The Australian Journal of
Agricultural and
Resource Economics

Journal of the Australian Agricultural and Resource Economics Society

The Australian Journal of Agricultural and Resource Economics, 56, pp. 521–541

Integrating spatial dependence into Stochastic Frontier Analysis

Francisco José Areal, Kelvin Balcombe and Richard Tiffin†

An approach to incorporate spatial dependence into stochastic frontier analysis is developed and applied to a sample of 215 dairy farms in England and Wales. A number of alternative specifications for the spatial weight matrix are used to analyse the effect of these on the estimation of spatial dependence. Estimation is conducted using a Bayesian approach and results indicate that spatial dependence is present when explaining technical inefficiency.

Key words: Bayesian, spatial dependence, spatial weight matrix, technical efficiency.

1. Introduction

Despite many economic phenomena being driven by spatial processes, spatial relationships have rarely been exploited in the economic literature before the late 1990s (Bockstael 1996; Anselin 2001). Disregarding spatial aspects of the data may produce inefficient or biased estimates and consequently, misleading inference (Anselin 2001). However, interest in spatial issues has increased recently. It was in the 1990s when there were the first calls for the introduction of spatial econometrics in agricultural economics (Bockstael 1996; Weiss 1996). Weiss (1996) stresses, as does Bockstael (1996), that economic processes such as agricultural production are spatial phenomena and factors such as yield, soil characteristics, landscape configurations and pest populations show spatial variability. Weiss (1996) calls for the use of spatial information in agricultural economics and points out that results obtained from incorporating spatial analysis into agricultural economics has implications for farm management and for agricultural and environmental policies. For instance, spatial information can reveal where fertiliser use is profitable and where it is counterproductive (Weiss 1996).

Two special issues have been devoted to the subject of spatial econometrics in agricultural economics journals in the recent years. Firstly, the special issue of Agricultural Economics (2002) and secondly the special issue of the Journal of Agricultural Economics (2007). Holloway and Lapar (2007) provide an excellent review of recent literature in which spatial econometrics techniques have been used. The authors focussed their review on studies dealing with spatial bio-economic modelling and land use modelling. They classified the

†Francisco José Areal (email: f.j.areal@reading.ac.uk), Kelvin Balcombe and Richard Tiffin are at Agricultural and Food Economics, University of Reading, Reading, UK.

doi: 10.1111/j.1467-8489.2012.00597.x

articles into two groups: those that explicitly use spatial econometric methods and those that use geographic information systems (GIS) techniques.

As for agricultural production, there are a number of potential sources of spatial dependence in efficiency, including soil quality, climatic conditions, socio-economic aspects and other location-specific attributes. For instance, spatial dependence in technical efficiency can be found because farmers in an area may emulate each other; it may be due to the level of infrastructure in the area or, it may be because of the climatic and topographic conditions of the area where the farm is located. All these are usually unobservable variables that may be spatially correlated. In other words, no matter what spatially correlated covariates are included in the production function, it is always possible that some will be omitted. This is also true for models that account for latent technical inefficiency. Although these models may include a set of covariates, more or less spatially correlated, that may help to explain the level of technical inefficiency, it is almost certain that some spatial effects will be omitted.

A number of models have been developed to account for spatial dependence such as the spatial autoregression (SAR) model (Whittle 1954; Mead 1967; Besag 1974; Ord 1975; Anselin 1988), the spatial error model (SEM) and its variant the higher order contiguity model or spatial Durbin model that allows for explanatory variables from neighbouring observations (LeSage 1999; Bell and Bockstael 2000). None of these models incorporate spatial dependence into technical efficiency analysis. Despite advances in the econometric application of spatial analysis (e.g. Anselin 1988; LeSage 1999) very little research can be found in the literature on how to incorporate spatial dependence into technical efficiency analysis (Druska and Torrace 2004; Schmidt *et al.* 2009). The efficiency literature usually considers spatial heterogeneity, where efficiency levels may differ depending on the location. Spatial dependence refers to the correlation between the efficiency levels of the farms and the efficiency levels of 'neighbouring farms'. Spatial heterogeneity in the technical efficiency literature is controlled (if controlled at all) by introducing dummy variables for political land divisions such as regions, counties and provinces. For example, Hadley (2006) introduced dummy variables to account for regional heterogeneity. The introduction of dummy variables has also been used to account for spatial heterogeneity in less favoured areas (Iraizoz *et al.* 2005; Hadley 2006).

We incorporate spatial dependence into technical efficiency analysis using an autoregressive specification of the inefficiency component of a compound error term. Our approach differs from Druska and Torrace (2004), which used an autoregressive specification in the error term to estimate the spatial dependence based on a standard fixed effects model. The work here also differ from Schmidt *et al.* (2009) which makes farm inefficiency dependent on a parameter that captures the unobserved spatial characteristics. Our work differs in both the specification of the model and the scope of the analysis. We directly integrate the unobserved spatial characteristics in the stochastic frontier model by specifying the inefficiency to be spatially autoregressive and including a parameter that measures the level of spatial dependence. Schmidt *et al.* (2009) examined the unobserved local characteristics in each municipality by incorporating them into the analysis assuming that (i) they follow a conditional autoregressive (CAR) prior distribution (i.e. they incorporate the assumption that neighbour municipalities have a similar level of unobserved local characteristics) or (ii) a Normal distribution (i.e. unobserved local characteristics are independent of the neighbours). Schmidt *et al.* (2009) examined unobserved spatial effects at relatively small levels (i.e. municipalities), whereas we take different specifications, some of which not restricted by political boundaries. By examining different spatial structures, we are able to discern how spatial dependence varies with different characterisations of neighbourhoods, which is an aspect that Schmidt *et al.* (2009) conclude is worth being investigated. Our approach enables us to obtain both the degree and significance of spatial dependence in the whole area studied for different characterisations of neighbouring farms. Schmidt *et al.* (2009) only provide information on the significance of spatial dependence at the municipality level.

The following sections are dedicated to the description of the data, – methodology and empirical approaches used for integrating spatial dependence into stochastic frontier analysis. The empirical section includes a description of the data used and along with the results the final section concludes.

2. Data

The analysis herein uses balanced panel data from the Farm Business Survey (FBS) for the years 2000–2005. A total of 215 dairy farms in England and Wales are included in the data set. The FBS includes a large amount of information related to farm enterprises. We classified farm output data into the following: (i) milk and other dairy products, (ii) leasing out quota and (iii) other products. Laspeyres and Paasche quantity indices were calculated to compute a Fisher's quantity index that aggregated the output in milk and other milk output into one variable and other products also into one variable. The base for price and output indices was calculated as the average of prices and outputs. With regard to inputs, we use utilised agricultural area (UAA) in ha; herd size (number of cows); labour (£); machinery and general farming costs (£), which includes contract work, machinery rental, machinery and equipment valuation, machinery and equipment repairs, vehicle fuel and oil, electricity, heating fuel for all purposes, water for all purposes, insurance excluding labour and farm buildings, bank charges professional fees, vehicle tax and other general farming costs; and livestock costs (concentrate feedstuff, coarse fodder, veterinary services and medicines).

Spatial information on farms was provided by the Department for Environment Food and Rural Affairs (Defra) as part of the FBS. The FBS provides information on the geographical location of the farm at a 10 km grid square level. Figure 1 shows the location of the FBS farms. A 10 square km block contains at least one farm in the data set and a maximum of four. This information was used to build a number of connectivity (or spatial weight) matrices that represent the relative spatial information.

3. Methods

3.1. The spatial weight matrix

Although the use of political land divisions in efficiency analysis may capture some effects associated with policies at regional, county or provincial levels, there may be factors such as climatic and topographic conditions which differ within those political divisions. To account for factors that may be present on a smaller or larger scale relative to a political division of the land, a quantification of the structure of spatial dependence needs to be introduced.

The spatial information about farms can be incorporated into a connectivity or spatial weight matrix (W). A connectivity matrix can be defined in different ways depending on the appropriate definition of a neighbourhood. Two questions usually arise when analysing technical efficiency: the structure of spatial dependence and the appropriate metric used to decide how close two farms are. Both issues are problematic in microdata environments where

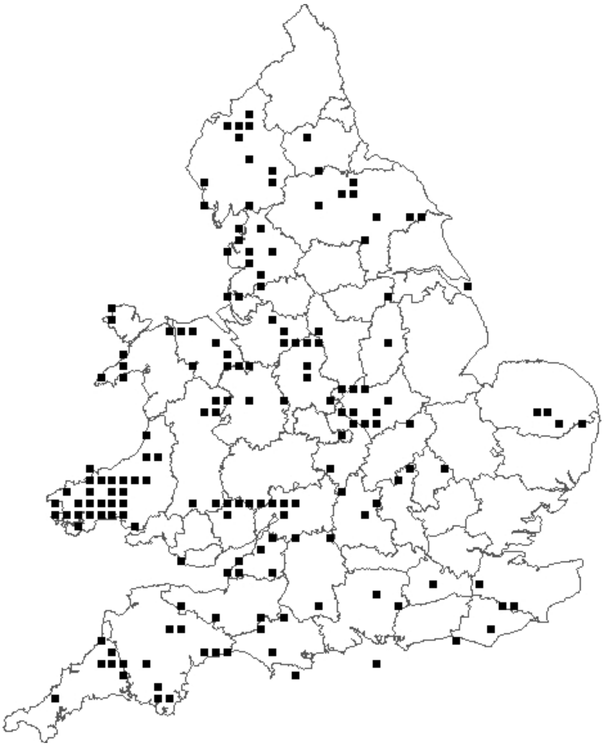

Figure 1 Farms location in 10 sq km blocks.

observations are scattered throughout a landscape (Bell and Dalton 2007; Holloway and Lapar 2007). There are two main ways in which the spatial weight matrix W can be specified.

The spatial weight matrix W is a symmetric matrix with the elements W_{ij} representing the distance or closeness of farm i with farm j.

W is usually row standardised such that the rows add up to one. This facilitates the interpretation of model coefficients. Let's consider four farms where farm 1 is close to farm 2; farm 2 is close to farms 1, 3 and 4; farm 3 is close to farms 2 and 4, while farm 4 is close to farms 2 and 3. The spatial weight matrix based on this spatial example could take the unstandardised form

$$W^U = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}. \tag{1}$$

The diagonal elements W^U_{ii} are set to 0 in order to preclude an observation of the efficiency for the i farm from directly predicting itself. The spatial weight matrix is row standardised so each element in the standardised matrix W, $W_{ij} = \frac{W^U_{ij}}{\sum_j W^U_{ij}}$, is between 0 and 1 as shown below

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0.33 & 0 & 0.33 & 0.33 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0.5 & 0 \end{pmatrix}. \tag{2}$$

Close proximity can have different interpretations. It can mean that two farms can be adjacent neighbours or neighbours within a given distance. For the latter, the elements of the W are given by: $W^U_{ij} = 1$ if $0 <$ distance between $i, j \leq s$ (s is the distance beyond which no dependence is assumed); otherwise $W^U_{ij} = 0$.

After standardising, we estimate the model $z = \rho Wz + \varepsilon$ where z is the vector of inefficiencies. The parameter ρ determines the correlation between the elements of z. If z referred to farms efficiency, then ρ would represent the correlation between individual farm efficiency and the mean efficiency of the neighbouring farms.

An alternative approach to the one shown earlier is the use of a spatial weight matrix based on distance (Anselin 2002). In this case, neighbours have different weights. Those with higher weights are closer in distance. Therefore, ρ determines the correlation between farm efficiency and adjusted by distance mean efficiency of neighbouring farms. This approach is also arbitrary in the sense that the cut-off distance is selected by the researcher. The distance weight matrix specified here is one of a power form

$$W_{ij} = \exp\left(-d_{ij}^2 / s^2\right), \quad (3)$$

where d_{ij} is the distance between a farm in location i and a farm in location j; s is the distance around a given observation over which other observations are likely to be dependent.

The cut-off distance chosen to determine the distance beyond which spatial effects are not relevant is a key issue. Bell and Bockstael (2000) found that their results were more sensitive to the specification of the neighbourhood in the spatial weight matrix (i.e. choice of the cut-off distance) than to the estimation technique used. They found that the spatial dependence estimate changed with the cut-off distance selected, increasing first at a small cut-off distance and falling afterwards as the cut-off distance was increased. They applied a higher order contiguity model and showed how spatial dependence diminishes with distance.

Roe *et al.* (2002) also estimated their models using different cut-off distances and highlighted that the appropriate cut-off distance is an empirical issue. Kim *et al.* (2003) used SAR and SEM hedonic price models to measure the benefits of air quality improvement. The spatial weight matrix was specified based on distances between district centroids with a cut-off distance of 4 km chosen after experimenting with a series of different cut-off distances. All these articles show that a cut-off distance exists where spatial dependence reaches a maximum.

3.2. Scope

Milk producer farm (in)efficiencies in England and Wales are studied in this study. Milk producers have an annual milk quota that partially binds production because producers can lease in and/or lease out milk during the production year. Therefore, we include, in the analysis, the fact that production is partially constrained by the annual quota Q that includes the initial quota $\pm$ quota bought/sold, leasing in quota *qui* and leasing out quota *quo*. Not accounting for such constraints may lead to wrongly attributing the effects of such constraints to the farmer being unsuccessful in optimising production (Färe *et al.* 1994).

If producers optimise their production by not wasting resources, this will lead them to operate near the edge of their production possibility set. However, there may be an array of motives that explain why not all producers are successful in optimising production. In this study, we focus on developing a way to explain technical inefficiency through spatial dependency. The departure point of any technical efficiency analysis is the definition of the production technology of a firm. This can be characterised in terms of a technology set, the output set of production technology and the production frontier.

3.3. Output distance function

We use a distance function approach because it describes technology in a way that allows efficiency to be measured for multi-input and multi-output enterprises (Coelli *et al.* 2005). An output distance function describes the degree to which a firm can expand its output given its input vector. We start from a producible output set, which is the set of all outputs that can be feasibly produced using the set of all inputs. The output set for production technology is defined as

$$\begin{aligned} P(x,Q) &= \left\{y \in R_{+}^{M}: x \text{ can produce } y \text{ given } y_1 = Q + qui - quo\right\} \\ &= \{y : (x,y) \in T\}, \end{aligned} \tag{4}$$

where y refers to all M outputs of the farm including milk (y_1), the leasing out of quota (quo) and other outputs, which take only positive real numbers R_{+}^{M}, and x refers to all K inputs used in the farm, which take only positive real numbers R_{+}^{K}, including the leasing in quota (qui) and the annual allocation of quota Q which includes the initial quota $\pm$ the amount of quota bought/sold in the current year. The output set is included within the technological set T.

The output distance function is defined on the output set $P(x,Q)$ as

$$D_O(x,y,Q) = \min\left\{\theta : \left(\frac{y}{\theta}\right) \in P(x,Q)\right\} \text{ for all } x \in R_{+}^{K}, \tag{5}$$

which means that the initial allocation of quota Q, the leasing in qui and leasing out quota quo are treated in the same way as conventional inputs (x) and outputs (y).

Assuming a translog functional form for the parametric distance function with M outputs and K inputs offers several attractive properties including flexibility, as well as making it easy to derive and permit the imposition of homogeneity, which makes it the preferred form in the literature (Lovell *et al.* 1994; Coelli and Perelman 1999; Brümmer *et al.* 2002, 2006).

$$\begin{aligned} \ln D_{Oi} = \alpha_0 + \sum_{m=1}^{M} \alpha_m \ln y_{mi} + \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M} \alpha_{mn} \ln y_{mi} \ln y_{ni} + \sum_{k=1}^{K} \beta_k \ln x_{ki} + \frac{1}{2} \\ \sum_{k=1}^{K}\sum_{l=1}^{K} \beta_{kl} \ln x_{ki} \ln x_{li} + \sum_{k=1}^{K}\sum_{m=1}^{M} \delta_{km} \ln x_{ki} \ln y_{mi}; i = 1,\ldots,n, \end{aligned} \tag{6}$$

where i denotes the ith farm in the sample; qui and Q are included in x as inputs; and quo is part of y as an output. Using linear homogeneity of the output distance function in outputs, Equation (6) can be transformed into an estimable regression model by normalising the function by one of the outputs (Lovell *et al.* 1994; Coelli and Perelman 1999; Brümmer *et al.* 2002, 2006; Orea 2002; O′Donnell and Coelli 2005). From Euler's theorem, homogeneity of degree one in output implies:

$$\sum_{m=1}^{M}\alpha_m + \sum_{m=1}^{M}\sum_{n=1}^{M}\alpha_{mn}\ln y_{ni} + \sum_{m=1}^{M}\sum_{k=1}^{K}\delta_{km}\ln x_{ki} = 1, \tag{7}$$

which will be satisfied if $\sum_{m=1}^{M}\alpha_m = 1$, $\sum_{m=1}^{M}\alpha_{mn} = 0$ for all n, and $\sum_{m=1}^{M}\delta_{km} = 0$ for all k. Substituting these constraints is equivalent to normalising by one of the outputs, which leads to the following expressions:

$$\ln D_O\left(\frac{y_i}{y_{2i}}, x\right) = \ln D_O\frac{1}{y_{2i}}(y_i, x) \tag{8}$$

and

$$-\ln y_2 = \alpha_0 + \sum_{m=1}^{M}\alpha_1\ln\frac{y_{mi}}{y_{2i}} + \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M}\alpha_{mn}\ln\frac{y_{mi}}{y_{2i}}\ln\frac{y_{ni}}{y_{2i}} + \sum_{k=1}^{K}\beta_k\ln x_{ki} + \frac{1}{2}$$
$$\sum_{k=1}^{K}\sum_{l=1}^{K}\beta_{kl}\ln x_{ki}\ln x_{li} + \sum_{k=1}^{K}\sum_{m=1}^{M}\delta_{km}\ln x_{ki}\ln\frac{y_{mi}}{y_{2i}} + \varepsilon_i + z_i, \tag{9}$$

where ε_i is a symmetric random error term that accounts for statistical noise and z_i is a non-negative random variable associated with technical inefficiency.

Monotonicity constraints involve constraints on functions of the partial derivatives of the distance function. As pointed out by O′Donnell and Coelli (2005), the elasticities of distance with respect to inputs and outputs are important derivatives.

$$\frac{\partial\ln D_O}{\partial\ln x_k} = \beta_k + \sum_{k=1}^{K}\beta_{kl}\ln x_{li} + \sum_{m=1}^{M}\delta_{km}\ln\frac{y_{mi}}{y_{2i}} \tag{10}$$

$$\frac{\partial\ln D_O}{\partial\ln y_m} = \alpha_m + \sum_{m=1}^{M}\alpha_{mn}\ln\frac{y_{ni}}{y_{2i}} + \sum_{k=1}^{K}\delta_{km}\ln x_{ki}. \tag{11}$$

For D_O to be nonincreasing in x, $\frac{\partial\ln D_O}{\partial\ln x_k} \leq 0$ while for D_O to be nondecreasing in y. The data were normalised so that each variable had a sample mean of one. This means that the monotonicity conditions can be expressed as $\alpha_m \geq 0$ and $\beta_k \leq 0$. It is worth noting that coefficient results have changed the sign, and therefore, the expected coefficients should be $\alpha_m \leq 0$ and $\beta_k \geq 0$.

We used the spatial information contained in the data set to create a number of specifications for the spatial weight matrix W that were used to investigate the effect of these on the results. One specification involved the introduction of a spatial connectivity matrix whose common specification was a $n \times n$ matrix W with elements $W_{ij} = 1$ for farms $j = 1, ..., n$ within a 10 square km grid to farm i and $W_{ij} = 0$ for those farms that were not close. Once W was row standardised, this effectively accounted for the average efficiency of the farms surrounding the farm within the 10 km square grid. Another alternative specification was where the spatial connectivity matrix W has elements $W_{ij} = 1$ for farms $j = 1, ..., n$ within the Government Office Regions (GOR) of farm i and $W_{ij} = 0$ for those farms that were located in the same GOR. Finally, four more alternatives were used by specifying a spatial distance matrix W with elements $W_{ij} = d_{ij}$ where d_{ij} is the Euclidean distance. The weight specification used was the power form (Eqn 3) and four cut-off distances were used ($s = 20$ km; $s = 100$ km; $s = 180$ km and $s = 240$ km). As pointed out earlier, defining how to quantify close proximity is arbitrary, that is, cut-off distance is decided by the investigator. The distance between farms is calculated using the Euclidean distance

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \tag{12}$$

where x_i, y_i are the coordinates of the points.

3.4. Estimation

A Bayesian procedure was used for estimation. We describe this procedure in the following section. We start with the standard stochastic output distance function model which is specified as

$$y_{it} = x_{it}\beta + \varepsilon_{it} - z_i, \tag{13}$$

wherey_{it} is a vector of the logarithm of milk and other milk products for each farm i in year t; x_{it} is a matrix of the logarithm of other outputs and inputs of the farm i in year t; β is a vector of parameters associated with the outputs and inputs of the farm to be estimated; ε_{it} is the random error and z_i represents the inefficiency of the ith farm. Stacking all the variables into matrices, we obtain

$$y = x\beta + \varepsilon - (z \otimes 1_T). \tag{14}$$

This standard model can be transformed to account for spatial dependence in the inefficiency term. The spatially dependent inefficiency term is

$$z = \rho Wz + \tilde{z}, \tag{15}$$

where W is a connectivity matrix that; is the spatial coefficient; and $\tilde{z}$ are latent variables whose distributional form is unknown. By plugging Equation (15) into Equation (14), we obtain the following expression

$$y = x\beta + \varepsilon - \left((I - \rho W)^{-1}\tilde{z}\right) \otimes 1_T. \tag{16}$$

The parameter ρ is assumed to be between 0 and 1.[1]

3.4.1. The conditional likelihood function.

The distributional assumptions determine the form of the likelihood function. Here, it is assumed that the prior distributions for the latent errors are normal and gamma distributed (Koop *et al.* 1995; Koop 2003). In this case, normality is assumed. Note that we have $i = 1, \ldots, n$ farms observed during T years ($t = 1, \ldots, T$). Here, $p()$ refers to the density and $p(|)$ is the conditional density.

$$p(y|\beta, h, \rho, \mu_z^{-1}, \tilde{z}) = \prod_{i=1}^{N} \frac{h^{\frac{T}{2}}}{(2\pi)^{\frac{T}{2}}} \exp\left(-h\frac{\varepsilon'\varepsilon}{2}\right) \tag{17}$$

noting that $p\left(y|\beta, h, \rho, \mu_z^{-1}, \tilde{z}\right) = p\left(y|\beta, h, \rho, \mu_z^{-1}, z\right) = p\left(y|\beta, h, \mu_z^{-1}, \tilde{z}\right)$.

Defining $\tilde{y} = \left[y + (I - \rho W)^{-1}\tilde{z} \otimes 1_T\right]$ the following expression is obtained

$$p\left(y|\beta, h, \mu_z^{-1}, \tilde{z}\right) \propto h^{\frac{TN}{2}} \exp\left[-\frac{h}{2}(\tilde{y} - x\beta)'(\tilde{y} - x\beta)\right]. \tag{18}$$

The expression above is of a standard form used for efficiency analysis (Koop *et al.* 1995; Koop 2003) with the spatial element being the extension of the model.

3.4.2. The priors

The likelihood function must be complemented with a prior distribution on the parameters $\left(\beta, h, \mu_z^{-1}, \rho\right)$ to conduct Bayesian inference. An independent Normal-Gamma prior is used for the coefficients in the production frontier and the error precision.

The distribution of the inefficiency term is determined by the distribution of z, which is a latent variable. We define $p\left(\tilde{z}|\mu_z^{-1}\right)$ instead of $p\left(z|\mu_z^{-1}\right)$, which is defined given ρ, W and $p\left(\tilde{z}|\mu_z^{-1}\right)$. The conditional distribution for the latent variable $\tilde{z}_i$ is

[1] We broke this assumption to evaluate the robustness of our results. They were found to be robust.

$$p\left(\tilde{z}_i|\mu_z^{-1}\right) = f_G\left(\tilde{z}_i|\alpha, \mu_z^{-1}\right) = \frac{\tilde{z}_i^{\alpha-1}}{\mu_z^j \Gamma(\alpha)} \exp\left(-\mu_z^{-1}\tilde{z}_i\right), \tag{19}$$

where $\Gamma(\cdot)$ is the gamma function; and $f_G\left(\tilde{z}_i|\alpha, \mu_z^{-1}\right)$ indicates the Gamma density with parameters α and μ_z^{-1}. This prior is commonly used in literature (van Den Broeck *et al.* 1994; Koop *et al.* 1995; Fernández *et al.* 2000). Assuming $\alpha = 1$, the inefficiency distribution is exponential and the inefficiency prior becomes

$$p\left(\tilde{z}_i|\mu_z^{-1}\right) \propto \exp\left(-\mu_z^{-1}\tilde{z}_i\right). \tag{20}$$

The prior for μ_z^{-1} is assumed to be gamma with parameters 2 and $-\ln(r^*)$ where r^* is the median of the prior distribution. Finally, the prior for ρ is assumed to be an indicator function.

$$f(\rho) = I(\rho \in [0, 1]). \tag{21}$$

The expression above is a uniform distribution and its applicability depends on the appropriate construction of the weight matrix. The indicator function $I(\cdot) = 1$ if $\rho \in [0, 1]$ or otherwise $I(\cdot) = 0$. This means that the parameter ρ that accounts for spatial dependence is expected to have a positive impact on the efficiency scores.

3.4.3. The joint posterior.

The joint posterior distribution can be broken down into as the multiplication of the conditional likelihood function and the priors. The joint posterior in terms of z is

$$p\left(\beta, h, \mu_z^{-1}, z, \rho|y\right) \propto p\left(y|\beta, h, \mu_z^{-1}, \tilde{z}\right) \times p(\beta) \times p(h) \times p\left(\tilde{z}|\mu_z^{-1}\right) \times p\left(\mu_z^{-1}\right) \times I(\rho \in [0, 1]). \tag{22}$$

3.4.4. The conditional posteriors.

The Gibbs sampler is based on conditional distributions that describe the probabilities of a combination of values for parameters of interest which are conditional on the observables. The use of conditional distributions facilitates obtaining posterior distributions of the parameters of interest. To estimate the model, it is useful to have the conditional distributions in order to employ the Gibbs sampling method (Geman and Geman 1984; Casella and George 1992). The conditional posterior for β is a Normal distribution after extracting the kernel for β from expression Equation (22).

$$p\left(\beta|h,\mu_z^{-1},\tilde{z},\rho,y\right) \sim N\left(b,\overline{V}\right). \tag{23}$$

As in Koop (2003), the conditional posterior density for h is

$$p\left(h|\beta,\mu_z^{-1},\tilde{z},\rho,y\right) \sim G\left(\overline{s}^{-2},\overline{v}\right). \tag{24}$$

To obtain the conditional posterior for μ_z^{-1}, it is more useful to use $\tilde{z}$ rather than z. The joint conditional posterior density for μ_z^{-1} and $\tilde{z}$ is the kernel from expression Equation (22) that involves μ_z^{-1} and $\tilde{z}$.

$$p\left(\tilde{z},\mu_z^{-1}|\beta,h,y,\rho\right) \propto \prod_{i=1}^{N} \exp\left(-h\frac{\varepsilon_i'\varepsilon_i}{2}\right) \times p\left(\tilde{z}|\mu_z^{-1}\right) \times p\left(\mu_z^{-1}\right) \tag{25}$$

from which the conditional posterior for μ_z^{-1} is

$$p\left(\mu_z^{-1}|\tilde{z},\beta,h,y,\rho\right) \sim G(m,\eta) \tag{26}$$

which is a Gamma distribution with parameters $m = \frac{N+1}{\sum_{i=1}^{N}\tilde{z}_i - \ln(r^*)}$ and $\eta = 2N+2$.

Recalling that z and $\tilde{z}$ are related as in expression Equation (15) the conditional posterior distribution for $\tilde{z}_i$ is

$$p(\tilde{z}_i|\beta,h,\mu_z,y,\rho) \propto \exp\left[-\frac{hT}{2}\left[z_i - \left(\bar{x}_i\beta - \bar{y}_i + \frac{\mu_z^{-1}}{Th}\right) + (\tilde{z}_i - z_i)\mu_z^{-1}\right]\right] \tag{27}$$

where $\overline{x}_i = \sum_{t=1}^{T}\frac{x_{i,t}}{T}$ and $\overline{y}_i = \sum_{t=1}^{T}\frac{y_{i,t}}{T}$.

Equation (27) is not of a recognisable distributional form. Therefore, a posterior simulator (i.e. a random number generator) needs to be used, such as a Metropolis–Hastings algorithm (Metropolis *et al.* 1953; Hastings 1970). We use a random walk algorithm proposal whereby a new set of $\tilde{z}_i$ are proposed using a Metropolis based on the posterior above. Given a new draw of $\tilde{z}_i$, then the entire z needs to be updated in each iteration using expression Equation (15).

To obtain the conditional posterior of ρ, the spatial problem can be represented in matrix form as

$$\left(y + \left((I-\rho W)^{-1}\tilde{z}\right) \otimes 1_T\right) - X\beta = \varepsilon. \tag{28}$$

It follows that the conditional posterior for ρ is

$$
\begin{aligned}
p\left(\rho|\beta, h, \mu_z^{-1}, y, \tilde{z}\right) &\propto \exp\left(-h\frac{\varepsilon'\varepsilon}{2}\right) \times p(\rho) \\
&= \exp\left(-h\frac{\varepsilon'\varepsilon}{2}\right) \times I(\rho \in (0,1)),
\end{aligned}
\tag{29}
$$

which provides the basis for the use of a second Metropolis–Hastings step. A random walk Metropolis–Hastings algorithm is used to draw ρ with probability of acceptance of the proposed ρ^* being

$$
P = \min\left(1, \frac{p\left(\rho^*|y, \beta, h, \mu_z^{-1}, \tilde{z}\right)}{p\left(\rho^{old}|y, \beta, h, \mu_z^{-1}, \tilde{z}\right)}\right). \tag{30}
$$

Recall that expressions Equations (29) and (30) are for the case that $\rho > 0$. In the case of $\rho = 0$ (i.e. there is no spatial component) note that $z = \tilde{z}$.

Classical approaches could be used to estimate our model. For example, it is possible that EM algorithms in the Max Likelihood context (or perhaps even GMM) could be used. However, the Markov Chain Monte Carlo (MCMC) approach to estimation means that we do not have to find explicit expressions for the likelihood function after the latent variables have been integrated out. The principle advantage of the MCMC approach is that the mapping of the latent distributions is an integral part of the estimation procedure.

4. Results

We expect the nature of the connectivity matrix will determine the results, and for this reason, we wish to explore alternative specifications for the weight matrix. We would expect ρ to increase with the cut-off distance for the spatial effects up to a distance and then decrease. We would expect the spatial dependence to be lower for small neighbourhoods because such areas may not include the whole area that has a spatial incidence on efficiency. In addition, we would expect that once we reach a given cut-off distance, the spatial effect should decrease indicating that the spatial dependence has a limit. Three spatial models for inefficiency were estimated, one where the weight or connectivity matrix is specified regarding neighbours as farms within a 10 km square grid (SM1); one where neighbours are those farms in the same GOR (SM2); and another where the connectivity matrix is specified as a distance matrix (SM3). The SM3 was estimated using four cut-off distances, 20, 100, 180 and 240 km (SM3-20; SM3-100; SM3-180 and SM3-240).

Results for the parameters associated with inputs and outputs of the production function are shown in Table 1 for models SM1 and SM2; Table 2 for models SM3-20, SM3-100, SM3-180 and SM3-240. All signs are as expected with the exception of the coefficient for the leasing quota in, which is negative,

Table 1 Slope parameters for models SM1 and SM2

	SM1		SM2	
	Coefficients	90% posterior	Coefficients	90% posterior
Intercept	−0.03	−0.10, 0.04	−0.04	−0.06, 0.18
Leasing quota out	−0.12	−0.18, −0.05	−0.11	−0.18, −0.05
Other output	−0.28	−0.33, −0.24	−0.29	−0.34, −0.25
UAA	0.05	0.00, 0.10	0.05	0.00, 0.12
Milk quota	0.39	0.28, 0.50	0.35	0.24, 0.46
Number of cows	0.42	0.31, 0.54	0.46	0.34, 0.58
Leasing quota in	−0.02	−0.06, 0.02	−0.02	−0.06, 0.02
Machinery & General costs	0.10	0.03, 0.18	0.10	0.02, 0.17
Labour costs	0.03	−0.03, 0.09	0.04	−0.03, 0.10
Livestock costs (per cow)	0.16	0.10, 0.22	0.18	0.12, 0.25

but the 90% coverage posterior region shows that there is no clear evidence that supports the belief that this coefficient is negative. The number of cows and milk quota allocated at the beginning of the year are the two most important inputs in terms of milk production, whereas the production of other outputs by the farm reduces the production of milk, holding everything else constant.

Figures 2 and 3 show the kernel distributions for farm efficiency of the posterior means of technical efficiency evaluated over all farms for models SM1, SM2 and SM3. Results suggest that the way in which the connectivity matrix is defined has an impact on the levels of efficiency obtained. The efficiency average is 0.86 when neighbours are considered to be those within a 10 km grid square and 0.78 when neighbours are considered to be those farms in the same region. Figure 3 shows smaller differences between the alternatives. The mean efficiencies are 0.84, 0.81, 0.80 and 0.80 for SM3-20, SM3-100, SM3-180 and SM3-240, respectively. Figure 4 shows the expected technical efficiency levels obtain from the different characterisations of W against the map of England and Wales at a 10 km grid square level. In those grid squares where more than one farm is located the result shown is the mean of the efficiency within that grid square. Results for the mean efficiencies were slightly higher for the SM1 model that is reflected in the map for SM1 which has darker squares than the other maps. We also calculated the percentage difference in absolute terms between the posterior means of TE for the case where spatial dependence is not taken into account and for the different models taking into account spatial dependence. Incorporating spatial dependence to the analysis leads to variations of TE estimates of the order of 6% with respect to farm TE estimates using a model where spatial dependence is not incorporated.

Regarding the results for the conditional posterior distribution for the spatial dependence parameter ρ, these are shown in Figures 5 and 6. Results for SM1 and SM2 are shown in Figure 5, whereas the four alternatives of MS2

Table 2 Slope parameters for models SM3-20, SM3-100, SM3-180 and SM3-240

	SM3-20		SM3-100		SM3-180		SM3-240	
	Coefficients	90% posterior	Coefficients	90% posterior	Coefficients	90% posterior	Coefficients	90% posterior
Intercept	0.02	−0.06, 0.11	0.17	−0.03, 0.51	0.21	−0.02, 0.57	0.19	−0.02, 0.52
Leasing quota out	−0.11	−0.18, −0.05	−0.11	−0.17, −0.05	−0.11	−0.17, −0.04	−0.11	−0.17, −0.04
Other output	−0.29	−0.34, −0.24	−0.29	−0.34, −0.25	−0.29	−0.34, −0.25	−0.29	−0.34, −0.25
UAA	0.06	0.00, 0.12	0.06	0.00, 0.12	0.06	0.00, 0.13	0.06	0.00, 0.12
Milk quota	0.35	0.24, 0.46	0.31	0.17, 0.43	0.29	0.15, 0.42	0.30	0.16, 0.43
Number of cows	0.45	0.34, 0.57	0.50	0.36, 0.65	0.51	0.37, 0.66	0.51	0.37, 0.66
Leasing quota in	−0.02	−0.05, 0.02	−0.01	−0.05, 0.02	−0.01	−0.05, 0.03	−0.01	−0.05, 0.02
Machinery & General costs	0.10	0.02, 0.18	0.09	0.00, 0.16	0.09	0.00, 0.17	0.08	0.00, 0.16
Labour costs	0.03	−0.04, 0.10	0.04	−0.03, 0.12	0.04	−0.03, 0.12	0.05	−0.03, 0.12
Livestock costs (per cow)	0.18	0.12, 0.25	0.20	0.13, 0.29	0.21	0.13, 0.29	0.20	0.13, 0.28

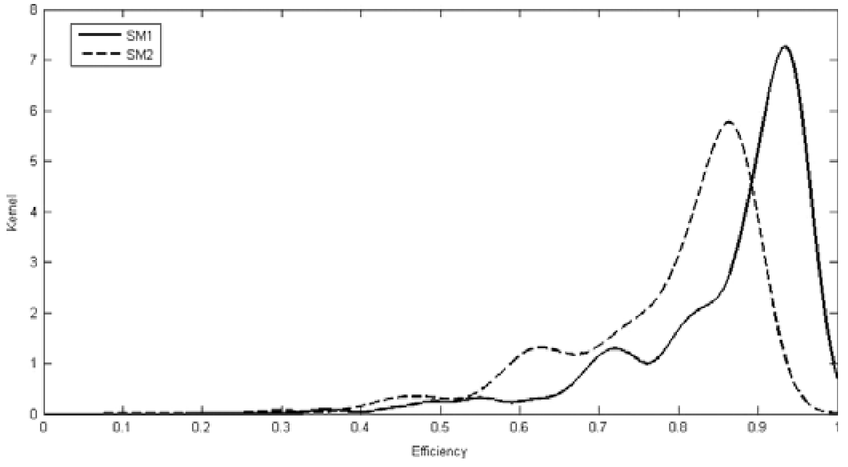

Figure 2 Kernel distributions of the posterior means of technical efficiency across all farms for SM1 and SM2.

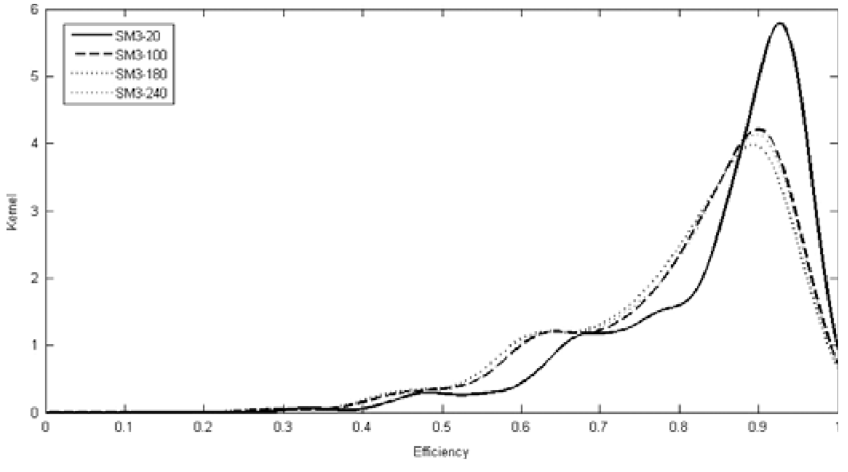

Figure 3 Kernel distributions of the posterior means of technical efficiency across all farms for SM3 models.

are shown in Figure 6. Spatial models SM1 and SM2 show similar results for the spatial parameter ρ with averages of 0.13 and 0.18, respectively, which suggests that efficiency is farm determined rather than spatially determined. The parameter ρ is 59.2% more likely to be higher using SM2 than SM1 which suggests that the spatial dependence is larger than just a 10 km square grid.

Models SM3 were run to investigate the effect of the cut-off distance chosen on the correlation ρ between efficiency and the adjusted by distance mean efficiency. Results show that the spatial dependence parameter ρ increases with the cut-off distance up to a point between 100 and 240 km and then decreases.

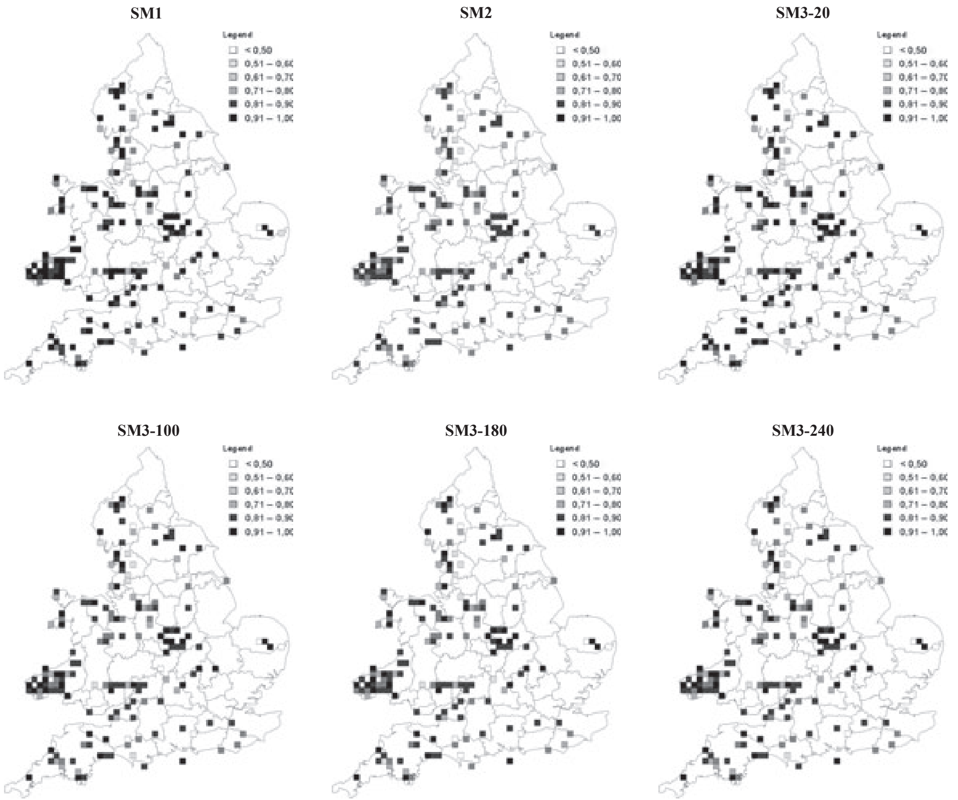

Figure 4 Efficiency estimates in England and Wales at 10 km grid square level.

The probability that ρ using SM3-180 is higher than using SM3-240 is 53%, whereas the probability that ρ using SM3-180 is higher than SM3-100 is 59%. These results indicate that the spatial parameter ρ may increase with the cut-off distance but will decrease once the cut-off reaches a distance between 100 and 240 km. These results are similar to those obtained by Bell and Bockstael (2000) where the spatial estimate increases and then falls. A reason for this is that a spatial matrix W which has been specified using a relatively small cut-off distance may not contain enough observations that help to obtain a good estimate of the mean efficiency in the neighbourhood. The spatial estimate will start to fall once farms that are not related in terms of efficiency with the farm of interest start to be included in the spatial distance weight matrix W. This will occur at a given distance. With regard to the mean of ρ- this is 0.14, 0.31, 0.35 and 0.34 for the 20, 100 and 180 and 240 km alternative models, respectively.

The relatively low values for ρ obtained by all models (i.e. efficiency depends mainly of farm management) can be seen in Figure 4. Apart from some clusters in the East Midlands region and North Yorkshire, no other clear large clusters can be seen on the maps. Differences were found between estimates of the spatial parameter ρ obtained by models where

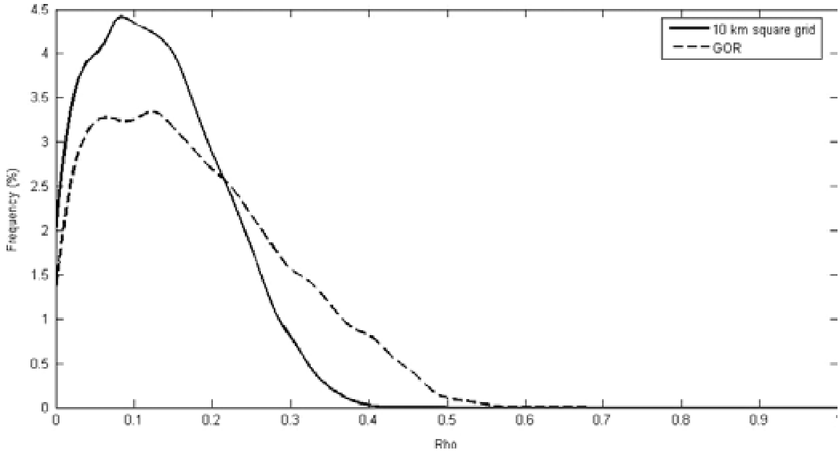

Figure 5 Estimates of the posterior distributions for ρ constructed from the Monte Carlo iterates: 10 km grid square versus GOR.

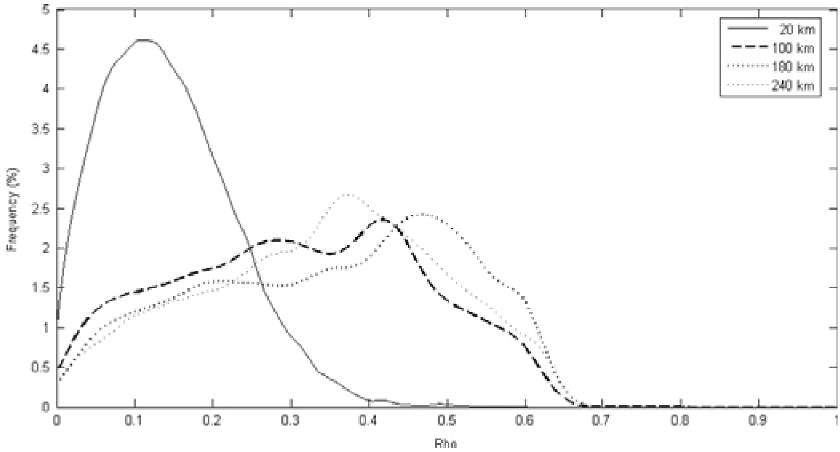

Figure 6 Estimates of the posterior distributions for ρ when s = 20 km; s = 100 km; s = 180 km and s = 240 km constructed from the Monte Carlo iterates.

neighbourhood comprehends farms within a relatively large area (i.e. between 0.31 and 0.35 for SM3-100, SM3-180 and SM3-240) and models where neighbourhood contains farms within a smaller area of between 10 and 20 grid square km (i.e. between 0.13 and 0.14 for SM1 and SM3-20). In other words, it is more likely that the efficiency level of a farm is closer to the mean efficiency level of neighbouring farms when neighbourhood is defined for areas larger than 20 grid square km than for areas equal or smaller than 20 grid square km. As it can be seen in the maps, there are areas where adjacent squares have different efficiency levels, which lead to low values of ρ if ‘small’ neighbourhoods are considered. By extending the cut-off area the efficiency level of a farm gets closer to the efficiency level of its neighbouring farms.

5. Conclusions

The work outlined in this article has shown how spatial dependence can be accounted for within a stochastic frontier model by specifying inefficiency to be spatially autoregressive and including a parameter that measures the level of spatial dependence. By examining different spatial structures, we also showed how spatial dependence varies with different characterisations of the neighbourhood. The application of these techniques gives insightful information on whether there is spatial dependence in technical efficiency and has implications for best implementing policies aiming at improving underperforming farms.

The results suggest that there is spatial dependence in technical efficiency in dairy farms in England and Wales, and not accounting for it may produce biased results for the efficiency distribution. Farm technical efficiency depends to some degree on where the farm is located, and therefore, policies aimed at improving efficiency should take this into account.

The results for the conditional posterior of the spatial dependence parameter ρ are sensitive to the specification of the spatial weight matrix. It may not be only due to whether we use a connectivity matrix or a distance based spatial matrix but also owing to the cut-off size chosen. Thus, results from the connectivity matrix raise the question of how big the size of the spatial effect is. Mean spatial dependence reaches its maximum over a 100 km distance from the farm. Therefore, an examination of how sensitive results are to the type of weight put to individual farms as well as to the cut-off size chosen must be conducted to deliver meaningful results.

When analysing spatial heterogeneity, there is not a strong reason to support this being analysed at the political division level. In fact, usually heterogeneity occurs owing to the geographical and climatic characteristics of the area, which do not necessarily coincide with the political divisions of the land. Therefore, it is not surprising that heterogeneity is not found at political division level and it should be analysed accordingly. The consequences of studying heterogeneity at the wrong spatial level may be important as policy decisions would be based on misleading information. For example, based on an analysis for which no heterogeneity is found between a number of regions the same policy may be applied for these regions. However, if heterogeneity is in fact present at other smaller or larger spatial levels a more appropriate policy would be to apply different policies within those regions or covering various regions. The results presented in this article are important for policy makers as they highlight that policies devoted to improving farm performance need not necessarily be applied at the national or regional level. Spatial dependency or heterogeneity may cross political borders or differ within the same political region. This represents a challenge to policy makers on how to implement policies at the 'right' geographical level. Governments would like to see production allocated to those areas where efficiency is higher and/or help to increase efficiency in those areas where efficiency can be improved. This article has shown that farm specific inefficiency associated with spatial dependence can be identified as well as identifying those farms which may need help in improving their performance. Most importantly, because farm efficiency was found to be spatially dependent this means that there are drivers behind technical efficiency that are correlated with where farms are located. Identification of these drivers can have a major impact on designing policies aimed at improving farm performance.

Future research should focus on developing ways to estimate the distance at which the dependence parameter reaches its maximum. This would be helpful to design more accurately the spatial level of policies that aim to improve farm efficiency. Once it has been identified that spatial dependence exists, research should concentrate on identifying and incorporating into the analysis potential explanatory factors for such spatial dependence.

References

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

Anselin, L. (2001). Spatial effects in econometric practice in environmental and resource economics, *American Journal of Agricultural Economics* 83, 705–710.

Anselin, L. (2002). Under the hood. Issues in the specification and interpretations of spatial regression models, *Agricultural Economics* 27, 247–267.

Bell, K.P. and Bockstael, N.E. (2000). Applying the Generalised-Moments estimation approach to spatial problems involving microlevel data, *The Review of Economics and Statistics* 82, 72–82.

Bell, K.P. and Dalton, T.J. (2007). Spatial econometric analysis in data-rich environments, *Journal of Agricultural Economics* 58, 487–501.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society*. *Series B (Methodological)* 36, 192–236.

Bockstael, N.E. (1996). Modeling economics and ecology: the importance of a spatial perspective, *American Journal of Agricultural Economics* 78, 1168–1180.

Brümmer, B., Glauben, T. and Thijssen, G. (2002). Decomposition of productivity growth using distance functions: the case of dairy farms in three European countries, *American Journal of Agricultural Economics* 84, 628–644.

Brümmer, B., Glauben, T. and Lu, W. (2006). Policy reform and productivity change in Chinese agriculture. Adistance function approach, *Journal of Development Economics* 81, 61–79.

Casella, G. and George, E.I. (1992). Explaining the Gibbs sampler, *The American Statistician* 46, 167–174.

Coelli, T.J. and Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: with application to European railways, *European Journal of Operational Research* 117, 326–339.

Coelli, T.J., Prasada Rao, D.S., O′Donnell, D.J. and Battese, G.E. (2005). *An Introduction to Efficieny and Productivity Analysis*. Springer Science + Business Media.Inc, New York, USA.

van Den Broeck, J., Koop, G., Osiewalski, J. and Steel, M.F.J. (1994). Stochastic frontier models: a Bayesian perspective, *Journal of Econometrics* 61, 273–303.

Druska, V. and Torrace, W. (2004). Generalized moments estimation for spatial panel data: indonesian rice farming, *American Journal of Agricultural Economics* 86, 185–198.

Färe, R., Grosskopf, S. and Lovell, C.A. (1994). *Production Frontiers*. Cambridge University Press, Cambridge.

Fernández, C., Koop, G. and Steel, M.F.J. (2000). A Bayesian analysis of multiple-output production frontiers, *Journal of Econometrics* 98, 47–79.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE transactions on pattern analysis and machine intelligence* 6, 721–741.

Hadley, D. (2006). Patterns in technical efficiency and technical change at the farm-level in England and Wales, 1982–2002, *Journal of Agricultural Economics* 57, 81–100.

Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* 57, 97–109.

Holloway, G. and Lapar, L. (2007). How big is your neighbourhood? Spatial implications of market participation among Filipino smallholders. *Journal of Agricultural Economics* 58, 37–60.

Iraizoz, B., Bardaji, I. and Rapun, M. (2005). The Spanish beef sector in the 1990s: impact of the BSE crisis on efficiency and profitability, *Applied Economics* 37, 473–484.

Kim, C.W., Phippa, T.T. and Anselin, L. (2003). Measuring the benefits of air quality improvement: a spatial hedonic approach, *Journal of Environmental Economics and Management* 45, 24–39.

Koop, G. (2003). *Bayesian Econometrics*. John Wiley & Sons Inc., Chicester, West Sussex.

Koop, G., Steel, M.F.J. and Osiewalski, J. (1995). Posterior Analysis of Stochastic Frontier models using Gibbs sampling, *Computational Statistics* 10, 353–373.

LeSage, J.P. (1999). *The Theory and Practice of Spatial Econometrics*. Available from URL: http://www.spatial-econometrics.com/html/sbook.pdf [accessed 17 May 2012].

Lovell, C.A., Richardson, S., Travers, P. and Wood, L.L. (1994). *Resources and Functionings: A New View of Inequality in Australia*. Springer-Verlag Press, Berlin.

Mead, R. (1967). A mathematical model for the estimation of inter-plant competition, *Boimetrics* 23, 189–205.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A. and Teller, E. (1953). Equations of state calculations by fast computing machines, *Journal of Chemical Physics* 21, 1087–1092.

O'Donnell, D.J. and Coelli, T.J. (2005). A Bayesian approach to imposing curvature on distance functions, *Journal of Econometrics* 126, 493–523.

Ord, K. (1975). Methods for models of spatial interaction, *Journal of the American Statistical Association* 70, 120–126.

Orea, L. (2002). Parametric decomposition of a generalized Malmquist productivity Index, *Journal of Productivity Analysis* 18, 5–22.

Roe, B., Irwin, E.G. and Sharp, J.S. (2002). Pigs in space Modeling the spatial structure of hog production in traditional and nontraditional production regions, *American Journal of Agricultural Economics* 84, 259–278.

Schmidt, A.M., Moreira, A.R.B., Helfand, S.M. and Fonseca, T.C.O. (2009). Spatial stochastic frontiers: accounting for unobserved local determinants of inefficiency, *Journal of Productivity Analysis* 31, 101–112.

Weiss, M.D. (1996). Precision farming and spatial economic analysis: research challenges and opportunities, *American Journal of Agricultural Economics* 78, 1275–1280.

Whittle, P. (1954). On stationary processes in the plane, *Biometrika* 41, 434–449.