The Impact of Drainage Management Technology in Agriculture: A Spatial Panel Data Model*

Benoît A. Delbecq[†], Raymond J.G.M. Florax[‡], Adela Nistor[§], Jason P. Brown[¶], and Jess Lowenberg-Deboer[‖]

June 1, 2009

PRELIMINARY VERSION, NOT FOR QUOTATION.

*We thank Paul Elhorst for his Matlab routines, suggestions and constructive comments. We also thank Nathan Utt and Roxanne Adeyua for their help in putting the data together. We gratefully acknowledge the United States Department of Agriculture-Cooperative State Research and Extension Service Grant 2004-51130-03111 entitled "Drainage Water Management Impacts on Watershed Nitrate Load, Soil Quality and Farm Profitability" for supporting this research.

†Benoît A. Delbecq, Department of Agricultural Economics, Purdue University, 403 W. State St. West Lafayette, IN 47907-2056 (bdelbecq@purdue.edu); Corresponding author.

‡Raymond J.G.M. Florax, Department of Agricultural Economics, Purdue University, 403 W. State St. West Lafayette, IN 47907-2056 (rflorax@purdue.edu) and Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.

§Adela Nistor, Department of Economics, Brock University, 500 Glenridge Avenue, St. Catharines, Ontario, Canada (anistor@brocku.ca).

¶Jason P. Brown, Department of Agricultural Economics, Purdue University, 403 W. State St. West Lafayette, IN 47907-2056 (brown151@purdue.edu).

‖Jess Lowenberg-Deboer, Department of Agricultural Economics, Purdue University, 403 W. State St. West Lafayette, IN 47907-2056 (lowenbej@purdue.edu).

The Impact of Drainage Management Technology in Agriculture: A Spatial Panel Data Model

Abstract

This paper shows that spatial panel data models can be successfully applied to an econometric analysis of farm-scale precision agriculture data. The application focuses on the estimation of the effect of controlled water drainage management equipment on corn yields at the farm level in Indiana. Utilizing field-level precision agriculture data, geographical information systems and spatial panel techniques, a yield response equation is estimated. The choice of a random effects spatial error model stems from the desire to disentangle the effects of spatial dependence from spatial heterogeneity and omitted variables by incorporating spatial and temporal dependence in the error term, while controlling for topography, weather and the controlled drainage treatment. More specifically, we employ the model developed by Baltagi et al. (2007) which encompasses spatial panel models previously formulated by Anselin (1988) and Kapoor et al. (2007) by allowing for a bi-component error term with both unit-specific and macro spatial dependence. Additionally we augment this model with yearly dummies as well as the corresponding interactions with the experiment variable to account for temporal heterogeneity and potential differential response to the drainage treatment over time. Using data from the Davis Purdue Agricultural Center collected over seven years, we are able to show that controlled drainage is invariably beneficial for the East side of the field. The West side story is a little different: the water management system is penalizing yield three out of four years.

JEL classification: C21, C23, Q15

Keywords: spatial econometrics; spatial panel data; corn yield; precision agriculture; water drainage management

1 Introduction

This paper applies some of the most recent spatial panel data techniques to agricultural yield monitor data. Specifically, we investigate an experiment using controlled water drainage technology and assess its impact on corn yields at the farm level in the state of Indiana, in the United States.

In terms of application we focus on the impact of using water drainage management infrastructure on corn yields. Apart from the potentially beneficial effect of drainage water management practice on yields, the use of the controlled drainage technology is also motivated by environmental concerns. Excess nutrients from anthropogenic sources increase algal production, causing eutrophication of coastal ecosystems. For instance, in the Midwest of the United States too much nitrate (N) load in surface waters from drained agricultural land creates negative environmental impacts in the Gulf of Mexico (Burkhart and James, 1999; Gilliam et al., 1999; Rablais et al., 2002). In the future, farmers may therefore be required to adopt technologies that have been demonstrated to reduce N loads to surface water, such as controlled drainage, also referred to as drainage water management. Controlled drainage restricts outflow during periods of the year when equipment operations are not required in the field (i.e., winter and midsummer). This may increase water available to crops in midsummer and thereby increase yields (Evans and Skaggs, 1996b). Drainage trials in small plots are difficult, as they require major investment in barriers to prevent water movement between plots, thus creating an unnatural situation that may not be representative of field conditions. For drainage trials, landscape experimental designs works well and the most cost effective way to collect yields from landscape designs is with yield monitors. The drainage water cases studied in this paper are motivated by the recognition that voluntary adoption of drainage water management by growers depends on the size of the yield increase (Evans and Skaggs, 1996b). In addition, existing incentive programs such as the United States Department of Agriculture (USDA) Environmental Quality Incentives Program (EQIP) require quantitative information on practice efficacy and on private benefits.

The use of spatial panel techniques is motivated by the fact that precision agriculture data are measured at such a low level of spatial aggregation that spatial correlation is endemic. It is not uncommon to observe levels of spatial autocorrelation as high as 0.8 or even 0.9 (e.g. Lowenberg-

DeBoer et al. (2006)). In addition, the precision agriculture literature shows that yield response can vary substantially from year to year (e.g. Cooper et al. (1991, 1992); Mejia et al. (2000)). The area is therefore perfectly suited to analyze spatial panel datasets, allowing for rather involved spatially autocorrelated error structures as well as appropriate parameter variation over time.

The organization of the remainder of this paper is as follows. In Section 2, we review the literature and pay particular attention to the precision agriculture studies that have investigated the yield impacts of water drainage management technology. In addition, we point out some studies in this field that have adopted a spatio-temporal perspective. Section 3.1 shows that these techniques are ideally suited for the analysis of the case study data that we obtained. The corn yield data are described in detail, both in terms of data sources and their characteristics. In Section 3.2 we introduce various possible specifications for the water drainage impact model, and provide details about the implied spatial and temporal correlation structure. We also outline the statistical tests that can be used to attain valid inferences regarding the underlying model specifications. The empirical results are presented in Section 4. Section 5 concludes and provides suggestions for future research.

2 Literature review

Baltagi et al. (2007) provide a review of recent economic studies applying spatial panel data models. There have been only a small number of studies that employed spatio-temporal regression analysis in the study of yield monitor data (Bongiovanni and Lowenberg-Deboer, 2002; Lambert et al., 2006; Liu et al., 2006; Nistor, 2007). Spatio-temporal prediction has became significantly important (Pace et al., 1998) and can contribute to a better understanding of complex phenomena studied in precision agriculture. Bullock and Lowenberg-DeBoer (2007) provide a recent review of studies using spatial econometric analysis techniques applied to precision agriculture data.

The 1996-2002 USDA Agricultural Resource Management Surveys (ARMS) found that adoption of precision agriculture continues to grow. They have also shown that yield monitors are the most common precision agriculture technology used on major field crops, especially by corn and soybean producers. As more combines are equipped with yield monitors and the amount of yield monitor data increases with each season, there is a growing need to determine how it can be used to best help farmers make management decisions. As Griffin et al. (2005) points out, precision agriculture has renewed farmers interest in on-farm planned comparisons. Yield monitor data can be collected on-the-go and planned on-farm comparisons implemented, harvested, and analyzed without interfering with crop production when appropriate precision technologies are used.

There are only a few studies of the effect of drainage management on average crop yields, and none of those address conditions in the Midwest of the U.S. Sipp et al. (1986), Cooper et al. (1991, 1992), Drury et al. (1997) and Fisher et al. (1999) documented yield increases with subirrigation, but results are less clear-cut without subirrigation, which is the relevant benchmark for the present paper. In another study, Evans and Skaggs (1996a) stated that in the longrun, controlled drainage would generate average yield increases of 2 to 5% above yields with a conventional drainage system for a crop such as corn (in North Carolina). Tan et al. (1998) reported a study in Ontario, CAN which showed that controlled drainage increased average soybean yields 12 to 14% above free-flow drainage in a conventional tillage system, but decreased yields in no-tillage systems. Nine out of 15 farmers involved in a central Illinois drainage management project said that they had higher yields with drainage management (Pitts, 2003). Not all studies however have shown yield benefits from controlled drainage. Sipp et al. (1986) in Illinois, Grigg et al. (2003) in Louisiana, and Fausey et al. (2004) in Ohio have shown no significant difference of crop yield between controlled and conventional drainage. Specifically, Sipp et al. (1986) stated there was no significant difference in corn and soybean yields when comparing controlled versus conventional drainage as long as drainage occurred (i.e. water was not completely held back).

All the above studies estimating the effect of controlled drainage on yields use small plot or whole-field data with the harvest from the combine harvester transferred to a weigh wagon, and subsequent analysis based on comparing treatment trials or performing an analysis of variance. In both cases, however, spatial econometric or spatial statistical techniques have not been used. Effectively what this implies is that it is a priori assumed that the distribution of yields across the field is homogenous and independent of location. Some of the earliest studies using yield monitor data have estimated and compared site-specific crop response functions using multivariate regression analysis (Buhasa et al., 1995; Davis et al., 1996; Malzer et al., 1996; Nielsen et al., 1997, 1999; Wendroth et al., 1998). Early applications relied on ordinary least squares (OLS), which does not account for heteroskedastic or correlated errors (e.g. Long et al. (1993); Kessler and Lowenberg-DeBoer (1999); Bakhsh et al. (2000)). However Kessler and Lowenberg-DeBoer (1999); Lambert et al. (2004); Anselin et al. (2004) among others show that this is an erroneous assumption and that yield monitor data are heterogenous and spatially correlated. A number of studies have considered crop response to nitrogen application by estimating spatial econometric models (Lambert et al., 2003; Bongiovanni and Lowenberg-Deboer, 2000, 2002; Liu et al., 2006).

Recentering the focus on analyzing the impact of controlled drainage on yields, we have identified two studies which model explicitly the spatial nature of yield monitor data. Brown (2006) applies spatial econometric techniques to cross-section yield monitor data in 2005 for four farms located in White, Montgomery and Randolph County in Indiana in order to study the economic feasibility of controlled drainage in the Cornbelt. Using spatial error regression models for the estimation of yields as a function of linear, quadratic and interaction terms including elevation, slope, distance to the nearest tile line and infrared soil color, Brown (2006) found that controlled drainage impacts yield in the range of 8 bu/acre to 29 bu/acre. Nistor (2007) proposes a framework to model crop sensor data over time by using the spatial fixed and random effects models, with an application focused on estimating the controlled drainage impact on farm profitability in the Cornbelt. Nistor (2007) found the decision to invest in controlled drainage technology to be supported for three of the four experimental farms, both with and without subsidy.

3 Methods and data

3.1 Data and specification

The empirical example in this paper is concerned with yield monitor data sampled from Davis Purdue Agricultural Center (DPAC), field W, located in Randolph County, Indiana. The yield data were collected with an AgLeader yield monitor linked to a global positioning system (GPS). The yield maps are created using the geographical coordinates, e.g., latitude and longitude recorded, and yields (bu/a). Anselin et al. (2004) provides a more complete description of the composition of the yield files. The yield measurement samples collected have been taken from the field surface with the locations considered as points or very small areas (see Griffin et al. (2005), for a more elaborate discussion). The data was cleaned using Yield Editor v1.02, a yield monitor data processor designed by USDA-ARS Cropping Systems and Water Quality Unit. Points were removed based on several "combine dynamics" criteria such as minimum and maximum yield, combine speed, start and end of pass delay and grain flow delay. Each one of these criterion is suceptible to create erroneous data points which need to be deleted for the data to be meaningful. The design of the controlled and conventional drainage experiments are created via digitization using the tile line maps. Because the points of the raw yield data were closer inside the row than from one row to another, the dataset was constructed by aggregating the data points into squares of width the average combine pass in order to provide data that are spatially balanced in all directions. Previous applications of this methodology can be found in Malzer et al. (1996), Mamo et al. (2003) and Anselin et al. (2004). The grid thus created was overlaid on the yield points after rotation by the corresponding field angle in order for the row of cells to follow the combine harvestor passes through the field and avoid mixing data from multiple passes. Each cell value, expressed in bushels per acre (bu/a or bu.a^{-1}), represented the average yield of all points contained within that square so that a yield map was created with a finite number of color scales easily identifiable to the viewer from many thousands of individual yield point values. This process was performed using the same grid each year, so that the grids are coincident, which permits the comparison of yields for different years in the "same" location. Because of the data cleaning mentioned earlier, some grid ended up not intersecting with any yield point for certain years. These cells where removed from the grid and therefore some data was lost in the process for the other years. The balanced design thus obtained allows for a spatial econometric approach using a weighting design (Anselin et al., 2004). Moreover, since the prediction error for the average values of yields within grids is smaller than the prediction error for any yield point prediction, the precision of the average yield estimator is higher than that of point estimator (Haining, 2003), although this procedure also introduces heteroskedasticity to a certain extent. Elevation point data with reference to the sea level, collected by topographic surveys performed by contractors for the farm, were interpolated using the Inverse Distance Weighted (IDW) power 1 method, so that a point data set was obtained with elevation across the whole field. Each cell was assigned the average of the elevation points that completely fell inside each cell and was converted with reference to the lowest elevation level in the field. This implies that the elevation in each grid cell equals the difference between the average elevation with respect to the sea level and the minimum average elevation (see figure 1). Field W at the DPAC farm is split in two halves, one to the East and one to the West of a drainage ditch. The controlled drainage project was started in 2005 but corn yield monitor data are available as far back as 1996. The experiment consisted in upgrading the conventional drainage system already in place by allowing for the depth of the water table to be adjusted by means of logs inserted at strategic points of the drain network. Only the southern half of the West and the northern half of the East were modified while the remainder of the field was left unchanged as a benchmark. Until 2004, field W was cultivated under a corn-soybeans rotation but 2005 marked a transition into corn monoculture. It resulted that corn was harvested in 1996, 1998 and from 2005 until 2008 on both sides of the field. In addition, data was collected in 2000 and 2002 for the East side and 2001 and 2003 for the West side. The difference in cultivation history motivated the separate analysis of the two sides.

Rainfall data over the growing season, taken as July to September, were obtained from the weather station located at 0.5-mile distance from Field W. The choice of the growing season period was determined by professional judgment of soil scientists and agricultural engineers involved in the project. Although this is unusual, in some years (2005, 2006) corn did not reach physiological maturity (i.e., the R6 growth stage when black layer forms at the tip of the kernels) before the end of September due to late planting (end of May, early June). This motivates the inclusion of the September rain data.

Heady and Dillon (1961) provide a review of algebraic functional forms for crop response estimation. The selection of variables and specification of the crop yield functional form are difficult because of lack of theoretical guidance in the agronomy and soil science literature, and the complexity of yield response (Swanson, 1963; Florax et al., 2002; Anselin et al., 2004). Nistor (2007) provides an elaborate overview of different functional forms that have been used in agronomy and

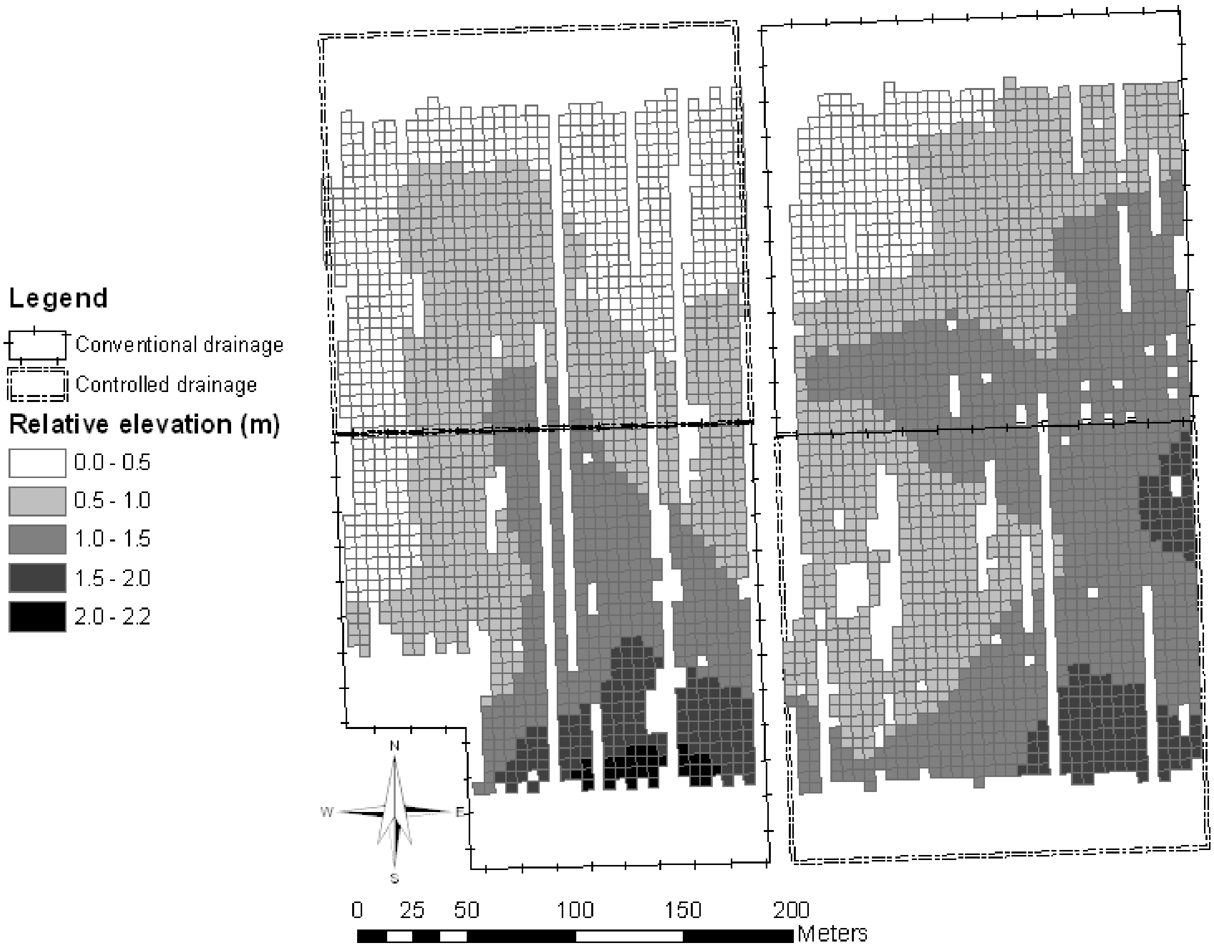

Figure 1: Elevation relative to the lowest point for each side (Davis, Field W)

soil science. For this application a simple linear form with interaction variables is chosen, because of the limited availability of data. For on-farm yield trials elevation and rainfall[1] are the most commonly available variables. Data that varies in time and space (e.g. annual soil tests, remotely sensed biomass) are sometimes available on research farms, but rarely for commercial fields like those used for the drainage trials.

Since the yield monitor data is a sample rather than a population (Griffin et al., 2005), the random effects (RE) model is more appropriate than the fixed effects (FE) model for the analysis of precision agriculture data. Nistor (2007) provides a discussion of the proper framework for precision agriculture data over time, and more specifically the fact that the spatial error model is more appropriate than the spatial lag model, because spatial autocorrelation is due to omitted variables rather than to the effect of corn yield grid cells on each other (Anselin et al., 2004; Lowenberg-DeBoer et al., 2006). In addition, temporal heterogeneity is much more important than spatial heterogeneity and should also be taken into account, since the yield response and the controlled drainage impact vary across the years (Bongiovanni and Lowenberg-Deboer, 2002; Nistor and Lowenberg-DeBoer, 2007). The lack of operational methods the extension of these models to the two-way framework in the literature, led us to consider temporal heterogeneity in the form of yearly dummy variables. Therefore, a random effects model allowing for spatially autocorrelated errors and extended to account for temporal heterogeneity would be the model of choice for estimation. Three such models exist in the literature: Anselin (1988) was the first to formally define a spatial random effects model, which was further extended by Kapoor et al. (2007) and Baltagi et al. (2007). Further details are given in the following section.

The controlled drainage treatment was incorporated in the model as a dummy variable which takes a value of one for the years and cells benefiting from the system. This dummy was interacted with the time dummies to account for likely differential impacts of controlled drainage across years. In other words, we expect the controlled drainage setup to either benefit, hurt or have no effect on yields in response to varying environmental conditions. The crop yield response to controlled drainage is different across years, with no yield benefit in years with insufficient rain, or a negative

[1] At the field level, precipitation is often a single value per year and hence cannot be used directly for purpose of estimation as it is perfectly collinear with the yearly dummies. It can however be included in interaction terms.

impact with very low field topography that would allow high enough water to have a detrimental effect (Nistor and Lowenberg-DeBoer, 2007). Because of the relationship between topographic attributes, soil properties and available water, the precipitation in the growing season is interacted with the topographic attributes that may influence crop yields (Kaspar et al., 2003). Finally, the interaction term between the drainage dummy and elevation was included since impact of controlled drainage vary with topography and controlled drainage does not affect yields the same across the field. With the inclusion of these interaction variables, the specification estimated reads as:

$$yield = \alpha + year\ \boldsymbol{\beta_1} + drain \times year\ \boldsymbol{\beta_2} + \gamma_1\ elev + \gamma_2\ drain \times elev + \delta\ rain \times elev \tag{1}$$

where $yield$ is the $NT \times 1$ vector of average yields, $year$ is the $NT \times T - 1$ matrix of time dummies, $drain$ is the $NT \times 1$ controlled drainage dummy, $rain$ is the $NT \times 1$ vector of July 5-Sept 5 precipitations, $elev$ is the $NT \times 1$ vector of elevations above the lowest point of the field, α is the intercept, N is the number of cells in the grid, and T is the number of years in the panel.

In view of the data aggregation procedure based on calculating average yields for data points included in each grid cell the actual left-hand side variable is average or expected yield rather than actual yields. This implies that the model provided in equation (1) is inherently heteroskedastic because the variance of the mean varies over grid cells. We therefore scale the left- and right-hand side of equation (1) by the standard error of the mean yield for each grid cell:

$$Std.Error(\overline{yield}_i) = \sqrt{\frac{\sigma_i^2}{n}} = \frac{\sigma_i}{\sqrt{n_i}} \;\; i = 1, \cdots, N \tag{2}$$

where $yield_i$ is the mean yield in grid cell i with mean $\overline{yield}_i$ and variance σ_i^2, and n_i is the number of yield points in grid cell i . Obviously this creates an operational problem in the case of grid cells for which only one data point is available. This occurs in about 5% of all grid cells. In those cases we have scaled the yield variable by the standard deviation of yield over the whole field (East or West) for each year.

3.2 Spatial panel models

The traditional panel data models used in applied research are the fixed effects (FE) and the random effects (RE) model (Baltagi, 2005). A panel data set consists of a sequence of observations repeated through time, on a set of units (e.g., individuals, firms, or countries). A panel data regression is different from a time-series or cross-section regression in that it considers both the temporal and the cross-sectional dimension. Panel data offer researchers extended modeling possibilities as compared to purely cross-sectional or time-series data, because they contain more information, more variability, less collinearity among the variables, more degrees of freedom, and hence the estimators are likely to be more efficient. Panel models also allow for the specification of more complicated behavioral hypotheses, including effects that cannot be addressed using pure cross-sectional or time-series data. For example, technical efficiency is better studied and modeled with panel data sets, because in cross-sectional models changes in technology cannot be identified, and in time series models the state of technology is assumed to be identical across cross-sectional units (Hsiao, 2003; Baltagi, 2005). An important advantage of panel data techniques is that they are better suited to identify and measure effects that are simply not detectable in pure cross-section or pure time-series data (Ben-Porath, 1973). Panel data can also reduce the effect of omitted variable bias by controlling for (unobserved) individual heterogeneity. Time-series and cross-section studies not controlling for this heterogeneity run the risk of obtaining biased results (Moulton, 1986, 1987).

Contemporaneous spatial dependence between observations at each point in time and spatial heterogeneity (i.e., parameter variation over space) may arise when panel data include a location component (Anselin, 1988; Elhorst, 2009). Spatial dependence may be incorporated into the model as spatial error autocorrelation or as a spatially lagged dependent variable, or a combination of both (Anselin and Hudak, 1992). These different specifications of spatial dependence have different implications for estimation and statistical inference. Estimating a model ignoring spatial error autocorrelation by means of Ordinary Least Squares (OLS) produces unbiased and consistent parameter estimates, but the OLS estimator loses the efficiency property. Erroneously omitting a spatially autocorrelated dependent variable from the explanatory variables causes the OLS estimator to be biased and inconsistent, except under special circumstances (Anselin, 1988).

Anselin et al. (2008) provides an overview of specifications and estimators available for spatial panel data. Building on conventional panel data models, the spatial random effects model (henceforth SEM-RE) described in Anselin (1988) comprises a bi-component error term split between an individual specific part and a residual part[2], the latter being spatial correlated. This model has recently been been extended by allowing the location specific component to be autocorrelated as well. First Kapoor et al. (2007) (henceforth KKP) assume that the spatial error autocorrelation is the same between the individual effects and the remainder errors, which is arguably a fairly strong and potentially unrealistic constraint. Baltagi et al. (2007) (henceforth BEP) go further by letting the spatial and remainder error components follow different spatial processes. The SEM-RE and KKP models being nested within the BEP model, Baltagi et al. (2007) proceed to derive a battery of Lagrange Multiplier and Likelihood Ratio tests for the proper specification of a random effects model with a spatial autoregressive error process. In the present study, we started from the more general model, i.e. the BEP model, and tested the results against the conventional RE, the SEM-RE and the KKP models which required estimation of all four models. Regarding the software resources for estimating panel spatial econometrics, the situation is still rather bleak Anselin et al. (2008). Aside from the estimation framework formulated and operationalized[3] by Elhorst (2009) for the SEM-RE model, there are no canned routine packages which are readily availble. For sake of consistency and completeness, the authors programmed all four estimations procedures in R based on the maximum likelihood framework described in Baltagi et al. (2007). A more detailed discussion of the models is presented in the remainder of this section.

Following Baltagi et al. (2007), the random effects model with spatially autocorrelated error

[2] The individual specific component can be considered as a random intercept

[3] Elhorst (2009) has made available on his website a battery of Matlab routines for estimation of Fixed Effects and Random Effects models with spatial lag or spatial error processes

components can be formulated for each time period t as:

$$\begin{aligned}\mathbf{y}_t &= \mathbf{X}_t\boldsymbol{\beta} + \mathbf{u}_t, \qquad t = 1, \ldots, T \\ \mathbf{u}_t &= \mathbf{u}_1 + \mathbf{u}_{2t} \\ \mathbf{u}_1 &= \rho_1 \mathbf{W}_N \mathbf{u}_1 + \boldsymbol{\mu}, \quad \boldsymbol{\mu} \overset{iid}{\sim} \mathcal{N}(0, \sigma_\mu^2) \\ \mathbf{u}_{2t} &= \rho_2 \mathbf{W}_N \mathbf{u}_{2t} + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\nu^2)\end{aligned} \tag{3}$$

where $\mathbf{y}_t$ is a vector of N observations of the dependent variable (at time t) and $\mathbf{X}_t$ is an $N \times K$ matrix of exogenous variables including an intercept with its corresponding vector of coefficients $\boldsymbol{\beta}$. The vector of errors $\mathbf{u}_t$ is composed of two spatially autocorrelated components, one time-invariant and unit-specific $\mathbf{u}_1$, and one time-varying $\mathbf{u}_{2t}$. The BEP model assumes that both spatial processes are subject to the same neighborhood structure, i.e. an $N \times N$ weights matrix $\mathbf{W}_N$, but it allows for different spatial autocorrelation parameters ρ_1 and ρ_2, respectively. The spatial weights matrix $\mathbf{W}_N$ is subject to the standard regularity conditions[4] and is normalized in such a way that its rows sum to one. It is further assumed that the elements of μ are independent across $i = 1, \ldots, N$, the elements of $\boldsymbol{\nu}_t$ are independent across i and t and they are also independent of each other.[5]

Stacking the model for each time period so that the slower index is time and rearranging equation (3) into a single equation yield the following reduced form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I}_{NT} - \rho_1 \mathbf{W}_{NT})\boldsymbol{\mu} + (\mathbf{I}_{NT} - \rho_2 \mathbf{W}_{NT})\boldsymbol{\nu} \tag{4}$$

where $\mathbf{y} = [\mathbf{y}_1', \ldots, \mathbf{y}_T']'$, $\mathbf{X} = [\mathbf{X}_1', \ldots, \mathbf{X}_T']'$, and $\boldsymbol{\nu} = [\boldsymbol{\nu}_1', \ldots, \boldsymbol{\nu}_T']'$. Because of the time invariance, $\boldsymbol{\mu} = \boldsymbol{\iota}_T \otimes \boldsymbol{\mu}$ where $\boldsymbol{\iota}_T$ is a vector of ones of dimension T, and $\mathbf{W}_{NT} = \mathbf{I}_T \otimes \mathbf{W}_N$. Finally, $\mathbf{I}_{NT}$ is the $NT \times NT$ identity matrix. The regularity conditions assumed on $\mathbf{W}_N$ ensure, among other things, that the two blockdiagonal matrices $\mathbf{A} = \mathbf{I}_{NT} - \rho_1 \mathbf{W}_{NT}$ and $\mathbf{B} = \mathbf{I}_{NT} - \rho_2 \mathbf{W}_{NT}$ are non-singular and hence invertible.

Baltagi et al. (2007) formalize the log likelihood function of their model as:

[4] See Kelejian and Prucha (1999) for a detailed description of these regularity conditions

[5] For a more complete description of the set of assumptions linked to the BEP model, Baltagi et al. (2007) provides more details.

$$LL_{BEP}(\boldsymbol{\beta}, \sigma_\nu^2, \sigma_\mu^2, \rho_\nu, \rho_\mu) = -\frac{NT}{2}\ln 2\pi - \frac{1}{2}\ln\det\left[T\sigma_\mu^2(\mathbf{A}'\mathbf{A})^{-1} + \sigma_\nu^2(\mathbf{B}'\mathbf{B})^{-1}\right] - \frac{T-1}{2}\ln\det\left[\sigma_\nu^2(\mathbf{B}'\mathbf{B})^{-1}\right] - \frac{1}{2}\mathbf{u}'\boldsymbol{\Omega}_u^{-1}\mathbf{u} \tag{5}$$

where $\mathbf{u} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ and $\boldsymbol{\Omega}_u^{-1}$ can be expressed as:

$$\boldsymbol{\Omega}_u^{-1} = \bar{\mathbf{J}}_T \otimes \left[T\sigma_\mu^2(\mathbf{A}'\mathbf{A})^{-1} + \sigma_\nu^2(\mathbf{B}'\mathbf{B})^{-1}\right]^{-1} + \frac{1}{\sigma_\nu^2}\left(\mathbf{E}_T \otimes (\mathbf{B}'\mathbf{B})\right) \tag{6}$$

where the elements of the $T \times T$ matrix $\bar{\mathbf{J}}_T$ are all equal to $\frac{1}{T}$ and $\mathbf{E}_T = \mathbf{I}_T - \bar{\mathbf{J}}_T$. Note that even though $\boldsymbol{\Omega}_u^{-1}$ is the inverse of an $NT \times NT$ matrix, its computation only involves the inverse of $N \times N$ matrices, making the maximum likelihood estimation computationally more manageable. Combining σ_μ^2 and σ_ν^2 into $\phi = \frac{\sigma_\mu^2}{\sigma_\nu^2}$, the log likelihood function in equation (6) can be concentrated over $\boldsymbol{\beta}$ and σ_ν^2:

$$LL_{BEP}^c(\rho_1, \rho_2, \phi) = -\frac{NT}{2}\ln 2\pi - \frac{NT}{2}\ln\hat{\sigma}_\nu^2(\rho_1, \rho_2, \phi) - \frac{1}{2}\ln\det\boldsymbol{\Sigma}_u(\rho_1, \rho_2, \phi) - \frac{NT}{2} \tag{7}$$

where

$$\begin{aligned}
\hat{\sigma}_\nu^2 &= \frac{\mathbf{u}(\hat{\boldsymbol{\beta}})'\boldsymbol{\Sigma}_u^{-1}\mathbf{u}(\hat{\boldsymbol{\beta}})}{NT} \\
\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\boldsymbol{\Sigma}_u^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}_u^{-1}\mathbf{y} \\
\boldsymbol{\Sigma}_u^{-1} &= \sigma_\nu^2\boldsymbol{\Omega}_u^{-1} \\
\ln\det\boldsymbol{\Sigma}_u &= \ln\det\left[T\phi(\mathbf{A}'\mathbf{A})^{-1} + (\mathbf{B}'\mathbf{B})^{-1}\right] + (T-1)\ln\det(\mathbf{B}'\mathbf{B})^{-1}
\end{aligned}$$

In the end, estimation of the BEP model requires the quasi-Newton non-linear maximization of equation (7) over its three parameters ρ_1, ρ_2 and ϕ. The authors define the parameter space for ρ_1 and ρ_2 as comprised between -1 and 1. However, following Anselin (1980) we allow the parameter optimization algorithm to span the interval $(\frac{1}{\omega_{min}}, \frac{1}{\omega_{max}})$ where ω_{min} and ω_{max} are the minimum and maximum eigenvalues of W_N respectively.[6] Since $\sigma_\mu^2 > 0$ and $\sigma_\nu^2 > 0$, then $\phi > 0$. As mentioned earlier, the BEP model encompasses all other RE models with spatial error processes. Indeed, the KKP model is easily obtained by setting $\rho_1 = \rho_2$, the SEM-RE corresponds to the case where $\rho_1 = 0$ and setting $\rho_1 = \rho_2 = 0$ yields the aspatial RE model. Baltagi et al. (2007) provide the corresponding Lagrange multiplier (LM) and Likelihood ratio (LR) tests along with their comprehensive derivation in the appendices. As of now, we only implemented the LR tests but we intend on adding the LM tests in future developments of this research. Likelihood ratio tests all follow the same principle, namely compare the value of the maximized log likelihoods of the restricted models to be tested (i.e. KKP, SEM-RE and RE) and the unrestricted alternative (i.e. BEP). The likelihood ration test can be defined as:

$$LR = -2(LL_u - LL_r) \quad \text{with} u = \text{BEP and } r = \text{KKP, SEM-RE, and RE}$$

Asymptotically, $LR \sim \mathcal{X}_q^2$ with $q = \#$ of restrictions.[7]

4 Data and results

4.1 Exploratory spatial data analysis

We can see from tables 1 and 2 that mean corn yield is fairly stable over time, except for 1996 (weed problems), 2002 (severe drought) and 2007. A quick comparison of the yields under controlled and conventional drainage after 2005 for both field would indicate that the former may generate higher yields with the exception of 2005 for the West and 2006 for the East. However the comparison based on average yield may be misleading because it does not take into account differences in topography, soils, microclimate and other factors between controlled drainage areas and those with free flowing drainage. A similar eyeball analysis of standard deviations would point toward a stabilization of yields by the water management system with exceptions in 2007 for the East and 2005 for the West.

[6] Given a row-normalized weights matrix W_N, ω_{max} is equal to 1.
[7] $q = 1$ for the KKP and SEM-RE models and $q = 2$ for the RE model.

(Controlled)	1996	1998	2002	2005	2006	2007	2008
Minimum	59	113	10	77	96	33	106
Maximum	131	189	105	221	219	154	238
Mean	99	144	45	174	172	107	192
SD	13	15	19	22	20	22	22
(Uncontrolled)							
Minimum	47	76	11	79	106	50	110
Maximum	131	200	98	237	218	163	239
Mean	97	145	51	154	175	107	192
SD	12	17	19	30	20	20	23
(Whole field)							
Minimum	47	76	10	77	96	33	106
Maximum	131	200	105	237	219	163	239
Mean	98	145	48	164	174	107	192
SD	13	16	19	28	20	21	23
rain (in)	3.60	4.03	2.53	5.67	3.78	8.24	5.38

Table 1: Corn yield ($bu.a^{-1}$) and precipitation - descriptive statistics, Davis, Field W, EAST

Controlled	1996	1998	2001	2003	2005	2006	2007	2008
Minimum	32	106	104	56	79	86	51	107
Maximum	119	200	227	184	206	220	151	239
Mean	81	151	177	137	150	167	110	196
SD	18	20	20	22	22	22	16	21
Uncontrolled								
Minimum	35	56	112	52	89	81	52	124
Maximum	121	208	232	188	209	216	141	236
Mean	89	138	175	123	156	155	104	189
SD	13	22	19	28	19	24	18	21
Whole field								
Minimum	32	56	104	52	79	81	51	107
Maximum	121	208	232	188	209	220	151	239
Mean	85	145	176	130	153	162	107	193
SD	16	22	19	26	21	24	18	21
rain (in)	3.60	4.03	4.96	7.33	5.67	3.78	8.24	5.38

Table 2: Corn yield ($bu.a^{-1}$) and precipitation - descriptive statistics, Davis, Field W, WEST

Referring to figure 1, one can observe that the field roughly slopes up when going South. This means that the Western part of the field has lower elevations where the controlled drainage system is in place, and the opposite is true for the Eastern part of the field. This will be have important repercussions on the effect of drainage trials on yields as will be discussed further in section 4.2.

Figure 2 shows one example of yield map. The year 2005 was chosen as an illustration for its "normality" in terms of yield minima, maxima and average. A direct observation of the map gives the impression that yields were higher in the Eastern half of the field. However, when looking more closely at the summary statistics, it becomes evident that this is only the case for the controlled sections of field (i.e. the South-East), while on average the free-flow areas are similar (with however, more extreme values in the East). Focusing our attention on the spatial distribution of yields in 2005, it would appear that there is some form of clustering of high (low) yield cells next to high (low) yield cells, which is referred to as positive autocorrelation.

Moran's I statistic provides the spatial econometrician with a tool to calculate the degree of spatial autocorrelation in a particular variable. In order to compute this statistic, a neighborhood structure, namely a weights matrix, needs to be defined. Spatial panel estimation requires massive computational power and the degree of connectivity of the weights matrix, i.e. the number of non-zero links, affects dramatically the feasibility of such estimation. Given the regular lattice nature of our data, a queen contiguity criterion of order one seems appropriate. Under this specification, grid cells are neighbors if they have a common border in the horizontal or vertical dimension, or if they share a common vertex, up to the one "band" of neighbors. When the weights matrix is row standardized, the spatially lagged yield variable, i.e. the product of the yield variable by the weights matrix, is the average of the yields in the neighboring grid cells. Inference on Moran's I can be done in multiple ways, one of which consists in permutating the locations and recomputing Moran's I a large number of time resulting in a distribution which the original estimated Moran's I can be compared to. We can see from table 3 that the sign of Moran's I statistic for yields is positive and highly significant so that high (low) values are surrounded by high (low) values in neighboring grids, indicating positive spatial correlation of yields.

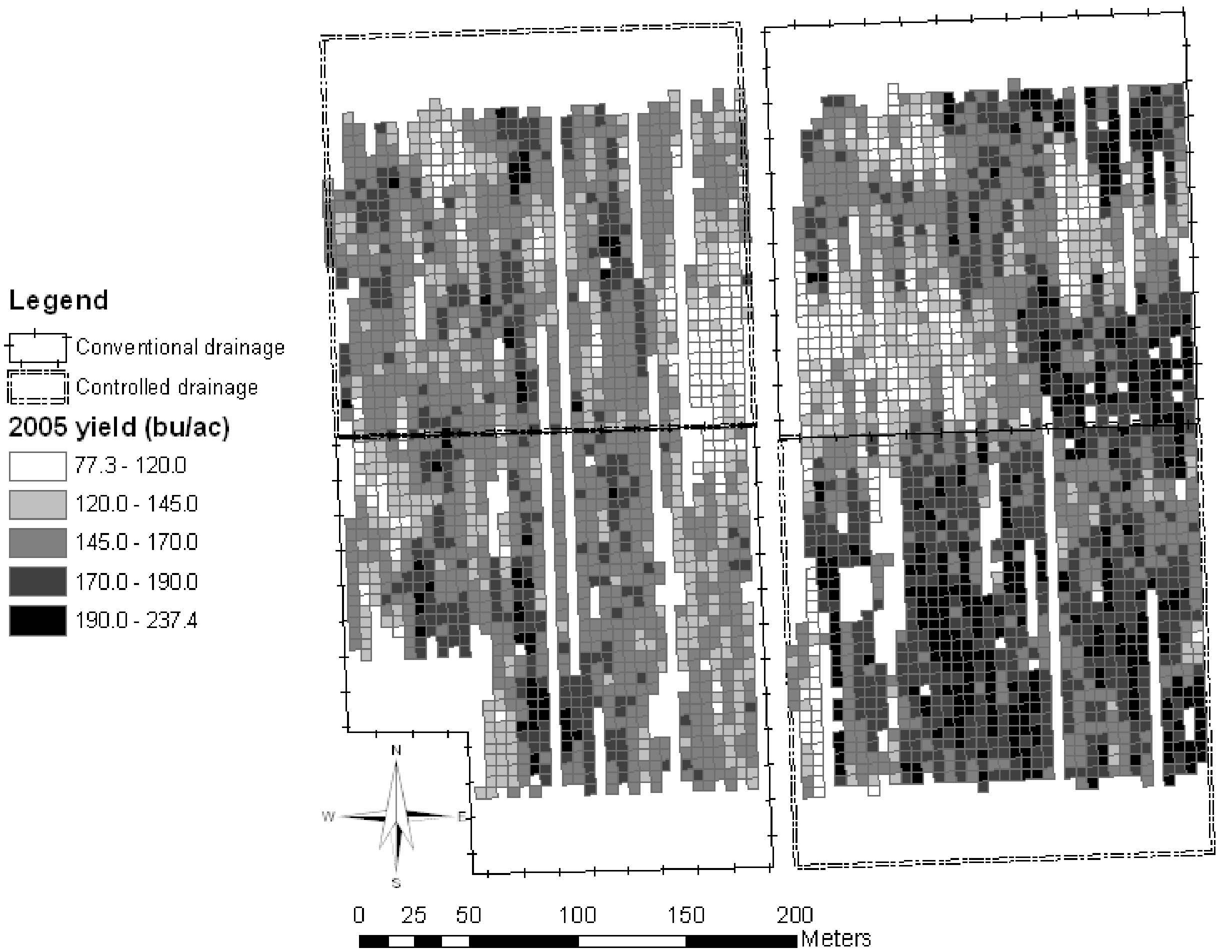

Figure 2: 2005 yield map (Davis, Field W)

	1996	1998	2001	2002	2003	2005	2006	2007	2008
EAST	0.48***	0.68***		0.62***		0.63***	0.56***	0.73***	0.62***
WEST	0.71***	0.83***	0.37***		0.62***	0.58***	0.62***	0.67***	0.61***

Table 3: Moran's I (yields), Davis, Field W

4.2 Regression results

Table 4 reports the estimation results for all four models: a-spatial RE, SEM-RE, KKP and BEP. The first thing to notice is that estimates do not vary in terms of signs or significance levels. There are minor changes in magnitude from one set of estimates to another. All three likelihood ratio tests reject their respective null hypothesis which indicates that the BEP model is the correct specification. Direction interpretation of the estimates is possible for the yearly dummies but marginal effects are a bit more complex when considering elevation and more importantly, the controlled drainage variable. Using 1996 as the reference, which recall was a bad year because of weed problems, it is comforting to see that all but one time dummy indicate higher yields. The severity of the 2002 drought is well displayed in these results with yields on average 78 bushels lower than in 1996.

Similarly, table 5 presents the estimation results for three out of the four models for the West side of Field W: a-spatial RE, SEM-RE and BEP. The reason for not reporting the KKP model is that the LR test for the null hypothesis $H0 : \rho_1 = 0$ fails to be rejected hereby rendering both the BEP and KKP models not relevant. Therefore, we will rely on the SEM-RE results for interpretation since the simple RE model is rejected by the relevant LR test. Still referring to 1996 as the benchmark, there are no negative estimates for the yearly dummies which is consistent with the observation of the data (2002 was not in corn on this half of the field).

The marginal effect of the water management system in year t ($t = 2005, \ldots, 2008$) is calculated as:

$$\left(\frac{\partial yield}{\partial D}\right)_t = \beta_{2t} \cdot year_t + \gamma_2 \cdot elev \tag{8}$$

This figure needs to be interpreted relatively to the conventional drainage since D is a dummy variable but this notion will remain implicit throughout our discussion of the results. This means that a positive (negative) marginal effect indicates that controlled drainage outperformed (was outperformed by) free-flow drainage. By plugging in the average elevation for the controlled part of each side of the field, it is possible to determine the average impact of the controlled drainage system on corn yields in a given year. When found not significant, a parameter was not included[8].

[8]This explains why the same marginal effects are shown in 2006 and 2007 for the East and 2005 and 2007 for the

yield	RE	SEM-RE	KKP	BEP
intercept	104.30 ***	104.21 ***	104.23 ***	104.33 ***
	(0.407)	(0.434)	(0.434)	(0.438)
year98	45.16 ***	45.85 ***	45.85 ***	45.84 ***
	(0.294)	(0.328)	(0.326)	(0.321)
year02	-75.43 ***	-78.31 ***	-78.24 ***	-77.92 ***
	(0.487)	(0.518)	(0.516)	(0.510)
year05	66.13 ***	68.39 ***	68.38 ***	68.25 ***
	(1.264)	(1.336)	(1.332)	(1.321)
year06	74.55 ***	73.25 ***	73.37 ***	73.35 ***
	(1.506)	(1.784)	(1.776)	(1.756)
year07	38.93 ***	40.79 ***	40.74 ***	40.56 ***
	(0.855)	(0.835)	(0.833)	(0.829)
year08	105.32 ***	109.83 ***	109.71 ***	109.26 ***
	(1.588)	(1.784)	(1.777)	(1.761)
elev	13.12 ***	14.10 ***	14.05 ***	13.76 ***
	(0.963)	(0.961)	(0.959)	(0.958)
year05 × D	15.62 ***	12.23 ***	12.22 ***	12.17 ***
	(1.744)	(1.842)	(1.837)	(1.828)
year06 × D	0.89	2.19	1.96	1.62
	(2.066)	(2.420)	(2.409)	(2.386)
year07 × D	1.35	0.90	0.89	0.84
	(1.096)	(1.123)	(1.12)	(1.114)
year08 × D	33.70 ***	32.57 ***	32.60 ***	32.72 ***
	(1.694)	(1.856)	(1.850)	(1.834)
elev × D	1.69 ***	1.72 ***	1.72 ***	1.74 ***
	(0.250)	(0.233)	(0.232)	(0.234)
elev × rain	-5.71 ***	-5.97 ***	-5.96 ***	-5.93 ***
	(0.209)	(0.202)	(0.202)	(0.201)
ρ_μ	—	—	0.527	0.949
ρ_ν	—	0.538	0.527	0.477
Log Likelihood	-63747.6 ***	-62644.0	-62631.6	-62565.0
LR vs BEP	2364.1 ***	157.4 ***	132.0 ***	—

Table 4: Estimation results for Davis field W - East

yield	RE	SEM-RE	KKP	BEP
intercept	76.80 ***	76.45 ***	—	76.45 ***
	(0.417)	(0.462)		(0.462)
year98	60.95 ***	62.43 ***	—	62.44 ***
	(0.399)	(0.432)		(0.432)
year01	101.89 ***	100.32 ***	—	100.33 ***
	(1.555)	(1.640)		(1.641)
year03	77.48 ***	78.26 ***	—	78.27 ***
	(1.894)	(1.998)		(1.998)
year05	89.90 ***	89.13 ***	—	89.17 ***
	(2.008)	(2.098)		(2.099)
year06	74.24 ***	75.03 ***	—	75.03 ***
	(2.304)	(2.730)		(2.731)
year07	63.06 ***	63.27 ***	—	63.28 ***
	(2.471)	(2.534)		(2.535)
year08	153.35 ***	153.31 ***	—	153.32 ***
	(0.872)	(0.907)		(0.908)
elev	30.36 ***	28.97 ***	—	29.00 ***
	(1.890)	(1.973)		(1.974)
year05 $\times$ D	1.89	1.41	—	1.37
	(3.056)	(3.344)		(3.346)
year06 $\times$ D	35.94 ***	37.46 ***	—	37.46 ***
	(2.864)	(3.190)		(3.191)
year07 $\times$ D	2.31	1.77	—	1.75
	(2.417)	(2.446)		(2.446)
year08 $\times$ D	13.766 ***	14.52 ***	—	14.51 ***
	(0.964)	(0.963)		(0.963)
elev $\times$ D	-33.76 ***	-33.90 ***	—	-33.87 ***
	(2.288)	(2.278)		(2.278)
elev $\times$ rain	-7.02 ***	-6.69 ***	—	-6.69 ***
	(0.469)	(0.486)		(0.486)
ρ_μ	—	—	—	-0.562
ρ_ν	—	0.568	—	0.57
Log Likelihood	-70651.1	-69101.7	—	-69100.7
LR vs BEP	3100.8 ***	1.91	—	—

Table 5: Estimation results for Davis field W - West

Δyield	2005	2006	2007	2008
EAST				
bu.ac^{-1}	14.1	2.0	2.0	34.7
%	8.1	1.2	1.9	18.1
WEST				
bu.ac^{-1}	-17.7	19.8	-17.7	-3.2
%	-11.8	11.9	-16.1	-1.6

Table 6: Average effect of controlled drainage on yields

Table 6 reports these figures:

Two principal conclusions can be drawn from these results. First, the average impact of controlled drainage on yields varies dramatically from one side to the other. It is so vastly different that all but one year's crop suffered from the presence of the water management system on the Western part of the field, while all years benefited from it for the East. Second the results display tremendous variability from year to year within each side. The average marginal effect ranges from 2 bu.a^{-1} in 2006 and 2007 to 34.7 bu.a^{-1} in 2008 for the East and from -17.7 bu.a^{-1} in 2005 and 2007 to 19.8 bu.a^{-1} in 2006 for the East. It is worth pointing out that the controlled drainage system performs simultaneously best in the Western part of field W and worst in the Eastern part. We are not able to explain these disparities at this moment, in particular the opposite responses of the drainage system to elevation between the two sides of the field, but we are working with agronomists to find an plausible interpretation for these results. One last point about these results is that they are generally higher than previous estimates found in the literature with the exception of Brown (2006). However, we would like to stress that the present research is the first known study to incorporate spatial dependence in combination with temporal heterogeneity in the analysis of this problem and as such has the potential to provide the most "accurante" estimates. In future developments of this research, we will calculate confidence intervals for the marginal effects.

Equation (8) shows clearly that the effect of the controlled drainage system depends directly on elevation. Figure 3 illustrates this fact visually for 2006:

Pay close attention to different scales of the marginal effects for both sides of the field. Not suprisingly, the impact of the controlled drainage on yield follows the elevation curves. More

West

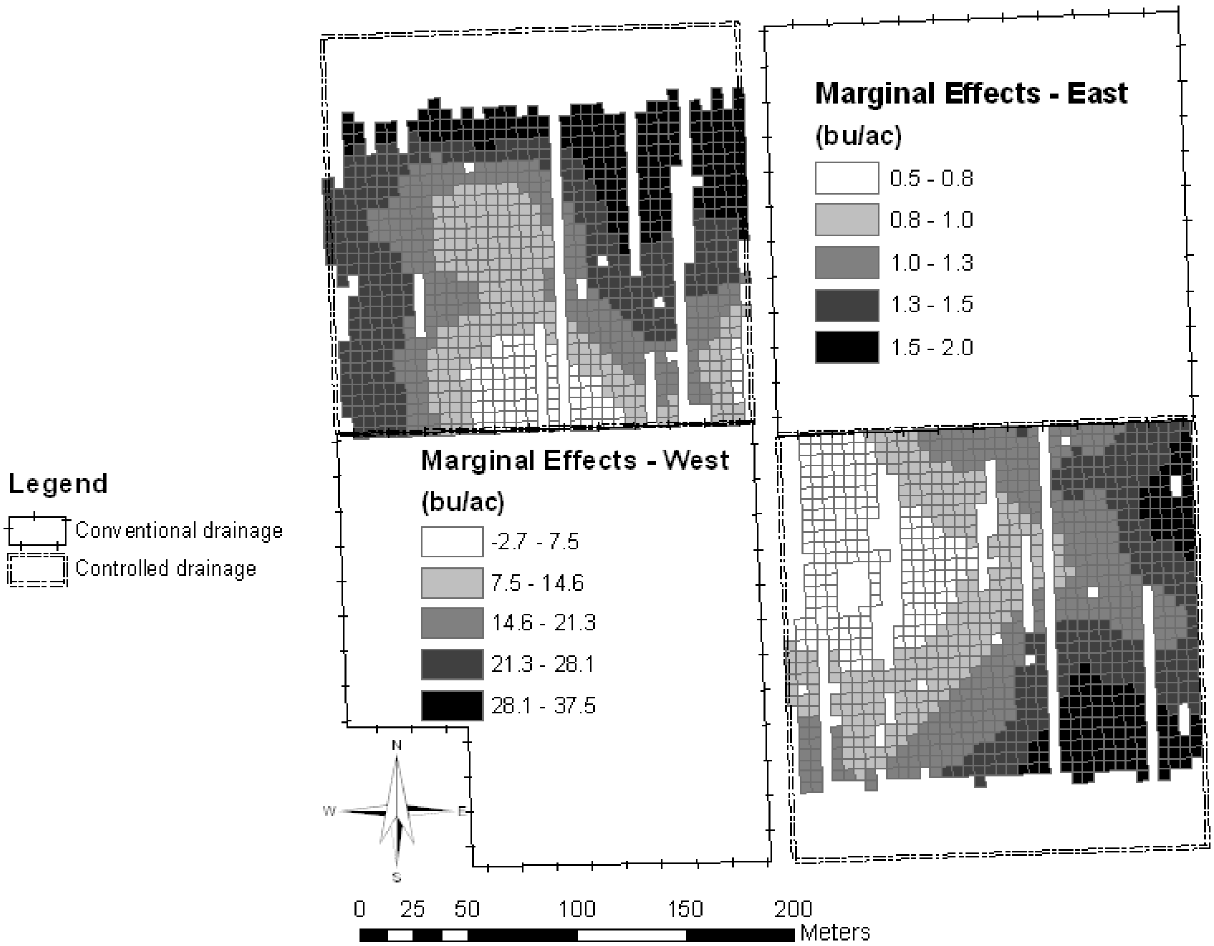

Figure 3: Effect of the water management system on corn yield (Davis, Field W, 2006)

interestingly, this map shows that even if the average marginal effect is positive and substantial in 2006 for the Western part of the field, there are parts of the field which actually experience a yield decrease due to the water management system.

5 Conclusions

This study shows that spatial panel data models can be applied to an econometric analysis of farm-scale precision agriculture information in data rich environments with independent variables that vary over time and space. The application deals with the assessment of the impact of controlled drainage technology on corn yields for two sides of one field in Indiana. Using field-level yield monitor data, the yield response equation is estimated using spatial panel econometric models, namely the spatial autoregressive error random effects model with both spatial and temporal heterogeneity incorporating spatial dependence in the error term, while controlling for the topography, weather and the controlled drainage treatment. The use of random effects allows for the disentanglement of the effects of spatial dependence from spatial heterogeneity and omitted variables, and thus, is necessary to properly investigate the yield response. The results show that the relationship between controlled drainage and corn yields is quite variable across years and fields and is directly related to elevation. While the controlled drainage is always better than its free-flow counterpart for the East side, it is outperformed three out of four years in the West. Evaluated at mean elevation in the relevant parts of the field, controlled drainage is credited with up to 34.4 bu/a in 2008 for the East but the increase is limited to a meager 2.0 bu/a in 2006 and 2007. On the other side of the ditch, the water management system is associated with a 19.8 bu/a rise in yields in 2006 but conventional drainage performed 17.7 bu/a better in 2005 and 2007.

This paper shows both results regarding controlled drainage impact on corn yields and a method of how to analyze precision agriculture data over time, by using GIS and spatial panel methods. Precision agriculture researchers can use the applied frameworks for modeling crop sensor data over time, to better evaluate the effect of various management practices and better understand the complex crop growth phenomena studied in precision agriculture. Regarding the implications for drainage management, the results have to be interpreted cautiously, due to drainage management issues. The experimental field was not under controlled drainage over the winter period, as environmental best practices would require (Frankenberger et al., 2006). More data is needed for more precise results. Inferences cannot be generalized to all the fields in the Midwest or beyond, since the analysis focuses on within field variations. Our efforts will focus on estimating similar models for three other sites in Indiana which are part of the same project. These three sites are farmer-owned and therefore suffer from a lesser data quality as well as shorter time periods hereby rendering spatial data analysis less performant but we hope to get interesting results nonetheless which would reinforce the present study.

References

Anselin, L. (1980). *Estimation Methods for Spatial Autoregressive Structures.* Regional Science Dissertation and Monograph Series, Cornell University, Ithaca, NY.

Anselin, L. (1988). Spatial econometrics: methods and models.

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.

Anselin, L. and S. Hudak (1992). Spatial econometrics in practice: A review of software options. *Regional Science and Urban Economics 22*(3), 509–536.

Anselin, L., J. Le Gallo, and H. Jayet (2008). Spatial panel econometrics. *Advanced Studies in Theoretical and Applied Econometrics 46*, 625.

Bakhsh, A., T. Colvin, D. Jaynes, R. Kanwar, and U. Tim (2000). Using soil attributes and GIS for interpretation of spatial variability in yield. *Transactions-American Society of Agricultural Engineers 43*(4), 819–828.

Baltagi, B. (2005). *Econometric Analysis of Panel Data.* John Wiley & Sons Ltd.

Baltagi, B., P. Effer, and M. Pfaffermayr (2007). A Generalized Spatial Panel Data Model with Random Effects. *Working paper*.

Baltagi, B., S. Heun Song, B. Cheol Jung, and W. Koh (2007). Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics 140*(1), 5–51.

Ben-Porath, Y. (1973). Labor-force participation rates and the supply of labor. *The Journal of Political Economy*, 697–704.

Bongiovanni, R. and J. Lowenberg-Deboer (2000). Nitrogen Management in Corn Using Site-Specific Crop Response Estimates From a Spatial Regression Model. In R. P., R. R., and W. Larson (Eds.), *Proceedings of the 5th International Precision Agriculture Conference, Madison, WI, 2002.*

Bongiovanni, R. and J. Lowenberg-Deboer (2002). Economics of nitrogen response variability over space and time: results from 1999-2001 field trials in Argentina. In *Proceedings of the 6th International Precision Agriculture Conference, Minneapolis, MN, 2002.*

Brown, J. (2006). Methodology for determining the economic feasibility of controlled drainage in the eastern corn belt. *Unpublished manuscript, MS Thesis, Purdue University*.

Buhasa, G., J. Apland, and D. Hicks (1995). *A regression Analysis of the Effects of Planting Date and Variety on Corn Yields in Minnesota.*

Bullock, D. 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.

Burkhart, M. and D. James (1999). Agricultural-nitrogen contributions to hypoxia in the Gulf of Mexico. *Journal of Environmental Quality 28*, 850–859.

Cooper, R., N. Fausey, and J. Streeter (1991). Yield potential of soybean grown under a subirrigation/drainage water management system. *Agronomy Journal 83*(5), 884.

Cooper, R., N. Fausey, and J. Streeter (1992). Effect of water table level on the yield of soybean grown under subirrigation/drainage. *Journal of production Agriculture 5*(1), 180–184.

Davis, J., G. Malzer, P. Copeland, J. Lamb, P. Robert, and T. Bruulsema (1996). Using yield variability to characterize spatial crop response to applied N. In *Proceedings of the Third International Conference on Site-Specific Management for Agricultural Systems.*

Drury, C., C. Tan, J. Gaynor, T. Oloya, I. Van Wesenbeeck, and D. McKenney (1997). Optimizing corn production and reducing nitrate losses with water table control-subirrigation. *Soil Science Society of America Journal 61*(3), 889.

Elhorst, J. (2009). Spatial Panel Data Models. *Handbook of Applied Spatial Analysis*, 2.

Evans, R. and W. Skaggs (1996a). Agricultural Water Mangement For Coastal Plain Soils. Technical Report AG 356, North Carolina Cooperative Extension Service.

Evans, R. and W. Skaggs (1996b). Operating controlled drainage and subirrigation systems. Technical Report AG 356, North Carolina Cooperative Extension Service.

Fausey, N., K. King, B. Baker, and R. Cooper (2004). Controlled drainage performance on Hoytville soil in Ohio. In *drainage VIII proceedings of the eighth International Symposium. R. Cook, ed. ASAE Publication*, pp. 84–88.

Fisher, M., N. Fausey, S. Subler, L. Brown, and P. Bierman (1999). Water table management, nitrogen dynamics, and yields of corn and soybean. *Soil Science Society of America Journal 63*(6), 1786–1795.

Florax, R., R. Voortman, and J. Brouwer (2002). Spatial dimensions of precision agriculture: a spatial econometric analysis of millet yield on Sahelian coversands. *Agricultural Economics 27*(3), 425–443.

Frankenberger, J., E. Kladivko, G. Sands, D. Jaynes, N. Fausey, M. Helmers, R. Cooke, J. Strock, K. Nelson, and L. Brown (2006). Questions and answers about drainage water management for the Midwest. *Purdue Extension WQ-44*.

Gilliam, J., J. Baker, and K. Reddy (1999). Water quality effects of drainage in humid regions, Chap. 24 in RW Skaggs and J. *Van Schilfgaarde (Eds.), Agricultural Drainage, Agronomy Monograph 38*, 801–830.

Griffin, T., R. Florax, and J. Lowenberg-DeBoer (2005). Yield Monitors and Remote Sensing Data: Sample Statistics or Population? *Site Specific Management Center, Purdue University*.

Griffin, T., D. Lambert, and J. Lowenberg-DeBoer (2005). Testing for appropriate on-farm trial designs and statistical methods for precision farming: A simulation approach. In *Proceedings of the 7th International Conference on Precision Agriculture and Other Precision Resources Management, ASA/SSSA/CSSA, Madison, Wisconsin.*

Grigg, B., L. Southwick, J. Fouss, T. Kornecki, and A. USDA (2003). Drainage system impacts on surface runoff, nitrate loss, and crop yield on a southern alluvial soil.

Haining, R. (2003). *Spatial data analysis: theory and practice.* Cambridge University Press.

Heady, E. and J. Dillon (1961). Agricultural production functions.

Hsiao, C. (2003). *Analysis of panel data.* Cambridge Univ Pr.

Kapoor, M., H. Kelejian, and I. Prucha (2007). Panel data models with spatially correlated error components. *Journal of Econometrics 140*(1), 97–130.

Kaspar, T., T. Colvin, D. Jaynes, D. Karlen, D. James, D. Meek, D. Pulido, and H. Butler (2003). Relationship between six years of corn yields and terrain attributes. *Precision Agriculture 4*(1), 87–101.

Kelejian, H. and I. Prucha (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 509–533.

Kessler, M. and J. Lowenberg-DeBoer (1999). Regression Analysis of Yield Monitor Data and Its Use in Fine Tuning Crop Decisions. In *Precision Agriculture: Proceedings of the 4th International Conference on Precision Agriculture.*

Lambert, D., J. Lowenberg-DeBoer, and R. Bongiovanni (2003). Spatial regression, an alternative statistical analysis for landscape scale on-farm trials: case study of variable rate nitrogen application in Argentina. In *Proceedings of the 6th International Conference on Precision Agriculture and Other Precision Resources Management, Minneapolis, MN, USA, 14-17 July, 2002*, pp. 828–842. American Society of Agronomy Madison, USA.

Lambert, D., J. Lowenberg-Deboer, and R. Bongiovanni (2004). A comparison of four spatial regression models for yield monitor data: A case study from Argentina. *Precision agriculture 5*(6), 579–600.

Lambert, D., J. Lowenberg-DeBoer, and G. Malzer (2006). Economic analysis of spatial-temporal patterns in corn and soybean response to nitrogen and phosphorus. *Agronomy Journal 98*(1), 43–54.

Liu, Y., S. Swinton, and N. Miller (2006). Is site-specific yield response consistent over time? Does it pay? *American Journal of Agricultural Economics 88*(2), 471–483.

Long, D., S. DeGloria, D. Griffith, G. Carlson, G. Nielsen, P. Robert, R. Rust, and W. Larson (1993). Spatial Regression Analysis of Crop and Soil Variability within an Experimental Research Field. In *Proceedings of the 1st Workshop on Soil Specific Crop Management.*

Lowenberg-DeBoer, J., T. Griffin, and R. Florax (2006). Use of cross regression to model local spatial autocorrelation in precision agriculture. *Site specific management center, Department of agricultural economics, Purdue University.*

Malzer, G., J. Copeland, J. Davis, J. Lamb, and P. Robert (1996). Spatial variability of profitability in site-specific N management. In W. L. P. Robert, R. Rust (Ed.), *Proceedings of the Third international conference on precision agriculture.* Soil science society of America, Madison, WI.

Mamo, M., G. Malzer, D. Mulla, D. Huggins, and J. Strock (2003). Spatial and Temporal Variation in Economically Optimum Nitrogen Rate for Corn Contrib. from the Dep. of Soil, Water, and Climate, Univ. of Minnesota, and the Minnesota Agric. Exp. Stn. *Agronomy Journal 95*(4), 958–964.

Mejia, M., C. Madramootoo, and R. Broughton (2000). Influence of water table management on corn and soybean yields. *Agricultural Water Management 46*(1), 73–89.

Moulton, B. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics 32*(3), 385–397.

Moulton, B. (1987). Diagnostics for group effects in regression analysis. *Journal of Business & Economic Statistics*, 275–282.

Nielsen, D., O. Wendroth, P. JIZRSCHIK, G. KUHN, and J. Hopmans (1997). Precision agriculture: challenges and opportunities of instrumentationand field measurements.

Nielsen, D., O. Wendroth, and F. Pierce (1999). Emerging concepts for solving the enigma of precision farming research. In *PC Robert et al. Precision Agriculture. Proc. 4th International Conference*, pp. 303–318.

Nistor, A. (2007). *The impact of controlled drainage adoption: a spatial panel model using yield monitor data*. Ph. D. thesis, Purdue University.

Nistor, A. and J. Lowenberg-DeBoer (2007). Drainage water management impact on farm profitability. *Journal of soil and water conservation 62*(6), 443–446.

Pace, R., R. Barry, J. Clapp, and M. Rodriquez (1998). Spatiotemporal autoregressive models of neighborhood effects. *The Journal of Real Estate Finance and Economics 17*(1), 15–33.

Pitts, D. (2003). Illinois drainage water management demonstration project. In *presentation at the Indiana Drainage Water Management Meeting, West Lafayette, IN.*

Rablais, N., R. Turner, and D. Scavia (2002). Beyond science into policy: Gulf of Mexico hypoxia and the Mississippi River. *Bio Science 52*, 129–142.

Sipp, S., W. Lembke, C. Boast, J. Peverly, M. Thorne, and P. Walker (1986). Water management of corn and soybeans on a claypan soil. *Transactions of the ASAE 29*, 780–784.

Swanson, E. (1963). The static theory of the firm and three laws of plant growth. *Soil Science 95*(5), 338–43.

Tan, C., C. Drury, M. Soultani, I. Wesenbeeck, H. Ng, J. Gaynor, and T. Welacky (1998). Effect of controlled drainage and tillage on soil structure and tile drainage nitrate loss at the field scale. *Water Science and Technology (United Kingdom) 38*, 103–110.

Wendroth, O., P. Jürschik, A. Giebel, and D. Nielsen (1998). Spatial statistical analysis of on-site-crop yield and soil observations for site-specific management. In *International Conference on Precision Agriculture*. ASACSSASSSA.