Individual vs. aggregate models of land use changes: Using spatial econometrics to improve predictive accuracy?

Jean-Sauveur AY[*,‡] Raja CHAKIR[†] Julie LE GALLO[‡]

March 28, 2014

Abstract

It is a widespread practice to estimate land use models at some aggregated scales but the consequences of such aggregations are rarely evaluated. This paper proposes an evaluation in terms of predictive accuracy, based on estimating a broad spectrum of individual and aggregated econometric models on the same dataset. Exploiting a detailed parcel-level dataset, we perform both short and long run predictions and compare them at the same 12×12 km aggregate scale of interest. In particular, we argue that data aggregation allows the application of spatial econometric tools. We show that modeling spatial autocorrelation can compensate for loss of information due to aggregation and, with well-designed predictors, can even outperform individual models. We provide a detailed analysis of the available predictors in the context of spatial econometrics and show how to extend them in a context of out-of-sample and counterfactual predictions. However, for predictions from counterfactual economic scenarios, aggregate models do not perform as well.

Keywords: Land use models, spatial econometrics, predictive accuracy, aggregate and individual data.

JEL Classifications: Q15, Q24, R1, C21.

[*]CNRS, UMR 7204 CESCO, 55 rue Buffon, 75005 PARIS, France. jsay.site@gmail.com

[†]INRA, UMR 210 Économie Publique, 16 rue Claude Bernard, 75005 PARIS, France.

[‡]Université de Franche-Comté, CRESE, 45D, Avenue de l'Observatoire, 25030 BESANÇON, France.

1 Introduction

Land Use Changes (LUC) produce significant economic and environmental outcomes with important implications for a wide variety of policy issues including food security, wildlife conservation, housing supply, carbon sequestration (Turner et al., 2007; Bateman et al., 2013). Given these large impacts and the potential for huge LUC in the future, prospective analysis requires a thorough understanding of how policies and economics affect land use patterns (Nelson et al., 2008; Lewis, 2010; Wu and Duke, 2014). In particular, accurate predictions about future land use need to integrate the complex spatiotemporal structure of human choices in relation to natural processes.

Thus, there is a real need for econometric models of land use for at least three reasons (Turner et al., 2007; Wu and Duke, 2014): first, to identify the key drivers of LUC; second, to predict expected LUC in response to projected changes in economic or climatic conditions; and third, to study the effects of public policies (taxes, subsidies) on LUC. Depending on the objective of the study, the choice of the appropriate specification for an econometric model of land use necessarily requires tradeoffs between data quality and availability, and computation costs.

With respect to data availability, land-use models can be classified into two general groups based on their use of aggregate or individual data.[1] Until recently, and due to the scarcity and cost of access to individual level data, most studies have been based on aggregated data for a region, a country, or other geographic scales. Most studies based on aggregated data specify the shares of land allocated to different land uses as a function of the explanatory variables, and use the logistic functional form (Plantinga, 1996; Plantinga et al., 1999; Chakir and Le Gallo, 2013). Land-use studies based on individual data are more recent and involve discrete choice models to explain the choice between two categories of land use (binomial probit or logit models) or several categories of land use (multinomial logit, nested logit or multinomial probit, Lewis and Plantinga, 2007; Lubowski et al., 2008; Chakir and Parent, 2009).

There are important trade-offs between aggregate and individual models although they may be quite complementary and can provide different insights into the determinants of land use changes and their effects on the environment. The choice of the scale is often dictated by the data available and also the objective of the study. Thus, if the goal is to make predictions about land uses across one or several countries with heterogenous raw data, an aggregate model is often preferable datasets. However, if the objective is to study more locally the effects of land use on biodiversity or water quality for instance, an individual data model is more relevant. Most socioeconomic variables are collected for administrative units rather than grid cells, making it more straightforward to apply the econometric models at the same administrative scale (Plantinga, 1996; Hardie and Parks, 1997). Furthermore, some economic data only make sense at the aggregate scale (e.g., commodity prices determined in national or global markets). However, a land-use-change model at coarse spatial resolution has limited value for ecological assessments, given that most ecological processes of interest, such as habitat suitability, dispersal, and spread of invasive species, operate at finer scales. Also, administrative boundaries and ecological boundaries rarely coincide which means that ecological conditions tend to vary substantially within each administrative unit, introducing further uncertainty into ecological assessments.

[1] See Irwin and Geoghegan (2001); Plantinga and Irwin (2006); Irwin (2010) for reviews of empirical methods of land-use modeling.

The issue of predictive accuracy of individual versus aggregate models has received little attention, especially in the literature on LUC. In a seminal paper, Grunfeld and Griliches (1960) have examined the relative power of micro and macro models for explaining the variability of the aggregate dependent variable and found that the aggregate equation may explain the aggregate data better than a combination of micro equations. Wu and Adams (2002) examine the issue in the context of predicting land allocation. They show that, even for linear prediction models, the decision over use of a micro or macro models to make aggregate predictions cannot be generally resolved by *a priori* reasoning.

At the center of the aggregation process, land use is inherently related to space. However, incorporating spatial dimension into land-use models presents several challenges related to econometric estimation, hypothesis testing and prediction (Anselin, 2007; Brady and Irwin, 2011). This is more challenging in the case of individual land-use models since the introduction of spatial dependence would render discrete choice models analytically intractable, and would require the use of simulation or Bayesian techniques (Fleming and Mae, 2004).[2] Given the size of our individual sample (around 500,000 observations) and our aim to study in what proportions the predictions from aggregated models can be improved by taking account of the spatial dimension in the econometric model, the introduction of spatial autocorrelation in a discrete choice model is beyond the scope.

The contribution of our paper is to show how the introduction of a spatial dimension in aggregated land use models enables better predictions than individual, non spatial models with higher numbers of observations. Our analysis is based on a detailed dataset of 514,074 individual plots of land use for the whole of continental France (about 675,000 km^2) observed annually over the period 1993–2003. These data are used to estimate individual plot-level discrete choice models of LUC and aggregate models of land use shares at the 12×12 km grid scale. We consider four categories of land use: arable, forest, pasture and urban. We estimate and compare different model specifications: two models at the individual scale and eight at the aggregate scale. Namely, we fit individual linear probability models, individual discrete multinomial logit models, aggregated logit-linearized share models estimated by ordinary least squares, aggregated logit-linearized share models estimated by geoadditive models, aggregate fractional model with multinomial and Dirichlet distributions and aggregated logit-linearized share models in a variety of spatial econometric specifications. We use the Root Mean Square Error (RMSE) to compare the prediction accuracy of these different land use models with different specifications and the different predictors proposed by Kelejian and Prucha (2007).

In section 2, we present the econometric models and the different formulae available to predict LUCs. Section 3 presents the data and section 4 presents the results for the estimated parameters, and the prediction accuracy of different models. The last sectionis the summary and conclusions.

[2]Other estimation procedures have also been proposed in the literature: EM method (McMillen, 1992), the generalized method of moments (Pinkse and Slade, 1998), the method of maximum pseudo-likelihood (Smirnov, 2010) and the method composite maximum likelihood Ferdous and Bhat (2013); Sidharthan and Bhat (2012). For a detailed review of the inclusion of spatial autocorrelation in discrete choice models see Fleming and Mae (2004) and Smirnov (2010).

2 Econometric models

2.1 Individual land use models

Following the literature (Stavins and Jaffe, 1990; Plantinga, 1996; Lubowski et al., 2006), we consider a risk-neutral landowner facing the choice of allocating a parcel of land of uniform quality among a set of alternative uses. We assume that landowners choose uses in order to maximize the present discounted value of the stream of expected net benefits from the land, and that landowners base their expectations of future land use profits on current and historic values of the relevant variables. Given these simplifying assumptions, the decision rule that emerges from the related dynamic optimization problem is to choose the use with the highest expected one-period return at time t, minus the current one-period expected opportunity cost for a specific use type.

In particular, a stylized landowner i chooses the use ℓ^*_{it} on a plot if this provides the highest utility from all uses that are possible. The following formula:

$$\ell^*_{it} = \arg\max_{\ell} \{u_{i\ell t}\} \tag{1}$$

Given that we do not observe data on all variables that might affect the landowner's returns to the different uses, the landowner's profit is written as a function to include both observed and unobserved components. Using a general random utility expression, the one-period expected net profit (utility) to the landowner on parcel i from use ℓ at time t as:

$$u_{i\ell t} = \beta x_{i\ell t} + \varepsilon_{i\ell t} \tag{2}$$

where $x_{i\ell t}$ is a vector of observed variables, β are parameters on each of these variables allowed to vary over time and $\varepsilon_{i\ell t}$ is a random error term.

As Train (2009) states, the two major implications of this framework – utilities are ordinal and only differences in utilities matter – are in accordance with the economic theory. Therefore, this discrete choice framework is fairly general, the strongest restrictions come from the parametrization of the utility functions necessary for their application to the data. On this latter point, we follow the empirical literature (Lubowski et al., 2008; Lewis, 2010; Ay et al., 2014) by considering the one-period vector of the returns $\mathbf{r}_{it}$ from the different land uses as the main explanatory variables driving utilities and, consequently, LUC. These monetary returns are completed by time-constant biophysical characteristics of land (noted $\mathbf{b}_i$ that represents land quality, topography, climate[3]) which are known to influence the returns from land. These variables are added separately from the economic returns because they are generally more precisely observed, for example, through digital elevation models. Pecuniary and non pecuniary conversion costs are also identified in the literature as important drivers of LUC, so we introduce lagged land uses $\mathbf{d}_{it'}, t' < t$ in the general specification of utility functions.

$$u_{i\ell t} = \mathbf{d}_{it'}^{\top}\boldsymbol{\gamma}_{\ell}^{D} + \mathbf{r}_{it}^{\top}\boldsymbol{\gamma}_{\ell}^{R} + \mathbf{b}_{i}^{\top}\boldsymbol{\gamma}_{\ell}^{B} + \varepsilon_{i\ell t}. \tag{3}$$

This specification restricts the actual utility to be free of uncertainty and irresistibility effects,[4] to have identical time preference and anticipation of the future in the population,

[3]We consider climate as constant despite the strong evidences provided by the Intergovernmental Panel on Climate Change. This is because it is not of special interest here, otherwise see Ay et al. (2014).

[4]See Schatzki, 2003 for an attempt to introduce them.

and to neglect non-pecuniary returns, liquidity constraints, interdependencies, spatial land constraints and transactions costs. Landowners are considered risk-neutral but intertemporal consistency is assured as this is important for pluri-annual land uses such as forestry and urban uses. In terms of specification, each element $d_{i\ell t'}$ in $\mathbf{d}_{it'}$ is 1 if the plot i is in use ℓ at t' and 0 otherwise. Because these columns sum to 1 in row, a modality-specific variable is dropped. The vector $\mathbf{r}_{it}$ contains in the row the L returns from different land-use. The vector $\mathbf{b}_i$ binds the K biophysical variables that are described in greater details in the data section.

Because all the sources of landowner's utility cannot be observed, an error term $\varepsilon_{i\ell t}$ is included in (3). The stochastic dimension of the model (and its predictions) is related only to these unobserved components of utilities and their associated densities. McFadden (1974) identifies three standard hypothesis about error terms that allow obtaining a multinomial logit model to be derived: independence, homoskedasticity and extreme value distribution (i.e., Gumbel). On the basis of these hypotheses, one can show that the probabilities of having the land use ℓ on i at t have simple closed forms, which correspond to the logit transformation of the deterministic part of the utility, $\overline{u}_{i\ell t} \equiv u_{i\ell t} - \varepsilon_{i\ell t}$:

$$p_{i\ell t} = \frac{\exp(\overline{u}_{i\ell t})}{\sum_{l=1}^{L} \exp(\overline{u}_{ilt})} = f\left(\mathbf{d}_{it'}, \mathbf{r}_{it}, \mathbf{b}_i; \boldsymbol{\Gamma}_\ell\right). \tag{4}$$

2.2 Aggregate land use models

There is an important literature on econometric aggregate land use models: Lichtenberg (1989), Stavins and Jaffe (1990), Wu and Segerson (1995) and Plantinga (1996), and Miller and Plantinga (1999) are the most significant papers. The underlying microeconomic theory is identical to that in the previous section, but individual choices are aggregated typically to estimate models of land use shares. This process of aggregation is generally considered as a loss of information through a drastic decrease in the number of observations. Usually, land use shares are specified as logistic functions (Wu and Segerson, 1995; Chakir and Le Gallo, 2013) which have the advantage of being empirically tractable thanks to the "logit-linear transformation" (Zellner and Lee, 1965). The observed shares of land use ℓ at the aggregate grid level g ($g = 1, ..., G$) in t is then expressed as ($\forall \ell = 1, \ldots, L$):

$$S_{g\ell t} = \frac{\exp\left(\mathbf{D}_{gt'}^\top \boldsymbol{\beta}_\ell^D + \mathbf{R}_{gt}^\top \boldsymbol{\beta}_\ell^R + \mathbf{B}_g^\top \boldsymbol{\beta}_\ell^B\right)}{\sum_{l=1}^{L} \exp\left(\mathbf{D}_{gt'}^\top \boldsymbol{\beta}_l^D + \mathbf{R}_{gt}^\top \boldsymbol{\beta}_l^R + \mathbf{B}_g^\top \boldsymbol{\beta}_l^B\right)}. \tag{5}$$

The meanings of these variables are the same as in the previous subsection, and capital letters here represent aggregate values. Aggregating the dummies $\mathbf{d}_{it'}$ consists of including land use shares in $t' < t$ as explanatory variables, still with a referecne modality. Through identification with (2) and (3), these aggregate share models can directly be estimated as fractional models, using pseudo maximum likelihood techniques (Gourieroux et al., 1984; Papke and Wooldridge, 1993, 2008; Mullahy, 2010). In parallel with these original techniques, we follow the current practices in noting that the natural logarithm of each observed share normalized by a reference share (here S_{glt}) is approximately equal to:[5]

[5] We choose the reference modality as the land use with the less number of shares equal to zero. Because it is still possible to have some zeros at the denominator, we add $\epsilon = .0001$ at the numerator and the denominator of (6). This is clearly an inconvenient but its effect will be evaluated by comparing with

$$\log(S_{g\ell t}/S_{glt}) \approx \mathbf{D}_{gt'}^{\top}\boldsymbol{\beta}_{\ell}^{D} + \mathbf{R}_{gt}^{\top}\boldsymbol{\beta}_{\ell}^{R} + \mathbf{B}_{g}^{\top}\boldsymbol{\beta}_{\ell}^{B} + \xi_{g\ell t} \;\; \forall \ell \neq l. \tag{6}$$

With L land use categories, the system has $L-1$ equations. The elements $\mathbf{D}_{gt'}$, $\mathbf{R}_{gt}$ and $\mathbf{B}_g$ in Equation 6 do not have an index ℓ since we use the same explanatory variables in all L equations. A Seemingly Unrelated Regressions approach could also be adopted (Considine and Mount, 1984). However, Chakir and Le Gallo (2013) show that estimating inter-equation correlations doesn't improve the predictive accuracy of the model. Therefore, to simplify the results, we skip this aspect in this paper.

Finally, space can be easily introduced in these models by including a smoothed function of the geographical coordinates of the grids' centroids in $\mathbf{B}_g$. This leads to semi-parametric Generalized Additive Models (GAM), estimated by penalized likelihood techniques (Hastie and Tibshirani, 1986; Wood, 2004). Because in this case, spatial autocorrelation is not modeled explicitly, we do not consider such models as being spatial econometric models but we include them in our comparative set.

2.3 Modeling spatial autocorrelation in aggregate models

There are various possible sources of spatial autocorrelation in LUC models. First, it might be the result of strategic interactions between neighboring individuals. Secondly, it might arise from measurement errors that spill across boundaries or be due to scale mismatch and the inherent need to integrate data for different scales. Third, it can arise from unobservable latent variables that are spatially correlated. The first explanation is particularly relevant for individual data, while the second affects models based on aggregated data. An econometric model that fails to include all the relevant spatial variables is adversely affected by its omissions. As stated in Chakir and Le Gallo (2013), these omitted variables could account for any specific bioclimatic regional characteristics (e.g. dairy production tends to take place in rainy areas while cereal production is located on plains) that are correlated over space. Moreover, regional agricultural systems are the outcome of spatially shapped historical and institutional determinants (e.g., the location of intensive livestock production is linked partly to infrastructure such as harbor facilities for importing soybeans, while vegetable production tends to be close to consumption centers).

The spatial econometric literature is extensive (Cliff and Ord, 1981; Anselin, 1988; LeSage and Pace, 2009; Anselin, 2010) and provides a number of ways to deal with spatial autocorrelation. Nevertheless, introducing spatial dependence in discrete choice models is still problematic econometrically, especially with high numbers of observations (Fleming and Mae, 2004; Smirnov, 2010). In particular, an important consequence of introducing spatial dimension in discrete choice models is the complex covariance structure due to heteroskedasticity. Moreover, it implies high dimension integrals in order to compute the likelihood function (Anselin, 2002). To avoid such complications associated with spatial autocorrelation in discrete choice models, in this paper we focus on introducing this spatial autocorrelation in the aggregate land use models only.

In the context of aggregated land use models, note $\widetilde{S}_{\ell t}^{l} \equiv \log(S_{\ell t}/S_{glt})$, $\mathbf{X}_{tt'} \equiv \left[\mathbf{D}_{t'} \mid \mathbf{R}_t \mid \mathbf{B}\right]$ and $\boldsymbol{\beta}_{\ell} \equiv \left[\boldsymbol{\beta}_{\ell}^{D} \mid \boldsymbol{\beta}_{\ell}^{R} \mid \boldsymbol{\beta}_{\ell}^{B}\right]$, where all vectors follow the same notations as in Equation 6 but stacked for the G grids, the most general spatial econometric model (MSAC) is written as:

other models. We consider this as a necessity of linearized logistic models, often used in the literature.

$$\widetilde{S}^l_{\ell t} \approx \rho_\ell \mathbf{W} \widetilde{S}^l_{\ell t} + \mathbf{X}_{tt'}\beta_\ell + \theta_\ell \mathbf{W}\mathbf{X}_{tt'} + \xi_{\ell t} \quad \text{with} \quad \xi_{\ell t} = \lambda_\ell \mathbf{W} \xi_{\ell t} + \eta_{\ell t} \tag{7}$$

still with $\forall \ell \neq l$ because l is the reference modality. The major modifications compared to previous aggregate models come from the inclusion of the $G \times G$ spatial weight matrix $\mathbf{W}$, which summarizes the connectivity structure of the observations. Once multiplied to a variable and if it is row-standardized, it contains the weighted average of the values of the neighbors of each observations. To avoid endogeneity problems, this matrix is often based on purely geographical considerations, such as borders or distances between observations. This model is sufficiently general that, for all land use $\ell \neq l$, the SARAR(1,1) model can be recovered with $\theta = 0$ (Kelejian and Prucha, 2007) (also called SAC by Bivand, 2002; Bivand et al., 2013), the spatial error model (SEM) can be recovered with $\rho = \theta = 0$, the spatial X model (SXM) with $\rho = 0$, the spatial autoregressive (SAR) model with $\theta = \lambda = 0$; and the spatial Durbin model (SDM) model can be recovered when $\lambda = 0$. The SDM, is the appropriate specification in the case of omitted variables (LeSage and Pace, 2009).

2.4 Performing predictions

2.4.1 On individual models

For the individual MNL models, the direct predictions (without changing exogenous variables) consist, for each plot i, of a fitted probability vector $\widehat{\mathbf{p}}_{it}$ of being in each use at t. Assuming $L = 4$ and that each observation counts for 100 ha (in anticipation of our application), the predicted probabilities can easily be converted into aggregate LUC. For example, consider a plot i which counts for 100 ha of annual crop in period t' and has a predicted probability vector for period t of $\widehat{\mathbf{p}}_{it} = (0.8, 0.15, 0.04, 0.01)$. This means that 80 ha are predicted to retain their land use, 15 ha will be converted to pasture, 4 ha to forest and 1 ha to urban. The aggregation of probabilities in terms of aggregate acreages (and aggregate shares) is operated by multiplying the probabilities by 100 and summing the results at the aggregate scale of interest. With this multinomial approach, the predicted acreages of each use are always positive and assured to sum to the national available land base. To evaluate the effect of these desirable prediction properties, we also estimate some linear probability models on individual data that do not take account of the discrete nature of land use choices but are less computationally intensive. Within this framework, counterfactual out of sample scenarios and policies are easily simulated. As it will be shown in the application, changing the values of $\mathbf{r}_{it''}$, $t'' > t$ allows to evaluate to corresponding raw LUC between t and t'' (Lubowski et al., 2008; Lewis, 2010).

2.4.2 On aggregate models

Obtaining predictions from aspatial aggregate models is immediate, by resolving the system described by Equation 6. For spatial econometric models, we perform predictions based on the work of Kelejian and Prucha (2007). They consider 5 predictors, from the more constrained in terms of information sets to the less constrained. They are written for each aggregate grid g, the first is:

$$\widehat{S}^{(1)}_{g\ell t} = (I - \rho_\ell \mathbf{W})^{-1}_{g.} \widetilde{\mathbf{X}}_{tt'}\phi_\ell \quad \text{with} \quad \phi_\ell = [\beta_\ell \mid \theta_\ell] \quad \text{and} \quad \widetilde{\mathbf{X}}_{tt'} = \left[X_{tt'} \mid WX_{tt'}\right] \tag{8}$$

This predictor KP1 can be computed with minimal information and for all prediction types: in-sample and out of sample. The other 4 predictors are written as:

$$\widehat{S}^{(2)}_{g\ell t} = \rho_\ell w_{g.}\tilde{S}_{lt} + \widetilde{\mathbf{X}}_{g.tt'}\phi_\ell + \frac{\text{cov}(\xi_{glt}, w_{g.}\tilde{S}_{g\ell t})}{\text{var}(w_{g.}\tilde{S}_{lt})}[w_{g.}\tilde{S}_{lt} - E(w_{g.}\tilde{S}_{lt})] \tag{9}$$

$$\widehat{S}^{(3)}_{g\ell t} = \rho_\ell w_{g.}\tilde{S}_{\ell t} + \widetilde{\mathbf{X}}_{g.tt'}\phi_\ell + \text{cov}(\xi_{glt}, \tilde{S}_{-g,\ell t})[\Omega(S_{-g,\ell t})]^{-1}[\tilde{S}_{-g,lt} - E(\tilde{S}_{-g,\ell t})] \tag{10}$$

$$\widehat{S}^{(4)}_{g\ell t} = \rho_\ell w_{g.}\tilde{S}_{\ell t} + \widetilde{\mathbf{X}}_{g.tt'}\phi_\ell \tag{11}$$

$$\widehat{S}^{(5)}_{g\ell t} = \widetilde{\mathbf{X}}_{gtt'}\phi_\ell + \lambda_\ell w_{g.}(\tilde{S}_{\ell t} - \widetilde{\mathbf{X}}_{tt'}\phi_\ell) \tag{12}$$

where $(I - \lambda_\ell \mathbf{W})^{-1}_{g.}$, $w_{g.}$ and $\mathbf{X}_{g.tt'}$ denote respectively the g^{th} row of $(I - \rho_\ell \mathbf{W})^{-1}$, W and $X_{tt'}$; $\tilde{S}_{-g,\ell t}$ represent the $G-1$ observations on $\tilde{S}_{\ell t}$ and:

$$\begin{aligned}
\text{cov}(\xi_{glt}, w_{g.}\tilde{S}_{glt}) &= \sigma^2_\xi \sigma^u_{g.}(I - \rho_\ell \mathbf{W}')^{-1} w'_{g.} \quad ; & E(\tilde{S}_{-g,lt}) &= S_{-g}(I - \rho_\ell \mathbf{W})^{-1}\widetilde{\mathbf{X}}_{tt'}\phi_\ell \\
E(w_{g.}\tilde{S}_{lt}) &= w_{g.}(I - \rho_\ell \mathbf{W}')^{-1}\widetilde{\mathbf{X}}_{lt}\phi_\ell \quad ; & \text{cov}(\xi_{glt}, \tilde{S}_{-g,lt}) &= \sigma^2_\xi \sigma^u_{g.}(I - \rho_\ell \mathbf{W}')^{-1}_{i.} S'_{-1} \\
\Omega(S_{-g,\ell t}) &= \sigma^2_\xi \tilde{S}_{-g,\ell t} \sum\nolimits^{\tilde{S}} \tilde{S}'_{-g,\ell t} \quad ; & \text{var}(w_{g.}\tilde{S}_{\ell t}) &= \sigma^2_\xi w_{g.} \sum\nolimits^{\tilde{S}} w'_{g.}
\end{aligned}$$

with $\sigma^u_{g.}$ being the g^{th} row of $\sum^u = (I-\lambda_\ell \mathbf{W})^{-1}(I-\lambda_\ell \mathbf{W}')^{-1}$ and $\sum^{\tilde{S}} = (I-\rho_\ell \mathbf{W})^{-1}\sum^u(I-\rho_\ell \mathbf{W}')^{-1}$.

All logit-transformed aggregate models have the desirable properties of positive acreages predictions and summing to one. Predictors KP2 and KP3 are unbiased, whereas predictors KP4 and KP5 are biased but are easier to compute and are more used in the literature. Note that while these last 4 predictors can be used in an in-sample framework, it is not possible to compute them when performing out of sample predictions since they include in their formulations the spatial lag of the dependent variable. In this case, we suggest some heuristic bypass approaches to approximate S_g, which are presented in details in the empirical application.

For each econometric model presented above, lagged land uses are included in the sets of explanatory variables. This practice is not frequent in the literature because of data availability but also endogeneity problems for studies based on several periods. Therefore, we extend the range of our comparisons by estimating similar models without lagged land use. Comparing the predictions from models with and without lagged land uses is not direct, as they correspond to different temporal horizons. In reference to the times series litterature (Box et al., 2013), when lagged endogenous variables are included, we interpret the outcomes as short run predictions. Inversely, models without lagged land use provide long term predictions. In a stationary world, the fact that long run predictions are the limit of the short run is easily demonstrable (see also LeSage and Pace, 2009 in the context of spatial model).

3 Data

3.1 Land use data

Data on land use are extracted from the TERUTI survey (AGRESTE, 2004), which is carried out every year by the statistical services of the French Ministry of Agriculture. It collects data on land use through the whole continental territory of France. It counts 514,074 points continuously geo-referenced and surveyed each year from 1992 to 2003. The survey uses a systematic area frame sampling with a two-stage sampling design. In the first stage, the total land area of France is divided into 12×12 km grids. For each of the 4,700 grids there are four aerial photographs which cover 3.5 km^2 each. In the second stage, on each photograph, a 6×6 grid determines 36 points (the area of each point is equal to 100 ha). On the basis of the detailed classification of land uses (81 items), we attribute to each plot a use among four more aggregate items:[6] arable crops (wheat, corn, sunflowers and perennial crop), pastures (a rather large definition: grassland, rangelands, productive fallows, moor), forests (both productive and recreational, including plantations and hedgerows) and urban areas (cities and exurban housing, and also roads, highways, airports, etc.) The following Table 1 presents the raw transitions 1993–2003.

Table 1: Raw land use transitions in %, TERUTI 1993–2003

$N = 514,074$	PASTURE	ARABLE	FOREST	URBAN	Sum
PASTURE	26.53	4.2	1.26	0.69	32.68
ARABLE	3.79	27.61	0.17	0.37	31.94
FOREST	0.56	0.13	29.03	0.15	29.87
URBAN	0.27	0.09	0.07	5.08	5.51
Sum	31.15	32.03	30.53	6.29	100

Table 1 shows that, in 2003, arable crops, pastures and forests each represented almost 30% of the continental France. It also shows that between 1992 and 2003, the area to pasture declined by almost 5%, while arable, forest and urban uses increased by 2%, 3% and 14% respectively. Worldwide (and in all other land use studies), land use presents a significant temporal inertia, which comes from conversion costs but also intertemporal decisions, land owner specializations, legislative constraints, etc.

As mentioned in the footnote 5, the presence of zeros in the denominator of the logit transformation is a limit of the logit transformation for aggregate modeling that is overridden by adding ϵ both in the numerator and the denominator. As Figure 5 and Figure 6 of the Appendix A.10 show, the logit transformation produces some mass probabilities around the value -7 but the distribution of the outcome is undoubtedly closer to that of a normal distribution than raw land use shares were.

3.2 Explanatory variables

The theoretical literature on land use suggests that the explanatory variables introduced in models include the net return to each land use and the distribution of land quality.

[6] We dropped from the data observations that concern salt marshes, ponds, lakes, rivers, marshes, wetlands, glaciers, eternal snow, wastelands, and moors, which accounted for about 7% of observations.

In this paper, these variables include: economic returns for each land use (computed for arable crops and pastures from land prices according to the Ricardian formula with an interest rate of 2%, similar to Ay et al., 2014) and population densities used as proxies for the economic returns from urban use. Finally, we include some biophysical attributes: slope, altitude, water holding capacity (WHC), and climate. The following Table 2 displays summary statistics for these variables aggregated at the grid scale without loss of generality.

Table 2: Summary statistics for explanatory variables

N=3,767	DESCRIPTION	MEAN	STD	MIN	MAX
Arable returns03	returns from arable crop (2003 euro)	183.500	89.178	0.000	1,210.599
Pasture returns03	returns from pasture (2003 euro)	126.083	74.393	0.000	619.683
Forest returns03	returns from forest (2003 euro)	88.914	131.145	0.000	792.223
POP03	urban pop density (hab/km^2)	3,109.910	17,929.310	51.639	819,298.800
Elevation	elevation (meters)	336.230	399.984	0.000	2,772.500
Slope	slope (degrees)	3.803	4.798	0.000	31.731
WHC	water holding capacity (mm)	131.031	49.295	13.000	343.193
Soil depth	soil depth (cm)	80.214	22.603	10.000	131.000
Precipitations	precipitations (mm/yrs)	871.268	200.217	359.672	1,988.323
Temperature	temperatures (degrees celsius)	11.528	1.947	-0.971	16.192
Humidity	relative humidity (%)	932.614	52.380	730.042	1,026.848
Radiation	solar radiation (J)	996.824	48.878	796.467	1,099.190

Data on land prices are available from the statistical services of the French Ministry of Agriculture. Yearly prices 1990–2005 are available for arable crops and pastures. For the other two land uses considered – forest and urban – the approximations of economic returns are computed differently and at different geographic scales. For the expected net returns from forest, we use data on wood raw production (in m^3), total forest area (in ha) and wood prices (in current euros per ha). We compute the expected returns from forest use by multiplying the aggregate production by its unitary price and dividing the result by the total forest area in each *département*. Urban returns are approximated by population densities for urban land use at the fine scale of the municipalities, based on the national census of the French population.

4 Results

4.1 Parameter estimates

Detailed results of different model specifications are provided in the appendix. The explanatory variables are scaled to obtain standardized parameters, and we report only here the results of the models for in-sample predictions, that is 1993–2003. Because of their proximity to the displayed models, the raw results from the Dirichlet estimations (close to the fractional FRA), the linear probabilities (close to the individual MNL), and the SAC and MSAC (close to SAR and SDM) are not reported but are available upon request.

On the one hand, we performed the estimation of individual MNL models using `nnet 7.3` on the `R` software. A critical aspect of such models is that the unobserved factors have to be uncorrelated over alternatives and periods, as well as having the same variance for all alternatives and periods. These assumptions, used to provide a convenient form for the choice probability, are not found to be restrictive (homoskedasticity cannot be rejected by a score test, p-value$= 0.283$). Moreover, these assumptions are associated with the classical restriction of Independence of Irrelevant Alternatives for which Hausman-McFadden specification tests were performed, with mixed evidence. The independence is not rejected for two uses: pasture and urban (p-values are respectively 0.001, 0.005 and 0.036) but is rejected for arable and forest at 5%. This means that the former choices can be dropped from the choice set without significant modification to the model (i.e., they are robust to the IIA restriction), a property that does not apply to the latter two choices. In the literature, use of nested multinomial logit is found not to change the results (Lubowski et al., 2008; Li et al., 2013).

On the other hand, we estimated the spatial econometric models using maximum likelihood through the `R` package `spdep`. Because we are interested in predictions, we do not run a detailed specification search, based on the specific-to-general or the general-to-specific approaches (see Florax et al., 2003, Elhorst, 2010 or Le Gallo, 2013 for reviews of these spatial specification searches). Instead, we estimate the full set of spatial models described in section 2.3 since spatial autocorrelation could arise from several sources. The summary measure of impacts, direct, indirect and total as defined in LeSage and Pace (2009), are not reported here but are available upon request. Globally, it appears that incorporating lagged land use (i.e. short-run models) strongly decreases the significance of the coefficients associated to the other variables or even renders then insignificant or with a counter-intuitive sign. However, as shown in Figure 9 and Figure 10 of Appendix A.12 that display the Moran scatterplots of regression residuals in the OLS and GAM models, it also allows to decrease or render spatial error autocorrelation insignificant. The spatial smoothed functions estimated by the GAMs are displayed in Figure 7 and Figure 8 of Appendix A.11. For the long run models without temporal lag, the regional specialisations of land use apprear clearly: arable crops for the south-east, forests for the south-west and urban areas around Paris, at the center-north. These contextual effects are intuitive and are still present (even if less marked) for the models with temporal lag.

Table 3 and Table 4 display the value of the spatial coefficients ρ and λ for respectively the long run and the short run models. Evidence of spatial autocorrelation is strong in all specifications, whether for the spatial error component or the spatial lag component. When a spatial lag of the dependent variable (SEM, SXM) and the spatial error coefficient models (SEM, SXM) are introduced separately, spatial autocorrelation appears to be positive but to different extents depending on the land use: land use shares in forest is the most spatially autocorrelated across specifications while urban use is the least spatially autocorrelated. In most general models (SAC, MSAC), some multicollinearity appears, with an instability of parameter according to the specification. In effect, for each model, the spatial coefficients have opposite signs indicating spurious compensation of the spatial effects between errors and lag. Finally, when comparing the long run and short run models (Table 3 *versus* Table 4), the extent of spatial autocorrelation is much less pronounced in the latter, and although the spatial lag coefficient remains positive in all specifications, only the spatial error coefficient is negative in most of the general SAC and MSAC specifications.

Table 3: **Spatial coefficients for long run models (i.e., without temporal lags)**

	Spatial Error component: λ			Spatial Lag component: ρ		
	AR	FO	UR	AR	FO	UR
SEM	0.6449**	0.7349**	0.4991**			
	(0.0183)	(0.0248)	(0.0217)			
SXM	0.626**	0.7019**	0.4902**			
	(0.0177)	(0.0158)	(0.0216)			
SAR				0.5654**	0.7017**	0.4586**
				(0.0171)	(0.0151)	(0.0209)
SDM				0.6205**	0.6944**	0.4877**
				(0.0174)	(0.0153)	(0.0215)
SAC	0.9093**	0.9306**	0.8166**	−0.6221**	−0.7208**	−0.6248**
	(0.0129)	(0.0086)	(0.0195)	(0.047)	(0.0426)	(0.0502)
MSAC	0.8995**	−0.7029**	−0.635**	−0.7909**	0.8991**	0.7958**
	(0.0114)	(0.0448)	(0.0602)	(0.044)	(0.0112)	(0.0215)

Table 4: **Spatial coefficients for short run models (i.e., with temporal lags)**

	Spatial Error component: λ			Spatial Lag component: ρ		
	AR	FO	UR	AR	FO	UR
SEM	0.1134**	0.3004**	0.2246**			
	(0.022)	(0.0295)	(0.0278)			
SXM	0.0473**	0.2404**	0.2**			
	(0.0131)	(0.0282)	(0.0288)			
SAR				0.1335**	0.1256**	0.1122**
				(0.0129)	(0.0087)	(0.0134)
SDM				0.0629**	0.2427**	0.2011**
				(0.0302)	(0.029)	(0.0287)
SAC	−0.1119**	0.1451**	0.1338**	0.1572**	0.1103**	0.0755**
	(0.0334)	(0.0324)	(0.0361)	NA	(0.0087)	(0.0179)
MSAC	−0.3827**	−0.0403	−0.3825**	0.3746**	0.2776**	0.48967**
	(0.0958)	(0.0418)	(0.0814)	(0.071)	(0.0428)	(0.0527)

4.2 Predictions results

The predictive accuracy of the models is compared statistically by computing the Root Mean Squared Errors (RMSE) for each model's predictions, based on comparing observed and predicted land use at the aggregate grid level. The comparisons are reported in the panels A, B, C and D of Table 5. They present respectively the in-sample and out of sample predictions for the models, and with and without lagged land uses.

The in-sample predictions consist of 2003 land use shares from the models fitted on the 1993–2003 time interval. Out of sample predictions consist of 2003 land use shares but fitted on the models estimated on the 1993–1998 time interval. The rows in the following tables are: REF, the reference RMSE (i.e. computed with national shares as predictors), OLS for Ordinary Least Squares, GAM for GeoAdditive Models, FRA for aggregate multinomial model, DIR for aggregate Dirichlet model, SEM for spatial error model, SXM for model with spatially-lagged explanatory variables, SAR for spatial autoregressive model, SDM for spatial Durbin model, SAC for the spatial error spatial autoregressive model, MSAC for the most general spatial model, Lpb for the linear probability model and Mnl for the individual multinomial model. The last two are estimated on individual data.

For in-sample predictions without lags, the predictors from the spatial econometric models are based on a full-information set and perform better than any other estimation techniques. The differences are relatively high, as it can be seen from the last columns reporting the RMSE means by rows.[7] Spatial models gains relative to OLS are half of the gains of OLS relatively to the benchmark. Thus, the effect is strong. In the same magnitudes, the GAM is in an intermediate position between the spatial and the aspatial models. For the aspatial models (both aggregate and individual) the predictive abilities are rather similar and the individual linear probability model is the worst. Note that the multicollinear models such as SAC and MSAC, perform the best, according to a well-known property that multicollinearity does not bias the predictions. Including lagged land uses for short run predictions drastically decreases the RMSE, and the differences between estimation techniques also decrease significantly. The spatial models perform best, but the performance of the GAM model is also quite similar. More importantly, the inclusion of temporal lag implies a loss of relative performance in the models (aggregate and individual) based on discrete outcomes: FRA, DIR, Lpb, Mnl.

For the out of sample predictions, a first counter-intuitive result is that, in some cases, the models perform better for the out-of-sample relatively to the previous in-sample. Because the predictors are not known for the out-of-sample, the full-information estimators cannot be implemented for spatial econometric models. As a consequence, their performance strongly decrease compared to other aspatial models. Also, the GAM presents the smallest RMSE. Including the lagged land use shares in the out of sample predictions does not change the previous results: GAM remains the most efficient. However, OLS estimation appears also as a good performer. Because of these results, we approximate full-information predictors for spatial econometrics models based on some approximations of S_{-g} to recover their relative performance as in the in-sample context.

Table 6 reports the in-sample and out of sample RMSE from the spatial econometric models, for the different KP predictors. Choosing a predictor without bias appears as an important choice because all the gain from spatial econometric models comes from this choice. Predictors KP2 and KP3 perform best, closely followed by the KP4 and KP5 when we choose KP4 for the models with spatial lag and KP5 for models with spatial errors. Because predictors KP2 to KP5 are not computable for out-of-sample predictions, we propose a heuristic solution. It consists of substituting S_{gt} by $S_{gt'}$ in formulas (9), (10), (11) and (12).[8] As the bottom panel of Table 6 shows, our heuristic computations of out-of-sample prediction allows some recovery of the gains in predictive ability through the introduction of space. Using KP2 and KP3 for out-of-sample prediction allows us to obtain RMSE close to 0.1 with long run models, and outperform the predictive abilities of the GAM and individual models. The results of the SAC and MSAC are interesting because they perform better with the good predictors but the gains from using more complex predictors are less marked.

[7] It is not surprising for a case with 4 land use categories to have a benchmark of 0.25.

[8] This choice seems natural from the inertia of the land use reported in Table 1 and would be applicable to every study about land use. For applications on economic processes that present less inertia, we can imagine using temporal projections of the variables of interest.

Table 5: Root Mean Square Errors for the models according to the predictive configurations: *The rows marked REF report the benchmark RMSE from constant predictions. The columns MEAN report the row means of RMSE. For spatial econometric models, the full-information predictor is used in in-sample and the KP1 (see (7)) is used in out of sample.*

A. In sample RMSE for models without lags

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
OLS	0.1581	0.1589	0.1773	0.0666	0.1467
GAM	0.134	0.146	0.1603	0.0641	0.1314
FRA	0.155	0.1491	0.1709	0.0618	0.1408
DIR	0.1558	0.1527	0.1735	0.0694	0.1436
SEM	0.1113	0.1198	0.1358	0.058	0.1102
SXM	0.1119	0.1205	0.1368	0.0578	0.1108
SAR	0.1135	0.1242	0.14	0.059	0.1133
SDM	0.1122	0.1206	0.1369	0.0579	0.1109
SAC	0.1005	0.1085	0.1227	0.0536	0.0997
MSAC	0.1006	0.1075	0.1253	0.0541	0.1004
Lpb	0.1629	0.1612	0.1756	0.0629	0.1477
Mnl	0.1573	0.1506	0.1732	0.0608	0.1424

B. In sample RMSE for models with temporal lags

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
OLS	0.0405	0.0379	0.0282	0.0182	0.0324
GAM	0.0396	0.0366	0.028	0.0181	0.0317
FRA	0.0516	0.0465	0.043	0.0295	0.0434
DIR	0.0541	0.0498	0.0461	0.0304	0.046
SEM	0.0392	0.0374	0.0272	0.0178	0.0316
SXM	0.0399	0.0385	0.0282	0.0179	0.0323
SAR	0.0395	0.037	0.0293	0.0178	0.032
SDM	0.0398	0.0384	0.0282	0.0178	0.0323
SAC	0.0394	0.0378	0.0289	0.0177	0.0321
MSAC	0.0392	0.0376	0.0281	0.0173	0.0318
Lpb	0.0572	0.0507	0.0269	0.018	0.0415
Mnl	0.0547	0.0485	0.0267	0.0175	0.0399

C. Out sample RMSE for models without lags

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
OLS	0.1675	0.1717	0.19	0.0677	0.1567
GAM	0.1342	0.1525	0.1645	0.0656	0.1347
FRA	0.1616	0.1556	0.1771	0.0633	0.1464
DIR	0.1625	0.1567	0.1737	0.0714	0.1468
SEM	0.1712	0.1692	0.186	0.0671	0.1558
SXM	0.1649	0.1743	0.1818	0.0673	0.1542
SAR	0.1661	0.1611	0.1795	0.0669	0.1502
SDM	0.1627	0.1626	0.1816	0.0652	0.1501
SAC	0.1832	0.1831	0.1904	0.072	0.1647
MSAC	0.1738	0.1965	0.181	0.0969	0.1665
Lpb	0.1669	0.1622	0.1804	0.0632	0.1506
Mnl	0.1633	0.1563	0.1778	0.0621	0.1471

D. Out sample RMSE for models with temporal lags

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
OLS	0.0317	0.0308	0.0202	0.0128	0.0251
GAM	0.0308	0.0309	0.0212	0.0133	0.0251
FRA	0.0478	0.0427	0.0416	0.0293	0.0409
DIR	0.0503	0.0461	0.0459	0.0299	0.0437
SEM	0.0319	0.0314	0.0203	0.0129	0.0254
SXM	0.0324	0.0329	0.021	0.0135	0.0262
SAR	0.0315	0.0318	0.0206	0.0129	0.0255
SDM	0.0324	0.033	0.0209	0.0134	0.0262
SAC	0.0316	0.0323	0.0206	0.0129	0.0257
MSAC	0.0355	0.04	0.0261	0.0155	0.0307
Lpb	0.0489	0.0441	0.0181	0.0126	0.0347
Mnl	0.048	0.0429	0.018	0.0124	0.034

Table 6: Means of the Root Mean Square Errors from different spatial predictors, long run models

	KP1	KP2	KP3	KP4	KP5
In Sample (i.e., $\mathbf{S}_{gt}$ is known)					
SEM	0.1649	0.1087		0.1649	0.1108
SXM	0.1617	0.1081		0.1617	0.1111
SAR	0.1638	0.1104		0.1152	0.1670
SDM	0.1580	0.1095		0.1116	0.1698
SAC	0.1844	0.1175		0.2716	0.1088
MSAC	0.1730	0.1730		0.1936	0.2739
Out Sample (i.e., $\mathbf{S}_{gt}$ is unknown)					
SEM	0.1703	0.1098		0.1619	0.1104
SXM	0.1650	0.1097		0.1519	0.1106
SAR	0.1616	0.1113		0.1130	0.1638
SDM	0.1565	0.1113		0.1109	0.1690
SAC	0.1736	0.1230		0.2590	0.1101
MSAC	0.1691	0.1096		0.1091	0.3510

When $\mathbf{S}_{gt}$ is unknown, we use our extension of KP.

In summary, it appears that the predictors used matters more than the specification for predictive accuracy. Hence, particular care should be taken in choosing the predictor. Results from the short run models are not reported due to less visible differences, the rankings of relative performance are the same. These results are available from the authors upon request.

4.3 Simulation results

Turning to the counter-factual simulations, we use the 1993–1998 models in out-of-sample predictions (i.e., for 2003) with the returns for pastures increased by 200 euros. This represents an example of a green policy which proposes incentives for extensive land use in order to increase water quality or biodiversity. We compute the differences in terms of aggregate pasture acreages in 2003 relatively to what is observed. For the spatial econometric models, both KP1 and KP2 (marked with a *) are reported in the following Table 7. Note that spatial predictions KP2 take into account both the direct and indirect effects of payments, so they imply no differences for the interpretation of changes to acreages at the aggregate scale.

In terms of aggregate changes in pasture acreages as a consequence of increasing their economic returns, there is a significant gap between individual and aggregate models when comparing short run and long run models. Predictions are relatively similar from a long term perspective (except for GAM and SEM which appear to strongly underestimate the effects) but in the short run, the aggregate models clearly underestimate the effects. We interpret this result as indicating an indisputable advantage of individual models (amount of information), which show significant effects of economics returns in short run models when lagged land use captures a lot of the effect. In contrast, no specific patterns emerge for the aggregate models in either the short run or long run. OLS seems to be in mid

Table 7: Simulation of 200 euros payments for pasture: acreages variations
The table reports the variations of pasture acreages for 2003 at the national scale on the basis of models estimated on the period 1993–1998. The t stat. are relative to the nullity of the correlation with the individual mnl model. The units are in thousand ha.

	LONG RUN			SHORT RUN		
	NET EFFECT	COR(mnl)	t stat.	NET EFFECT	COR(mnl)	t stat.
OLS	+ 714.5	− 0.018	− 1.155	+ 41.82	+ 0.131	+ 8.084
GAM	+ 126.4	− 0.212	− 13.32	+ 30.52	+ 0.243	+ 15.36
FRA	+ 651.3	+ 0.061	+ 3.727	− 3.467	− 0.292	− 18.72
DIR	+ 568.8	+ 0.049	+ 3.020	+ 13.12	+ 0.271	+ 17.30
SEM	+ 298.6	+ 0.074	+ 4.547	+ 45.05	+ 0.128	+ 7.938
SXM	+ 713.5	+ 0.224	+ 14.08	− 13.12	− 0.028	− 1.775
SAR	+ 667.3	+ 0.300	+ 19.13	+ 91.00	+ 0.012	+ 0.735
SDM	+ 885.7	+ 0.169	+ 10.51	+ 51.13	+ 0.001	+ 0.071
SAC	+ 382.1	+ 0.025	+ 1.558	+ 18.10	+ 0.023	+ 1.400
MSAC	+ 285.8	+ 0.016	+ 9.992	− 27.25	− 0.033	− 2.011
mnl	+ 756.2	1.000		+ 166.19	1.000	

position, between the spatial econometric models that sometimes perform well (SXM in the long run and SAR in the short run) but appears really contrasted with the individual model. Aggregate FRA and Dirichlet perform less badly in relation to predictive accuracy terms, at least in the long run.

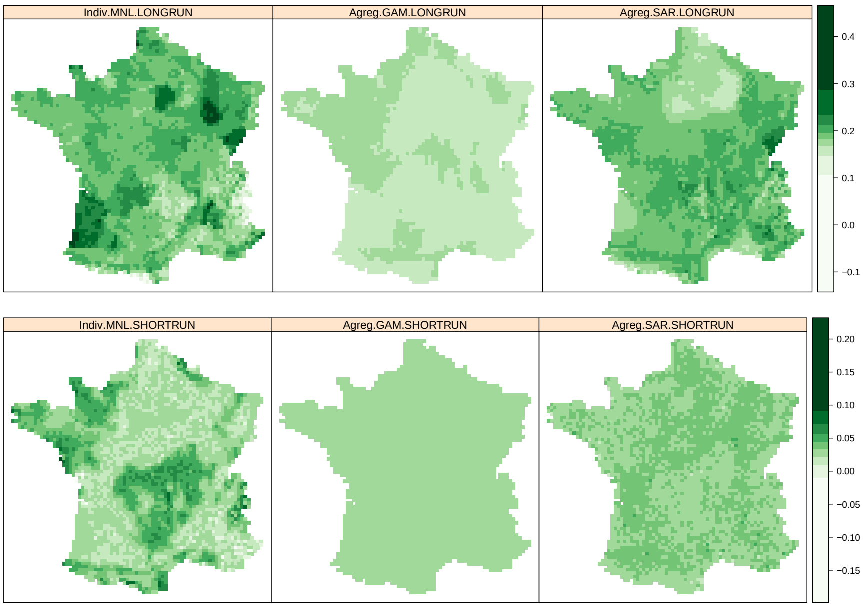
We compare the spatial patterns of the simulations of each aggregate model with the results of the individual mnl in columns 3–4 and 6–7, which report the correlations and the associated t statistics for each of them. Note first that the correlation coefficients are relatively low although they are mostly significant. In the long run models, the spatial patterns of new pasture acreages in OLS and GAM are negatively correlated with the simulations for the individual model, and its is bigger and significant for the GAM. Compared to the out-of-sample predictive performance of GAM, it seems to induce antagonism between predictive accuracy and capacity to mimic the individual model. Among the spatial models, the SAR presents the highest significant correlation coefficients: 0.3. For short run simulations, the Dirichlet model performs the best with 0.27 and outperforms the spatial econometric models. These simulation results are confirmed by the following maps of Figure 1 showing the intuitive consistency of the simulation from the individual models relatively to some aggregate models.

5 Conclusion

It is generally recognized that economic behaviors are more accurately analyzed using individual models. The increasing availability of individual land use data is allowing the estimation of individual models. However, modeling at an aggregate scale is still common because of complexity of individual discrete choice models or because the aggregate scale corresponds to the outcome of interest. It is generally assumed that aggregation of these individual micro relationships yields better predictions than more aggregate models.

In this paper, we compared the predictive abilities of different land use model specifications at the individual and aggregate levels. More specifically, we showed how the

Figure 1: Simulation of 200 euros payments for pasture: spatial patterns *The maps report both long run (top panel) and short run (bottom panel) predictions, for the individual MNL (left), the GeoAdditive model (middle) and the spatial autoregressive (right). The units are in thousand ha.*



introduction of a spatial dimension in aggregate models matters for improving their predictions related to aggregate changes in land use. Our results show that: (i) introducing spatial autocorrelation in aggregate grid-level models improves their predictive accuracy and even outperforms individual models if unbiased predictors are used, (ii) a specification including lagged land use as explanatory variable in the aggregated as well in the individual models, outperforms any other specification where only economic and bioclimatic variables are included, (iii) in terms of policy simulation, individual models perform better than aggregate models.

Our findings show that it may not be worth using individual land use data when the only objective is to predict aggregate land use. This result corroborates the findings in Grunfeld and Griliches (1960) that show "*aggregation is not necessarily bad if one is interested in the aggregates*". By taking advantage of the progress made in spatial econometrics tools, we show how the introduction of spatial autocorrelation in aggregated land use models allow more precise predictions than individual models. However, individual land use data are needed for simulation purpose if the focus is impact of land use changes on greenhouse gas emissions or other local environmental issues such as biodiversity loss or ground-water pollution.

6 Acknowledgements

The authors acknowledge the contributors of the `R` functions used for this research: Roger Bivand et al. for `spdep`, `sp`, `rgdal` and `maptools` ; Marek Hlavac for `stargazer` ; Marco J. Maier for `DirichletReg` ; Simon Wood for `mgcv` and Brian Ripley for `nnet`. The `R` functions used to perform predictions from spatial econometric model and the associated documentation are freely available from `jsay.github.com`. The paper has benefited from comments from Workshop Spatial Econometrics and Statistics (Avignon) and the VII Conference of Spatial Econometrics Association (Washington). We particularly thank Benoît Delbecq, Ghislain Geniaux and Catherine Baumont. We would also like to thank the *Service de la Statistique et de la Prospective du Ministère de l'Agriculture* for providing the data on land use from the TERUTI survey, and data on land prices. We finally thank Alexandra Niedzwiedz for providing data on wood prices and forest harvests.

References

AGRESTE (2004). L'utilisation du territoire en 2003 - nouvelle série 1992 à 2003. *Chiffres et Donnees - Serie Agriculture* 157: 406–414.

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.

Anselin, L. (2007). Spatial econometrics in RSUE: Retrospect and prospect. *Regional Science and Urban Economics* 37: 450–456.

Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science* 89: 3–25.

Ay, J.-S., Chakir, R., Doyen, L., Jiguet, F. and Leadley, P. (2014). Integrated models, scenarios and dynamics of climate, land use and common birds. *Accepted Climatic Change* .

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., Soest, D. van and Termansen, M. (2013). Bringing ecosystem services into economic decision-making: Land use in the United Kingdom. *Science* 341: 45–50.

Bivand, R. (2002). Spatial econometrics functions in R: Classes and methods. *Journal of Geographical Systems* 4: 405–421.

Bivand, R. S., Pebesma, E. and Gomez-Rubio, V. (2013). *Applied Spatial Data Analysis with R*. second Edition, UseR! Series, Springer.

Box, G. E., Jenkins, G. M. and Reinsel, G. C. (2013). *Time series analysis: Forecasting and control*. John Wiley & Sons.

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: 487–509.

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

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

Cliff, A. D. and Ord, J. K. (1981). *Spatial Processes : Models and Applications*. London, UK: Pion.

Considine, T. J. and Mount, T. D. (1984). The use of linear logit models for dynamic input demand systems. *The Review of Economics and Statistics* : 434–443.

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

Ferdous, N. and Bhat, C. R. (2013). 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.

Fleming, M. M. and Mae, F. (2004). Techniques for estimating spatially dependent discrete choice models. *In Advances in Spatial Econometrics, Anselin Luc and Raymond Florax, eds. Springer-Verlag, Heidelberg.* .

Florax, R. J., 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.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica: Journal of the Econometric Society* : 681–700.

Grunfeld, Y. and Griliches, Z. (1960). Is aggregation necessarily bad? *The Review of Economics and Statistics* 42: pp. 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: 299–310.

Hastie, T. and Tibshirani, R. (1986). Generalized additive models. *Statistical science* : 297–310.

Irwin, E. G. (2010). New directions for urban economic models of land use change: incorporating spatial dynamics and heterogeneity. *Journal of Regional Science* 50: 65–91.

Irwin, E. G. and Geoghegan, J. (2001). Theory, data, methods: Developing spatially explicit economic models of land use change. *Agriculture, Ecosystems & Environment* 85: 7–24.

Kelejian, H. H. and Prucha, I. R. (2007). The relative efficiencies of various predictors in spatial econometric models containing spatial lags. *Regional Science and Urban Economics* 37: 363–374.

Le Gallo, J. (2013). Cross-section spatial regression models. *Handbook of Regional Science, Fischer, Manfred M., Nijkamp, Peter (Eds.)* .

LeSage, J. and Pace, R. (2009). *Introduction to Spatial Econometrics*. CRC Press Boca Raton FL.

Lewis, D. J. (2010). An economic framework for forecasting land-use and ecosystem change. *Resource and Energy Economics* 32: 98–116.

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* Vol. 71, No. 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 Economics* 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.

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

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

Miller, D. J. and Plantinga, A. J. (1999). Modeling land use decisions with aggregate data. *American Journal of Agricultural Economics* 81(1): 180–194.

Mullahy, J. (2010). Multivariate Fractional Regression Estimation of Econometric Share Models. NBER Working Papers 16354, National Bureau of Economic Research, Inc.

Nelson, E., Polasky, S., Lewis, D. J., Plantinga, A. J., Lonsdorf, E., White, D., Bael, D. and Lawler, J. J. (2008). Efficiency of incentives to jointly increase carbon sequestration and species conservation on a landscape. *Proceedings of the National Academy of Sciences* 105: 9471–9476.

Papke, L. E. and Wooldridge, J. (1993). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates.

Papke, L. E. and Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics* 145: 121–133.

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

Plantinga, A. and Irwin, E. (2006). Overview of empirical methods. *Economics of Rural Land-Use Change. Bell, K.P., Boyle, K.J., and Rubin, J., eds., Ashgate Publishing* .

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: 1082–1091.

Schatzki, T. (2003). Options, uncertainty and sunk costs: An empirical analysis of land use change. *Journal of Environmental Economics and Management* 46: 86–105.

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

Smirnov, O. A. (2010). Modeling spatial discrete choice. *Regional Science and Urban Economics* 40: 292 – 298, doi:http://dx.doi.org/10.1016/j.regsciurbeco.2009.09.004, <ce:title>Advances In Spatial Econometrics</ce:title>.

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.

Train, K. (2009). *Discrete Choice Methods with Simulation*, *Second Edition*. Cambridge University Press.

Turner, B. L., Lambin, E. F. and Reenberg, A. (2007). The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences* 104: 20666–20671.

Wood, S. N. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association* 99.

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.

Wu, J. and Duke, J. M. (2014). *The Oxford Handbook of Land Economics*. Oxford University Press, USA.

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: 1033–1047.

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

A Supporting Information (not for publication)

A.1 Raw results from OLS

Table 8: Linear OLS models of land use on 1993–2003

	Arable Share		Forest Share		Urban Share	
	long run	short run	long run	short run	long run	short run
ARlog93		0.900***		−0.003		0.016
		(0.020)		(0.010)		(0.012)
FOlog93		0.008		0.937***		0.007
		(0.013)		(0.017)		(0.014)
URlog93		−0.033**		−0.013		0.847***
		(0.014)		(0.011)		(0.021)
scale(Arable returns03)	0.510***	0.041**	0.272***	0.012	0.397***	0.060***
	(0.042)	(0.019)	(0.036)	(0.011)	(0.033)	(0.015)
scale(Pasture returns03)	−0.331***	−0.027	−0.325***	−0.030**	−0.234***	−0.045***
	(0.036)	(0.017)	(0.032)	(0.014)	(0.032)	(0.015)
scale(Forest returns03)	−0.078**	0.018	0.525***	0.039***	0.116***	−0.014
	(0.035)	(0.019)	(0.036)	(0.014)	(0.029)	(0.017)
scale(POP03)	−0.239**	−0.043	−0.053	−0.013	0.141	0.016
	(0.121)	(0.068)	(0.127)	(0.023)	(0.300)	(0.034)
scale(Elevation)	−1.452***	−0.189***	−0.754***	−0.139***	−0.859***	−0.108**
	(0.100)	(0.059)	(0.104)	(0.026)	(0.098)	(0.048)
scale(Slope)	−0.429***	−0.135**	0.450***	0.069***	0.017	0.038
	(0.083)	(0.054)	(0.073)	(0.014)	(0.077)	(0.028)
scale(WHC)	0.378***	0.085***	−0.287***	0.014	−0.026	−0.017
	(0.054)	(0.028)	(0.056)	(0.019)	(0.047)	(0.023)
scale(Soil depth)	−0.260***	−0.052*	0.255***	−0.026	0.051	0.006
	(0.053)	(0.028)	(0.055)	(0.019)	(0.049)	(0.023)
scale(Precipitations)	−0.568***	−0.091***	0.040	−0.032***	−0.104***	−0.023
	(0.035)	(0.022)	(0.030)	(0.009)	(0.032)	(0.014)
scale(Temperature)	0.167**	−0.082*	0.151	−0.021	−0.194**	0.039
	(0.084)	(0.046)	(0.093)	(0.018)	(0.084)	(0.033)
scale(Humidity)	−0.003	−0.102***	−0.119*	−0.048***	−0.319***	−0.035
	(0.062)	(0.032)	(0.065)	(0.013)	(0.070)	(0.023)
scale(Radiation)	−0.354***	0.025	−0.650***	−0.018	0.243***	0.019
	(0.074)	(0.037)	(0.081)	(0.021)	(0.078)	(0.034)
Constant	−0.615***	−0.097***	−0.177***	0.060**	−1.815***	−0.082**
	(0.025)	(0.034)	(0.023)	(0.029)	(0.023)	(0.039)
Observations	3,767	3,767	3,767	3,767	3,767	3,767
R^2	0.663	0.911	0.229	0.919	0.359	0.852
Adjusted R^2	0.662	0.911	0.227	0.918	0.357	0.851

Note: *p<0.1; **p<0.05; ***p<0.01.
Reference modality= Pastures, scaled variables, HC robust standard errors.

A.2 Raw results fom GAM

Table 9: GeoAdditive models of land use on 1993–2003

	Arable Share		Forest Share		Urban Share	
	long run	short run	long run	short run	long run	short run
ARlog93		0.881***		−0.006		0.013
		(0.010)		(0.006)		(0.008)
FOlog93		−0.004		0.912***		−0.004
		(0.010)		(0.006)		(0.008)
URlog93		−0.031***		−0.015**		0.837***
		(0.010)		(0.006)		(0.008)
scale(Arable returns03)	0.403***	0.032*	−0.018	−0.018	0.245***	0.045***
	(0.035)	(0.019)	(0.031)	(0.012)	(0.032)	(0.016)
scale(Pasture returns03)	−0.126***	−0.020	−0.037	−0.016	−0.106***	−0.041***
	(0.033)	(0.018)	(0.029)	(0.011)	(0.030)	(0.015)
scale(Forest returns03)	−0.068*	0.011	0.053	0.021*	0.044	0.022
	(0.041)	(0.020)	(0.037)	(0.013)	(0.037)	(0.018)
scale(POP03)	−0.180***	−0.042***	−0.026	−0.014*	0.141***	0.012
	(0.023)	(0.013)	(0.021)	(0.008)	(0.021)	(0.011)
scale(Elevation)	−1.036***	−0.062	−0.594***	−0.120***	−0.731***	−0.168***
	(0.118)	(0.066)	(0.105)	(0.039)	(0.108)	(0.055)
scale(Slope)	−0.700***	−0.202***	0.453***	0.062***	0.057	0.059**
	(0.062)	(0.034)	(0.055)	(0.021)	(0.056)	(0.029)
scale(WHC)	0.375***	0.062**	−0.233***	0.002	0.0002	−0.013
	(0.051)	(0.028)	(0.046)	(0.017)	(0.047)	(0.024)
scale(Soil depth)	−0.383***	−0.059**	0.097**	−0.030*	−0.057	−0.010
	(0.050)	(0.028)	(0.044)	(0.017)	(0.046)	(0.023)
scale(Precipitations)	−0.486***	−0.084***	0.211***	−0.003	−0.134***	−0.034*
	(0.039)	(0.021)	(0.035)	(0.013)	(0.035)	(0.018)
scale(Temperature)	0.414***	0.025	0.188*	−0.002	0.152	−0.006
	(0.114)	(0.061)	(0.101)	(0.037)	(0.104)	(0.051)
scale(Humidity)	0.028	−0.090**	0.324***	0.022	−0.031	0.040
	(0.067)	(0.036)	(0.060)	(0.022)	(0.061)	(0.030)
scale(Radiation)	−0.118	0.044	−0.442***	0.0002	0.237***	0.070
	(0.097)	(0.051)	(0.086)	(0.031)	(0.088)	(0.043)
Constant	−0.615***	−0.109***	−0.177***	0.047***	−1.815***	−0.107***
	(0.023)	(0.023)	(0.020)	(0.014)	(0.020)	(0.019)
Observations	3,767	3,767	3,767	3,767	3,767	3,767
Adjusted R^2	0.716	0.913	0.426	0.921	0.418	0.855
UBRE	1.932	0.595	1.509	0.208	1.599	0.399
Note:	*p<0.1; **p<0.05; ***p<0.01.					
	Reference= Pastures, scaled variables, bivariate smooth function of coordinates, see subsection A.11					

A.3 Raw results from FRA fractional

Table 10: Aggregate FRA fractional models of land use on 1993–2003

	Long Run arable share	 forest share	 urban share	Short Run arable share	 forest share	 urban share
ARBLE93				2.899***	−0.165	−0.786**
				(0.230)	(0.249)	(0.346)
PSTUR93				−2.929***	−2.933***	−3.862***
				(0.223)	(0.215)	(0.342)
FORST93				−0.396	3.256***	−1.224***
				(0.250)	(0.207)	(0.368)
URBAN93				−1.354	−0.910	5.226***
				(0.960)	(0.940)	(0.923)
scale(Arable returns03)	0.498***	0.321***	0.357***	0.050	0.034	0.103
	(0.064)	(0.062)	(0.091)	(0.073)	(0.066)	(0.096)
scale(Pasture returns03)	−0.298***	−0.339***	−0.242***	−0.016	−0.039	−0.047
	(0.056)	(0.061)	(0.082)	(0.062)	(0.067)	(0.088)
scale(Forest returns03)	0.025	0.355***	0.094	0.036	0.044	0.010
	(0.058)	(0.052)	(0.082)	(0.061)	(0.058)	(0.088)
scale(POP03)	−0.495***	−0.065	0.090	−0.036	0.001	−0.010
	(0.129)	(0.073)	(0.056)	(0.085)	(0.047)	(0.046)
scale(Elevation)	−0.889***	−0.538***	−0.671**	−0.419**	0.028	−0.091
	(0.193)	(0.125)	(0.274)	(0.209)	(0.144)	(0.284)
scale(Slope)	−0.387**	0.321***	0.081	−0.207	−0.049	−0.033
	(0.163)	(0.087)	(0.203)	(0.168)	(0.097)	(0.210)
scale(WHC)	0.335***	−0.283***	0.112	0.006	−0.002	−0.031
	(0.097)	(0.102)	(0.151)	(0.104)	(0.110)	(0.162)
scale(Soil depth)	−0.203**	0.260***	0.026	0.001	0.002	0.003
	(0.095)	(0.099)	(0.150)	(0.102)	(0.106)	(0.159)
scale(Precipitations)	−0.410***	0.081*	−0.111	−0.067	−0.043	−0.027
	(0.068)	(0.048)	(0.096)	(0.073)	(0.056)	(0.103)
scale(Temperature)	0.135	0.032	−0.357*	0.107	0.001	−0.015
	(0.150)	(0.116)	(0.216)	(0.159)	(0.129)	(0.227)
scale(Humidity)	−0.064	−0.192**	−0.578***	0.063	−0.047	−0.099
	(0.115)	(0.086)	(0.162)	(0.122)	(0.096)	(0.172)
scale(Radiation)	−0.164	−0.358***	0.508**	−0.232	0.006	0.129
	(0.142)	(0.112)	(0.203)	(0.151)	(0.128)	(0.218)
Constant	−0.355***	−0.080*	−1.621***			
	(0.062)	(0.046)	(0.078)			
Akaike Inf. Crit.	8,545.113	8,545.113	8,545.113	7,634.633	7,634.633	7,634.633

Note: *p<0.1; **p<0.05; ***p<0.01.
Reference= Pastures, scaled variables, corrected standard errors.

A.4 Raw results from SEM

Table 11: Spatial Error Models of land use on 1993–2003

	Arable Share		Forest Share		Urban Share	
	long run	short run	long run	short run	long run	short run
ARlog93		0.889***		−0.009		0.013
		(0.009)		(0.006)		(0.008)
FOlog93		0.006		0.920***		−0.001
		(0.010)		(0.006)		(0.008)
URlog93		−0.030***		−0.014**		0.842***
		(0.010)		(0.006)		(0.008)
scale(Arable returns03)	0.464***	0.050***	0.031	0.010	0.323***	0.063***
	(0.045)	(0.018)	(0.043)	(0.012)	(0.038)	(0.016)
scale(Pasture returns03)	−0.204***	−0.031*	−0.135***	−0.031***	−0.173***	−0.047***
	(0.049)	(0.017)	(0.047)	(0.012)	(0.039)	(0.015)
scale(Forest returns03)	−0.087*	0.016	0.339***	0.044***	0.116***	−0.005
	(0.051)	(0.016)	(0.053)	(0.011)	(0.038)	(0.015)
scale(POP03)	−0.152***	−0.042***	−0.026	−0.014*	0.124***	0.014
	(0.025)	(0.013)	(0.022)	(0.008)	(0.023)	(0.011)
scale(Elevation)	−1.065***	−0.191***	−0.531***	−0.140***	−0.830***	−0.119***
	(0.099)	(0.045)	(0.090)	(0.029)	(0.086)	(0.039)
scale(Slope)	−0.448***	−0.140***	0.570***	0.071***	0.061	0.044
	(0.066)	(0.032)	(0.059)	(0.020)	(0.059)	(0.027)
scale(WHC)	0.310***	0.084***	−0.195***	0.006	0.017	−0.013
	(0.061)	(0.028)	(0.055)	(0.018)	(0.054)	(0.024)
scale(Soil depth)	−0.213***	−0.049*	0.144***	−0.016	−0.013	0.006
	(0.061)	(0.028)	(0.055)	(0.018)	(0.054)	(0.024)
scale(Precipitations)	−0.510***	−0.095***	0.076	−0.032***	−0.139***	−0.027*
	(0.052)	(0.018)	(0.052)	(0.012)	(0.041)	(0.016)
scale(Temperature)	0.494***	−0.069*	0.422***	−0.004	−0.082	0.041
	(0.110)	(0.040)	(0.107)	(0.027)	(0.089)	(0.035)
scale(Humidity)	0.067	−0.095***	0.140*	−0.041**	−0.272***	−0.036
	(0.083)	(0.030)	(0.082)	(0.020)	(0.067)	(0.027)
scale(Radiation)	−0.267**	0.016	−0.613***	−0.030	0.245***	0.017
	(0.114)	(0.038)	(0.113)	(0.026)	(0.088)	(0.034)
Constant	−0.639***	−0.099***	−0.194***	0.049***	−1.814***	−0.097***
	(0.059)	(0.024)	(0.069)	(0.016)	(0.040)	(0.021)
Observations	3,767	3,767	3,767	3,767	3,767	3,767
σ^2	1.656	0.594	1.250	0.203	1.491	0.394
Akaike Inf. Crit.	12,891.050	8,769.077	11,936.960	4,771.993	12,373.050	7,246.211
Wald Test (df = 1)	1,247.921***	26.557***	880.688***	103.971***	527.874***	65.066***
LR Test (df = 1)	917.587***	12.524***	1,399.780***	96.056***	435.836***	58.226***

Note: *p<0.1; **p<0.05; ***p<0.01
scaled variables. Reference= Pastures

A.5 Raw results from SXM

Table 12: Spatial X Models of land use on 1993–2003

	Arable Share		Forest Share		Urban Share	
	long run	short run	long run	short run	long run	short run
ARlog93		0.834***		−0.019***		0.006
		(0.011)		(0.006)		(0.009)
FOlog93		−0.009		0.897***		−0.019**
		(0.011)		(0.006)		(0.009)
URlog93		−0.019*		−0.017***		0.836***
		(0.010)		(0.006)		(0.009)
scale(Arable returns03)	0.352***	0.077**	−0.054	−0.054***	0.171***	0.048*
	(0.056)	(0.035)	(0.049)	(0.020)	(0.054)	(0.029)
scale(Pasture returns03)	−0.032	0.004	−0.010	0.034	−0.022	−0.029
	(0.068)	(0.042)	(0.060)	(0.024)	(0.066)	(0.034)
scale(Forest returns03)	−0.035	0.011	0.074	0.043	0.066	0.124***
	(0.093)	(0.057)	(0.081)	(0.033)	(0.090)	(0.047)
scale(POP03)	−0.132***	−0.016	−0.020	−0.010	0.123***	0.011
	(0.025)	(0.016)	(0.022)	(0.009)	(0.024)	(0.013)
scale(Elevation)	−0.857***	−0.053	−0.512***	−0.092**	−0.844***	−0.133**
	(0.105)	(0.067)	(0.093)	(0.038)	(0.101)	(0.054)
scale(Slope)	−0.432***	−0.154***	0.578***	0.076***	0.046	0.068**
	(0.067)	(0.042)	(0.059)	(0.024)	(0.063)	(0.034)
scale(WHC)	0.238***	0.047	−0.188***	−0.027	0.013	−0.001
	(0.064)	(0.040)	(0.057)	(0.023)	(0.062)	(0.033)
scale(Soil depth)	−0.180***	−0.014	0.132**	0.013	−0.044	−0.001
	(0.063)	(0.039)	(0.055)	(0.023)	(0.060)	(0.032)
scale(Precipitations)	−0.200**	−0.020	0.197***	0.005	−0.155*	−0.066
	(0.083)	(0.051)	(0.073)	(0.030)	(0.080)	(0.042)
scale(Temperature)	1.017***	0.283***	0.307**	0.041	0.379**	0.021
	(0.161)	(0.101)	(0.141)	(0.059)	(0.156)	(0.083)
scale(Humidity)	−0.225	−0.148*	0.209*	0.020	−0.062	0.023
	(0.138)	(0.084)	(0.120)	(0.049)	(0.133)	(0.069)
scale(Radiation)	−0.277	−0.080	−0.546***	−0.013	0.176	0.108
	(0.176)	(0.108)	(0.153)	(0.063)	(0.170)	(0.089)
Constant	−0.638***	−0.116***	−0.191***	0.074**	−1.814***	0.001
	(0.055)	(0.043)	(0.061)	(0.029)	(0.039)	(0.040)
Observations	3,767	3,767	3,767	3,767	3,767	3,767
σ^2	1.616	0.572	1.244	0.197	1.476	0.390
Akaike Inf. Crit.	12,802.470	8,650.725	11,900.890	4,673.891	12,353.180	7,225.390
Wald Test (df = 1)	1,251.796***	12.987***	1,982.371***	72.671***	514.239***	48.084***
LR Test (df = 1)	905.610***	2.284	1,271.983***	65.701***	421.610***	46.372***

Note: *p<0.1; **p<0.05; ***p<0.01
scaled variables. Reference= Pastures coefficients from spatially lagged explanatory variables are not reported

A.6 Raw results from SAR

Table 13: Spatial Autoregressive Regressions of land use on 1993–2003

	Arable Share		Forest Share		Urban Share	
	long run	short run	long run	short run	long run	short run
ARlog93		0.854***		−0.010*		0.013*
		(0.010)		(0.005)		(0.008)
FOlog93		−0.006		0.890***		−0.007
		(0.012)		(0.007)		(0.012)
URlog93		−0.026***		−0.018***		0.830***
		(0.010)		(0.006)		(0.009)
scale(Arable returns03)	0.297***	0.017	0.069***	−0.005	0.242***	0.034**
	(0.028)	(0.015)	(0.024)		(0.029)	(0.015)
scale(Pasture returns03)	−0.145***	−0.002	−0.110***	−0.010***	−0.132***	−0.029**
	(0.026)	(0.005)	(0.023)	(0.003)	(0.024)	(0.013)
scale(Forest returns03)	−0.040	0.028*	0.170***	−0.0001	0.067***	−0.016
	(0.025)	(0.017)	(0.022)		(0.024)	(0.019)
scale(POP03)	−0.164***	−0.037***	−0.026	−0.011	0.113***	0.010
	(0.022)	(0.013)	(0.018)	(0.008)	(0.021)	(0.011)
scale(Elevation)	−0.652***	−0.069	−0.460***	−0.132***	−0.564***	−0.063*
	(0.075)	(0.043)	(0.057)	(0.024)	(0.076)	(0.038)
scale(Slope)	−0.309***	−0.116***	0.357***	0.069***	0.029	0.045*
	(0.051)	(0.030)	(0.039)	(0.019)	(0.067)	(0.027)
scale(WHC)	0.197***	0.053**	−0.146***	0.026	−0.027	−0.021**
	(0.046)	(0.026)	(0.039)	(0.016)	(0.034)	(0.010)
scale(Soil depth)	−0.131***	−0.028	0.117***	−0.038**	0.031	0.006
	(0.046)	(0.026)	(0.039)	(0.017)	(0.056)	
scale(Precipitations)	−0.248***	−0.038**	−0.005	−0.040***	−0.063*	−0.014
	(0.030)	(0.017)		(0.010)	(0.037)	(0.012)
scale(Temperature)	0.064	−0.090**	0.072*	−0.027	−0.143**	0.050***
	(0.078)	(0.036)	(0.040)	(0.026)	(0.060)	(0.017)
scale(Humidity)	−0.094*	−0.117***	0.034	−0.026	−0.209***	−0.015**
	(0.057)	(0.027)	(0.028)	(0.019)	(0.050)	(0.006)
scale(Radiation)	−0.157**	0.042	−0.296***	0.012	0.143**	−0.011
	(0.071)	(0.035)	(0.043)	(0.027)	(0.063)	
Constant	−0.275***	−0.036	−0.058***	0.053***	−0.982***	0.081***
	(0.024)	(0.023)	(0.019)	(0.013)	(0.043)	(0.029)
Observations	3,767	3,767	3,767	3,767	3,767	3,767
σ^2	1.721	0.580	1.265	0.201	1.513	0.396
Akaike Inf. Crit.	12,962.830	8,684.350	11,939.190	4,694.791	12,403.390	7,243.951
Wald Test (df = 1)	1,091.723***	106.356***	2,162.109***	207.793***	479.396***	69.676***
LR Test (df = 1)	845.807***	97.251***	1,397.558***	173.258***	405.499***	60.486***
Note:	*p<0.1; **p<0.05; ***p<0.01					
	scaled variables. Reference= Pastures					

A.7 Raw results from SDM

Table 14: Spatial Durban Models of land use on 1993–2003

	Arable Share		Forest Share		Urban Share	
	long run	short run	long run	short run	long run	short run
ARlog93		0.831***		−0.021***		0.005
		(0.009)		(0.006)		(0.013)
FOlog93		−0.010		0.893***		−0.021**
				(0.006)		(0.010)
URlog93		−0.019*		−0.017***		0.834***
		(0.010)		(0.005)		(0.008)
scale(Arable returns03)	0.342***	0.079***	−0.119	−0.055***	0.148**	0.048**
	(0.050)	(0.027)		(0.017)	(0.058)	(0.024)
scale(Pasture returns03)	0.005	0.004	0.042	0.033***	−0.015	−0.029***
	(0.014)			(0.004)	(0.030)	(0.011)
scale(Forest returns03)	−0.031	0.011	−0.039	0.044	0.036	0.126***
	(0.048)			(0.037)	(0.052)	(0.048)
scale(POP03)	−0.100***	−0.016	−0.011	−0.011	0.115***	0.009
	(0.029)			(0.008)	(0.021)	(0.014)
scale(Elevation)	−0.768***	−0.052	−0.476***	−0.094	−0.831***	−0.137*
	(0.111)	(0.085)	(0.097)		(0.120)	(0.076)
scale(Slope)	−0.443***	−0.155	0.603***	0.076***	0.055	0.067
	(0.070)		(0.058)	(0.012)	(0.098)	(0.048)
scale(WHC)	0.226***	0.047	−0.165	−0.027	0.028	−0.003
	(0.070)				(0.063)	
scale(Soil depth)	−0.176***	−0.014	0.106	0.014	−0.065	0.001
	(0.065)				(0.067)	(0.002)
scale(Precipitations)	−0.203***	−0.022	0.239***	0.006	−0.129*	−0.068***
	(0.055)		(0.052)		(0.074)	(0.026)
scale(Temperature)	1.086***	0.286	0.376***	0.050	0.399***	0.009
	(0.160)		(0.138)		(0.119)	
scale(Humidity)	−0.211	−0.147***	0.301***	0.026	−0.026	0.033
	(0.136)	(0.018)	(0.036)		(0.060)	(0.065)
scale(Radiation)	−0.206	−0.080	−0.541***	−0.019	0.189	0.113
	(0.171)		(0.158)			(0.073)
Constant	−0.242***	−0.109***	−0.058***	0.061	−0.929***	0.013
	(0.023)	(0.019)	(0.017)		(0.044)	(0.028)
Observations	3,767	3,767	3,767	3,767	3,767	3,767
σ^2	1.619	0.571	1.236	0.197	1.476	0.389
Akaike Inf. Crit.	12,803.420	8,648.834	11,865.140	4,671.029	12,349.980	7,223.800
Wald Test (df = 1)	1,278.793***	4.324**	2,047.633***	69.960***	516.272***	49.006***
LR Test (df = 1)	904.656***	4.175**	1,307.732***	68.563***	424.817***	47.962***

Note: *p<0.1; **p<0.05; ***p<0.01
on scaled variables. Reference= Pastures, coefficients from spatially lagged explanatory variables are not reported

A.8 Raw results from individual MNL

Table 15: Individual mnl models on 1993–2003

	Long Run			Short Run		
	arable share	forest share	urban share	arable share	forest share	urban share
U93PSTUR				−1.861***	−3.032***	−3.590***
				(0.008)	(0.013)	(0.017)
U93ARBLE				1.592***	−3.120***	−2.548***
				(0.009)	(0.035)	(0.025)
U93FORST				−1.477***	3.939***	−1.217***
				(0.043)	(0.019)	(0.041)
U93URBAN				−1.245***	−1.315***	2.865***
				(0.054)	(0.059)	(0.028)
Arable returns03	0.495***	0.332***	0.391***	0.288***	0.170***	0.252***
	(0.005)	(0.005)	(0.008)	(0.007)	(0.012)	(0.013)
Pasture returns03	−0.269***	−0.308***	−0.257***	−0.143***	−0.237***	−0.199***
	(0.005)	(0.005)	(0.007)	(0.006)	(0.012)	(0.013)
Forest returns03	0.006	0.335***	0.070***	0.034***	0.181***	−0.049***
	(0.005)	(0.004)	(0.007)	(0.006)	(0.010)	(0.013)
POP03	−0.615***	−0.122***	0.120***	−0.262***	−0.047***	0.046***
	(0.013)	(0.008)	(0.005)	(0.013)	(0.008)	(0.005)
Elevation	−0.903***	−0.224***	−0.533***	−0.616***	−0.153***	−0.275***
	(0.012)	(0.007)	(0.017)	(0.017)	(0.019)	(0.029)
Slope	−0.224***	0.148***	0.034***	−0.136***	0.141***	−0.005
	(0.009)	(0.005)	(0.011)	(0.012)	(0.012)	(0.019)
WHC	0.262***	−0.238***	0.091***	0.157***	−0.089***	0.009
	(0.008)	(0.008)	(0.012)	(0.010)	(0.020)	(0.022)
Soil depth	−0.162***	0.204***	0.019	−0.082***	0.077***	0.031
	(0.007)	(0.008)	(0.012)	(0.010)	(0.019)	(0.022)
Precipitations	−0.453***	0.078***	−0.122***	−0.324***	0.018*	−0.091***
	(0.005)	(0.004)	(0.008)	(0.008)	(0.010)	(0.014)
Temperature	0.088***	0.027***	−0.331***	0.022	−0.083***	−0.125***
	(0.011)	(0.008)	(0.016)	(0.015)	(0.020)	(0.028)
Humidity	−0.058***	−0.240***	−0.549***	−0.005	−0.394***	−0.407***
	(0.009)	(0.006)	(0.012)	(0.012)	(0.016)	(0.022)
Radiation	−0.066***	−0.208***	0.496***	−0.103***	0.172***	0.390***
	(0.011)	(0.009)	(0.016)	(0.015)	(0.022)	(0.029)
Constant	−0.286***	−0.060***	−1.629***			
	(0.005)	(0.004)	(0.007)			
Akaike Inf. Crit.	1,160,067.000	1,160,067.000	1,160,067.000	413,591.400	413,591.400	413,591.400
Note:	*p<0.1; **p<0.05; ***p<0.01					
	on scaled variables. Reference= Pastures					

A.9 Maps at the aggregate scale

Figure 2: Aggregated land use shares in 2003

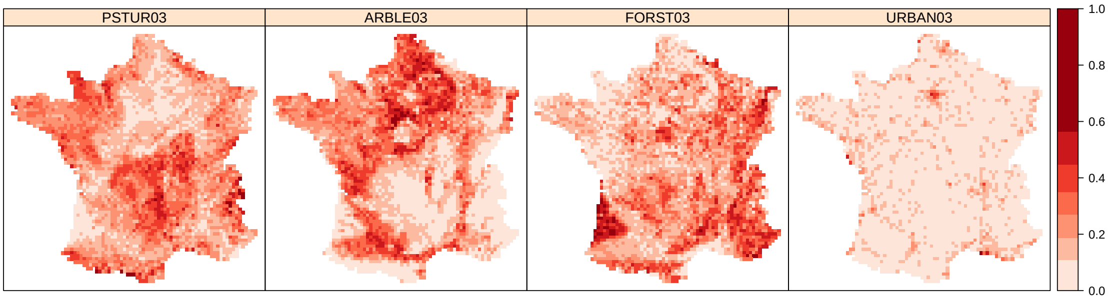


Figure 3: Aggregated land use variations on 1993–2003, in km^2

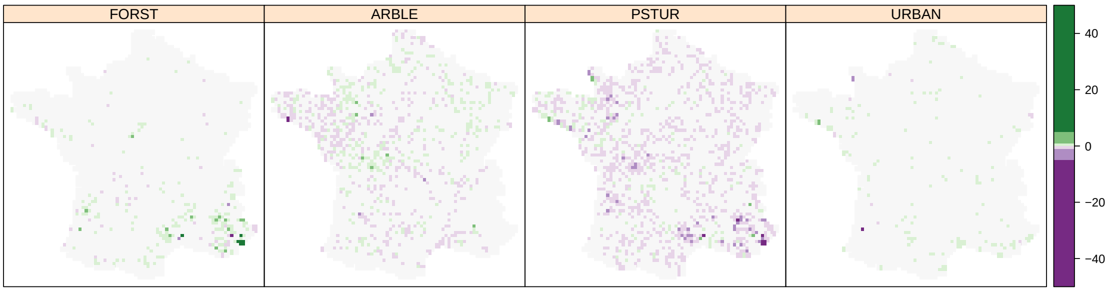
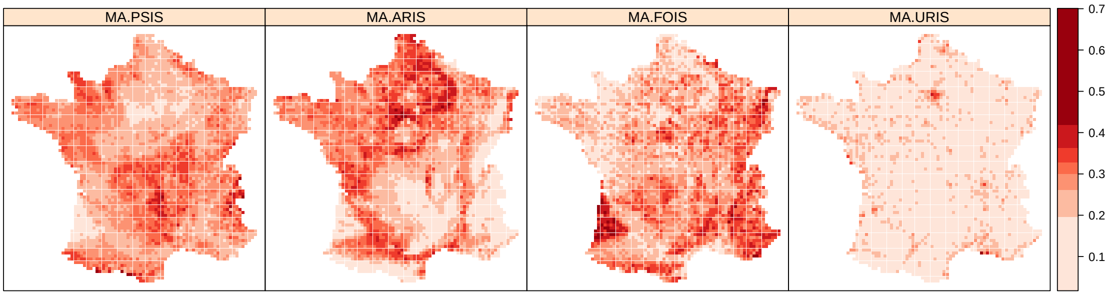


Figure 4: Out of sample 2003 predictions from individual mnl



A.10 Aggregate outcome variables

Figure 5: Raw distribution of 1998 aggregate land use shares

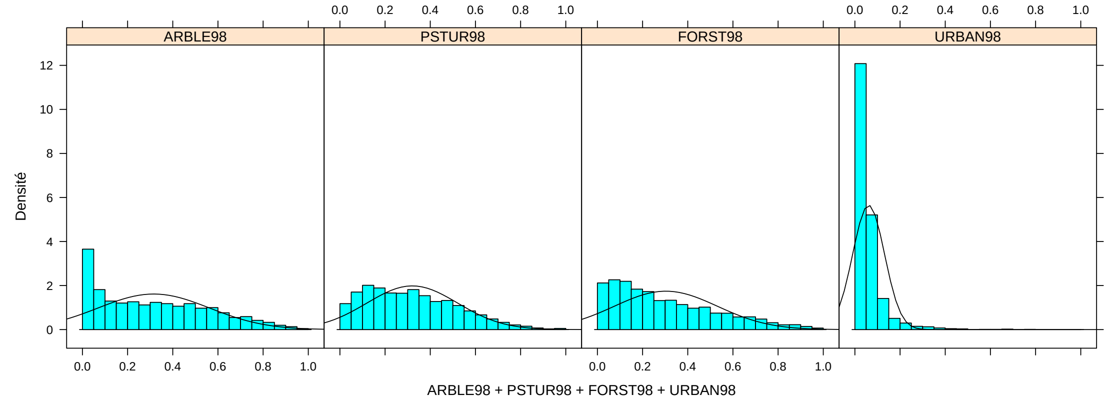
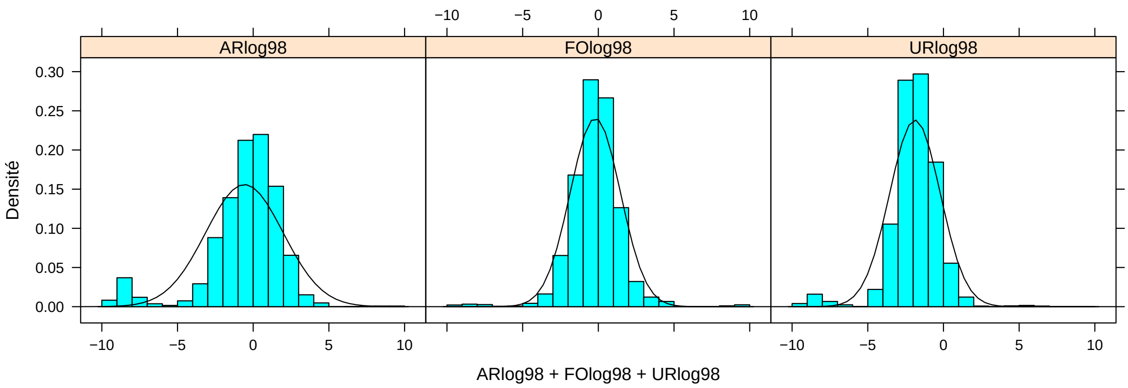


Figure 6: Linearized distribution of 1998 aggregate land use shares



A.11 Spatial Smoothing Functions

Figure 7: Semi-parametric smoothing functions of geographical coordinates: without temporal lags

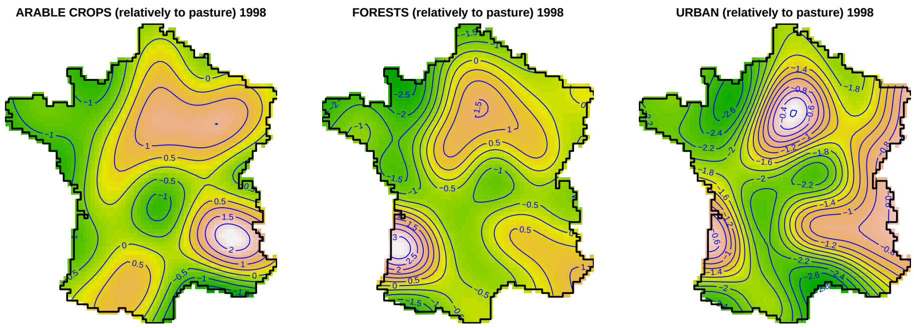
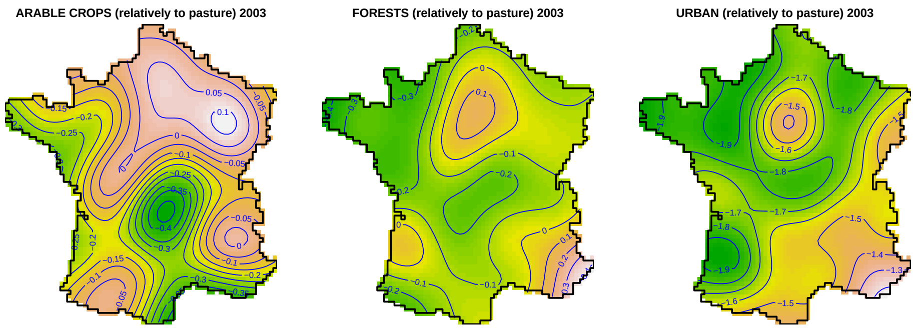


Figure 8: Semi-parametric smoothing functions of geographical coordinates: with temporal lags



A.12 Morans' *I* on residuals

Figure 9: Morans' *I* from OLS and GAM without temporal lags

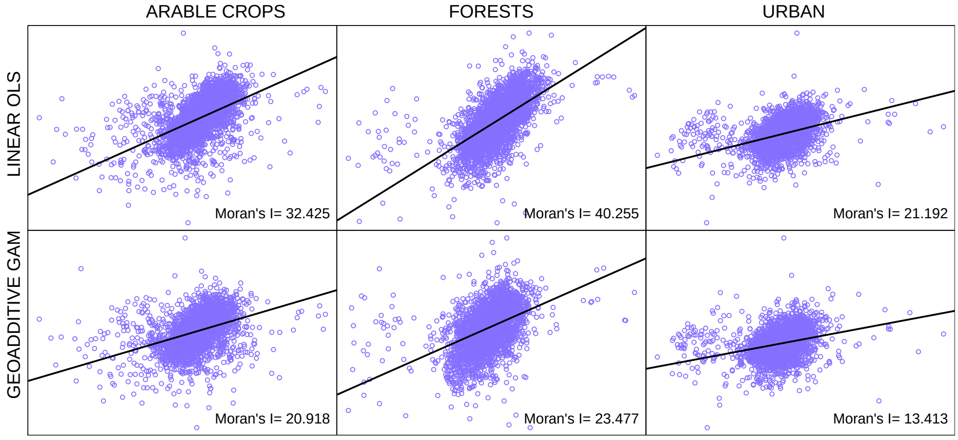


Figure 10: Morans' *I* from OLS and GAM with temporal lags