils data collected from a cotton tillage experiment are used. Four cotton tillage treatments were replicated five times at the University of Arizona's Maricopa Agricultural Center. The analysis compares an aspatial regression model estimated as ordinary least squares (OLS) to spatial regression methods. The aspatial model estimated as OLS is mathematically identical to analysis of variance (ANOVA) with continuous covariates but estimated with regression techniques. Specific attention is paid to including local spatial autocorrelation patterns induced by the design of the experiments that provides results for higher-order models allowing for the simultaneous presence of local and global spatial autocorrelation processes. The hypothesis is that spatial regression that models local or global autocorrelation processes can provide more reliable results than aspatial models. Specific hypotheses include 1) cross regressive models aimed at addressing local spatial correlation facilitate estimation of treatment differences by explicitly modeling treatment edge effects and 2) model specification with the proposed local spatial spillover via cross regression is as useful as more elaborate spatial regression models at discerning treatment effects in field-scale experimentation.

Precision agriculture: practice and research

Precision agriculture builds on the use of modern information technology in agriculture, including global navigation satellite systems (GNSS) and geographical information systems (GIS). Information-intensive precision agriculture technologies continue to be adopted at the farm level with 42% of corn and 45% of soybean acreage harvested in 2005 and 2006, respectively, using yield monitors (Schimmelpfennig and Ebel, 2011). However, percent of cotton acres with harvesters equipped with yield monitors are reported to range from 10 to 20 percent of the total U.S. cotton crop (Griffin, 2010; Daystar et al., 2017; Hellerstein et al., 2019). Adoption of cotton yield monitors are envisaged to follow similar patterns as corn and soybean.

In yield monitoring experiments, spatial analysis techniques have been used to improve the reliability of farm management decisions. Cotton farmers with GNSS-equipped yield monitors report that on-farm experimentation is the number one use of the technology (Griffin, 2010). Spatial statistical techniques have been developed primarily in geostatistics and geography (Cressie, 1993) where an emphasis on modeling induced the emergence of the field of spatial regression (Anselin et al., 2004; Florax and van der Vlist, 2003). Site-specific production functions, i.e. variable rate application, have been estimated using spatial regression (Lambert et al., 2006; Hurley et al., 2005; Ruffo et al., 2006; and Bullock et al., 2009). For cotton, spatial analysis can help growers and those that advise them to cope with the large plots required by aerial application, field scale tillage equipment, and spatial patterns created by irrigation or natural soil factors. The finer scale swath width of cotton harvesters allows for greater spatial detail and flexibility in analysis than grain harvesters. Suspect data points, outliers, or even entire cotton rows may be removed from the analysis leaving an adequate number of observations for properly planned experiments.

Several publications have described on-farm comparison and field-scale research in mechanized agriculture (Bramley et al., 1999; Griffin et al., 2007; Knighton, 2001; Nafziger, 2003; Whelan et al., 2003; Wittig and Wicks, 2001) and the economic ramifications when replications, treatments, or site years are reduced (Young et al., 2004). These methodologies for on-farm comparisons were derived from small plot designs developed in the early twentieth century for the technology available at that time. Concurrent publications recommend designs such as strip or split planter trials to accommodate variability across the field. Some studies have taken on-farm comparisons a step further by integrating precision agriculture technologies to measure variability and record yield data (Adams and Cook, 2000; Anselin et al., 2004; Brouder and Nielsen, 2000; Cook et al., 2013; Knight and Pettitt, 2003; Lark and Wheeler, 2003; Liu et al., 2015; Lowenberg-DeBoer et al., 2003; Lyle et al., 2003; Nielsen, 2000; Whelan et al., 2003). Anselin et al. (2004), Florax et al. (2002), Griffin et al. (2008), Lambert et al. (2004), Lambert et al. (2006), and Liu et al. (2015) used spatial regression models to analyze site specific field data. This paper extends the work of Griffin et al. (2005) and builds upon Florax et al. (2002), Lowenberg-DeBoer et al. (2003), Hurley et al. (2001), Lambert et al. (2004), Liu et al. (2015), and Anselin et al. (2004) by applying spatial statistical and spatial regression techniques to field-scale experimental designs in cotton production by applying proposed cross-regressive models of Griffin and Lowenberg-DeBoer (2019). While most of the above studies used binary dummy variables to model terrain and soils, Florax et al. (2002), Liu et al. (2015) and Griffin et al. (2005) used continuous covariates. Griffin et al. (2005) and Vories et al. (2015) used Boolean Euclidean distance weights matrices, and Liu et al. (2005) used inverse distance weights, while others used a first-order queen contiguity weights matrix (Velandia et al., 2008; Liu et al., 2015). "Queen contiguity" links a data point to all contiguous polygons vertical, horizontal and diagonal, i.e. includes neighboring observations that share a side including those of zero length. Although the first-order queen matrix may have been the appropriate spatial interaction structure for those particular datasets, it is not likely to be universally the most appropriate due to limited connectedness. In many ways this analysis of large block designs provides a complement to the on-farm research designs suggested by Bullock and colleagues (e.g. Bullock and Mieno, 2019).

Those designs focus on using precision agricultural technology to treat relatively small plots when accurate control is possible (e.g. fertilizer rates, plant population).

It has long been known that crop yields vary spatially, even over small areas. Fisher reported that one acre of wheat in 1910 at Rothamsted was harvested in 500 small plots, with yield varying by approximately 30% from the mean (Fisher, 1931). Field heterogeneity is not randomly but systematically distributed, with plots near one another more alike than plots farther apart (Fisher, 1931; Littell et al., 1996; Tobler, 1970). Reducing experimental unit sizes, i.e. plot size, has traditionally counteracted this pattern of spatial autocorrelation, until it could be assumed that the experimental units were homogeneous. In addition, randomization and replication were used with entire replicates placed such that no spatial autocorrelation was assumed to exist within the replicate (Fisher, 1926). Data on soils, topography or other field characteristics are used with spatial analysis to help explain patterns. Inferences drawn on the basis of regression models estimated by ordinary least squares (OLS) are, however, compromised when spatial autocorrelation is present in the data (Anselin, 1988). If the systematic spatial patterns in variability can be appropriately analyzed, farmers can have more confidence in results from experiments they conduct at landscape scale on their farms.

Precision agriculture technologies such as GNSS-equipped yield monitors and others provide many geo-referenced observations per acre at relatively low cost. From the standpoint of classical statistics, one of the key problems with precision agriculture data, and particularly yield monitor data, is the inherent spatial autocorrelation. Spatial heterogeneity can also be present in the data or even be induced by spatial autocorrelation. Spatial regression methods are useful for modeling spatial autocorrelation (Anselin, 1988; Cressie, 1993). Lambert et al. (2002) identified several types of spatial statistics appropriate for analyzing spatially autocorrelated yield monitor data. Anselin et al. (2004), Florax et al. (2002) and Hurley et al. (2001) used spatial statistics to analyze data from designs derived from small plot statistics. Anselin et al. (2004), Florax et al. (2002), and Lambert et al. (2004) corrected for spatial heteroskedasticity using groupwise heteroskedasticity models with and without spatial regimes. Lowenberg-DeBoer et al. (2003) suggested that large block limited replication designs may be appropriate if spatial statistics are used. Cressie (1993) wrote that randomization and replication were not always possible especially for landscape scale ecological and environmental science.

This idea can be extended to the agricultural field sciences with precision agriculture as an example of a new set of technologies. Cressie (1993, p. 249) also observes "classical experimental designs of agricultural field trials ignore the spatial position of the treatment in the design". By taking spatial variation into account, the researcher can obtain unbiased rather than biased as well as more efficient estimates (Cressie, 1993; Duby et al., 1977; Wilkinson et al., 1983; Martin, 1986; Besag and Kempton, 1986; Grondona and Cressie, 1991).

Many university extension systems provide regional recommendations for input use under general agricultural practices. Incorporating on-farm comparison results can often make better farm management decisions. Urcola and Lowenberg-DeBoer (2007) report most commercial Corn Belt farmers do some planned comparisons each season. Most of these comparisons are large block, split field or paired field designs (Cook et al., 2018). Farmers base their decisions on average yield per block or field, paying little attention to within field variability or reliability indicators. On-farm comparison data seem to be most important for farmers who use yield monitors. Cotton farmers are expected to increasingly conduct similar experiments (Griffin, 2010).

Traditional agronomic cotton on-farm comparisons use strip plot designs intended to reduce heterogeneity within experimental units. Those strip plot protocols are based on classic small plot experimental designs such as randomized complete blocks, Latin squares and split plots require intensive planning, management, labor, and human capital efforts during planting and harvesting operations (Piepho et al., 2011). Field activities associated with planting and harvesting (Griffin and Barnes, 2016) are the most critical to the success of the farm operation, causing the value of farmer's management time and labor to be at a premium, thus discouraging implementation of classical experimental designs (Griffin et al., 2014). Familiar experimental designs are often costly and cumbersome, interfering with production logistics. Even though on-farm comparison designs derived from small plot research, such as strip or split planter trials, reduce time requirements compared to classical randomized complete block designs, the perceived benefits of research may still not overcome resource and time costs (Lowenberg-DeBoer et al., 2003).

For instance, there are logistical problems associated with strip trials. For split planter trials on a farm with a six-row cotton picker, filling every six planter boxes with a different variety, seed treatment, furrow insecticide or fungicide potentially leads to human error. When a change of treatments is made, filling planter boxes with small quantities of seed and cleaning boxes for successive varieties hinders planting operations. Minor planter alignment problems may lead to seeding rate, seeding depth, or row spacing issues that impact yield response. With larger acreage farms, the person planning may not be the person planting, potentially leading to communication and coordination problems. From the viewpoint of the analyst, it is a complex and tedious task to keep treatments and cotton picker passes in line.

Moreover, timing and application of inputs for cotton production complicate implementation of on-farm comparisons. In general, cotton farmers apply more inputs than grain farmers. In addition to variety, fertilizer, herbicide, and planting time insecticide treatments commonly used by grain farmers, cotton producers might wish to compare mid-season application of insecticide, growth regulator, or defoliant products. Aerial applications of those mid-season inputs are quite common in cotton such that strip trials are difficult to implement. Furrow irrigation is commonly practiced in cotton such that important differences in the amount of water plants receive from one end of a field to the other can occur (Adamsen et al., 2000).

Many of the problems associated with small plot, strip designs, or specific factors affecting cotton on-farm comparisons could be eliminated if larger experimental units could be used for on-farm comparison designs. Many farmers already conduct planned comparison experiments on single non-replicated large blocks, particularly with new varieties to guide seed decisions in subsequent years. This effort to compare treatments indicates farmers are interested in conducting on-farm comparisons and willing to implement large single block designs. They choose experimental designs for which the cost (mainly in terms of time) is acceptable relative to the perceived benefit. For instance, instead of cleaning planter boxes and taking the time to refill with selected varieties or different types of treated seed, with large block designs planter boxes are filled with the same product and then treatments could be changed during normal reloading times; and areas of the field with mixed seed due to transitioning between treatments can be tagged with a dummy variable (Griffin et al., 2006). Aside from calibration (Vories et al., 2019), collecting yield monitor data requires little extra time during harvest season. Other than planning and analysis in the off-season, large block designs require minimal additional effort during planting or harvest compared to no comparisons. Large blocks also offer the advantage of being less sensitive to human and mechanical error or treatment edge effects, especially from drift of aerial applied pesticides and mobility of pests.

The cotton spindle harvester yield monitor has a distinct benefit over the combine yield monitor because pickers can collect and record row-wise data. Combine yield monitors aggregate data across combine heads, which may be approximately 18.3 meters (60 feet) wide or more with grain cut from the ends of the head entering the yield monitor several seconds after grain cut from middle of head. Cotton picker yield monitors subsequently avoid aggregation and combine dynamics problems that grain yield monitors are vulnerable to as suggested by Lark and Wheeler (2003). To generate usable data, grain growers need to be very careful that harvest passes match exactly the pattern of planter or input application equipment to avoid mixing yields from different treatments. Large blocks offer the benefit of manageable treatment edge effects. If treatment edge effects are thought to exist, yield monitor points near treatment borders can be excluded from analysis.

Treatment edge effects are not to be confused with boundary value edge effects commonly identified in spatial statistics (Anselin, 1988). In practice, real data sets often show observable areas in which a spatial pattern can be identified, but this area is a part of a larger, partly unobserved area in which the underlying process operates. This is similar to the missing starting period observation in time series analysis, although spatial data may contain many more edge effects due to being multidirectional rather than unidirectional. The essential difficulty is that unobserved data and/or processes outside the sampled dataset may interact with data within the observed data. Since the out-of-sample data are not observed, it is difficult to account for edge effects, and currently no satisfactory treatment for spatial edge effects is available. Here, spatial effects arising from treatment edge effects are explicitly modeled via methods proposed by Griffin and Lowenberg-DeBoer (2019) by applying their proposed cross regressive models to experimental design effects rather than relative elevation terrain position.

An exhaustive search of the literature yielded one existing study that attempted to model field-scale treatment edge effects with spatial processes. Bongiovanni et al. (2007) hypothesized relatively narrow widths of treatment strips may induce non-constant variance across observations such as within the same strips. Testing whether treatment strips were a source of error dependence required an additional spatial weights matrix. Their estimation of the model specification with a weighting matrix correlated observations belonging to the same treatment strip. They reported spatial autocorrelation and groupwise heteroskedasticity were induced in wheat yield response to applied nitrogen due to experimental design, i.e. treatment strips. Rather than modeling treatment edge effects as cross regressive variables, they explicitly modeled as a spatial error process with separate spatial weights. Specification of spatial weights matrix identified observations from the same treatment strip rather than identifying the proportion of neighboring observations near treatment edges (Bongiovanni et al., 2007).

Data collection and data filtering

The field study was conducted at the University of Arizona's Maricopa Agricultural Center 40 km south of Phoenix in 2002. Two soil series dominated the field; Mohall (fine-loamy, mixed hyperthermic Typic Haplargid) and Casa Grande (fine-loamy, mixed hyperthermic Typic Natrargids), Arizona's state soil. These sodic-saline alluvial soils formed in the floodplain of the Santa Cruz River. The 6-hectare (15-acre) precision-leveled field was planted to cotton (*Gossypium hirsutum L.,* cv Delta Pine 448B). Operations included a conventional and three alternative tillage treatments: CONVENTIONAL OR CONV (shred, disk, rip, disk, list), ROTOVATOR or Rot (shred, Rotovate), Sundance or Sun (shred, root pull, rip/list), and Pegasus or Peg (single combined operation), in five large block replications and implemented in a randomized complete block design (Figure 1). The individual treatment strip size ranged from 165 m (540 feet) long in the north to 174 m (570 feet) in the south. Treatment blocks were approximately 40 m (130 feet) or 40 cotton rows wide for Conventional and 11 m (36 feet) or 12 cotton rows wide for the three reduced tillage treatments.

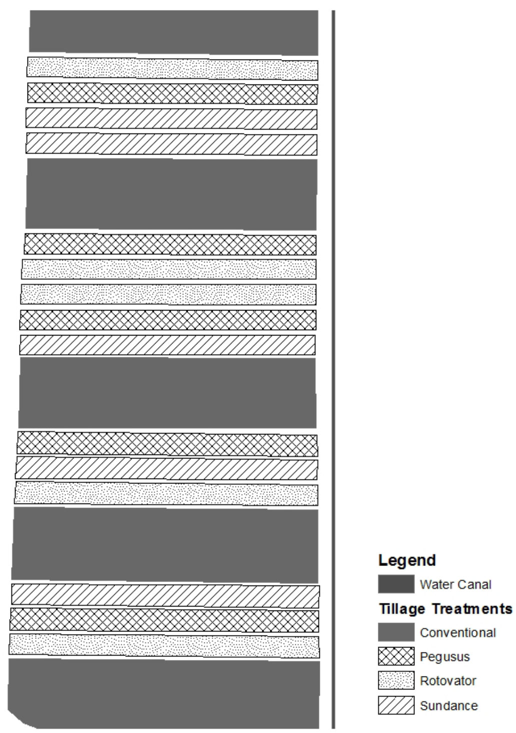

Figure 1. Spatial distribution of the experimental design with four different tillage treatments.

Soil clay content was derived from 2,508 EM38 measurements taken approximately 3.8 m (12.5 feet) apart within each transect and approximately 7 m (23 feet) between transects. Data from EM38 are similar to electrical conductivity (EC) measurements taken from devices such as Veris' Mobile Sensor Platform measuring the resistance of electrical flow through the soil (Corwin and Lesch, 2003). Although electrical conductivity does not directly affect plant growth, an indirect measure of factors that may affect productivity is provided. The EM38 data was calibrated with additional soil samples analyzed in the laboratory and correlated to soil clay content (Triantafilis and Lesch, 2005). Yield data from a four-row cotton picker was collected by optical flow sensors in each vacuum chute by an AGRIplan system and aggregated across the rows before logging a total of 12,824 points in 2002. Yield data points were recorded approximately one meter apart within the row and aggregated over four rows. A weigh boll buggy was used to monitor the calibration of the cotton picker yield monitor.

Table 1. Parameters, criteria, and number of points deleted in yield data filtering.

Parameter	Criteria	Points deleted[a]
Maximum velocity (kph)	4.8	528
Minimum velocity (kph)	0.6	1,447
"Smooth" velocity (%)	0.15	1,062
Maximum yield (kg ha^{-1})	11,209	276
Minimum yield (kg ha^{-1})	1,681	1,988
Standard deviation filter	3	2,112

[a] Points deleted are not cumulative, i.e. the "same" point can be deleted by multiple criteria.

Subsequently, cotton lint yield data were filtered and cleaned for potentially erroneous data with Yield Editor (Sudduth and Drummond, 2007) using parameters set as presented in Table 1 following the procedures set forth by Griffin et al. (2007). The treatment dummies and distance from the irrigation water source was appended to each EM38 soil data point. A 4-m noncontiguous buffer was created around each EM38 point, and a simple average of yield data points within this buffer were assigned to the EM38 point for the purpose of assigning dependent data points (yield) to the location of explanatory variable data points (soil data) in the statistical analysis. A 4 m buffer was used because it was slightly less than the distance between rows of differing treatments blocks so that yield data from adjacent treatments would not be included in the yield estimate. Dummy variables for tillage treatment were added to the respective data points. Of the total 2,508 EM38 points, 57 did not have yield data within 4 meters or were not assigned to a treatment thus omitted from the analysis, leaving 2,451 total observations in the final dataset.

Spatial interaction structure

In applications of spatial regression techniques to precision agriculture, spatial spillover effects have exclusively been modeled as global rather than local processes. Local spatial spillovers exist with only immediately adjacent observations while 'global' refers to each location in the field being linked to any other location in the field (Anselin, 1988; Anselin, 2003). Global linkage processes are inherent to, for instance, the frequently used spatial autoregressive error model:

$$y = X\beta + \varepsilon, \quad \varepsilon = \lambda W\varepsilon + \mu, \tag{1}$$

where **y** is an n x 1 vector of observations on the dependent variable, **X** is an n by k matrix of explanatory variables, **β** is an k by 1 vector of regression coefficients, , $\boldsymbol{\varepsilon}|\boldsymbol{X} \approx i.i.d. N\left(0, \sigma_\varepsilon^2\ {}_n\mathbf{I}_n\right)$, and **μ** an independently and identically distributed (i.i.d.) error term. Substitution and rewriting shows the presence of a spatial multiplier term $(I - \xi W)^{-1}$, where ξ represents the spatial parameter. As a result of the inverse term, each location is linked to any other location when $\xi \neq 0$, irrespective of whether they are linked through the specification of the weights matrix. For the spatial error model, rewriting leads to the specification:

$$y = X\beta + (I - \lambda W)^{-1}\mu$$

$$(I - \lambda W)(y - X\beta) = \mu$$

$$y = \lambda Wy + X\beta - \lambda WX\beta + \mu \, , \tag{2}$$

which embeds a set of k–1 nonlinear constraints commonly known as the Durbin or common factor model (Mur and Angulo, 2006). Further rewriting gives:

$$y = (I - \lambda W)^{-1}(X\beta - \lambda WX\beta + \mu) \tag{3}$$

Equation (3) embeds a series of nonlinear constraints and it contains spatially lagged exogenous variables in addition to the exogenous variables itself (Anselin, 2003).

Although both specifications have been overwhelmingly popular in spatial regression precision agriculture analyses, they are unlikely to be the data generating process. The spatial lag model *a priori* assumes that the spatial correlation extends to the whole spatial system. The spatial error model implicitly includes *a priori* nonlinear constraints on the local spatial autocorrelation, and it is based on the assumption that the global and local correlation intensity is the same. The spatial error process can be characterized by the autoregressive (AR) or the moving average (MA) error process resulting in global and local spillovers, respectively. Therefore, the spatial autoregressive (SAR) was estimated. It should be noted that although spatial moving average (SMA) models were desirable (Baltagi et al. 2019; Dogan and Taspinar, 2013; Fingleton, 2008a, 2008b), these have not been developed into readily available statistical software. In addition, the specification of the spatial weights matrix is by definition the same for the local and the global autocorrelation processes. The implementation of different tillage treatments by its very design induces local spatial correlation patterns in the data. This represents an omitted variable problem that should ideally be solved before testing for residual spatial correlation.

Florax and de Graaff (2004, page 42) present an example of omitted variable problem based on the "true" model $y = \alpha + \beta x + \gamma Wx + \mu$, where μ is the usual iid error term with mean zero to exemplify this point. If autocorrelated exogenous variables are omitted, the actual regression becomes, $y = \alpha + \beta x + \varepsilon$, where $\varepsilon = \mu + \gamma Wx$, but now $\mathrm{E}(\varepsilon) = W \cdot \mathrm{E}(x) = m \neq 0$, with m symbolizing the omitted variable bias. The covariance between the residuals at locations i and j, where i and j are not first- or second- order neighbors, equals:

$$\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = \mathrm{E}\big((\varepsilon_i - m)(\varepsilon_j - m)\big) = \mathrm{E}(\varepsilon_i \varepsilon_j) - m^2 \qquad (4)$$

where

$$\mathrm{E}(\varepsilon_i \varepsilon_j) = \mathrm{E}\big([\mu_i + \rho(Wx)_i][\mu_j + \rho(Wx)_j]\big) = \rho^2 (Wx)_i (Wx)_j \qquad (5)$$

such that the residuals comprising the omitted variable tend to be correlated, regardless if topologically invariant or of relative spatial arrangement. Therefore, evaluating omitted spatially autocorrelated exogenous variables with spatial misspecification tests may not adequately detect failures of the model (Florax and de Graaff, 2004, page 42).

To compare local and global spillovers, separate spatial interaction structures were chosen for use in the separate regression models. Two row-standardized spatial weights matrices are constructed, one for local effects referred to as W_1 and one for global effects referred to as W_2. Both weights matrices are constructed such that $w_{ii} = 0$, $w_{ij} > 0$ for observations considered neighbors, and $w_{ij} = 0$ for non-neighbors. The W_1 matrix was selected on the basis that only immediate neighboring observations are of interest, so a first-order queen criterion was constructed in R with spdep contributed package (Bivand et al, 2013; Bivand and Wong, 2018). Since the experimental data is such that there are two transects of data in each reduced tillage treatment and four transects of data in the CONVENTIONAL tillage treatments, a first-order queen criterion may include neighbors from differing treatments. As a result, the cross-regressive terms W_1D_T, where D_T is dummy variable with ones for a specific treatment *T*, captures local dependence and spatial externalities (Arbia, 2014) arising from experimental treatment edge effects that would otherwise have been missed. Pre-multiplying D_T by W_1 creates spatial weighted average of the specific treatment dummy variable on all neighbors specified by the weights matrix such that W_1D_T is bound by 0 and 1. The cross-regressive term W_1D_T equals 1 when all neighbors are of the treatment in question and less than 1 if a portion of neighbors are of a differing treatment.

An inverse-distance criterion is selected for W_2 and the matrix was created using the spdep contributed package to R (R Core Team, 2019). The inverse distance weights matrix was chosen because it implies a smooth distance decay of spatial correlation, up to an empirically determined relevant distance band.

The full model specification includes cotton lint YIELD as the dependent variable, and percent clay content (CLAY) and its square, distance to the water source (DIST) and its square, tillage treatment dummies and their spatially weighted average using weights matrix W_1, and interaction terms of tillage treatment and CLAY, and CLAY and DIST. Table 2 provides an overview of the explanatory variables.

Table 2. Description of explanatory variables.[a]

Variable	Description
CONSTANT	Intercept
CLAY	Percent clay content of the soil
CLAY2	Square of CLAY
CLAYDIST	Interaction term between CLAY and DIST
DIST	Distance of the soil point to the irrigation canal
DIST2	Square of DIST
PEG	Dummy variable for PEGASUS
ROT	Dummy variable for ROTOVATOR
SUN	Dummy variable for SUNDANCE
PEGC	Interaction term of PEG and CLAY
ROTC	Interaction term of ROT and CLAY
SUNC	Interaction term of SUN and CLAY
W_1CLAY	Spatially weighted average of CLAY using W_1
W_1PEG	Spatially weighted average of PEG using W_1
W_1ROT	Spatially weighted average of ROT using W_1
W_1SUN	Spatially weighted average of SUN using W_1

[a] W_1 is a standardized first-order contiguity matrix based on the queen criterion.

An inverse distance weights matrix with a 75-meter band was selected as the appropriate spatial interaction structure for the spatial regression portion for this particular data. Underspecifying the weights matrix causes greater estimation errors than over-specifying indicating that distances larger than thought appropriate are preferred to distance less than appropriate especially for tests of spatial autocorrelation (Florax and Rey, 1995). Summary measures with respect to the connectivity structure implied by both weights matrices are provided in Table 3. The inverse-distance weights matrix assumes a considerably greater spatial range than the first-order contiguity matrix.

Table 3. Connectivity data for the different spatial weights matrices.

	First-order queen, W_1	75 m inverse-distance, W_2
Dimension	2,451	2,451
Nonzero links	14,324	1,182,982
Nonzero weights (%)	0.24	19.70
Average weight	0.17	0.002
Average number of links	5.84	482.65

Largest root (eigenvalue)	1.00		1.00	
Smallest root (eigenvalue)	–0.57		–0.23	
# and frequency of neighbors	# Neighbors	Frequency	# Neighbors	Frequency
	2	3	177 - 232	36
	3	49	233 - 288	91
	4	119	289 - 344	224
	5	329	345 - 400	365
	6	1696	401 - 456	327
	7	178	457 - 512	331
	8	67	513 - 568	321
	9	8	569 - 623	384
	10	2	624 - 678	372

Empirical analysis

The whole field YIELD average was 5,908 kg ha^{-1} (5,271 lbs ac^{-1}) with a standard deviation of 556 kg (1,225 lbs). Seed cotton yields ranged from a minimum of 1884 kg ha^{-1} (1,681 lbs ac^{-1}) to a maximum of 10,094 kg ha^{-1} (9,006 lbs ac^{-1}) (Table 4). Soil CLAY content ranged from a minimum of 7.9% to a maximum of 31.6% with a mean of 23.2% (Table 4). Moran's I test statistics were 0.45 and 0.58 for YIELD and CLAY, respectively, indicating that the dependent and continuous explanatory variables are spatially autocorrelated (Clift and Ord, 1981; Anselin, 1988).

Table 4. Descriptive statistics for variables.

	Descriptive Statistics				**Wald test on normality**		**Randomization Assumption**		
	Mean	**S.D.**	**Min**	**Max**	**Wald**	**Prob**	**Moran's I**	**z-value**	**S.D.**
Yield (kg ha^{-1})	5905	1375	1884	10094	32.15	0.0000	0.45	251.95	0.0018
Clay (%)	23.2	4.4	7.9	31.6	325.37	0.0000	0.58	325.59	0.0018

A general moments (GM) estimator is more appropriate than the traditional maximum likelihood (ML) estimator for site-specific data because site-specific data tends to be large sample size data and normality is unexpected. Maximum likelihood estimation uses an eigenvalues computation of the Jacobian matrix which loses numerical precision beyond 1,000 observations. Although GM is not as efficient in general as ML, this limitation is overcome by the large sample size of site-specific data. General moments estimation can be conducted for very large data sets of several thousand observations without the assumption of normality.

In addition to estimating the aspatial model, a cross-regressive model utilizing ***WX*** plus traditional spatial models are estimated. The cross-regressive model is estimated as OLS and is intended to account for local spatial externalities (Anselin, 2003) due to treatment edge effects of neighboring observations and is given as:

$$y = X\beta + W_1 X\gamma + \varepsilon \qquad (6)$$

Where $\boldsymbol{W_1X}$ is the cross-regressive term as described above, $\boldsymbol{\gamma}$ is the vector of regression coefficients on the cross-regressive term, and the other terms as already defined (Griffin and Lowenberg-DeBoer, 2019). A traditional spatial model was estimated using $\boldsymbol{W_2}$ to account for global spatial externalities.. The spatial error model has spatially autocorrelated errors and is similar to the traditional aspatial model with the exception that the error term $\boldsymbol{\varepsilon}$ is spatially autocorrelated and given as:

$$y = X\beta + \varepsilon \text{ where } \varepsilon = \lambda W_2 \varepsilon + \mu \qquad (7)$$

where λ is the spatial autoregressive parameter, $\boldsymbol{\mu}$ is the new vector of errors and the remaining terms as previously defined. The spatial error model more concisely written is:

$$y = X\beta + (I - \lambda W_2)^{-1} \mu \qquad (8)$$

The spatial error model indicates that the spatial process is present in the whole system of data while the cross-regressive model only addresses local spatial externalities (Anselin, 2003). Since the spatial covariance structure realized from the traditional spatial models relates all locations in the dataset to all other locations, these models are said to be global (Anselin, 2003).

RESULTS AND DISCUSSION

As previously stated, the full regression model specification includes cotton lint YIELD as the dependent variable with explanatory variables including percent clay content (CLAY) and its square, distance to the water source (DISTANCE or DIST) and its square, tillage treatment dummy, an interaction term of tillage treatment and CLAY, and an interaction term between CLAY and DISTANCE (Table 2). The square of clay was included because the relationship is expected to be quadratic, i.e. it is known that water holding capacity is directly correlated with clay content. However, at high clay content levels may prevent soil aeration and subsequent crop growth. The square of distance was included because it is expected that plants close to the irrigation canal may have over applications of water to irrigate the opposite side of the field which may not have received adequate water or any water at all. Interaction terms of clay with the tillage treatment dummies were included in the model to allow each tillage treatment to be estimated with its individual slope coefficient, allowing each tillage system to induce a differing cotton yield response to clay content. Interactions between distance and clay were included to capture pooling effects of irrigation water due to changes in soil surface properties and differing crop residue. The full model can be presented as:

$$\text{YIELD} = \text{CLAY} + \text{CLAY}^2 + \text{DIST} + \text{DIST}^2 + \text{CLAY*DIST} + \text{TRT}_i + \text{TRT}_i\text{*CLAY} \qquad (9)$$

where TRT_i = PEG, ROT, SUN

In this study, CLAY and the three alternative tillage treatment dummy variables were chosen as the cross-regressive terms for inclusion in the local spatial correlation model. If localized spatial externalities exist in the system, an omitted variable problem occurs unless these effects are modeled. In traditional aspatial or traditional spatial regression models, these localized spatial effects are overlooked. The $\boldsymbol{W_1}$CLAY, $\boldsymbol{W_1}$PEG, $\boldsymbol{W_1}$ROT, and $\boldsymbol{W_1}$SUN variables were added to account for local differences in clay content and tillage treatments. These variables exist in the explanatory $\boldsymbol{X}$ variables just as with aspatial model, plus as spatially weighted lag cross-regressive terms using the first order queen contiguity matrix $\boldsymbol{W_1}$. The $\boldsymbol{W_1}$CLAY term accounts for localized differences in the soil clay content. The $\boldsymbol{W_1}$PEG, $\boldsymbol{W_1}$ROT, and $\boldsymbol{W_1}$SUN variables account for potential treatment edge effects induced by tillage treatments. Since PEGASUS, ROTOVATOR, and SUNDANCE are binary dummy variables, the value of the $\boldsymbol{W_1}$PEG, $\boldsymbol{W_1}$ROT, and $\boldsymbol{W_1}$SUN are bounded between zero and one, inclusive, and can be presented as:

$$\text{YIELD} = \text{CLAY} + \text{CLAY}^2 + \text{DIST} + \text{DIST}^2 + \text{CLAY}*\text{DIST} + \text{TRT}_i + \text{TRT}_i*\text{CLAY} + \boldsymbol{W_1}*\text{CLAY} + \boldsymbol{W_1}*\text{TRT}_i \qquad (10)$$

where TRT_i = PEG, ROT, SUN

Lagrange Multiplier (LM) results arising from using the 75-meter inverse distance matrix (referred to as $\boldsymbol{W_2}$) were examined for guidance in choosing a spatial model (Table 5). Lagrange Multiplier and Robust LM tests indicated that spatial autocorrelation was in both the dependent variable (lag) and the error term (Table 5). A spatial error model (Anselin, 1988), spatial lag (Anselin, 1988), and a higher order model simultaneously correcting for both lag and error autocorrelation (Kelejian and Prucha, 1998) were conducted to account for spatial autocorrelation in the data. Aspatial diagnostics and regression results for OLS, local spatial correlation (LSC), and spatial autoregressive error model (SEM) are presented in Table 5.

Table 5. Estimated results for aspatial, local spatial and global spatial models, n=2,451

	OLS	**LSC**	**SEM**
CONSTANT	-901.16 (-2.285)	-897.456 (-2.268)	-1004.47 (-2.630)
CLAY	286.32 (7.424)	219.213 (5.154)	310.51 (8.594)
CLAY2	-1.09 (-1.128)	-0.897 (-0.926)	-1.93 (-2.151)
CLAY_DIST	-1.28 (-10.925)	-1.190 (-10.069)	-1.12 (-9.697)
DIST	38.58 (12.747)	36.401 (11.930)	36.03 (11.768)
DIST2	-0.01 (-1.575)	-0.014 (-1.617)	-0.02 (-2.254)
PEG	-262.97 (-0.952)	-745.138 (-2.482)	-303.72 (-1.187)
ROT	872.71 (3.137)	319.496 (1.080)	665.15 (2.565)
SUN	-497.59 (-1.921)	-1124.59 (-3.726)	-468.61 (-1.925)
PEGC	-32.22 (-2.808)	-32.316 (-2.842)	-30.39 (-2.863)
ROTC	-68.39 (-5.747)	-70.118 (-5.954)	-59.14 (-5.334)
SUNC	-16.39 (-1.484)	-15.424 (-1.379)	-20.44 (-1.978)
W_1CLAY		59.897 (4.461)	
W_1PEG		529.458 (3.289)	
W_1ROT		764.588 (4.922)	
W_1SUN		770.511 (4.899)	
λ			0.329 (93.766)
LM Error	28056.64	29459.79	
Robust LM Error	24834.77	25625.99	
LM Lag	3221.88	3836.83	
Robust LM Lag	0.02	3.03	
AIC	40862.6	40804.9	38763.7

[a]In parentheses t-values are reported for OLS and LSC, and z-values for the spatial models.

The variable CLAY was significant in all four regression models, but CLAY2 was not significant in any model. Distance to irrigation water source was significant for every model but the distance squared term was significant only for the two traditional spatial models. The PEG and SUN treatment dummy variables were significant for only the cross-regressive model while the ROT treatment dummy was significant for all the models except the cross-regressive model. The CLAY by PEG and CLAY by ROT interaction terms were significant for all models while the clay by SUN interaction term was significant only for the three traditional spatial models. All four cross-regressive terms, $\boldsymbol{W_1}$CLAY, $\boldsymbol{W_1}$PEG, $\boldsymbol{W_1}$ROT, and $\boldsymbol{W_1}$SUN, were significant for the cross-regressive model.

As expected from *a priori* agronomic information, the LM and Robust LM tests for the OLS residuals indicates that the spatial error model dominates the spatial lag model for this dataset, although the LM test for spatial lag was significant (Table 5). Based on spatial diagnostics and a priori conceptual understanding, only the spatial error process models were run. The estimated spatial autoregressive term λ is significant at 0.329 for the spatial error model.

A key contribution of spatial models is that it clarifies the effect of soil clay on tillage choice. In this dataset, spatial models more clearly demonstrated the yield superiority of CONVENTIONAL tillage treatment across a wider range of soil clay levels, and it clarified the role of soil clay content levels in choice of alternative tillage systems. Probably because the clay variable explained substantial proportion of yield variability by essentially absorbing spatial structure, the yield response to clay content of all four tillage systems were similar under OLS estimation while were substantially different when spatial structure of dependent yield and independent explanatory variables were explicitly modeled.

If the relationships in the 2002 data were confirmed in subsequent seasons, a grower who wanted to use reduced tillage systems for soil conservation or other reasons might decide on a field-specific tillage plan. Varying tillage within fields is unlikely with current equipment because it would complicate logistics. But fields where soil clay content is low might be managed differently from those which have generally higher clay contents. Tillage effects may also be related to other soil and landscape properties such as slope, aspect, or organic matter.

CONCLUSIONS

This study provides comparisons and an example for the potential for modeling local and global spatial externalities of precision agriculture data, particularly in cotton tillage experiments. Given the spatial effects present at field scales, aspatial analyses were misspecified because the assumption of independent errors was violated, and the soil clay variable absorbed spatial structure effects. Diagnostic tests on OLS residuals indicated spatial error was preferred to spatial lag. These techniques for modeling localized spatial externalities in topographical attributes are being modeled for crops grown in rolling terrain. Both hypotheses were supported. Modelling local spatial spillovers led to improved farm management decisions in combination with the limited replication strip trial data farmers currently collect.

ACKNOWLEDGEMENTS

The authors appreciate Robert “Bob” Roth and staff at the Maricopa Agricultural Center, University of Arizona, Maricopa, Arizona, USA where field research was conducted. The research purpose was to demonstrate the usefulness of spatial analysis methods for precision agriculture data. Results are from a single year at a single location; and tillage system rankings are not intended as generalized knowledge across regions, growers, or systems, but rather as a demonstration of alternative methods of on-farm experimentation and data analysis. Mention of specific suppliers of hardware and software is for informative purposes only and does not imply endorsement. This work is dedicated to the memory of Raymond J.G.M. Florax, a leader in spatial econometrics applied to agriculture and one of the lead author's mentors.

REFERENCES

Adams, M.L. and S.E. Cook. 2000. On-farm Experimentation: Application of Different Analytical Techniques for Interpretation. P.C. Robert et al., ed. In: proceedings of the 6th International Conference on Precision Agriculture. ASA/CSSA/SSSA Madison, WI.

Adamsen, F.J., Hunsaker, D.J., Barnes, E.M., Clemmens, A.J., and Bautista, E. 2000. Surface irrigation and precision crop management. In Proceedings of the Fifth International Conference on Precision Agriculture. ASA-CSSA-SSSA, 677 South Segoe Road, Madison, WI 53711, USA.

Anselin, L. 1988. Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Drodrecht, Netherlands.

Anselin, L. 2003. Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2):153-166

Anselin, L., R. Bongiovanni and J. Lowenberg-DeBoer. 2004. A Spatial Econometric Approach to the Economics of Site-Specific Nitrogen Management in Corn Production. *American Journal of Agricultural Economics*, 86(3): 675-687.

Arbia, G. 2014. *A Primer for Spatial Econometrics with Applications in R*. Palgrave MacMillan New York, NY.

Baltagi, B.H., Fingleton, B., and Pirotte, A. 2019. A time-space dynamic panel data model with spatial moving average errors. *Regional Science and Urban Economics*, 76: 13-31.

Besag, J.E. and Kempton, R.A. 1986. Statistical analysis of field experiments using neighboring plots. *Biometrics*, 42:231-251.

Bivand, R.S. and Wong, D.W. S. 2018 Comparing implementations of global and local indicators of spatial association. *TEST*, 27(3):716-748.

Bivand, R.S., Pebesma, E., and Gomez-Rubio, V. 2013. Applied spatial data analysis with R, Second edition. Springer, NY.

Bongiovanni, R.G., Robledo, C.W., Lambert, D.M. 2007. Economics of site-specific nitrogen management for protein content in wheat. *Computers and Electronics in Agriculture*, 58(1):13-24

Bramley, R., S. Cook, M. Adams, and R. Corner. 1999. Designing your own on-farm experiments: How precision agriculture can help. CSIRO Land and Water.

Brouder, S., and Nielsen, R. 2000. On-Farm Research," in Precision Farming Profitability, J. Lowenberg-DeBoer and K. Erickson, eds., Purdue University Agricultural Research Programs, p. 103-112.

Bullock, D.S. and J. Lowenberg-DeBoer. 2007. Using Spatial Analysis to Study the Values of Variable Rate Technology and Information. *Journal of Agricultural Economics*, 58(3): 517–535.

Bullock, D.S. and Mieno, T. 2019. "The economic value of on-farm experimentation," in Precision Agriculture '19, edited by John Stafford, Wageningen Academic Publishers, Wageningen, Netherlands, 2019, p. 817-823.

Bullock, D.S., Ruffo, M.L., Bullock, D.G., and Bollero, G.A. 2009. The Value of Variable Rate Technology: An Information-Theoretic Approach. *American Journal of Agricultural Economics,* 91(1):209-223.

Cliff, A. and Ord, J., Spatial Processes, Models and Applications. London: Pion, 1981.

Cook, S.E., Cock, J., Oberthür, T. and Fisher, M. 2013. On-Farm Experimentation. *Better Crops*, 97(4):17-20.

Cook, S.E., Lacoste, M., Evans, F., Ridout, M, Gibberd, M, and Oberthür T. 2018. An On-Farm Experimental philosophy for farmer-centric digital innovation. Proceedings of the 14th International Conference on Precision Agriculture June 24 - 27, 2018 Montreal, Quebec, Canada.

Corwin, D.L., and Lesch, S.M. 2003. Application of Soil Electrical Conductivity to Precision Agriculture: Theory, Principles, and Guidelines. *Agronomy Journal* 95:455–471.

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

Daystar, J.S., Barnes, E.M., Hake, K., and Kurtz, R. 2017. Sustainability trends and natural resource use in U.S. cotton production. *BioResources* 12(1):362-392.

Dogan, O., and Taspinar, S. 2013. GMM estimation of spatial autoregressive models with moving average disturbances. *Regional Science and Urban Economics*, 43:903-926.

Duby, C., Guyon, X., and Prum, B. 1977. The precision of different experimental designs for a random field. *Biometricka*, 64:59-66.

Fingleton, B. 2008a. A Generalized Method of Moments Estimator for a Spatial Panel Model with an Endogenous Spatial Lag and Spatial Moving Average Errors. *Spatial Economic Analysis*, 3(1)27-44.

Fingleton, B. 2008b. A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. *Empirical Economics*, 34(1): 35-57.

Fisher, R.A. 1926. The Arrangement of Field Experiments. *Journal of the Ministry of Agriculture of Great Britain*, 33:503-513.

Fisher, R.A. 1931. Principles of plot experimentation in relation to the statistical interpretation of the results. Report of a Conference on The Technique of Field Experiments held at Rothamsted on 7 May, 1931, pp 11-13.

Florax, R.J.G.M., and Rey. S. 1995. The Impacts of Misspecified Spatial Interaction in Linear Regression Models. In L. Anselin, R.J.G.M Florax (eds), New Directions in Spatial Econometrics.

Florax, R.J.G.M., and Van der Vlist, A.J. 2003. Spatial Econometric Data Analysis: Moving Beyond Traditional Models. *International Regional Science Review*, *26*(3):223–243.

Florax, R.J.G.M., Voortman, R.L., and Brouwer, J. 2002. Spatial dimensions of precision agriculture: a spatial econometric analysis of millet yield on Sahelian coversands. *Agricultural Economics* 27:425-443.

Florax, R.J.G.M., and de Graaff, T. 2004. The Performance of Diagnostic Tests for Spatial Dependence in Linear Regression Models: A Meta-Analysis of Simulation Studies. pp 29-66. Advances in Spatial Econometrics Methodology, Tools and Applications, Anselin, L., Florax, R.J.G.M., and Rey, S.J. (Editors). Advances in Spatial Science. Springer. 513 pages.

Griffin, T.W. 2010. The Spatial Analysis of Yield Data. In M. Oliver (Ed.) Geostatistical Applications for Precision Agriculture. Springer. 295p.

Griffin, T.W. and Barnes, E. 2017. Available Time to Plant and Harvest Cotton across the Cotton Belt. *Journal of Cotton Science*, 21(1):8-17.

Griffin, T.W., Brown, J.P., and Lowenberg-DeBoer, J. 2007. Yield Monitor Data Analysis Protocol: A Primer in the Management and Analysis of Precision Agriculture Data. Version 2 June 2007.

Griffin, T.W., Dobbins, C.L., Vyn, T., Florax, R.J.G.M. and Lowenberg-DeBoer, J. 2008. Spatial analysis of yield monitor data: Case studies of on-farm trials and farm management decision-making. *Precision Agriculture*, 9(5):269-283.

Griffin, T., Fitzgerald, G., Lambert, D., Lowenberg-DeBoer, J., Barnes, E., and Roth, R. 2005. Testing appropriate statistical methods for on-farm cotton research using precision farming, Proceedings, Beltwide Cotton Conferences, New Orleans, LA, 2005, pp. 383-392.

Griffin, T.W., Florax, R.J.G.M, and Lowenberg-DeBoer, J. 2006. Field-Scale Experimental Designs and Spatial Econometric Methods for Precision Farming: Strip-Trial Designs for Rice

Production Decision Making. Southern Agricultural Economics Association Annual Meeting, Orlando, FL, February 2006.

Griffin, T.W., and Lowenberg-DeBoer, J.M. 2019. Modeling local spatial correlation of terrain attributes in landscape-scales using spatial cross regressive variables. *Precision Agriculture*. 10.1007/s11119-019-09702-5

Griffin, T.W., Mark, T.B., Dobbins, C.L., and Lowenberg-DeBoer, J. 2014. Estimating Whole Farm Costs of Conducting On-farm Research: A Linear Programming Approach. *International Journal of Agricultural Management*. 4(1):21-27.

Grondona, M.O., and Cressie, N. 1991. Using spatial considerations in the analysis of experiments. *Technometrics*, 33:381-392.

Hellerstein, D., Vilorio, D. and Ribaudo, M. 2019. Agricultural Resources and Environmental Indicators, 2019. EIB-208, U.S. Department of Agriculture, Economic Research Service, May 2019.

Hurley, T., Kilian, B. and H. Dikici. 2001. The Value of Information for Variable Rate Nitrogen Applications: A Comparison of Soil Test, Topographical, and Remote Sensing Information. Selected Paper, AAEA Annual Meeting, Chicago, IL, August 5-8, 2001.

Hurley, T., Malzer, G., and Kilian, B., 2004. Estimating site-specific crop response functions: a conceptual framework and geostatistical model. *Agronomy Journal* 96:1331-1343.

Hurley, T.M., Oishi, K., and Malzer, G.L. 2005. Estimating the Potential Value of Variable Rate Nitrogen Applications: A Comparison of Spatial Econometric and Geostatistical Models. *Journal of Agricultural and Resource Economics*, 30(2):231-249.

Kelejian, H.H., and Prucha, I.R. 1998. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *The Journal of Real Estate Finance and Economics,* 17(1):99-121

Knight, S., and Pettitt, T. 2003. The design and analysis of experiments using yield monitoring technology. In "Solutions for a better environment". Proceedings of the 11th Australian Agronomy Conference, 2-6 Feb. 2003, Geelong, Victoria. Australian Society of Agronomy.

Knighton, R.E. 2001. Setting Up On-Farm Experiments. In Site-Specific Management Guidelines SSMG-17.

Lambert, D., Lowenberg-DeBoer, J., and R. Bongiovanni. 2002. Spatial Regression, an Alternative Statistical Analysis for Landscape Scale On-farm Trials: Case Study of Variable Rate N Application in Argentina. P.C. Robert et al., ed. In: proceedings of the 6th International Conference on Precision Agriculture. ASA/CSSA/SSSA Madison, WI.

Lambert, D.M., Lowenberg-DeBoer, J., Bongiovanni, R., 2004. A comparison of four spatial regression models for yield monitor data: a case study from Argentina. *Precision Agriculture*, 5:579–600.

Lambert, D.M., J. Lowenberg-DeBoer and G.L. Malzer. 2006. Economic Analysis of Spatial-Temporal Patterns in Corn and Soybean Response to Nitrogen and Phosphorus. *Agronomy Journal* 98(1):43-54.

Lark, M. and Wheeler, H.C. 2003, Experimental and analytical methods for studying within-field variation of crop responses to inputs. Proceedings of the 4th European Conference on Precision Agriculture, J. Stafford & A. Werner, editors, Wageningen Academic Publishers, Netherlands, 2003.

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. 1996. SAS System for Mixed Models. The SAS Institute Inc., Cary, North Carolina.

Liu, Z., T.W. Griffin, T.L. Kirkpatrick, W.S. Monfort. 2015. Spatial Econometric Approaches to Site-Specific Nematode Management Strategies. *Precision Agriculture,* 16(5):587-600

Liu, Z., Griffin, T.W., and Kirkpatrick, T.L. 2014. Statistical and Economic Techniques for Site-specific Nematode Management. *Journal of Nematology*, 46(1):12-17.

Liu, Y., S.M. Swinton, and N.R. Miller. 2006. Is Site-Specific Yield Response Consistent over Time? Does it Pay? *American Journal of Agricultural Economics,* 88(2):471-483

Lowenberg-DeBoer, J., Lambert, D.M., and Bongiovanni, R. 2003. Appropriate On-Farm Trial Designs for Precision Farming. Proceedings of the 4th European Conference on Precision Agriculture, J. Stafford & A. Werner, editors, Wageningen Academic Publishers, Netherlands.

Lyle, G, Wong, M.T.F., Norris, P., and Adams, M. 2003. Using Precision Agriculture to fine tune paddock management: A case study with the Yuna Farm Improvement Group in WA. In "Solutions for a better environment". Proceedings of the 11th Australian Agronomy Conference, 2-6 Feb. 2003. Geelong, Victoria. Australian Society of Agronomy.

Marchant, B., Rudolph, S., Roques, S., Kindred, D., Gillingham, V., Welham, S., Coleman, C., Sylvester-Bradley, R. 2019. Establishing the precision and robustness of farmers' crop experiments. *Field Crops Research*, 230(1):31-45

Martin, R.J. 1986. On the design of experiments under spatial correlation. *Biometrika*, 73:247-277.

Mur, J. and Angulo, A. 2006. The Spatial Durbin Model and the Common Factor Tests. *Spatial Economic Analysis,* 1(2):207-226

Nafziger, E. 2003. On-Farm Research. Chapter 21, Illinois Agronomy Handbook, University of Illinois, Urbana-Champaign, 2003.

Nielsen, R. 2000. Opportunities for On-Farm Variety Performance Testing Using GPS Enabled Technologies, in J. Lowenberg-DeBoer and K. Erickson, Eds, Precision Farming Profitability, Purdue University, Agricultural Research Program, 2000, p. 12-18.

Piepho, H.P., Richter, C., Spilke, J., Hartung, K., Kunick, A., Thöle, H. 2011. Statistical aspects of on-farm experimentation. *Crop and Pasture Science*, 62(9):721-735

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ruffo, M.L., Bollero, G.A., Bullock, D.S., Bullock, D.G. 2006. Site-specific production functions for variable rate corn nitrogen fertilization. *Precision Agriculture* 7(5):327-342.

Schimmelpfennig, D., and Ebel, R. 2011. On the Doorstep of the Information Age: Recent Adoption of Precision Agriculture. United States Department of Agriculture.

Sudduth, K.A., and Drummond, S.T. 2007. Yield Editor: Software for Removing Errors from Crop Yield Maps. *Agronomy Journal* 99:1471-1482

Tobler, W.R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography,* 46:234-240.

Triantafilis, J., and Lesch, S. 2005. Mapping Clay content variation using electromagnetic induction techniques. *Computers and Electronics in Agriculture* 46(1):203-237.

Urcola, H.A., and Lowenberg-DeBoer, J. 2007. A stochastic dominance method for incorporating yield monitor data into the hybrid and variety decisions of Argentinean farmers. *Computers and Electronics in Agriculture*, 58(1):4-12.

Velandia, M, Rejesus, R.M., Bronson, K., and Segarra, E. 2008. Economics of Management Zone Delineation in Cotton Precision Agriculture. *The Journal of Cotton Science,* 12:210-227

Vories, E.D., Stevens, W.E., Sudduth, K.A., Drummond, S.T., Benson, N.R. 2015. Impact of Soil Variability on Irrigated and Rainfed Cotton. *The Journal of Cotton Science,* 19:1-14

Vories, E.D., Jones, A.S., Meeks, C.D., and Stevens, W.E. 2019. Variety Effects on Cotton Yield Monitor Calibration. *Applied Engineering in Agriculture,* 35(3):345-354

Whelan, B.M., McBratney, A.B., and Stein, A. 2003. On-Farm Field Experiments for Precision Agriculture. Proceedings of the 4th European Conference on Precision Agriculture, J. Stafford & A. Werner, editors, Wageningen Academic Publishers, Netherlands, 2003.

Wilkinson, G.N., Eckert, S.R., Hancock, T.W., and Mayo, O. 1983. Nearest neighbor (NN) analysis with field experiments. *Journal of the Royal Statistical Society* B, 45:151-178.

Wittig, T.A. and Z.W. 2001. Wicks III, Simple On-Farm Comparisons. In Site-Specific Management Guidelines SSMG-18.

Young, D.L., Kwon, T.J., and Young, F.L. 2004. Downsizing an Agricultural Field Experiment Alters Economic Results: A Case Study. *Review of Agricultural Economics*, 26(2):255-264.