UNITE MIXTE DE RECHERCHE
EN ECONOMIE PUBLIQUE

JOINT RESEARCH UNIT IN
PUBLIC ECONOMICS

2020
CAHIERS DE RECHERCHE
WORKING PAPERS
2020/01

Spatial autocorrelation in econometric land use models: an overview

R. Chakir; J. Le Gallo

UMR Economie Publique
Avenue Lucien Brétignières – 78850 Grignon
16 rue Cl. Bernard – 75005 Paris
Tel. +33 (0)1 30 81 53 30
Fax. +33 (0)1 30 81 53 68
https://www6.versailles-grignon.inra.fr/economie_publique/

Spatial autocorrelation in econometric land use models: an overview

Raja Chakir[1] & Julie Le Gallo[2]
[1] *Université Paris-Saclay, INRAE, AgroParisTech, Economie Publique, France.*
[2] *CESAER UMR1041, Agrosup Dijon, INRAE, Université de Bourgogne Franche-Comté, 21000, Dijon, France*

Abstract. This chapter provides an overview of the literature on econometric land use models including spatial autocorrelation. These models are useful to analyze the determinants of land use changes and to study their implications for the environment (carbon stocks, water quality, biodiversity, ecosystem services). Recent methodological advances in spatial econometrics have improved the quality of econometric models allowing them to identify more precisely the determinants of land use changes and make more accurate land use predictions. We review the current state of the literature on studies which account explicitly for spatial autocorrelation in econometric land use models or in the environmental impacts of land use.

1 Introduction

Land use plays a vital role in many major societal issues: food security (Verburg et al., 2013), preservation of biodiversity and ecosystem services (Foley, 2005), climate change mitigation (Lal, 2004) and the achievement of many Sustainable Development Goals (Gao and Bryan, 2017). Land use choices are the result of complex decision-making processes related to the local and global biophysical and socioeconomic drivers. The researcher faces two central and related questions: "what drives land use change?" and "what are the (environmental and socio-economic) impacts of land use change on stakeholders and the whole society?". The answers to these questions are crucial for the design of public policies related to how to feed the growing world population and avoid unwanted land use effects on the environment.

Various disciplines (economics, statistics, geography, land use science) have developed a range of empirical land use modeling approaches, using either aggregate or individual data. However, most of this work pays little attention to spatial autocorrelation (SA) in modelling land use although spatial interdependence is prevalent in all economic decisions in general and in land use decisions in particular. As a result, "standard" statistical and econometric methods, which assume independent observations, are inappropriate. More generally, taking account of the spatial dimension in econometric models involves two effects: spatial heterogeneity and SA. Spatial heterogeneity is the spatial differentiation of variables and behaviors in space and usually does not require specific econometric methods. Switching models, semi-parametric modeling of coordinates or clustered robust inference can handle this effect appropriately. Conversely, SA refers to the lack of independence among geographic observations. It measures the degree of similarity between an attribute in one location and the same attribute in neighboring locations (Anselin, 1988). Unlike temporal autocorrelation, SA is multidimensional requiring a specialized set of techniques, which are not simple extensions of two-dimensional time series methods. In this chapter, we focus on SA in econometric land use models.

There is a growing body of work on econometric modeling of land use. These studies address the determinants of land use and land use change and their impacts on water quality (Bockstael, 1996), deforestation (Chomitz and Gray, 1996), carbon sequestration costs (Lubowski et al., 2006) and habitat fragmentation (Lewis and Plantinga, 2007). Before the 1990s, econometric land use studies that explicitly introduced SA of observations were relatively rare as the presence of SA makes discrete choice models analytically intractable and requires use of computationally expensive Bayesian techniques or simulation estimation methods (Fleming, 2004). Thus, most land use studies and especially those based on individual data, avoid thorough treatment of spatial effects or use *ad hoc* procedures aimed at reducing the negative consequences of ignoring them[1].

Although land use studies taking explicit account of SA have increased (Brady and Irwin, 2011), they remain relatively scarce (Ay et al., 2017; Chakir and Le Gallo, 2013; Li et al., 2013; Sidharthan and Bhat, 2012; Ferdous and Bhat, 2012; Chakir and Parent, 2009). Most econometric land use models in papers published in high quality journals still tend either to ignore SA, or use *ad hoc* methods to deal with it (*e.g.* Irwin et al., 2003; Carrion-Flores and Irwin, 2004; Lubowski et al., 2008; Fezzi and Bateman, 2011). This is because SA raises several issues related to econometric estimation, hypothesis testing, and prediction – especially in the case of discrete choice models (Billé and Arbia, 2019).

Then, the aim of this chapter is to present the state of the art in the literature on econometric land use models and to show how methodological developments in spatial econometrics have been introduced into these models. We point out that this is not an exhaustive review; rather the objective is to highlight the main contributions to econometric land use models and their methodological advances. Our literature reviews departs from those provided by Brady and Irwin (2011), which summarize the econometric challenges of spatial models a in land use and hedonic model context, Plantinga (2015), who focuses on methods for integrating economic land-use and biophysical models and Chakir (2015), who reviews methodological developments in spatial econometrics that have been introduced into land use models. The main goal of this literature review is to summarize the studies which include SA explicitly in land use models or in models of the environmental impacts of land use[2].

The remainder of the chapter is organized as follows. First, we provide some general considerations related to the econometric modeling of land use (section 2). Then we focus respectively on land use share (section 3) and discrete choice (section 4) models. Section 5 shows how SA enhances models that focus on the impact of land use on various environmental outcomes. Section 6 concludes and highlights some directions for further research.

2 Econometric land use models

Most econometric land use studies are based on the classical theory which considers that land use activities are chosen to maximize land rents and that rents vary with land characteristics,

---

[1] Ignoring spatial effects can result in biased and/or inefficient parameter estimates or assessment of statistical significance (Anselin, 1988).

[2] We did a literature search for articles adopting an explicit spatial econometric approach to land use issues. Then, among these articles, we chose those that we considered the most important either from a methodological point of view or from the point of view of the environmental impacts of land uses.

in particular soil fertility (Ricardo, 1817) and location (von Thunen, 1875). Yet, other factors might influence land use decisions for a given land parcel: socioeconomic factors (input and output prices) and policy variables (taxes and subsidies). The extent and significance of these determinants are analyzed in two broad categories of models: aggregate land use models which use aggregate (county level, state level, etc.) data, and individual land use models which are based on parcel-level or sample plot data. Table 1 provides a summary of some papers that provide econometric modeling of land use employing both aggregate and individual data.

Paper	Land use categories	Model	Spatial	Journal
Aggregate land use share studies				
Alig (1986)	crops, 3 types de forest, pasture and urban	land-use share	no	FS
Lichtenberg (1989)	7 crops	land-use share	no	AJAE
Stavins and Jaffe (1990)	crops, forest	land-use share	no	AER
Wu and Segerson (1995)	6 cultures	land-use share	no	AJAE
Plantinga (1996)	agriculture to forest	land-use share	no	AJAE
Plantinga et al. (1999)	agriculture, forest and urban/other land use	land-use share	no	AJAE
Hardie and Parks (1997)	agriculture, forest, urban/other use	land-use share	no	AJAE
Plantinga and Ahn (2002)	crops, forest	land-use share	no	JARE
Chakir and Le Gallo (2013)	agriculture, forest, urban and other use	land-use share	yes	EE
Marcos-Martinez et al. (2017)	19 land use categories	land-use share	yes	LUP
Chakir and Lungarska (2017)	agriculture, forest, urban and other use	land-use share	yes	SEA
Marcos-Martinez et al. (2017)	extensive grazing, pastures, cereals, annuals, perennials	land-use share	yes	LUP
Amin et al. (2019)	Deforestation		yes	JEEM
Individual discrete choice studies				
McMillen (1989)	farm, residential	multinomial logit	no	LE
Bockstael (1996)	urbanisation	probit	no	AJAE
Chomitz and Gray (1996)	deforestation	multinomial logit	no	WBER
Claassen and Tegene (1999)	culture, pasture	probit	no	ARER
Carrion-Flores and Irwin (2004)	urbanisation	probit	no	AJAE
Lubowski et al. (2006)	crops, pasture, forest, urban, *range* and CRP	nested logit	no	JEEM
Chakir and Parent (2009)	agriculture, forest, urban and other uses	multinomial probit	yes	PIRS
Wang and Kockelman (2009)	4 levels of urbanisation	ordered probit	yes	PIRS
Ferdous and Bhat (2012)	4 levels of urbanisation	ordered probit	yes	JGS
Sidharthan and Bhat (2012)	urban, commercial, industrial and non-developped	multinomial probit	yes	GA
Li et al. (2013)	farm, forest, grass, water, urban, unused	multinomial probit	yes	LE
Bhat et al. (2015)	commercial, industrial, residential, underdeveloped	multiple discrete-continuous probit	yes	JRS
Carrión-Flores et al. (2018)	commercial, industrial, residential, parks, agriculture	multinomial logit	yes	RSUE

Table 1: Example of econometric individual and aggregate, spatial and aspatial land use studies. (AER: American Economic Review, AJAE: American Journal of Agricultural Economics, ARER: Agriculture and Ressource Econ. Review, EE: Ecological economics, GA: Geographical Analysis, FS: Forest Science, JARE: J. of Agri and Resource Econ. JEEM: Journal of Environmental Economics and Management, JGS: Journal of Geographical Systems, JRS: Journal of Regional Science, LE: Land Economics, LUP: Land Use Policy, PIRS: Papers in regional Science, RSUE: Regional Science and Urban Economics, SEA: Spatial Economic Analysis, WBER: World Bank Economic Review)

Aggregate and individual land use models are complementary and provide different insights into the determinants of land use and land use changes, and their environmental effects. The choice between an aggregate and individual land use model often depends on data availability and the objective of the study. If the objective is to make land use predictions at the scale of one or a group of countries (such as European countries), an aggregate data model is required. If the objective is to study the effects of land use on biodiversity or water quality, a model based on individual data is more relevant. Both approaches have drawbacks.

On the one hand, aggregate data limits the capacity to explain the effects of heterogeneous physical characteristics such as soil quality on land use choices. Because the data are aggregated to units such as the county, intra-county variations in soil quality are ignored. Moreover, while aggregate data can be useful to study global issues (changes in land use shares within a region), the results are of limited use for policy making related to the spatial organization of land use in a region, or local issues related to biodiversity, water quality, or urbanization.

On the other hand, one of the frequent difficulties related to modeling land use at the individual level is the lack of "good" explanatory variables or their scale incompatibilities. Although geophysical explanatory variables such as slope, altitude and soil quality, are increasingly available and at very fine resolution, economic variables (rents, conversion costs and prices) are either not available or observable only at aggregate scales. To compensate for this lack of data, empirical models often use proxies for rents at more or less aggregated scales. Another disadvantage of individual level land use models is related to the difficulty involved in estimating discrete choice models in the multinomial case. This difficulty is accentuated if SA is included in the specification.

In relation to this latter issue, SA in land use choices tends not to be included in theoretical frameworks but added ex post in the empirical specification. In land use modeling, SA can stem from two sources. First, it can arise from spillovers among the error terms due to omitted spatial variables affecting land-use decisions such as weather or soil quality. A spatial error model or spatial robust inference allows to control for these omitted variables provided that they are not correlated with the observables. Second, it can arise from spillovers among land use decisions or spatial interaction relationships in the land use choices. This might be due for example, to the neighboring plots being owned by the same landowner, or to shared information which induces forest or agricultural clustering and landowners adopting the same technology based on shared learning. In this case, a spatial autoregressive model would account for these spatial interactions.

In the case of aggregate data, logarithmic transformation on land use shares implies linear equations that can easily be estimated. Therefore, SA in the case of land use models can be estimated using spatial models in the linear case (section 3). Conversely, in most cases of individual data (section 4), the presence of SA tends to make discrete choice models analytically intractable and requires use of simulation estimation methods or Bayesian techniques (Smith and LeSage, 2004). Other estimation procedures have been proposed in the literature: the expectation-maximization method (McMillen, 1992), the generalized method of moments (GMM) (Pinkse and Slade, 1998) and the composite maximum likelihood method (Sidharthan and Bhat, 2012; Ferdous and Bhat, 2012). For a detailed review of the inclusion of SA in discrete choice models see Fleming (2004), Smirnov (2010) and Billé and Arbia (2019).

Considering SA also sheds new light on the issue of prediction. Comparing individual and aggregate models with respect to their predictive accuracy is an ongoing and still open issue with mixed evidence. The seminal paper by Grunfeld and Griliches (1960) examined the relative power of individual (micro) and aggregate (macro) models for explaining aggregate outcomes and found that an aggregate model often performs better. In the context of land use models, Wu and Adams (2002) show that even in the case of linear models, the choice between the micro and macro scales to make aggregate predictions cannot generally be resolved by *a priori* reasoning. Ay et al. (2017) show that introducing SA in aggregate land use models provides better predictions than using individual aspatial models with higher numbers of observations. This suggests that there might be little to be gained from using individual land use data if the sole objective is to predict land use at the aggregate spatial resolution.

Some studies choose none of these modeling approaches and resort instead to *ad hoc* methods to circumvent the problems related to estimating discrete choice models in the presence of SA (De Pinto and Nelson, 2007). These models are summarized below:

- **Spatial sampling**: Most early studies in the land-use literature simply purge the data of SA using a spatial sampling technique which allows construction of a data sample without neighbors. This is a fairly widespread practice: Nelson and Hellerstein (1997); Carrion-Flores and Irwin (2004); Irwin et al. (2003); Irwin and Bockstael (2004); Lewis and Plantinga (2007); Lubowski et al. (2008); De Pinto and Nelson (2009); Fezzi et al. (2015)
- **Introduction of latitude and longitude as explanatory variables**: Nelson et al. (2001); Muller and Zeller (2002) claim to account for SA by using two additional explanatory variables representing the latitude and longitude of each observation. While this type of correction is likely to be useful if the spatial effect is caused by an unobserved variable which varies linearly between regions, it captures spatial heterogeneity rather than capturing the SA as claimed by the authors;
- **The introduction of spatially shifted geophysical variables**: Nelson et al. (2001) and Munroe et al. (2002) use spatial shifts (i.e. weighted averages of values in neighboring locations) of geophysical variables such as soil type, slope and vegetation index as exogenous variables. A possible justification for using these types of variables is that they account for the direct influence of the environment on land-use decisions in a particular location.

While useful, these methods cannot control for substantive SA, an issue to which we turn in the next two sections.

3 Linear land use models

The objective of most studies using aggregate data is to identify the determinants of land use shares. Most U.S. econometric studies use the county scale and land use data derived generally from federal sources such as agricultural census. The most common method is to specify county land use shares as a logistic function. Examples of studies that use this method include Lichtenberg (1989), Plantinga (1996) and Hardie and Parks (1997). While the authors attempt to explain the factors that influence the share of land allocated to a particular land use, other aggregate data studies try to explain changes in land use shares in an area (Stavins and Jaffe, 1990; Plantinga and Ahn, 2002). All these studies ignore SA. More recent studies taking explicit account of SA have been conducted at the French level by Chakir and Le Gallo (2013), Ay et al. (2017) and Chakir and Lungarska (2017) who estimate aggregate land use share models at the department level, 12×12 km and 8×8 km grid cells respectively.

3.1 Land use share models

Although all econometric studies are based on the same economic theory, several variants of theoretical land allocation models have been proposed (Lichtenberg (1989); Stavins and Jaffe (1990); Plantinga (1996); Hardie and Parks (1997)). We present here a fairly simple version of these models based on Wu and Segerson (1995)'s static model where the landowner n_i ($n_i = 1, ...N$) in the region i ($i = 1, ..., I$) is assumed to be risk neutral and maximizes his expected profit from the use k ($k = 1, ...K$) on quality land j ($j = 1, ..., J$), at time t ($t = 1, ...T$), denoted $\pi_{jk}(x(t, n_i), a_{jk}(t, n_i), n_i)$, where $x(t, n_i)$ is a vector of the exogenous variables such as prices, costs and other economic variables and $a_{jk}(t, n_i)$ is the area of land of quality j allocated to use k. For each quality of land, the landowner chooses the area $a_{jk}(t, n_i) \geq 0$ that maximizes his total profit:

$$\sum_{k=0}^{K} \pi_{jk}(x(t, n_i), a_{jk}(t, n_i), n_i) \quad \text{subject to} \sum_{k=0}^{K} a_{jk}(t, n_i) = A_j(t, n_i) \tag{1}$$

where $A_j(t, n_i)$ is the total surface of available quality land j. The resolution of the optimization program (1) gives the optimal area $a^*_{jk}(x(t, n_i), A_j(t, n_i), n_i)$ allocated to each use k for each quality of the land j at time t. The optimal share of land allocated to the use k is:

$$s_k(x(t, n_i), t, n_i) = \frac{1}{A_j(t, n_i)} \sum_j a^*_{jk}(t, n_i) \tag{2}$$

The optimal uses derived from the theoretical model for each owner should be aggregated to match the scale of the observed data. In practice, the available data are the shares of land uses at an aggregate resolution (county, region, municipality). The land use share k ($k = 1, ..., K$) in the region i at time t is written as:

$$s_{ikt} = p_{ikt} + \varepsilon_{ikt} = \frac{e^{\beta'_k X_{it}}}{\sum_{j=1}^{K} e^{\beta'_j X_{it}}} + \varepsilon_{ikt} \quad \forall i = 1, \ldots, I, \forall k = 1, \ldots, K \text{ and } \forall t = 1, \ldots, T \tag{3}$$

where p_{ikt} is the expected share of land allocated for use k in the i region at time t. The observed land use share at time t, s_{ikt} may differ from the optimal land use share due to possible hazards such as climate or policy shocks. These elements, of zero average, are captured by the error term ε_{ikt}. X_{it} are the explanatory variables and β'_k are the associated coefficients.

As in Wu and Segerson (1995) and Plantinga et al. (1999), most aggregate studies specify land use shares in the logistic functional form for three reasons: first, this functional form allows predicted land use shares to stay between 0 and 1, second, this specification is parsimonious in terms of parameters, and third, logarithmic transformation allows use of linear equations which are easily estimated. This transformation[3] has been proposed by Zellner and Lee (1965) and, applied to land use choices, it allows to write the logarithm of each use share normalized

[3] This transformation corresponds to the Additive Log Ratio transformation (ALR) in the literature on composition data in statistics, see Aitchison (1986) for more details.

by a given share as follows:

$$\widetilde{y}_{ikt} = ln(s_{ikt}/s_{ikt}) = \beta_k^{'} X_{it} + u_{ikt} \quad \forall i = 1, \ldots, I, \forall k = 1, \ldots, K \text{ et } \forall t = 1, \ldots, T \qquad (4)$$

where u_{ikt} is the transformed error term. Model (4) has $K - 1$ equations that are Seemingly Unrelated Regressions (SUR) and can be estimated by methods accounting for correlations between the error terms associated with each equation.

3.2 Spatial autocorrelation in linear models

In linear specification models SA is handled by the inclusion of spatially lagged variables, that is weighted averages of the observations of "neighbors" of a given location. These spatially lagged variables can be used as the dependent variable (spatial auto-regressive SAR models), explanatory variables (spatial cross regressive SLX models), or the error terms (SEM) or any combination of these options which results in a range of spatial models (Elhorst, 2010). For instance, the spatial autoregressive combined (SARAR) model accounts simultaneously for autocorrelation in the error term and for spatial associations of the dependent variable. The spatial Durbin model (SDM) is a combination of SAR and SLX and can be reduced to SLX (LeSage and Pace, 2009), while the spatial Durbin error model (SDEM) integrates all the elements of the SLX and the SEM. Finally, the general nesting spatial (GNS) model combines the SARAR and the SLX models. Until the early 2000s, most empirical spatial econometrics studies were interested mainly in two specifications: SAR and SEM. Specifications accounting for richer and combined forms of SA are now more commonly estimated. For more details on the taxonomy of linear SA models for cross-section data see Elhorst (2014).

The choice of the best spatial specification can be made based on theory or by applying statistical tests to different models. The literature proposes several strategies, the most common being either the so-called classical strategy starting from the simplest "specific to general" model, the most general model going from "general to specific". Florax et al. (2003) compare these strategies and show that the classical approach gives the best results in terms of identifying the best specification and most precisely estimated parameters but LeSage and Pace (2009) argue that choice of the best specification should start with the SDM. Elhorst (2010) proposes a mix of these two strategies.

Model	Model	Interpretation
SEM	$\widetilde{y} = X\beta + \varepsilon$ and $\varepsilon = \lambda W \varepsilon + u$	Unobserved omitted variables follow a spatial pattern, data measurement errors
SAR	$\widetilde{y} = \rho W \widetilde{y} + X\beta + \varepsilon$	LU for one location is determined jointly with that of neighbors
SLX	$\widetilde{y} = X\beta + WX\gamma + \varepsilon$	LU for one location is determined by the explanatory variables of neighbors
SDM	$\widetilde{y} = \rho W \widetilde{y} + X\beta + WX\gamma + \varepsilon$	A combination of SLX and SAR and can be reduced to SLX
SARAR	$\widetilde{y} = \rho W \widetilde{y} + X\beta + \varepsilon$ and $\varepsilon = \lambda W \varepsilon + u$	A combination of SEM and SAR
SDEM	$\widetilde{y} = X\beta + WX\gamma + \varepsilon$ and $\varepsilon = \lambda W \varepsilon + u$	A combination of SEM and SLX
GNS	$\widetilde{y} = \rho W \widetilde{y} + X\beta + WX\gamma + \varepsilon$ and $\varepsilon = \lambda W \varepsilon + u$	A combination of SLX and SARAR

Table 2: Summary table of the estimated linear land use (LU) spatial model specifications (Chakir and Lungarska, 2017). ρ is the spatial autoregressive coefficient, λ the SA coefficient, γ and β represent a vector of unknown parameters to be estimated. W is a nonnegative $n \times n$ matrix describing the spatial configuration or arrangement of the units in the sample.

3.3 Example of spatial land studies with linear models

This section provides some examples of aggregate land use studies which take account of SA.

Some works include SA in order to improve the specification and understanding of what drives land use change. For instance, Meyfroidt and Lambin (2008) analyze the causes of reforestation in Vietnam during the 1990s on a national scale, and test emerging forest transition theories on the same scale. They build a reforestation spatial lag regression model using census and geographic data at a fine level of aggregation for the whole country. Their results show that forest land distribution affects forests not just in the focal district but also in neighboring districts. This observation can be interpreted in terms of a diffusion process: early and successful implementation of the policy in some districts may have facilitated its rapid adoption by neighboring districts.

Marcos-Martinez et al. (2017) estimate the determinants of land-use in Australia's intensive agricultural region during the period 1992–2010. They estimate land-use shares with spatial error and random effects combined with variance decomposition analysis to identify the statistical significance, direction and magnitude of the observed associations between land-uses and their drivers. Their results show that improved transportation infrastructure, zoning regulations and mechanisms to reduce exposure to farm debt and climate variability risks have significant impacts on the configuration of the Australian agricultural landscape.

Amin et al. (2019) analyze whether protected areas are efficient instruments to fight deforestation in Brazilian Amazonia. They estimate a dynamic SDM and assess the impact of different types of protected areas (integral protected areas, sustainable protected areas, indigenous lands) on deforestation. The results differ according to the type of protected area: i) integral protected areas and indigenous lands reduce deforestation; ii) sustainable use areas do not contribute to reducing deforestation; and iii) the spillover effects generated by integral protected areas and indigenous lands lead to a reduction in the deforestation in their vicinities.

Two studies focus on prediction in spatial land use share models. Chakir and Le Gallo (2013) make a methodological contribution to the literature by controlling for both unobservable individual heterogeneity and SA in an aggregate land-use model. Their study was conducted on a panel of land use data at the French departments NUTS3 scale, observed between 1992 and 2003. The authors were interested in the relationship between four land uses (agriculture, forest, urban, and other) and their potential economic and demographic determinants. The econometric model consists of a system of three equations with a panel dimension and SA in the errors associated to each equation. Thus, their econometric model is a SUR model with random individual effects and autoregressive spatial structure of the error term. The model was estimated using the FGLS (Feasible Generalized Least Square) estimation method proposed by Baltagi and Pirotte (2011) for SUR-SEM-RE (Seemingly Unrelated Regressions-Spatial Error Model-Random Effects) model estimations. Their results are of three orders: first, controlling for both unobservable individual heterogeneity and SA yields the best predictions relative to any other specification in which SA and/or individual heterogeneity are ignored. Second, taking into account the correlations between the error terms in the different equations does not seem to improve prediction performance. Third, ignoring individual heterogeneity introduces substantial loss of prediction accuracy.

Chakir and Lungarska (2017) estimate land-use-share models for France at a homogeneous (8×8 km) grid scale for five land use classes - agriculture, pasture, forest, urban and other. They investigate the determinants of land use shares using economic, physical and demographic explanatory variables. They model SA between grid cells and compare prediction accuracy and estimated elasticities for the different spatial model specifications (OLS (ordinary least square), SLX, SEM, SAR, SDM, SDEM, SARAR, GNS). They compare these spatial specifications using three rent proxies: farmers' revenues, land prices and shadow land prices derived from a mathematical programming model. Their comparison is based on several criteria: quality of economic explanation (significance of agricultural rents and their marginal impacts), prediction quality (NRMSE), specification tests (LM tests), and goodness of fit (log-likelihood, R2, AIC). The test results show that the SDM, SDEM, SARAR and GNS models should be considered. According to the goodness-of-fit (pseudo-R2, log-likelihood and AIC) and prediction quality criteria, GNS is the specification that best fits their data. In a context of aggregate land use, the existence of autocorrelation is due mainly to spatially correlated errors – essentially a data measurement problem. This applies especially to their case since they use artificially constructed grids, and different scales for the explanatory variables and land use data. Their results show also that including SA in land use share models improves the quality of the predictions which confirms the results in the previous aggregate land use literature.

4 Discrete choice land use models

When using individual (parcel or plot) data, the land use variable is generally a categorical variable so that estimating land use patterns on individual data usually requires a discrete choice framework. Discrete choice models are based on McFadden (1974)'s random utility theory which states that the landowner decides to switch from one use to another if the expected net revenues exceed the revenues from the original use.

4.1 Individual choice land use model

This section presents the theoretical land use model based on individual data as in Lubowski et al. (2008). We assume that the landowner chooses the land use of a plot based on the costs and benefits associated with each possible use. For example, the landowner chooses land use k at time t if:

$$R_{kt} - rC_{jkt} > R_{jt} \quad \forall j, k = 1, ..., K \quad \text{and} \quad \forall t = 1, ..., T \tag{5}$$

where R_{jt} and R_{kt} represent the discounted expected net benefits at time t of a unit of land for uses j and k, respectively, C_{jkt} is the marginal cost of converting a unit of land from use j to use k at time t ($C_{jjt} = 0$) and r is the discount rate.

In order to estimate the determinants of land use econometrically, the theoretical model suggests comparing the benefits and costs of converting land from one use to another at each date. To move to the econometric specification, land use conversion revenues and costs are rewritten as functions of the observed and unobserved variables. Thus the utility U_{ikt} of the owner of parcel i with land use k at time t is written as follows:

$$U_{ikt} = \beta x_{ikt} + \epsilon_{ikt} \quad \forall i = 1, ...N, \quad \forall k = 1, ..., K \quad \text{and} \quad \forall t = 1, ..., T, \tag{6}$$

where x_{ikt} are the observed explanatory variables, β the vector of parameters to be estimated and ϵ_{ikt} are the error terms which take account of the unobserved variables that might affect the landowner's utility.

We assume that the owner has a choice between K land use categories for each parcel at each date. The landowner chooses the optimal land use for his or her plot by comparing the utilities associated to each land use category. If $y_{it} = 1, 2, ...K$; is the landowner's land use choice for the parcel i at time t, we obtain:

$$y_{it} = k, \text{if} \quad U_{ikt} \geq \max U_{ijt} \quad \forall i = 1, ...N, \quad \forall j, k = 1, ..., K \quad \text{and} \quad \forall t = 1, ..., T, \qquad (7)$$

Thus, the probability that the parcel i is allocated to the use k at the time t is written as:

$$P(y_{ikt} = 1) = Pr[U_{ikt} \geq maxU_{ijt}] \qquad (8)$$

for all $j = 1, ..., K$ with $y_{ikt} = 1$ if k is the observed use and 0 otherwise; U_{ikt} is the utility associated with land use k.

Since estimation of discrete choice models in the multinomial case is dimensionally constrained, some studies are limited to two use categories and use a probit model in the binary case (Bockstael, 1996; Kline and Alig, 1999; Irwin and Bockstael, 2002). Other studies estimate a multinomial logit model because of its computational simplicity (Chomitz and Gray, 1996; Nelson and Hellerstein, 1997; Nelson et al., 2001) which involves the questionable assumption of independence of irrelevant alternatives (IIA). Finally, a nested logit model could be a good alternative if the alternatives can be partitioned into several subsets.

4.2 Spatial autocorrelation in discrete choice models

SA can be accounted for in the discrete choice land-use model (equations (6) - (8)), by including the spatially lagged variables or the error terms. Simplifying the notations and removing the subscripts, the general nonlinear nesting model (GNNM) can be written as follows:

$$U = \rho WU + X\beta + WX\gamma + \varepsilon, \text{ and } \varepsilon = \lambda W\varepsilon + u \qquad (9)$$

where WU is the shifted utility function for the weight matrix W, ρ is the autoregressive spatial parameter which indicates the magnitude of the interaction between the latent variables U, γ, like β, is a vector of the unknown parameters to be estimated, λ is the parameter of the intensity of the SA between the residuals, and u is a classical error term such as $u \sim iid(0, \sigma^2 I)$. The GNNM model presented in eq. (9) becomes a SEM model if $\rho = 0$ and $\gamma = 0$, and becomes a SAR model if $\lambda = 0$ and $\gamma = 0$. In contrast to the linear case, the spatially lagged variable in the SAR model is not observable. For example, in the case of a land use model it is the utility associated to the profitability of neighboring plots and not the observed land use which should define the utility function of the landowner (Anselin, 2002).

In the case where the error terms ε follow a normal distribution, estimation of the probit-SAR model raises two problems. On the one hand, heteroskedasticity makes the classical estimators inconsistent. On the other hand, estimation of a probit-SAR requires computation of a likelihood function with $N-1$ (where N is the number of observations) integrals which makes maximum likelihood estimation impossible. This second difficulty applies also to the logit model case (Anselin, 2002). Several approaches have been proposed in the literature to deal with these estimation problems, including simulation estimation (Geweke et al., 1994) or Bayesian (LeSage, 2000) methods able to deal with the computation of multidimensional integrals of the likelihood function. Other estimation procedures have been proposed to cope with the problems associated to the introduction of SA in the case of discrete choice models. These include the expectation-maximization method (McMillen, 1992), the GMM (Pinkse and Slade, 1998), the maximum pseudo-likelihood method (Smirnov, 2010) and finally the method of maximum approximate composite marginal likelihood (CML) (Sidharthan and Bhat, 2012). For detailed reviews of SA in discrete choice models see Fleming (2004) and Smirnov (2010). Simulation estimation and Bayesian methods have been employed only recently to deal with the computational problems associated to considering SA in discrete choice models. Because these methods are still relatively expensive to implement their use in the land use literature remains limited. Table 3 provides an overview of these studies.

Model	Estimation method	Example
spatial autoregressive logit	Bayesian	Blackman et al. (2008)
ordered probit	Bayesian	Wang and Kockelman (2009)
multinomial probit	Bayesian	Chakir and Parent (2009)
random parameter logit	Max simulated likelihood	Lewis et al. (2011)
multinomial probit	Max approximate CML	Sidharthan and Bhat (2012)
ordered probit	Max CML	Ferdous and Bhat (2012)
multinomial logit	GMM	Li et al. (2013)
multiple discrete-continuous probit	Max CML	Bhat et al. (2015)
conditional parametric probit	Max Locally Weighted log-Likelihood estimator	McMillen and Soppelsa (2015)
multinomial logit	GMM	Carrión-Flores et al. (2018)

Table 3: Summary table of the estimated spatial discrete choice models

4.3 Examples of spatial land use studies with discrete choice models

To tackle the complexities induced by SA in discrete choice models for land use, some papers resort to Bayesian methods, for example Wang and Kockelman (2009) who estimate an ordered probit spatial dynamic model using satellite land cover data. Chakir and Parent (2009) also use a Bayesian approach to estimate land use determinants in a multinomial probit econometric model which accounts for both unobservable individual heterogeneity and SA in errors. They analyze the determinants of land based on a panel of 3,130 points in the Rhône department in France between 1992 and 2003. It appears that land use changes are indeed influenced by unobserved factors in neighboring plots. Finally, Blackman et al. (2008) estimate a bayesian heteroskedastic SAR logit model of land cover for a shade-grown coffee region in southern Mexico. Their results show that all other things being equal plots close to large cities are less likely to be cleared which contrasts to the pattern usually observed in natural forests. They also find that belonging to a coffee-marketing cooperative, farm size, and certain soil types are associated to tree cover while proximity to a small town center is associated to forest clearing. This study is extended in Blackman et al. (2012) who estimate a SAR probit model.

Other papers use variants of maximum simulated likelihood. Lewis et al. (2011)[4] estimate a random parameter logit model to take account of the non-observed space-time components of the willingness to pay. This specification makes it possible to take account of spatial heterogeneity rather than SA and also allows consideration of heteroskedasticity via a block variance-covariance matrix with individual effects which depend on space. It is a kind of SA but with no spatial structure and with a matrix of weights as in spatial models. In the spatial econometrics literature, CML has become a popular approach for estimating spatial probit models and has been used to model land use. For instance, Ferdous and Bhat (2012) analyze changes in the intensity of urban land use taking account of both the spatial dimension and temporal dynamics. Their econometric model is an ordered probit estimated using CML. The results show that ignoring the presence of autocorrelation and spatial heterogeneity introduces important bias and that ignoring spatial heterogeneity is more serious than ignoring lagged spatial dynamics. Sidharthan and Bhat (2012) use maximum approximate CML (MACML) to estimate a multinomial probit-type land-use model with SA between plots and spatial heterogeneity.

Finally, rather than tackling the spatial autoregressive coefficient directly as in the previous papers, McMillen and Soppelsa (2015) estimate a conditional parametric spatial probit model imposing far less structure on the data than conventional parametric models. They illustrate the approach using data on 474,170 individual lots in the City of Chicago. Their results suggest that simple functional forms are not appropriate for explaining the spatial variation in residential land use across the entire city. Similarly, Carrión-Flores et al. (2018) propose a GMM spatial estimator for a multinomial logit model with spatial lag dependence. The model is linearized to avoid the repeated matrix inversion required for the full GMM estimation. The linearization breaks up the estimation procedure into two simple steps: a standard multinomial logit model with no SA followed by a two-stage least squares (TSLS) estimation of the linearized model which accounts for SA. This model is applied to estimate land use conversion in the rural-urban fringe for four different land uses (agricultural, residential, industrial and commercial). The results show a positive SA of about 0.36—a result consistent with the widely-accepted idea that land-use conversion is a spatial process.

5 Land use and its impacts on the environment

Land use is considered as one of the main drivers of global changes to nature, which endanger numerous species or cause their extinction and compromise the supply of ecosystem services (ES) which are important for humans (Millenium Ecosystem Assessment, 2005). The protection of ES is emerging as a major concern alongside climate change issues (IPCC, 2019) and biodiversity conservation (IPBES, 2019). This has resulted in land use becoming a growing concern for policy makers as means of protecting ecosystems (Bateman et al., 2013). There is

[4] Several attempts in the literature introduce spatial dependence in multinomial models but, except for Lewis et al. (2011) to the best of our knowledge, they have not been used in the land use literature.

a large literature estimating the effects of land use on various ES: water quality (Fezzi et al., 2015), carbon sequestration (Lubowski et al., 2006) and biodiversity (Polasky et al., 2008).

In this context, accounting for SA when studying the impacts of land use on ES is a major issue. Research shows that including SA in species distribution models improves model fit and prediction accuracy (Record et al., 2013) and that, ignoring SA can produce inaccurate results (Kühn, 2007). Below, we review a selection of those studies that model SA explicitly to estimate the impacts of land use on the environment.

5.1 Land use and ES

The relationships between land use and ES is complex. For instance, some land uses such as intensive agriculture could have negative impacts on ecosystems while others could contribute to the provision of many ES. For example, tropical forests are an example of a supplier of ES at various scales. At the local scale, these services include wood, secondary forest products, pollination, etc. More generally, they sequester large amounts of carbon which regulates the global climate (IPCC, 2019). In addition, the productivity of some land uses such as agriculture are dependent on ecosystems such as biological pest control, soil fertility and pollination. Thus, degradation of these ecosystems constitutes a serious threat to the long-term agricultural productivity growth. Below, we provide two examples of spatial studies dealing with this link.

Chen et al. (2020) employ an integrated spatial panel approach to examine the geographic variations and spatial determinants of the ES balance in the middle reaches of the Yangtze River urban agglomerations (MRYRUA) in China. They analyze the spatiotemporal evolution features of landscape patterns and the supply of, demand for and balance among ES and landscape pattern metrics for the period 1995 to 2015. The results indicate that construction land in the MRYRUA has increased continuously, while farmland has decreased. Counties with higher ES supply and balance indices are concentrated primarily in mountainous areas, while the indices of ES demand in the three smaller urban agglomerations, plains areas, counties surrounding major cities and along major traffic routes are higher. SA and spatial spillover effects of the ES balance index are observed in the MRYRUA. Population density and road density are negatively associated to an ES balance. Landscape pattern metrics are also statistically significant, either positive or negative. The findings suggest that both drivers and spillover effects should be accounted for when considering integrative ecosystem management and land use sustainability measures in urban agglomerations. Both have important implications for urban planning and decision-making related to development and ES.

Klemick (2011) uses cross-sectional farm survey data to estimate the value of fallow ES in shifting cultivation in one region in the Brazilian Amazon. The objective is to test whether it provides economically significant local externalities which might justify forest conservation from a local perspective. The author estimates a production function to determine the contributions to agricultural income of on-farm and off-farm forest fallow. Soil quality controls, instrumental variables and spatial econometric approaches help address issues of endogeneity and variation in unobservable factors over space. The results suggest that Bragantina farmers generally allocate land between cultivation and fallow efficiently taking account of beneficial spillovers. This finding does not necessarily imply that farmers intentionally internalize the value of these services but might suggest that private land tenure plays a role in promoting sustainable land management given the different findings from other studies of shifting cultivation in common property tenure regimes which identify overexploitation of fallow biomass.

5.2 Land use and water quality

There is a large literature on the effects of land use on water quality and freshwater biodiversity. Most of these papers ignore SA. Here, we provide some examples of studies that model SA explicitly in a study of land use and water quality.

Most studies show that forest areas have a positive impact on water quality compared to intensive agriculture, livestock and urban areas. For example, Abildtrup et al. (2015) analyze the economic impacts of land use on the cost of drinking water supply, taking account of both the organizational choice of water supply and spatial factors in the same model. They estimate a model for the choice of management type and for the price of water, accounting for the potential dependence of the error terms between equations, as well as between neighboring water services. They estimate a sample selection model adapted to a spatial context that is, allowing for spatial lags and spatial error processes. The model is applied to data from the French Vosges department. The results show spatial interactions related to the characteristics of neighboring water services but no SA of the error terms in the management choice equation, or in prices. They show that forest land cover significantly reduces water supply costs at the large but not the local scale.

Induced land use adaptation on fresh water biodiversity is analyzed by Bayramoglu et al. (2020). They study the links between land use (agriculture, pasture, forest and urban environment) and the fish based index (FBI) an indicator of the ecological state of surface water measured for various French rivers observed between 2001 and 2013. They estimate two models: a spatial econometric model of land use and a spatial panel statistical model of the FBI. Their results indicate that adapting land use to climate change is reducing the biodiversity of freshwater in France. Furthermore, rivers located in regions with intensive agriculture and pastures are associated to lower freshwater biodiversity than those in forest regions. Simulations show that climate change will exacerbate these negative impacts through changes to land use. They show how two policies for regulating the level of fertilizers in agriculture and carrying capacity in grasslands could help improve freshwater biodiversity and cope with the adverse effects of land use and climate change.

5.3 Land use and climate change

The interactions between land use and climate are complex (IPCC, 2019). First, land use and land practices affect the global concentration of greenhouse gases (Houghton, 2003). Second, while land use change is an important driver of climate change, a changing climate can lead to changes in land use. For example, farmers might convert pasture to crop land which has higher economic returns under changing climatic conditions. Third, spatially heterogeneous land use activities have important impacts on local weather (Feddema et al., 2005). Fourth, land use changes could play an important role in mitigating climate change either by increasing carbon sequestration or by reducing greenhouse gas emissions. This could be achieved by adopting land uses such as afforestation or preservation of permanent pasture (Pielke, 2005).

Land use adaptation to climate change could exacerbate the adverse impacts of land use on the environment. For example, Lungarska and Chakir (2018) show that in France, climate change will reduce forest areas which could increase greenhouse gas emissions. They estimate a spatial econometric land use model and simulate the impacts of two IPCC climate change scenarios (A2 and B1, horizon 2100) and a mitigation policy in the form of a tax on greenhouse gas emissions (0 to 200 euros / tCO2) aimed at reducing agricultural greenhouse gas emissions. They show that both climate change scenarios lead to an increase in agricultural area at the expense of forests. Greenhouse gas mitigation policies reduce expansion of agriculture, and therefore could counteract the consequences of climate change on land use. Taking account of land use adaptations to climate change makes it possible to reduce abatement costs in the agricultural sector.

6 Conclusion

The objective of this review was to summarize the literature on econometric land use modelling and show how SA can be accounted for in these models. Despite the recent advances in econometric land-use models, several research directions remain to be explored and several issues need to be addressed concerning data, theories and empirical models.

First, there is a frequent lack of data to construct relevant explanatory variables implied by theoretical models. In particular, land rents are described in the theoretical model as among the main decision variables related to land use or land use change but are unobservable in the case of agricultural or urban use. In the case of forestry use, these rents are even more difficult to calculate. More research is needed along these lines, and especially to investigate the question of the links between land price and land rent, drawing on the work of Randall and Castle (1985) and Goodwin et al. (2003).

Second, more investigation is needed into scale issues in land-use studies. For example, most economic variables refer to administrative units rather than grids which makes it easier to estimate econometric models at the same administrative scale (such as department, municipality or small agricultural regions in the French case). However, a land-use model with aggregate spatial resolution is less relevant for assessing the local ecological effects of land uses. Ecological issues such as habitat quality or dispersion of species operate on fine scales. Ecological conditions vary considerably within each administrative unit, introducing additional uncertainty for ecological assessments.

Third, in addition to the spatial dimension, it would be interesting to incorporate the dynamic dimension explicitly in econometric land use models (Epanchin-Niell et al., 2017). Methodological advances in the specification and estimation of spatio-temporal panel models are one of the difficulties related to spatial econometrics as noted in Arbia (2011). The estimation methods developed by Ferdous and Bhat (2012); Sidharthan and Bhat (2012) seem promising as alternatives to the computing intensive Bayesian or simulation methods.

Fourth, all the models presented here assume implicitly that the spatial weight matrix is exogenous. If spatial units refer to individual landowners making land use choices, these choices might be influenced substantially by the choices of peers with whom they choose to be linked in which case the weight matrix becomes endogenous. Identification of endogenous peer effects and how to disentangle them from exogenous effects and correlated effects in networks has been studied extensively[5].The way landowners form networks and how these affect land use decisions are of considerable interest to understand the drivers of these decisions.

Finally and related to this issue, structural models should be further developed to study the links between land use and land use changes, and their effects on the environment for example on GHG emissions and biodiversity. The advantage of a structural approach is that it makes more explicit assumptions about observable and unobservable variables. The structural approach also makes it possible to unambiguously account for the endogeneity of prices and the feedbacks that determine the market equilibrium (Timmins and Schlenker, 2009). The aim is to propose a theoretical economic model which includes the farmer's decisions about crop rotations, choice of inputs (fertilizers), land allocation between agricultural and grassland uses, and herd size and composition. This would be quite challenging and would force a limited focus on a subset of these decisions (Kaminski et al., 2013).

Addressing these issues would help to improve the quality of econometric land-use models. Developing accurate models is important for policy making to allow for more accurate predictions about land use and future changes and more accurate measurement of the effects of these changes on natural resources (biodiversity, water quality, soil quality and air quality).

Acknowledgments *The authors are grateful for helpful comments from the editors and the two anonymous reviewers. Raja Chakir acknowledges the support of the Agence Nationale de la Recherche as part of the "Investments d'Avenir" Programme within STIMUL (Scenarios Towards integrating multi-scale land use tools) flagship project (LabEx BASC; ANR-11-LABX-0034) and Cland Institut de convergence (ANR-16-CONV-0003).*

References

Abildtrup, J., Garcia, S., and Kere, E. N. (2015). Land use and drinking water supply: A sample selection model with spatial dependence. *Revue d'Économie Régionale et Urbaine*, 2015/1-2:321–342.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall, London.

Alig, R. (1986). Econometric analysis of the factors influencing forest acreage trends in the southeast. *Forest Science*, 32(1):119–134.

Amin, A., Choumert-Nkolo, J., Combes, J.-L., Motel, P. C., Kéré, E. N., Ongono-Olinga, J.-G., and Schwartz, S. (2019). Neighborhood effects in the brazilian amazônia: Protected areas and deforestation. *Journal of Environmental Economics and Management*, 93:272–288.

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

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

[5] See Hsieh et al. (2019) for a recent paper on the specification and estimation of network formation and network interaction and applications of this literature to land use issues can be found in Isaac and Matous (2017) and Baird et al. (2016).

Arbia, G. (2011). A lustrum of sea: Recent research trends following the creation of the spatial econometrics association (2007-2011). *Spatial Economic Analysis*, 6(4):377–395.

Ay, J.-S., Chakir, R., and Gallo, J. L. (2017). Aggregated versus individual land-use models: Modeling spatial autocorrelation to increase predictive accuracy. *Environmental Modeling & Assessment*, 22:129–145.

Baird, J., Jollineau, M., Plummer, R., and Valenti, J. (2016). Exploring agricultural advice networks, beneficial management practices and water quality on the landscape: a geospatial social-ecological systems analysis. *Land Use Policy*, 51:236–243.

Baltagi, B. H. and Pirotte, A. (2011). Seemingly unrelated regressions with spatial error components. *Empirical Economics*, 40:5–49.

Bateman, I. J., Harwood, A. R., Mace, G. M., Watson, R. T., Abson, D. J., Andrews, B., Binner, A., Crowe, A., Day, B. H., Dugdale, S., Fezzi, C., Foden, J., Hadley, D., Haines-Young, R., Hulme, M., Kontoleon, A., Lovett, A. A., Munday, P., Pascual, U., Paterson, J., Perino, G., Sen, A., Siriwardena, G., van Soest, D., and Termansen, M. (2013). Bringing ecosystem services into economic decision-making: Land use in the united kingdom. *Science*, 341:45–50.

Bayramoglu, B., Chakir, R., and Lungarska, A. (2020). Impacts of land use and climate change on freshwater ecosystems in France. *Environmental Modeling & Assessment*, 25(2):147–172.

Bhat, C. R., Dubey, S. K., Alam, M. J. B., and Khushefati, W. H. (2015). A new spatial multiple discrete-continuous modeling approach to land use change analysis. *Journal of Regional Science*, 55(5):801–841.

Billé, A. G. and Arbia, G. (2019). Spatial limited dependent variable models: a review focused on specification, estimation, and health economics applications. *Journal of Economic Surveys*, 33(5):1531–1554.

Blackman, A., Albers, H. J., Avalos-Sartorio, B., and Murphy, L. C. (2008). Land cover in a managed forest ecosystem: Mexican shade coffee. *American Journal of Agricultural Economics*, 90(1):216–231.

Blackman, A., Avalos-Sartorio, B., and Chow, J. (2012). Land Cover Change in Agroforestry: Shade Coffee in El Salvador. *Land Economics*, 88(1):75–101.

Bockstael, N. E. (1996). Modeling economics and ecology: The importance of a spatial perspective. *American Journal of Agricultural Economics*, 78(5):1168–1180.

Brady, M. and Irwin, E. (2011). Accounting for spatial effects in economic models of land use: Recent developments and challenges ahead. *Environmental & Resource Economics*, 48(3):487–509.

Carrion-Flores, C. and Irwin, E. G. (2004). Determinants of residential land-use conversion and sprawl at the rural-urban fringe. *American Journal of Agricultural Economics*, 86(4):889–904.

Carrión-Flores, C. E., Flores-Lagunes, A., and Guci, L. (2018). An estimator for discrete-choice models with spatial lag dependence using large samples, with an application to land-use conversions. *Regional Science and Urban Economics*, 69:77–93.

Chakir, R. (2015). L'espace dans les modèles économétriques d'utilisation des sols: enjeux méthodologiques et applications empiriques. *Revue d'Economie Regionale Urbaine*, 1/2(1):59–82.

Chakir, R. and Le Gallo, J. (2013). Predicting land use allocation in france: A spatial panel data analysis. *Ecological Economics*, 92(0):114–125.

Chakir, R. and Lungarska, A. (2017). Agricultural rent in land-use models: comparison of frequently used proxies. *Spatial Economic Analysis*, 12(2-3):279–303.

Chakir, R. and Parent, O. (2009). Determinants of land use changes: A spatial multinomial probit approach. *Papers in Regional Science*, 88(2):327–344.

Chen, W., Chi, G., and Li, J. (2020). The spatial aspect of ecosystem services balance and its determinants. *Land Use Policy*, 90:104263.

Chomitz, K. M. and Gray, D. A. (1996). Roads, land use, and deforestation: A spatial model applied to Belize. *World Bank Economic Review*, 10(3):487–512.

Claassen, R. and Tegene, A. (1999). Agricultural land use choice: A discrete choice approach. *Agricultural and Resource Economics Review*, 28(1):26–36.

De Pinto, A. and Nelson, G. C. (2007). Modelling deforestation and land-use change: Sparse data environments. *Journal of Agricultural Economics*, 58(3):502–516.

De Pinto, A. and Nelson, G. C. (2009). Land use change with spatially explicit data: A dynamic approach. *Environmental and Resource Economics*, 43:209–229.

Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, 5(1):9–28.

Elhorst, J. P. (2014). *Spatial econometrics: from cross-sectional data to spatial panels*. Springer.

Epanchin-Niell, R., Kuwayama, Y., and Walls, M. (2017). Spatial-dynamic complexities of climate challenge for rural areas: Integrating resource and regional economic insights. *American Journal of Agricultural Economics*, 99(2):447–463.

Feddema, J. J., Oleson, K. W., Bonan, G. B., Mearns, L. O., Buja, L. E., Meehl, G. A., and Washington, W. M. (2005). The importance of land-cover change in simulating future climates. *Science*, 310(5754):1674–1678.

Ferdous, N. and Bhat, C. R. (2012). A spatial panel ordered-response model with application to the analysis of urban land-use development intensity patterns. *Journal of Geographical Systems*, 15:1–29.

Fezzi, C. and Bateman, I. J. (2011). Structural agricultural land use modeling for spatial agro-environmental policy analysis. *American Journal of Agricultural Economics*, 93(4):1168–1188.

Fezzi, C., Harwood, A. R., Lovett, A. A., and Bateman, I. J. (2015). The environmental impact of climate change adaptation on land use and water quality. *Nature Climate Change*, 5(3):255–260.

Fleming, M. M. (2004). Techniques for estimating spatially dependent discrete choice models. *In R. Florax and L. Anselin (eds.), Advances in Spatial Econometrics: Methodology, Tools, and Applications. Berlin: Springer-Verlag.*, pages 145–167.

Florax, R. J. G. M., Folmer, H., and Rey, S. J. (2003). Specification searches in spatial econometrics: the relevance of Hendry's methodology. *Regional Science and Urban Economics.*, 33:557–579.

Foley, J. A. (2005). Global consequences of land use. *Science*, 309(5734):570–574.

Gao, L. and Bryan, B. A. (2017). Finding pathways to national-scale land-sector sustainability. *Nature*, 544(7649):217–222.

Geweke, J., Keane, M., and Runkle, D. (1994). Alternative computational approaches to inference in the multinomial probit model. *The Review of Economics and Statistics*, 76 (4):609–632.

Goodwin, B. K., Mishra, A. K., and Ortalo-Magne, F. N. (2003). What's wrong with our models of agricultural land values? *American Journal of Agricultural Economics*, 85(3):744–752.

Grunfeld, Y. and Griliches, Z. (1960). Is aggregation necessarily bad? *The Review of Economics and Statistics*, 42(1):1–13.

Hardie, I. W. and Parks, P. J. (1997). Land use with heterogeneous land quality: An application of an area base model. *American Journal of Agricultural Economics*, 79(2):299–310.

Houghton, R. A. (2003). Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000. *Tellus B*, 55(2):378–390.

Hsieh, C.-S., Lee, L.-F., and Boucher, V. (2019). Specification and estimation of network formation and network interaction models with exponential probability distribution. *SSRN*.

IPBES (2019). Global assessment report on biodiversity and ecosystem services. Technical report, IPBES.

IPCC (2019). IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse gas fluxes in Terrestrial Ecosystems. Technical report.

Irwin, E. and Bockstael, N. (2002). Interacting agents, spatial externalities and the evolution of residential land use patterns. *Journal of Economic Geography*, 2:331–54.

Irwin, E. G., Bell, K. P., and Geoghegan, J. (2003). Modeling and managing urban growth at the rural-urban fringe: A parcel-level model of residential land use change. *Agricultural and Resource Economics Review*, 32(1):83–102.

Irwin, E. G. and Bockstael, N. E. (2004). Land use externalities, open space preservation, and urban sprawl. *Regional Science and Urban Economics*, 34(6):705–725.

Isaac, M. and Matous, P. (2017). Social network ties predict land use diversity and land use change: a case study in ghana. *Regional Environmental Change*, 17:1823–1833.

Kaminski, J., Kan, I., and Fleischer, A. (2013). A structural land-use analysis of agricultural adaptation to climate change: A proactive approach. *American Journal of Agricultural Economics*, 95(1):70–93.

Klemick, H. (2011). Shifting cultivation, forest fallow, and externalities in ecosystem services: Evidence from the eastern amazon. *Journal of Environmental Economics and Management*, 61(1):95–106.

Kline, J. D., J. D. and Alig, R. J. (1999). Does land use planning slow the conversion of forest and farm lands? *Growth and Change*, 30:3–22.

Kühn, I. (2007). Incorporating spatial autocorrelation may invert observed patterns. *Biodiversity letter*, 13:66–69.

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, 304(5677):1623–1627.

LeSage, J. and Pace, R. (2009). Introduction to spatial econometrics. *CRC Press*.

LeSage, J. P. (2000). Bayesian estimation of limited dependent variable spatial autoregressive model. *Geographical Analysis*, 32:19–35.

Lewis, Plantinga, Nelson, and Polasky (2011). The efficiency of voluntary incentive policies for preventing biodiversity loss. *Ressource and Energy Economics*, (07):1–38.

Lewis, D. J. and Plantinga, A. J. (2007). Policies for habitat fragmentation: Combining econometrics with gis-based landscape simulations. *Land Economics*, 83(19):109–127.

Li, M., Wu, J., and Deng, X. (2013). Identifying drivers of land use change in china: A spatial multinomial logit model analysis. *Land Economics*, 89:632–654.

Lichtenberg, E. (1989). Land quality, irrigation development, and cropping patterns in the northern high plains. *American Journal of Agricultural Economics*, 71(1):187–194.

Lubowski, R., Plantinga, A., and Stavins, R. (2008). What drives land-use change in the united states? a national analysis of landowner decisions. *Land Econmics*, 84(4):551–572.

Lubowski, R. N., Plantinga, A. J., and Stavins, R. N. (2006). Land-use change and carbon sinks: Econometric estimation of the carbon sequestration supply function. *Journal of Environmental Economics and Management*, 51:135–152.

Lungarska, A. and Chakir, R. (2018). Climate-induced land use change in France: Impacts of agricultural aaptation and climate change mitigation. *Ecological Economics*, 147:134–154.

Marcos-Martinez, R., Bryan, B. A., Connor, J. D., and King, D. (2017). Agricultural land-use dynamics: Assessing the relative importance of socioeconomic and biophysical drivers for more targeted policy. *Land Use Policy*, 63:53–66.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior. chap. 2 in Frontiers in Econometrics*. New York: Academic Press.

McMillen, D. and Soppelsa, M. E. (2015). A conditionally parametric probit model of microdata land use in Chicago. *Journal of Regional Science*, 55(3):391–415.

McMillen, D. P. (1989). An empirical model of urban fringe land use. *Land Economics*, 65(2):138–145.

McMillen, D. P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science*, 32(3):335–348.

Meyfroidt, P. and Lambin, E. F. (2008). The causes of the reforestation in Vietnam. *Land Use Policy*, 25(2):182–197.

Millenium Ecosystem Assessment (2005). Ecosystems and human well-being: synthesis. Technical report, Island Press, Washington, DC.

Muller, D. and Zeller, M. (2002). Land use dynamics in the central highlands of Vietnam: a spatial model combining village survey data with satellite imagery interpretation. *Agricultural Economics*, 27(3):333–354.

Munroe, D. K., Southworth, J., and Tucker, C. M. (2002). The dynamics of land-cover change in western honduras: Exploring spatial and temporal complexity. *Agricultural Economics*, 27(3):355–369.

Nelson, G. C., Harris, V., and Stone, S. W. (2001). Deforestation, land use, and property rights: Empirical evidence from darien, panama. *Land Economics*, 77 (2):187–205.

Nelson, G. C. and Hellerstein, D. (1997). Do roads cause deforestation? using satellite images in econometric analysis of land use. *American Journal of Agricultural Economics*, 79(1):80–88.

Pielke, R. A. (2005). Land use and climate change. *Science*, 310(5754):1625–1626.

Pinkse, J. and Slade, M. E. (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85(1):125–154.

Plantinga, A., Mauldin, T., and Miller, D. (1999). An econometric analysis of the costs of sequestering carbon in forests. *American Journal of Agricultural Economics*, 81:812–24.

Plantinga, A. J. (1996). The effect of agricultural policies on land use and environmental quality. *American Journal of Agricultural Economics*, 78(4):1082–1091.

Plantinga, A. J. (2015). Integrating economic land-use and biophysical models. *Annu. Rev. Resour. Econ.*, 7(1):233–249.

Plantinga, A. J. and Ahn, S. (2002). Efficient policies for environmental protection: An econometric analysis of incentives for land conversion and retention. *Journal of Agricultural and Resource Economics*, 27(1):128–145.

Polasky, S., Nelson, E., Camm, J., Csuti, B., Fackler, P., Lonsdorf, E., Montgomery, C., White, D., Arthur, J., Garber-Yonts, B., Haight, R., Kagan, J., Starfield, A., and Tobalske, C. (2008). Where to put things? spatial land management to sustain biodiversity and economic returns. *Biological Conservation*, 141:1505–1524.

Randall, A. and Castle, E. N. (1985). Land resources and land markets. In Kneese, A. V. and Sweeney, J. L., editors, *Handbook of Natural Resource and Energy Economics*, volume 2, chapter 13, pages 571–620. Elsevier.

Record, S., Fitzpatrick, M. C., Finley, A. O., Veloz, S., and Ellison, A. M. (2013). Should species distribution models account for spatial autocorrelation? a test of model projections across eight millennia of climate change. *Global Ecology and Biogeography*, 22:760–771.

Ricardo, D. (1817). The principles of political economy and taxation. *London: John Murray, Albemarle-Street,*.

Sidharthan, R. and Bhat, C. R. (2012). Incorporating spatial dynamics and temporal dependency in land use change models. *Geographical Analysis*, 44(4):321–349.

Smirnov, O. A. (2010). Modeling spatial discrete choice. *Regional Science and Urban Economics*, 40:292–298.

Smith, T. E. and LeSage, J. P. (2004). A bayesian probit model with spatial dependencies. *Advances in Econometrics, Volume 18.*

Stavins, R. N. and Jaffe, A. B. (1990). Unintended impacts of public investments on private decisions: The depletion of forested wetlands. *American Economic Review*, 80(3):337–352.

Timmins, C. and Schlenker, W. (2009). Reduced-form versus structural modeling in environmental and resource economics. *Annual Review of Resource Economics*, 1(1):351–380.

Verburg, P. H., Mertz, O., Erb, K.-H., Haberl, H., and Wu, W. (2013). Land system change and food security: towards multi-scale land system solutions. *Current Opinion in Environmental Sustainability*, 5(5):494–502.

von Thunen, J. H. (1875). Der isolierte staat in beziehung auf landwirthschaft und nationalokonomie; hrsg. von h. schumacher-zarchlin.

Wang, X. and Kockelman, K. M. (2009). Application of the dynamic spatial ordered probit model: Patterns of land development change in Austin, Texas. *Papers in Regional Science*, 88(2):345–365.

Wu, J. and Adams, R.-M. (2002). Micro versus macro acreage response models: Does site-specific information matter? *Journal-of-Agricultural-and-Resource-Economics.*, 27(1):40–60.

Wu, J. and Segerson, K. (1995). The impact of policies and land characteristics on potential groundwater pollution in wisconsin. *American Journal of Agricultural Economics*, 77(4):1033–1047.

Zellner, A. and Lee, T. (1965). Joint estimation of relationships involving discrete random variables. *Econometrica*, 33:382–94.