Aggregated *versus* individual land-use models: Modeling spatial autocorrelation to increase predictive accuracy

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

Abstract

The objective of this paper is to compare the predictive accuracy of individual and aggregated econometric models of land-use choices. We argue that modeling spatial autocorrelation is a comparative advantage of aggregated models due to the smaller number of observation and the linearity of the outcome. The question is whether modeling spatial autocorrelation in aggregated models is able to provide better predictions than individual ones. We consider a complete partition of space with four land-use classes: arable, pasture, forest, and urban. We estimate and compare the predictive accuracies of individual models at the plot level (514,074 observations) and of aggregated models at a regular 12 $\times$ 12 km grid level (3,767 observations). Our results show that modeling spatial autocorrelation allows to obtain more accurate predictions at the aggregated level when the appropriate predictors are used.

Keywords: Discrete choice modeling, spatial econometrics, predictive accuracy, scale effects.

JEL Classifications: Q15, Q24, C21.

Running title: Predictive accuracy of land use models.

*INRA, UMR 1041 CESAER, 25 boulevard Dr Petitjean, 21000 DIJON, France. jsay@dijon.inra.fr
†Economie Publique, INRA, AgroParisTech, Université Paris-Saclay, 78850 Thiverval-Grignon, France.
‡AgroSup Dijon, UMR 1041 CESAER, 25 boulevard Dr Petitjean, 21000 DIJON, France.

1 Introduction

Land-Use Changes (LUC) have significant economic and environmental impacts with implications for a wide variety of policy issues including food security, wildlife conservation, housing supply, and carbon sequestration [65, 12]. Given these impacts and the expected LUC in the next decades, prospective analysis requires a thorough understanding of how economic mechanisms and policy decisions affect LUC patterns [50, 40, 71].

LUC econometric models can be classified in two general groups based on their use of aggregated or individual data.[1] Due to the scarcity and cost of access to individual level data, most studies in the literature have been based on aggregated data for a region, a country, or other geographic scales [56, 55, 17]. Recent studies increasingly rely on individual data and consider discrete plot-level choices as the outcome of interest [43, 41, 7]. Thanks to individual data availability and to methodological advances, the estimation of individual LUC models has become easier. However, aggregated LUC models are still appealing as the current trend in LUC modeling is definitely towards global models at very large scales [37]. Well-known examples where global LUC models are needed are tropical deforestation, agricultural expansion, intensification, and food security [25].

The comparative advantages between individual and aggregated models in terms of predictive accuracy remains an open question with mixed evidence in the literature. In a seminal paper, [27] have examined the relative power of individual (micro) and aggregated (macro) models for explaining aggregated outcomes and found that the aggregated model is often better. [70] examine this issue in the context of predicting LUC. They show that, even for linear models, the choice between micro or macro scale to make aggregate predictions cannot generally be resolved by *an priori* reasoning. [58] show that modeling spatial autocorrelation is the most effective way to predict LUC at the individual scale, while limited to small study areas or small densities of observations. Here, we consider modeling spatial autocorrelation as a comparative advantage of ag-

[1] In the literature, "individual data" corresponds both to "sample plot" and "parcel level" data (see for example [54]). Our "individual" data are "sample plot" data, also called "micro level data" by [66] or "disaggregate data" by [13]. For more reviews of land-use modeling, see [30, 29].

gregated models due to their smaller number of observation and the linearity of their outcome. The question is whether this possibility of using more advanced econometric tools could provide better predictions than aspatial individual models.[2] These comparisons are made at the same aggregated scale, both in terms of in- and out-of-sample predictive accuracy. In the sake of exhaustivity, we also differentiate LUC models according to their temporal horizons, namely short run and long run models (respectively called land-use change models and land allocation models by [57])

With respect to the spatial dimension, the vast majority of past studies assumes spatial independence of land-use choices, both at aggregated and individual scales. Recent exceptions include: [18, 61, 17, 22, 41]. Incorporating spatial autocorrelation into land-use models raises several issues related to econometric estimation, hypothesis testing and prediction [5, 16]. This is even more challenging in the case of individual multinomial land-use models since the introduction of spatial dependence renders discrete choice models analytically intractable. The estimation of such models requires the use of simulation or Bayesian techniques [23].[3] Consequently, we consider in this paper only spatially independent models at the individual scale but we introduce spatial autocorrelation in some of the aggregated models.

This paper contributes to the LUC literature in two ways. Firstly, we explicitly introduce a Ricardian framework to formalize the possibility of using land price as a measure of the economic returns from land. This provides a consistent framework to compare individual, discrete choice models and aggregated, log-linearized models. Moreover, this highlights the fundamental role of conversion costs between land uses in differentiating short run and long run LUC models. Secondly, we show how introducing spatial autocorrelation in aggregated models enables better predictions than individual, aspatial models with higher number of observations. This is a very important result for

[2]In the following of this paper, we call spatial models those that model spatial autocorrelation explicitly. Conversely, we call aspatial models those that can include spatial effects without explicitly estimating spatial autocorrelation.

[3]Other estimation procedures have also been proposed in the literature: EM method [47], the generalized method of moments [53], the method of maximum pseudo-likelihood [62] and the method composite maximum likelihood [22, 61]. For a detailed review of the inclusion of spatial autocorrelation in discrete choice models see [23] and [62].

empirical work in LUC modeling as we show that it may not be worth using individual land use data when the sole objective is to predict aggregated land use.

The paper is organized as follows. In section 2, we set up the micro-economic Ricardian framework to motivate our empirical specifications. In section 3, we present the in-sample and out-of-sample formulas available to predict LUC. In section 4, we present the data and in section 5 the results both in terms of estimated parameters and predictive accuracies. The last section 6 reports a summary and the conclusions.

2 Econometric models of LUC

2.1 Individual models

Following the consensus of the econometric literature about LUC [63, 56, 44], we consider a risk-neutral landowner facing the choice of allocating a parcel of land of uniform quality to a use ℓ among a set of L alternatives. Conditionally on previous land use l, this stylized landowner i chooses at time t the use ℓ_{it}^* that provides the highest utility. Accordingly, this choice is driven by the following program:

$$\ell_{it}^* = \arg\max_{\ell} \left\{ u_{il\ell t} \;\middle|\; \ell_{i(t-1)} = l \right\}. \tag{1}$$

The current utility depends on previous land use l because of conversion costs of changing land use that lead to temporal autocorrelation of individual choices [44, 40]. As [64] states, the two major implications of this random utility framework – utilities are ordinal and only differences in utilities matter – are in accordance with the standard economic theory. Therefore, this discrete choice framework is fairly general, the strongest restrictions come from the parametrization of the utility functions that is necessary for the application to the data. If the utility from land use is the present discounted value of the stream of expected net benefits from the land, [56] shows that landowners choose the use with the highest expected one-period return at time t, minus the current one-period opportunity cost of conversion. This reads as: $u_{il\ell t} = \mathrm{E}(r_{i\ell t}) - c_{l\ell}$, where E is the expectation operator evaluated at t, $r_{i\ell t}$ is the one-period return associated to land use ℓ for i at t and $c_{\ell l}$ is the cost of conversion from use l to use ℓ. We consider these bilateral conversion costs to be constant in time and independent from the attributes of the plot i.[4] By noting $d_{il(t-1)}$ a dummy variable equals to 1 if the plot i is in use l at $(t-1)$ and 0 otherwise, we can simplify the utilities $\forall i, \ell, t$:

$$u_{i\ell t} = \sum\nolimits_{l=1...L} d_{il(t-1)} u_{il\ell t} = \mathrm{E}(r_{i\ell t}) - \sum\nolimits_{l=1...L} d_{il(t-1)} c_{l\ell}. \tag{2}$$

In contrast to the above random utility framework, which is shared by most of the econometric LUC literature, the relevant data used to proxy expected one-period returns are very heterogeneous in previous studies. It is clear that obtaining precise data about expected returns for each use and each plot of land is challenging.[5] Here, following [7], we match data about land price with the Ricardian formula to proxy the expected returns.[6]

For each land use ℓ, the Ricardian formula states that the observed land price $\overline{r}_{\ell t}$ at t for a given land plot is equal to the net present value of all expected future returns. We note τ the discount factor and κ_ℓ the expected growth rate of return from use ℓ as such $\mathrm{E}(r_{\ell t}) = \kappa_\ell \times r_{\ell(t-1)}$. This leads to a proportional relationship between land price and expected one-period return:

$$\overline{r}_{\ell t} = \sum_{s=0}^{\infty} \frac{\mathrm{E}(r_{\ell(t+s)})}{(1+\tau)^s} = \frac{1+\tau}{1+\tau-\kappa_\ell} \mathrm{E}(r_{\ell(t+1)}). \tag{3}$$

This shows that land prices can be used to substitute expected returns in the above specification (2) of utility, without ensuring that returns are perfectly observed. Typically, available data contain more precise biophysical variables (land quality, topography, climate) that might also affect the landowner's returns, as they are some non-economic determinants of utility. Hence, we specify the expected returns as $\mathrm{E}(r_{i\ell t}) =$

[4] These assumptions can be relaxed in the empirical part by including interactions between explanatory variables or by specifying random coefficients.

[5] For the United States, [45] makes an important data gathering effort to construct county-level returns for crops, pastures, forests and urban. The other studies use more partial information about returns, in very heterogeneous ways depending on the research questions.

[6] We argue that data about land price are in general available at fine spatial scales, it is at least true for our case study of France.

$\mathbf{b}_i^\top \boldsymbol{\gamma}_\ell^B + \overline{r}_{g(i)\ell t}\gamma_\ell^R + \varepsilon_{i\ell t}$ where $\mathbf{b}_i$ is a vector of perfectly observed biophysical variables, $g(i)$ is the unit corresponding to i in the scale g of land price data availability, and $\varepsilon_{i\ell t}$ represents the random deviations from the average values due to unobserved variables. Substituting this approximation of expected net returns in Equation 2 allows us to obtain the following reduced form for utilities:

$$u_{i\ell t} = \mathbf{d}_{i(t-1)}^\top \boldsymbol{\gamma}_\ell^D + \overline{\mathbf{r}}_{g(i)t}^\top \boldsymbol{\gamma}_\ell^R + \mathbf{b}_i^\top \boldsymbol{\gamma}_\ell^B + \varepsilon_{i\ell t},\ \ell = 1, \dots, L. \tag{4}$$

The vector $\mathbf{d}_{i(t-1)}$ binds the indicator functions $d_{i\ell(t-1)}$ for $\ell = 1 \dots L$ such that bilateral conversion costs are identified through the associated parameter vector $\boldsymbol{\gamma}_\ell^D$. Neglecting one-shot conversion costs amounts to consider utility at a longer run time horizon, such specification without lagged land use are called land allocation models by [57]. The $(L \times 1)$ vector $\overline{\mathbf{r}}_{g(i)t}$ contains the L land prices corresponding to the different land uses, and the $(K \times 1)$ vector $\mathbf{b}_i$ binds the K biophysical variables described in greater details in the data section. The vectors $\boldsymbol{\gamma}_\ell^D$, $\boldsymbol{\gamma}_\ell^R$ and $\boldsymbol{\gamma}_\ell^B$ are the unknown vectors of parameters to be estimated of respective dimensions $(L \times 1)$, $(L \times 1)$ and $(K \times 1)$. By identification with Equation 3, we have $\gamma_\ell^R = 1 - (\kappa_\ell/(1+\tau))$ in order to identify (up to a constant discount factor τ) the expected growth rates κ_ℓ through the utility parameters $\boldsymbol{\gamma}_\ell^R$ associated to land prices.

The stochastic dimension of this framework is only related to the unobserved components $\varepsilon_{i\ell t}$ and their associated densities. [46] considers three standard assumptions about error terms that allow one to obtain a multinomial logit model: independence, homoskedasticity and extreme value distribution. On the basis of these assumptions, one can show that the probabilities of observing 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, $\overline{u}_{i\ell t} \equiv u_{i\ell t} - \varepsilon_{i\ell t}$:

$$p_{i\ell t} = \frac{\exp(\overline{u}_{i\ell t})}{\sum_{l=1}^{L} \exp(\overline{u}_{ilt})} \tag{5}$$

2.2 Aggregated models

There is an important literature on econometric LUC models estimated on aggregated data: [42, 63, 72, 56, 48] are the most significant papers. The underlying micro-economic theory is identical to that in the previous section, but individual choices are aggregated in order to estimate land-use shares models instead of discrete choice models. This process of aggregation is generally considered as a loss of information through a drastic decrease in the number of observations. A direct aggregation of land use from the parcel i to the grid g implies considering land use shares $S_{g\ell t}$. The land use share $S_{g\ell t}$ is computed as the share of parcels in the grid g with land use ℓ at time t.

Based on the same economic rationales than individual models, the observed shares of land use ℓ in t for $g = 1, \ldots, G$ are generally expressed as ($\forall \ell = 1, \ldots, L$):

$$S_{g\ell t} = \frac{\exp\left(\mathbf{S}_{g(t-1)}^{\top}\boldsymbol{\beta}_{\ell}^{D} + \overline{\mathbf{R}}_{gt}^{\top}\boldsymbol{\beta}_{\ell}^{R} + \mathbf{B}_{g}^{\top}\boldsymbol{\beta}_{\ell}^{B}\right)}{\sum_{l=1}^{L}\exp\left(\mathbf{S}_{g(t-1)}^{\top}\boldsymbol{\beta}_{l}^{D} + \overline{\mathbf{R}}_{gt}^{\top}\boldsymbol{\beta}_{l}^{R} + \mathbf{B}_{g}^{\top}\boldsymbol{\beta}_{l}^{B}\right)}. \tag{6}$$

The meanings and dimensions of these variables are the same as in the previous subsection, capital letters represent aggregated values (averaged from individual data) and vectors $\boldsymbol{\beta}_{\ell}^{D}$, $\boldsymbol{\beta}_{\ell}^{R}$ and $\boldsymbol{\beta}_{\ell}^{B}$ are the unknown vectors of parameters to be estimated, of respective dimensions $(L \times 1)$, $(L \times 1)$ and $(K \times 1)$. Aggregating the dummy vector $\mathbf{d}_{i(t-1)}$ consists of computing land use shares $\mathbf{S}_{g(t-1)}$ from previous $(t-1)$ period as explanatory variables. The elements $\overline{\mathbf{R}}_{gt}$ and $\mathbf{B}_{g}$ are not indexed by ℓ since we use the same explanatory variables in all equations.

When estimating aggregated LUC models, one needs to handle two specific issues which arise for dependent variables as shares or proportions. The first issue concerns the bounded nature of shares in which zeros and ones may appear. The second issue concerns the adding-up constraint as the land use shares have to sum up to one. The most common strategy in the literature is to specify the shares as logistic functions [72, 17]. This has the advantage of being linearly tractable thanks to the logit-linear transformation [73].[7]

[7] We have also considered two other specifications that have been proposed in the literature to deal with this problem: the fractional logit model proposed by [52] and the fractional Dirichelet model pro-

In terms of logistic shares, we note $\widetilde{S}_{g\ell t} \equiv \log(S_{g\ell t}/S_{glt})$ the natural logarithm of each observed share ℓ normalized by a reference land use l. The aggregated land use share model is approximately:[8]

$$\widetilde{S}_{g\ell t} \approx \widetilde{S}_{g\ell(t-1)}\beta_\ell^D + \overline{\mathbf{R}}_{gt}^\top \boldsymbol{\beta}_\ell^R + \mathbf{B}_g^\top \boldsymbol{\beta}_\ell^B + \xi_{g\ell t} \;\; \forall \ell \neq l. \tag{7}$$

With L land use categories, the system has $L-1$ equations. Again, the elements $\overline{\mathbf{R}}_{gt}$ and $\mathbf{B}_g$ do not have an index ℓ since we use the same explanatory variables in all equations. A Seemingly Unrelated Regressions approach could also be adopted to allow for correlated errors between equations of the system [20], but [17] show that allowing for inter-equation correlations does not improve the predictive accuracy of such models. Therefore, to simplify the results, we skip this aspect and estimate the linearized Equation 7 by Ordinary Least Squares (OLS) first. The spatial dimension of the observations can easily be introduced in these aggregated 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 [28, 69] also used in the literature [68, 67]. Because in this case, spatial autocorrelation is not modeled explicitly, we do not consider such models as being spatial models but we nevertheless include them in our comparative set.

2.3 Spatial autocorrelation

The spatial econometric literature is extensive [19, 4, 39, 6] and provides a number of ways to deal with spatial autocorrelation. Nevertheless, introducing spatial dependence in discrete choice models is still problematic econometrically, especially for high numbers of observations [23, 62, 60]. Indeed, an important consequence of introducing spatial autocorrelation in discrete choice models is the complex covariance structure

posed by [49]. As they do not perform better that the specifications included in the paper, we do not report the results but they are available from the authors upon request.

[8]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 (7). This is a minor drawback that can be visually evaluated from Appendix A.9. More rigorously, it will be also evaluated by comparing the predictions with those from other models, as we consider this as a necessity for estimating linearized logistic models, always used in the literature.

due to heteroskedasticity and the necessity to linearize the objective functions [35, 41]. Moreover, it implies high dimension integrals in order to compute the likelihood function [3] and relies on complex optimization algorithms, dependent on starting values and tolerances of the fixed-point iterations for the generalized method of moments [36]. To avoid such complications associated with spatial autocorrelation in discrete choice models, we focus in this paper on introducing spatial autocorrelation in the aggregated land use models only. By design of our research question, the possibility of modeling spatial autocorrelation is a comparative advantage of aggregated models, and this advantage is evaluated in terms of predictive accuracy.

For individual models, wrongly omitting spatial autocorrelation has two consequences that potentially impact predictive accuracy. This depends on the type of autocorrelation. In the case of omitted spatial lag of the endogenous variable, the estimated coefficients are biased. In the case of omitted spatial errors, the estimated coefficients are inconsistent. However, because modelling spatial autocorrelation is challenging for the reason mentioned above, we propose an evaluation of these negative consequences by comparison with aggregate spatial econometric models.

In the context of aggregated LUC models, we introduce spatial autocorrelation through three additional terms corresponding to three spatial mechanisms that will be formally tested with model estimation [39]. Let $\mathbf{W}$ be a $G \times G$ spatial weight matrix, which summarizes the spatial connectivity structure of the aggregated observations. Because this matrix is row-standardized, it returns the weighted average of the values of the neighbors of each observations once multiplied to a variable. We first consider the most general spatial econometric model applied to aggregated land-use shares (that we call SMC for spatial autoregressive mixed conditional) that can be written as ($\forall \ell \neq l$):

$$\begin{aligned} \widetilde{S}_{\ell t} \approx \rho_\ell \mathbf{W} \widetilde{S}_{\ell t} + \beta_\ell^D \widetilde{S}_{\ell(t-1)} + \theta_\ell^D \mathbf{W} \widetilde{S}_{\ell(t-1)} \\ + \overline{\mathbf{R}}_t \boldsymbol{\beta}_\ell^R + \mathbf{W} \overline{\mathbf{R}}_t \boldsymbol{\theta}_\ell^R + \mathbf{B} \boldsymbol{\beta}_\ell^B + \mathbf{W} \mathbf{B} \boldsymbol{\theta}_\ell^B + \lambda_\ell \mathbf{W} \xi_{\ell t} + \eta_{\ell t} \end{aligned} \tag{8}$$

The first included spatial terms (in the RHS of the first row) are related to simultaneous and time-lagged interactions that lead land use at t to depend upon neighboring land uses at t and $t-1$ respectively. Hence, we include both $\rho_\ell \mathbf{W}\widetilde{S}_{\ell t}$ and $\theta_\ell^D \mathbf{W}\widetilde{S}_{\ell(t-1)}$ in the model, where $\widetilde{S}_{\ell t}$ and $\widetilde{S}_{\ell(t-1)}$ are $(G \times 1)$ vectors containing the G observations for $\widetilde{S}_{glt}$ and $\widetilde{S}_{gl(t-1)}$. The parameters $|\rho_\ell| < 1$ and $|\theta_\ell^D| < 1$ represent the intensity of, respectively, simultaneous and time-lagged spatial lag dependence. The second term is related to spatial error autocorrelation specified from Equation 7 as following: $\xi_{\ell t} = \lambda_\ell \mathbf{W}\xi_{\ell t} + \eta_{\ell t}$, with $|\lambda_\ell| < 1$ denoting the strength of spatial error dependence with $\eta_{\ell t}$ an iid error term. The third and last spatial terms are related to the influence from neighborhood characteristics, and is modeled by adding the spatially lagged exogenous variables: $\mathbf{W}\overline{\mathbf{R}}_t\theta_\ell^R$ and $\mathbf{W}\mathbf{B}\theta_\ell^B$ in the regression functions, where $\overline{\mathbf{R}}_t$ is a $(G \times L)$ matrix containing the G observations for $\overline{\mathbf{R}}_{gt}$ and $\mathbf{B}$ is a $(G \times K)$ matrix containing the G observations for $\mathbf{B}_g$ over K explanatory biophysical variables.

This model is sufficiently general so that a large range of dynamic spatial econometric models are nested. For instance, the spatial autoregressive conditional (SAC) model can be recovered with $\theta_\ell^D = \boldsymbol{\theta}_\ell^R = \boldsymbol{\theta}_\ell^B = 0$ (14, 15, also called SARAR(1,1) by 32), the spatial error model (SEM) can be recovered with $\theta_\ell^D = \boldsymbol{\theta}_\ell^R = \boldsymbol{\theta}_\ell^B = \rho_\ell = 0$, the spatial X model (SXM) with $\theta_\ell^D = \rho_\ell = 0$, the spatial autoregressive (SAR) model with $\theta_\ell^D = \boldsymbol{\theta}_\ell^R = \boldsymbol{\theta}_\ell^B = \lambda_\ell = 0$, and the spatial Durbin model (SDM) model can be recovered when $\theta_\ell^D = \lambda_\ell = 0$. Following [39], the SDM model is particularly suited to alleviate the impact of spatially autocorrelated omitted variables. This point is particularly attractive here because aggregated explanatory variables suffer of measurement errors by averaging, which is a special case of omitted variables. Based on maximum likelihood, we estimate all of these implied spatial specifications in the empirical part of the paper, we will also test the relative importance of the three sources of spatial autocorrelation and compare their implications in terms of predictive accuracy.

Finally, note that all the previously described models admit the temporal lag of the outcome variables at their right hand sides, they consider explicitly land use changes between $t-1$ and t. Predicting the long run equilibrium (or steady state) of land use is also of interest in applied works, and is of social and political concern. The long run equivalents of previous models correspond to the limit of the expectation of the endogenous variable when t tends to infinity. Using a similar argument to that of [39], it is possible to show that they correspond to a model similar to Equation 8 without the term $\widetilde{S}_{\ell(t-1)}$ as explanatory variable.

3 Performing predictions

For both individual and aggregated models, we perform in-sample and out-of-sample predictions. In our context, the first case consists in predicting the value of the dependent variable for the last year belonging to the sample used to estimate the models (1993–2003). For the second case, we restrict the sample on 1993–1998 and predict the value of the dependent variable for the same year 2003 that is not used to estimate the models. This is the most interesting case for policy implications due to the time lag usually required to produce and analyse the data. Researchers are rarely interested in predicting land use from a year of observation to the following year. A 5 year interval appears as a balanced choice, close to typical practices. In both cases, in-sample and out-of-sample predictions are performed for short-run and long-run models.

Introducing a panel structure in the models by using the annual observations between 1993–1998 and 1998–2003 raises the classical issue of the choice between fixed and random effects, in addition to the choice between spatial and aspatial models. But this additional choice is more difficult to evaluate in terms of predictive accuracy as we choose to do in this paper. In a fixed-effect model, the biophysical variables (elevation, slope, water holding capacity) will be "swept away" by the within estimator and the associated coefficients will not be identified. This implies that we need to substitute the well-used biophysical variables by unobserved heterogeneity identified from dummy variables to predict land use. In addition, out of sample predictions are not possible with year fixed effects. Models with random effects allow the inclusion of time-invariant variables but impose the absence of correlation between the individual random effects and the explanatory variables included in the model. This assumption is rarely verified in most empirical applications.

3.1 From individual models

For the individual MNL models, the predictions 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 assuming that each observation counts for 100 ha (in anticipation of our empirical application), the predicted probabilities can easily be converted into aggregated LUC. For example, consider a plot i which counts for 100 ha of annual crop in period $t-1$ 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 land-use shares is operated by multiplying the probabilities by 100 and summing the results at the aggregated scale of interest.

With the MNL approach, the predicted acreages of each use are always positive and assured to sum to the national available land base. We also estimate some linear probability models on individual data that do not take account of the discrete nature of LUC but are computationally less intensive. Within this framework, short run out-of-sample predictions for the next period are easily simulated. As it will be used in the application, putting the observed land-use dummies $\mathbf{d}_{it}$ in the regression equation (5) and changing the values of $\overline{\mathbf{r}}_{g_i t}$ to $\overline{\mathbf{r}}_{g_i(t+1)}$ allows one to obtain the vector $\widehat{\mathbf{p}}_{i(t+1)}$ of predictions and compute the aggregated land-use shares to be compared.

3.2 From aspatial aggregated models

Aspatial aggregated predictors are based on short-run and long-run models without spatial lag variables and without spatial error terms, i.e., Equation 6 and Equation 7. Their form are described in the first part of the following Table 1 both for the in-sample and the out-of-sample case. They only involve the knowledge of the current observations for the explanatory variables (for both short and long run models) and the previous values of the endogenous variable (for the short-run model). As the residuals are neither temporally nor spatially autocorrelated, these simple predictors are the best linear unbiased predictors. More specifically, we compute two aspatial predictors : (i) aspatial predictor on a simple land-use share model estimated by OLS and (ii) aspatial predictor on a land-use share model including a smoothed function of the geographical coordinates of the grids' centroids (GAM).

3.3 From spatial aggregated models

The issue of prediction in spatial econometric models has gained considerable attention in the last decade. For instance, [8] and [11] derive the Best Linear Unbiased Predictor (BLUP, see 26) for static spatial panel data models with random effects and [10] derive the BLUP for a dynamic spatial panel data model with random effects. When it comes to compare the predictive accuracies of models, [9] find that for 1 year ahead forecasts of the US states' demand for liquor, estimators taking into account spatial correlation and heterogeneity across states perform the best. [2] forecast employment in 50 Spanish provinces and show that a dynamic spatial lag panel data model outperforms a non-spatial dynamic panel model and that it is only slightly dominated by a seasonal ARIMA model. Based on spatial dynamic panel models, [34] make multi-step forecasts of the annual growth rates of the real GDP for 16 German länder and show that spatial effects substantially improve the forecast performance. Finally, [59] forecast unemployment levels for German labour-market district with a spatial GVAR model. Again, spatial models lead to better results compared to non-spatial ones, all at an aggregated scale.

The first category of spatial predictors that we use is labeled "reduced" spatial predictors. They are based on the range of spatial models described in Equation 8 by writing their reduced forms (i.e., factorizing the outcomes variable and putting all the endogenous terms on the LHS):

$$\begin{aligned}\widetilde{S}_{\ell t} \approx (I-\rho_\ell \mathbf{W})^{-1}\Big(&\beta_\ell^D \widetilde{S}_{\ell(t-1)} + \theta_\ell^D \mathbf{W}\widetilde{S}_{\ell(t-1)} \\ &+ \overline{\mathbf{R}}_t \boldsymbol{\beta}_\ell^R + \mathbf{W}\overline{\mathbf{R}}_t \boldsymbol{\theta}_\ell^R + \mathbf{B}\boldsymbol{\beta}_\ell^B + \mathbf{W}\mathbf{B}\boldsymbol{\theta}_\ell^B + \lambda_\ell \mathbf{W}\xi_{\ell t} + \eta_{\ell t}\Big)\end{aligned} \tag{9}$$

The reduced spatial predictors in the in-sample and out-of-sample cases are de-

Table 1: The predictors for aggregated models of LUC. We note $\mathbf{X}_t\boldsymbol{\beta} \equiv \overline{\mathbf{R}}_t\boldsymbol{\beta}_\ell^R + \mathbf{B}\boldsymbol{\beta}_\ell^B$. The table shows the different predictors used both for in-sample and out-of-sample predictions, and for the different time and scale structures.

TYPE	IN-SAMPLE	OUT-OF-SAMPLE
	Aspatial predictors	
LONG RUN	$\widehat{\widetilde{S}}_{\ell t} = \mathbf{X}_t\widehat{\boldsymbol{\beta}}$	$\widehat{\widetilde{S}}_{\ell(t+1)} = \mathbf{X}_{t+1}\widehat{\boldsymbol{\beta}}$
SHORT RUN	$\widehat{\widetilde{S}}_{\ell t} = \widehat{\beta}_\ell^D \widetilde{S}_{\ell(t-1)} + \mathbf{X_t}\widehat{\boldsymbol{\beta}}$	$\widehat{\widetilde{S}}_{\ell(t+1)} = \widehat{\beta}_\ell^D \widetilde{S}_{\ell t} + \mathbf{X}_{t+1}\widehat{\boldsymbol{\beta}}$
	Reduced spatial predictors	
LONG RUN	$\widehat{\widetilde{S}}_{\ell t} = (I - \widehat{\rho}_\ell \mathbf{W})^{-1}\mathbf{X}_t\widehat{\boldsymbol{\beta}}$	$\widehat{\widetilde{S}}_{\ell(t+1)} = (I - \widehat{\rho}_\ell \mathbf{W})^{-1}\mathbf{X}_{t+1}\widehat{\boldsymbol{\beta}}$
SHORT RUN	$\widehat{\widetilde{S}}_{\ell t} = (I - \hat{\rho}_\ell \mathbf{W})^{-1}(\widehat{\beta}_\ell^D \widetilde{S}_{\ell(t-1)} + \widehat{\theta}_\ell^D \mathbf{W}\widetilde{S}_{\ell(t-1)} + \mathbf{X}_t\widehat{\boldsymbol{\beta}})$	$\widehat{\widetilde{S}}_{\ell(t+1)} = (I - \widehat{\rho}_\ell \mathbf{W})^{-1}(\widehat{\beta}_\ell^D \widetilde{S}_{\ell t} + \widehat{\theta}_\ell^D \mathbf{W}\widetilde{S}_{\ell t} + \mathbf{X}_{t+1}\widehat{\boldsymbol{\beta}})$
	Structural spatial predictors	
LONG RUN	$\widehat{\widetilde{S}}_{\ell t} = \widehat{\rho}_\ell \mathbf{W}\widetilde{S}_{\ell t} + \mathbf{X}_t\widehat{\boldsymbol{\beta}} + \widehat{\lambda}_\ell \mathbf{W}(\widetilde{S}_{\ell t} - \mathbf{X}_t\widehat{\boldsymbol{\beta}})$	$\widehat{\widetilde{S}}_{\ell(t+1)} = \widehat{\rho}_\ell \mathbf{W}\widetilde{S}_{\ell t} + \mathbf{X}_{t+1}\widehat{\boldsymbol{\beta}} + \widehat{\lambda}_\ell \mathbf{W}(\widetilde{S}_{\ell t} - \mathbf{X}_t\widehat{\boldsymbol{\beta}})$
SHORT RUN	$\widehat{\widetilde{S}}_{\ell t} = \widehat{\rho}_\ell \mathbf{W}\widetilde{S}_{\ell t} + \widehat{\beta}_\ell^D \widetilde{S}_{\ell(t-1)} + \widehat{\theta}_\ell^D \mathbf{W}\widetilde{S}_{\ell(t-1)} + \mathbf{X}_t\widehat{\boldsymbol{\beta}} + \widehat{\lambda}_\ell \mathbf{W}(\widetilde{S}_{\ell t} - \mathbf{X}_t\widehat{\boldsymbol{\beta}})$	$\widehat{\widetilde{S}}_{\ell(t+1)} = (\widehat{\rho}_\ell + \widehat{\theta}_\ell^D)\mathbf{W}\widetilde{S}_{\ell t} + \widehat{\beta}_\ell^D \widetilde{S}_{\ell t} + \mathbf{X}_{t+1}\widehat{\boldsymbol{\beta}} + \widehat{\lambda}_\ell \mathbf{W}(\widetilde{S}_{\ell t} - \mathbf{X}_t\widehat{\boldsymbol{\beta}})$

scribed in the second part of Table 1. As there is no temporal autocorrelation in the residuals, these predictors are BLUP for the out-of-sample case. Predictors of similar form have been used in [34] and [59]. They also correspond to the first unbiased predictor suggested by [31].[9] In the short-run case, these predictors necessitate information on the temporal lagged values of the endogenous variables for all the sample units under consideration while in the long run case, they only necessitate the observations of the explanatory variables.

The third category of predictors that we used are “structural” spatial predictors as they are based on the range of spatial models in structural form described in Equation 8. Their form in the in-sample and out-of-sample case are described in the third part of Table 1. The last term of these spatial predictors aims at accommodating spatial error autocorrelation when it is present. Note that in the in-sample case and when $\lambda_\ell = 0$, they correspond to the "trend-signal-noise" predictors used in the geo-statistical literature [14] and use information on the current spatially lagged endogenous variables. In the out-of-sample case, implementing these structural spatial predictors is not possible as they necessitate observations on the future spatially lagged values of the dependent variables, which are not available. As a consequence, we adopt a heuristic solution that consists in replacing these future spatially lagged values of the dependent variable by the current spatially lagged values of the dependent variable. Our justification in the specific case of LUC models is that land use presents strong temporal autocorrelation due to conversion costs.

[9] The case figure considered by [31] ("leave-one-out" predictors) and also by [51] ("ex-sample predictors") is different from ours as they are concerned with a particular case of prediction: the case where for a cross-section of observations, part of the observations on the dependent variable is missing and needs to be predicted.

4 Data

4.1 Land use data

Our data about land use are extracted from the TERUTI survey [1], which was carried out every year between 1993 and 2003 by the statistical services of the French Ministry of Agriculture [18, 17]. It contains data about land use through the whole continental territory of France and counts 550,903 land plots surveyed. 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 3,767 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 (each point is representative of 100 ha at NUTS 2 level). On the basis of the detailed classification of land uses (81 items), we attribute to each plot a use among four more aggregated items:[10] 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 4 presents the raw transitions 1993–2003, from the rows to the columns.

Table 4 shows that, in 2003, arable crops, pastures and forests each represented almost 30% of the continental France. It also shows that between 1993 and 2003, the pastures area declined by almost 5%, while arable, forest and urban uses increased by 2%, 3% and 14% respectively. As in the other studies, LUC presents a strong temporal autocorrelation, which comes from conversion costs but also inter-temporal decisions, land owner specializations, legislative constraints, etc.

The aggregation from these individual data is operated at the first stage of the sampling design, at the 12 $\times$ 12 km regular grid scale. As mentioned in footnote 8, the presence of zeros in the denominator is a drawback of the logit transformation

[10]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. Our final sample counts N=514,074 points.

for aggregate modeling that can be dealt with by adding a small value ϵ both in the numerator and the denominator. As Figure 4 and Figure 5 of the Appendix 8.3 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.

4.2 Explanatory variables

The theoretical literature on LUC and the micro-economic framework set out in the previous section suggest that the potential explanatory variables are the net returns from each land uses. As presented in subsection 2.1, we use current land prices to proxy the net returns to each land use according to the Ricardian formula.

Data on land prices are obtained from the statistical services of the French Ministry of Agriculture. Yearly prices 1990–2005 are available for arable crops and pastures. A central feature of these data is that the ministry only use sales between farmers and with a sure agricultural purpose to estimate yearly prices. So, it seems natural to assume that these prices correspond well to Equation 3, even if landowners are allowed to change the land use. For the other two non-agricultural land uses considered – forest and urban – the approximations of economic returns are computed differently. 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. Finally, we include some biophysical attributes: slope, altitude, water holding capacity (WHC), and climate. The following Table 5 displays summary statistics for these variables aggregated at the grid scale but they are initially available at the individual scale.

5 Results

5.1 Specifications

Our comparative set includes a wide spectrum of econometric LUC models of the literature. We estimate a total of 7 types of models at different scales, and with different spatial and temporal structures. Moreover, each specification is estimated two times. On the one hand, the short run models (with time-lagged land use as explanatory variable) are estimated on the 1993–1998 and 1993–2003 periods, and on the other hand, the long run models (without time-lagged land use) are estimated on 1998 and 2003 cross-sections. For the models estimated on 1998 and 1993–1998 periods, the predicted year (2003) is not used in the estimations. Hence, such predictions are considered out of sample.

To keep the models comparable, we use the same specifications for the effects of explanatory variables.[11] We include the explanatory variables (land prices and biophysical variables) additively, jointly with dummies about previous land use for individual short run models and previous land use share for aggregated short run models. We maintain the assumption of homogeneous conversion costs, again to ensure the comparability between the models: while nothing preclude applied researchers to include interaction or polynomials terms, we do not see any reason for our results to be dependent on this choice.

Detailed estimation results (coefficients and standard errors) based on maximum likelihood of different model specifications are provided in the Supporting Information 9 (SI). The explanatory variables are scaled to obtain standardized parameters, and we report in SI only the results of the models estimated over the period 1993–2003 (i.e., those used to make in-sample predictions). Because of their proximity to the displayed models, the raw results from the linear probabilities (close to the individual MNL), and the SAC and SMC (close to SAR and SDM) are not reported but are available upon

[11]One can consider that, because of the higher number of observations, the individual models can contain additional terms such as interactions between variables or polynomials. To keep our comparative exercise simple, we choose to not consider this possibility of more complex regression functions as a comparative advantage of individual models.

request.

5.2 Parameter estimates

5.2.1 Aspatial models

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 probabilities, 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) 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. However, in the literature, it has been been found that the use of nested multinomial logit does not change the main results [43, 41].

By comparing the significance of coefficients from aspatial models (Table 9 to 12 in the SI 9), a first important result is the strong effects of time-lagged land uses in short run models, indicating strong conversion costs and strong inertia between LUC. On aggregated models, OLS and GAM, the short run models present globally some R^2 close to 0.9 where the OLS long run models have respectively 0.66, 0.23 and 0.26 for arable, forest and urban land uses. Including geographical coordinates in the GAM increase substantially the R^2. For the long run models without temporal lag, the regional specializations of land use appear 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.

For the long run models (both individual and aggregated) and whatever the land use considered, we see that land prices are very significant and have the expected signs. We interpret these results as a quite robust validation of our Ricardian framework. An interesting point is the time-scale interaction for the effects of the economic variables of expected returns. From the long run model to the short run, the land prices keep their significance for individuals models but are not longer significant in the aggregated models. It is clearly related to the number of observation as this effect of the loss of significance is related to aggregation. This can be considered as an expression of multicolinearity in short run aggregated models that is not present in the short run individual models because they use more observations [33].

5.2.2 Spatial models

We estimated the spatial econometric models using maximum likelihood through the `R` package `spdep`. To avoid endogeneity problems, the spatial weight matrix $\mathbf{W}$ is based on purely geographical considerations, we use queen contiguity of order one for all models. 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 24, 21 or 38 for reviews of these spatial specification searches). Instead, as we do not have any theoretical prior as to the form taken by spatial autocorrelation, we estimate the full set of spatial models described in subsection 2.3 since spatial autocorrelation could arise from several sources. The summary measure of impacts (direct, indirect and total as defined in 39) are not reported here but are available upon request. Globally, it still 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.

Table 6 and Table 7 in the subsection 8.1 display the values 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 the spatial lag of the dependent variable (SAR, SDM) 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, SMC), 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 (between Table 6 and Table 7 in Appendix A.2), 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 SMC specifications.

5.3 Predictive accuracy

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 in 2003 at the aggregated grid level. The comparisons are reported in the panels A, B, Table 2 for in-sample predictions and in panels A, B, Table 3 for out-of-sample predictions. Each panel presents respectively the long run and the short run predictions for different specifications and predictors.

The general patterns of predictive accuracy are threefold: (i) the short run models present smaller RMSE than long run ones, (ii) the individual models do not present substantial smaller RMSE than aspatial aggregated models, and (iii) spatial aggregated models perform better than individual models in the long run, if the structural predictors are used. From these both in-sample and out-of-sample predictions, the advantage of using individual models is not found in terms of predictive accuracy whereas the advantage of using spatial econometrics techniques is clearly present.

Looking in more details the results of Table 2 and Table 3, it appears that using the structural predictors for the aggregated spatial models perform better than any other estimation techniques. The SAC and SMC models are exceptions but they still have

Table 2: In-sample Root Mean Square Errors for the different predictors: *The rows marked REF report the benchmark RMSE from constant predictions (i.e., the national shares for each plot). The columns MEAN report the row means of RMSE. OLS is for Ordinary Least Squares, GAM for Generalized Additive Models, 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, SMC 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.*

A. RMSE for predictors from long run models

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
Aggregate aspatial models					
OLS	0.1581	0.1589	0.1773	0.0666	0.1467
GAM	0.134	0.146	0.1603	0.0641	0.1314
Aggregated spatial models					
Reduced predictors					
SEM	0.1615	0.1603	0.1801	0.0669	0.1489
SXM	0.1555	0.1581	0.1741	0.0661	0.1448
SAR	0.1562	0.1566	0.1777	0.0664	0.1457
SDM	0.1540	0.1533	0.1678	0.0647	0.1410
SAC	0.1853	0.1889	0.1944	0.0699	0.1678
SMC	0.1602	0.1790	0.1848	0.0949	0.1588
Structural predictors					
SEM	0.1113	0.1198	0.1358	0.0580	0.1102
SXM	0.1119	0.1205	0.1368	0.0578	0.1108
SAR	0.1135	0.1242	0.1400	0.0590	0.1133
SDM	0.1122	0.1206	0.1369	0.0579	0.1109
SAC	0.1545	0.1734	0.1767	0.0881	0.1524
SMC	0.1571	0.1874	0.1684	0.0871	0.1547
Individual aspatial models					
LPB	0.1629	0.1612	0.1756	0.0629	0.1477
MNL	0.1573	0.1506	0.1732	0.0608	0.1424

B. RMSE for predictors from short run models

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
Aggregated aspatial models					
OLS	0.0405	0.0379	0.0282	0.0182	0.0324
GAM	0.0396	0.0366	0.028	0.0181	0.0317
Aggregated spatial models					
Reduced predictors					
SEM	0.0411	0.0383	0.0288	0.0183	0.0329
SXM	0.0406	0.0386	0.0288	0.0182	0.0328
SAR	0.0403	0.0379	0.0298	0.0179	0.0327
SDM	0.0405	0.0385	0.0286	0.0180	0.0327
SAC	0.0403	0.0381	0.0294	0.0179	0.0326
SMC	0.0411	0.0391	0.0294	0.0210	0.0336
Structural predictors					
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	0.0370	0.0293	0.0178	0.0320
SDM	0.0398	0.0384	0.0282	0.0178	0.0323
SAC	0.0394	0.0379	0.0294	0.0178	0.0323
SMC	0.0433	0.0468	0.0329	0.0270	0.0383
Individual aspatial models					
LPB	0.0572	0.0507	0.0269	0.018	0.0415
MNL	0.0547	0.0485	0.0267	0.0175	0.0399

Table 3: Out-of-sample Root Mean Square Errors for different predictors: *The rows marked REF report the benchmark RMSE from constant predictions (i.e., the national shares for each plot). The columns MEAN report the row means of RMSE. OLS is for Ordinary Least Squares, GAM for Generalized Additive Models, 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, SMC 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.*

A. RMSE for predictors from long run models

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
Aggregated aspatial models					
OLS	0.1675	0.1717	0.19	0.0677	0.1567
GAM	0.1342	0.1525	0.1645	0.0656	0.1347
Aggregated spatial models					
Reduced predictors					
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
SMC	0.1738	0.1965	0.181	0.0969	0.1665
Structural predictors					
SEM	0.1120	0.1215	0.1368	0.0588	0.1112
SXM	0.1124	0.1262	0.1398	0.0588	0.1135
SAR	0.1140	0.1295	0.1410	0.0597	0.1153
SDM	0.1124	0.1267	0.1399	0.0589	0.1137
SAC	0.1582	0.1702	0.1838	0.0964	0.1558
SMC	0.1642	0.1835	0.1781	0.0981	0.1596
Individual aspatial models					
LPB	0.1669	0.1622	0.1804	0.0632	0.1506
MNL	0.1633	0.1563	0.1778	0.0621	0.1471

B. RMSE for predictors from short run models

	PSTUR03	ARBLE03	FORST03	URBAN03	MEAN
REF	0.2376	0.2765	0.2589	0.2262	0.2505
Aggregated aspatial models					
OLS	0.0317	0.0308	0.0202	0.0128	0.0251
GAM	0.0308	0.0309	0.0212	0.0133	0.0251
Aggregated spatial models					
Reduced predictors					
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
SMC	0.0355	0.04	0.0261	0.0155	0.0307
Structural predictors					
SEM	0.0310	0.0308	0.0196	0.0129	0.0248
SXM	0.0320	0.0330	0.0206	0.0135	0.0261
SAR	0.0310	0.0314	0.0205	0.0129	0.0252
SDM	0.0320	0.0330	0.0206	0.0135	0.0261
SAC	0.0311	0.0319	0.0204	0.0130	0.0253
SMC	0.0499	0.0682	0.0458	0.0239	0.0495
Individual aspatial models					
LPB	0.0489	0.0441	0.0181	0.0126	0.0347
MNL	0.048	0.0429	0.018	0.0124	0.034

RMSE that are comparable to those of the other models. The differences are relatively high, as it can be seen from the last columns reporting the RMSE means by rows. The gains from the spatial models relative to OLS represent half the gains of OLS relatively to the benchmark REF from constant predictions (i.e., the mean share for each unit). Thus, this means that the effect is strong. With the same magnitude, the GAM is in an intermediate position between the spatial and the aspatial models. For the aspatial models (both aggregated 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 SMC, perform the best, according to a property that multicollinearity does not bias the predictions [33]. 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 (aggregated and individual) based on discrete outcomes: LPB and MNL. These results are confirmed visually by the maps presented in the appendix 8.2 and 8.4.

6 Conclusion

It is widely acknowledged that micro-economic behaviors are more accurately analyzed using individual data and models. However, with the recent development of individual land-use models, aggregated models are still appealing as they ease estimation with more advanced econometric techniques and are less data and computational demanding. Nevertheless, the comparative advantages of individual models in terms of predictive accuracy has remained an open question that was investigated in this paper.

We have compared the predictive accuracies of a wide spectrum of econometric models of land use at different scales (individual and aggregated), spatial (with and without spatial autocorrelation) and temporal horizons (short term and long term). More specifically, we have showed how the introduction of spatial autocorrelation in aggregated models matters for improving their predictions related to aggregated LUC. Our results suggest firstly that introducing spatial autocorrelation in aggregated models improves their predictive accuracy and even outperforms individual models if appropriate predictors are used. Secondly, we show that 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 biophysical variables are included.

Our findings show that it may not be worth using individual land-use data when the only objective is to predict aggregate land use. By taking advantage of the progress made in spatial econometrics tools, applied researchers can use the newly available structural predictors to improve the quality of their LUC predictions. We hope that our results give a chance for spatial econometrics to be more frequently used to perform predictions, instead of mainly focused on causal inference. The results of our study may not hold for other data and models even if they corroborates findings of previous literature that aggregation is not necessarily bad. We have to mention that the strong temporal autocorrelation obtained (and the high predictive accuracies of short run models) could be attributed to the fact that we model LUC in a developed country with limited economic growth. This temporal autocorrelation is probably less strong in developing countries with higher economic growth. Nevertheless, we can imagine that our main results about using spatial autocorrelation to improve predictive accuracy is still valid, even if this conjecture has to be empirically verified on other datasets.

7 Acknowledgements

The research leading to these results has received funding from the European Union by the European Commission within the Seventh Framework Programme in the frame of RURAGRI ERA-NET under Grant Agreement 235175 TRUSTEE (ANR–13–RURA–0001–01) and the French Agence Nationale de la Recherche through the ModULand project (ANR–11–BSH1–005). The authors only are responsible for any omissions or deficiencies. Neither the TRUSTEE project and any of its partner organizations, nor any organi-

zation of the European Union or European Commission are accountable for the content of this research.

References

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

[2] Anna Angulo and F.J. Trivez. The impact of spatial elements on the forecasting of spanish labour series. *Journal of Geographical Systems*, 12:155–174, 2010.

[3] L. Anselin and W.K.T. Cho. Spatial effect and ecological inference. *Political Analysis*, 10:276–297, 2002.

[4] Luc Anselin. *Spatial Econometrics : Methods and Models*. Kluwer Academic Publishers, Dordrecht., 1988.

[5] Luc Anselin. Spatial econometrics in RSUE: Retrospect and prospect. *Regional Science and Urban Economics*, 37(4):450–456, 2007.

[6] Luc Anselin. Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1):3–25, 2010.

[7] Jean-Sauveur Ay, Raja Chakir, Luc Doyen, Frédéric Jiguet, and Paul Leadley. Integrated models, scenarios and dynamics of climate,land use and common birds. *Climatic Change*, 126:13–30, 2014.

[8] Badi Baltagi and Dong Li. Prediction in the spatially autocorrelated error component model. *Econometric Theory*, 15:2: 259, 1999.

[9] Badi Baltagi and Dong Li. Prediction in the panel data model with spatial correlation: the case of liquor. *Spatial Economic Analysis*, 1:175–195, 2006.

[10] Badi H. Baltagi, Bernard Fingleton, and Alain Pirotte. Estimating and forecasting with a dynamic spatial panel data model. *Oxford Bulletin of Economics and Statistics*, 76(1):112–138, 2014.

[11] B.H Baltagi, G. Bresson, and A. Pirotte. Forecasting with spatial panel data. *Computational Statistics and Data Analysis*, 56:3381–3397, 2012.

[12] Ian J. Bateman, Amii R. Harwood, Georgina M. Mace, Robert T. Watson, David J. Abson, Barnaby Andrews, Amy Binner, Andrew Crowe, Brett H. Day, Steve Dugdale, Carlo Fezzi, Jo Foden, David Hadley, Roy Haines-Young, Mark Hulme, Andreas Kontoleon, Andrew A. Lovett, Paul Munday, Unai Pascual, James Paterson, Grischa Perino, Antara Sen, Gavin Siriwardena, Daan van Soest, and Mette Termansen. Bringing ecosystem services into economic decision-making: Land use in the United Kingdom. *Science*, 341(15):45–50, 2013.

[13] Kathleen P. Bell and Elena G. Irwin. Spatially explicit micro-level modelling of land use change at the rural-urban interface. *Agricultural Economics*, 27(3):217–232, November 2002.

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

[15] Roger S. Bivand, Edzer Pebesma, and V. Gomez-Rubio. *Applied Spatial Data Analysis with R*. second Edition, UseR! Series, Springer, 2013.

[16] Michael Brady and Elena Irwin. Accounting for spatial effects in economic models of land

use: Recent developments and challenges ahead. *Environmental & Resource Economics*, 48(3):487–509, 2011.

[17] Raja Chakir and Julie Le Gallo. Predicting land use allocation in France: A spatial panel data analysis. *Ecological Economics*, 92(0):114–125, 2013.

[18] Raja Chakir and Olivier Parent. Determinants of land use changes: A spatial multinomial probit approach. *Papers in Regional Science*, 88(2):327–344, 06 2009.

[19] A. D. Cliff and J. K. Ord. *Spatial Processes : Models and Applications*. Pion, London, UK, 1981.

[20] Timothy J Considine and Timothy D Mount. The use of linear logit models for dynamic input demand systems. *The Review of Economics and Statistics*, pages 434–443, 1984.

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

[22] Nazneen Ferdous and Chandra R Bhat. A spatial panel ordered-response model with application to the analysis of urban land-use development intensity patterns. *Journal of Geographical Systems*, 15(1):1–29, 2013.

[23] Mark M. Fleming and Fannie Mae. Techniques for estimating spatially dependent discrete choice models. *In Advances in Spatial Econometrics, Anselin Luc and Raymond Florax, eds. Springer-Verlag, Heidelberg*, 2004.

[24] Raymond J.G.M Florax, Hendrik Folmer, and Sergio J Rey. Specification searches in spatial econometrics: The relevance of hendry's methodology. *Regional Science and Urban Economics*, 33(5):557–579, 2003.

[25] J. A. Foley and al. Global consequences of land use. *Science*, 309:570–574, 2005.

[26] Arthur S Goldberger. Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association*, 57(298):369–375, 1962.

[27] Yehuda Grunfeld and Zvi Griliches. Is aggregation necessarily bad? *The Review of Economics and Statistics*, 42(1):pp. 1–13, 1960.

[28] Trevor Hastie and Robert Tibshirani. Generalized additive models. *Statistical science*, pages 297–310, 1986.

[29] Elena G. Irwin. New directions for urban economic models of land use change: incorporating spatial dynamics and heterogeneity. *Journal of Regional Science*, 50(1):65–91, 2010.

[30] Elena G. Irwin and Jacqueline Geoghegan. Theory, data, methods: Developing spatially explicit economic models of land use change. *Agriculture, Ecosystems & Environment*, 85(1-3):7–24, June 2001.

[31] Harry H Kelejian and Ingmar R Prucha. The relative efficiencies of various predictors in spatial econometric models containing spatial lags. *Regional Science and Urban Economics*, 37(3):363–374, 2007.

[32] H.H. Kelejian and I.R. Prucha. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40:509–533, 1999.

[33] P. Kennedy. *A Guide to Econometrics*. Blackwell, 2003.

[34] A. Kholodilin, B. Siliverstovs, and S. Kooth. A dynamic panel approach to the forecasting of the gdp of german lŁnder. *Spatial Economic Analysis*, 3:195–207, 2008.

[35] Thomas Klier and Daniel P. McMillen. Clustering of auto supplier plants in the United States. *Journal of Business & Economic Statistics*, 26(4), 2008.

[36] Christopher R Knittel and Konstantinos Metaxoglou. Estimation of random-coefficient demand models: Two empiricists' perspective. *Review of Economics and Statistics*, 96(1), 2012.

[37] Eric F Lambin and Patrick Meyfroidt. Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences*, 108(9):3465–3472, 2011.

[38] Julie Le Gallo. Cross-section spatial regression models. *Handbook of Regional Science, Fischer, Manfred M. (Eds.), Nijkamp, Peter (Eds.)*, 2014.

[39] J. LeSage and R.K. Pace. *Introduction to Spatial Econometrics*. CRC Press Boca Raton FL, 2009.

[40] David J. Lewis. An economic framework for forecasting land-use and ecosystem change. *Resource and Energy Economics*, 32(2):98–116, 2010.

[41] Man Li, JunJie Wu, and Xiangzheng Deng. Identifying drivers of land use change in China: A spatial multinomial logit model analysis. *Land Economics*, 89(4):632–654, 2013.

[42] Erik Lichtenberg. Land quality, irrigation development, and cropping patterns in the northern high plains. *American Journal of Agricultural Economics*, Vol. 71, No. 1:187–194, 1989.

[43] Ruben Lubowski, Andrew Plantinga, and Robert Stavins. What drives land-use change in the United States? A national analysis of landowner decisions. *Land Economics*, 84(4):551–572, 2008.

[44] Ruben N. Lubowski, Andrew J. Plantinga, and Robert N. Stavins. Land-use change and carbon sinks: Econometric estimation of the carbon sequestration supply function. *Journal of Environmental Economics and Management*, 51:135–152, 2006.

[45] Ruben Noah Lubowski. Determinants of land-use transitions in the united states: Econometric analysis of changes among the major land-use categories. *PhD dissertation, Harvard University Cambridge, Massachusetts*, 2002.

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

[47] Daniel P McMillen. Probit with spatial autocorrelation. *Journal of Regional Science*, Vol. 32, number 3:335–348, 1992.

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

[49] John Mullahy. Multivariate fractional regression estimation of econometric share models. NBER Working Papers 16354, National Bureau of Economic Research, Inc, September 2010.

[50] Erik Nelson, Stephen Polasky, David J. Lewis, Andrew J. Plantinga, Eric Lonsdorf, Denis White, David Bael, and Joshua J. Lawler. Efficiency of incentives to jointly increase carbon sequestration and species conservation on a landscape. *Proceedings of the National Academy of Sciences*, 105(28):9471–9476, 2008.

[51] R. Kelley Pace and James P. LeSage. Spatial econometric models: Prediction. In Shashi Shekhar and Hui Xiong, editors, *Encyclopedia of Geographical Information Science*. Springer-Verlag, 10.1007/978-0-387-35973-1, 2008.

[52] Leslie E Papke and Jeffrey Wooldridge. Econometric methods for fractional response variables with an application to 401 (k) plan participation rates, 1993.

[53] Joris Pinkse and Margaret E. Slade. Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85(1):125–154, July 1998.

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

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

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

[57] Maksym Polyakov and Daowei Zhang. *Land Use Dynamics along Urban-Rural Gradient: A Comparison of Modeling Approaches*, chapter 3, pages 33–46. In: Global Change and Foresty, Ed: J. Gan, S. Grado and I. A. Munn. Nova Science Publishers, 2010.

[58] Richard D. Robertson, Gerald C. Nelson, and Alessandro De Pinto. Investigating the predictive capabilities of discrete choice models in the presence of spatial effects. *Papers in Regional Science*, 88(2):367–388, 2009.

[59] N. Schanne, R. Wapler, and A. Weyh. Regional unemployment forecasts with spatial interdependence. *International Journal of Forecasting*, 26:908–926, 2010.

[60] Kurt E Schnier and Ronald G Felthoven. Accounting for spatial heterogeneity and autocorrelation in spatial discrete choice models: implications for behavioral predictions. *Land economics*, 87(3):382–402, 2011.

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

[62] Oleg A. Smirnov. Modeling spatial discrete choice. *Regional Science and Urban Economics*, 40(5):292 – 298, 2010.

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

[64] Kenneth Train. *Discrete Choice Methods with Simulation*, volume Second Edition. Cambridge University Press, 2009.

[65] Billie L Turner, Eric F Lambin, and Anette Reenberg. The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences*, 104(52):20666–20671, 2007.

[66] Peter H. Verburg, Paul P. Schot, Martin J. Dijst, and A. Veldkamp. Land use change modelling: current practice and research priorities. *GeoJournal*, 61:309–324, 2004.

[67] Kathrine von Graevenitz and Toke Emil Panduro. An alternative to the standard spatial econometric approaches in hedonic house price models. *Land Economics*, 91(2):386–409, 2015.

[68] Yiyi Wang, Kara M Kockelman, and Xiaokun Cara Wang. Understanding spatial filtering for analysis of land use-transport data. *Journal of Transport Geography*, 31:123–131, 2013.

[69] Simon N Wood. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99(467), 2004.

[70] JunJie Wu and Richard M Adams. Micro versus macro acreage response models: Does site-specific information matter? *Journal of Agricultural and Resource Economics.*, 27(1), 2002.

[71] JunJie Wu and Joshua M. Duke. *The Oxford Handbook of Land Economics*. Oxford University Press, USA, 2014.

[72] JunJie Wu and Kathleen Segerson. The impact of policies and land characteristics on potential groundwater pollution in wisconsin. *American Journal of Agricultural Economics*,

77(4):1033–1047, 1995.

[73] A Zellner and T.H. Lee. Joint estimation of relationships involving discrete random variables. *Econometrica*, 33:382–94, 1965.

Table 4: 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

Note: 1993 land use in rows and 2003 land use in columns

Table 5: 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	17,929	51.639	819,298
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

8 Appendix (for publication)

8.1 Spatial coefficient from aggregated models

8.2 Observed land use in 2003 and long run predictions

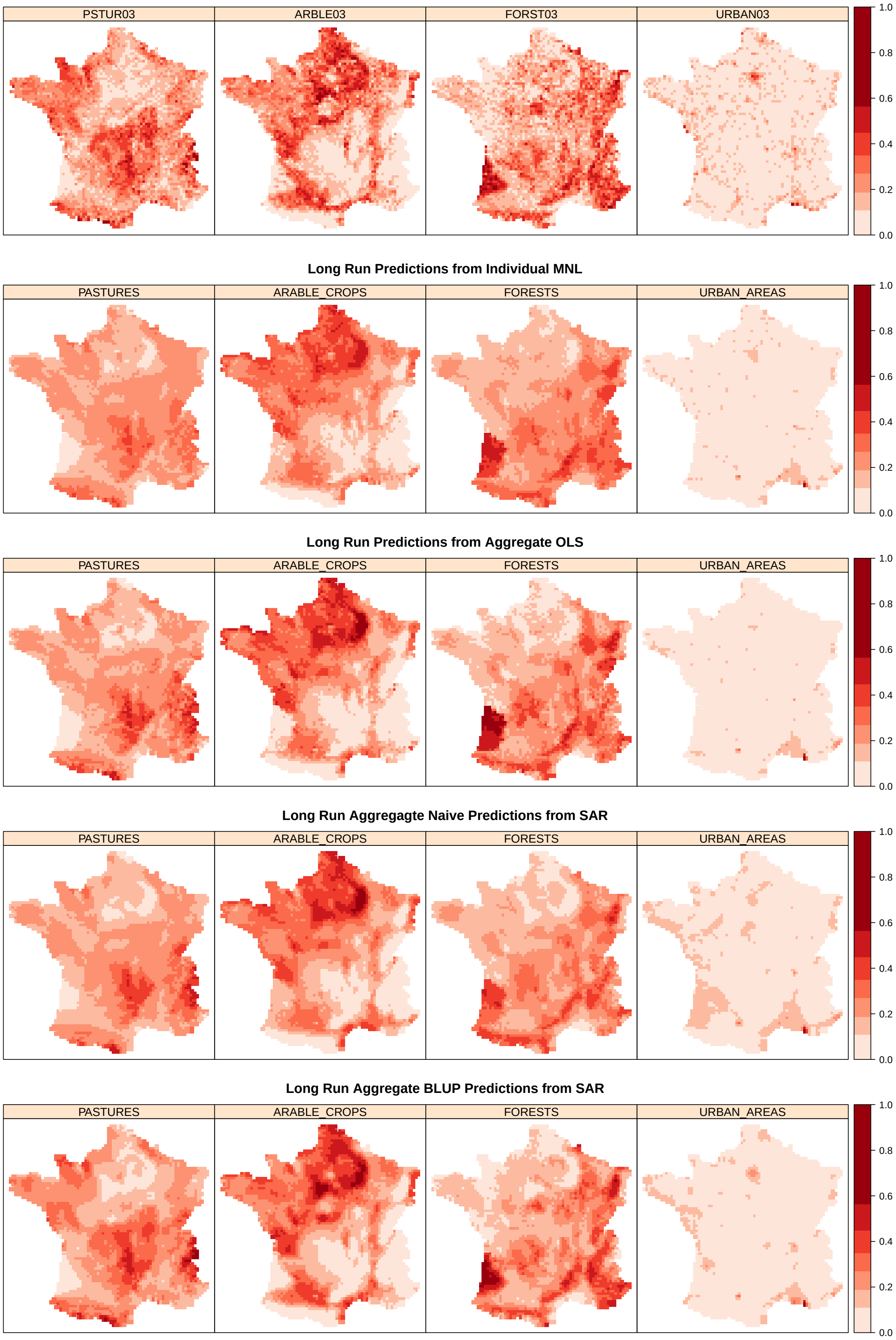
8.3 Aggregated outcome variables

Table 6: Spatial coefficients for long run models

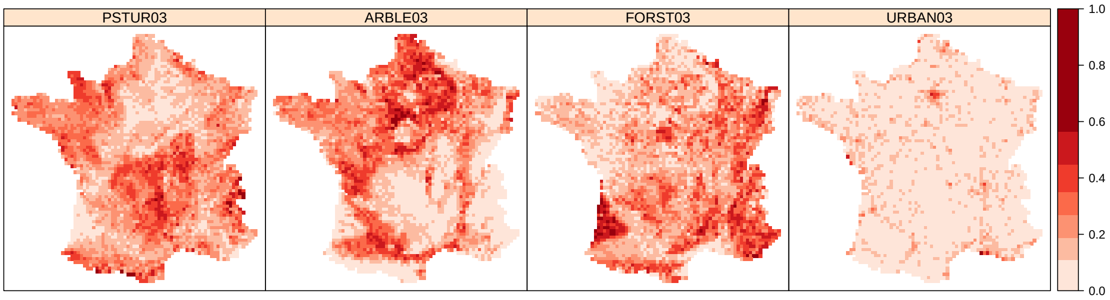
	Spatial Error components: λ_ℓ^*			Spatial Lag components: ρ_ℓ^*		
	Arable crop	Forest	Urban	Arable crop	Forest	Urban
SEM	0.6449^{**}	0.7349^{**}	0.4991^{**}			
	(0.0183)	(0.0248)	(0.0217)			
SXM	0.626^{**}	0.7019^{**}	0.4902^{**}			
	(0.0177)	(0.0158)	(0.0216)			
SAR				0.5654^{**}	0.7017^{**}	0.4586^{**}
				(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)
SMC	0.8995^{**}	-0.7029^{**}	-0.635^{**}	-0.7909^{**}	0.8991^{**}	0.7958^{**}
	(0.0114)	(0.0448)	(0.0602)	(0.044)	(0.0112)	(0.0215)

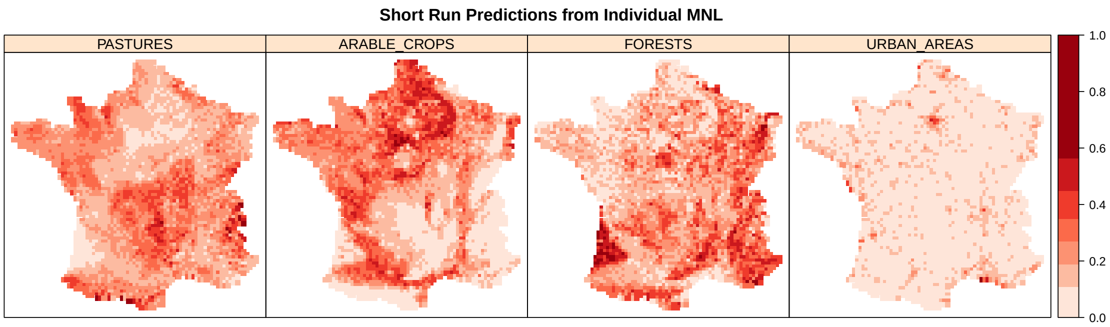
Table 7: Spatial coefficients for short run models

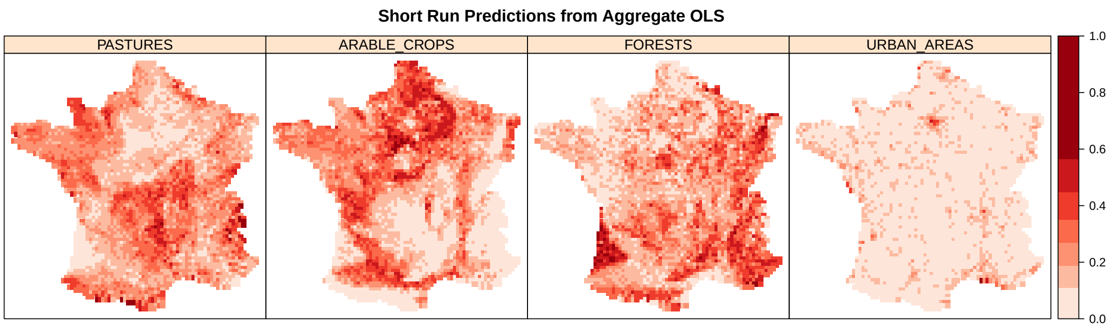
	Spatial Error components: λ_ℓ			Spatial Lag components: ρ_ℓ		
	Arable crop	Forest	Urban	Arable crop	Forest	Urban
SEM	0.1134^{**}	0.3004^{**}	0.2246^{**}			
	(0.022)	(0.0295)	(0.0278)			
SXM	0.0473^{**}	0.2404^{**}	0.2^{**}			
	(0.0131)	(0.0282)	(0.0288)			
SAR				0.1335^{**}	0.1256^{**}	0.1122^{**}
				(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)	(0.0132)	(0.0087)	(0.0179)
SMC	-0.3827^{**}	-0.0403	-0.3825^{**}	0.3746^{**}	0.2776^{**}	0.48967^{**}
	(0.0958)	(0.0418)	(0.0814)	(0.071)	(0.0428)	(0.0527)

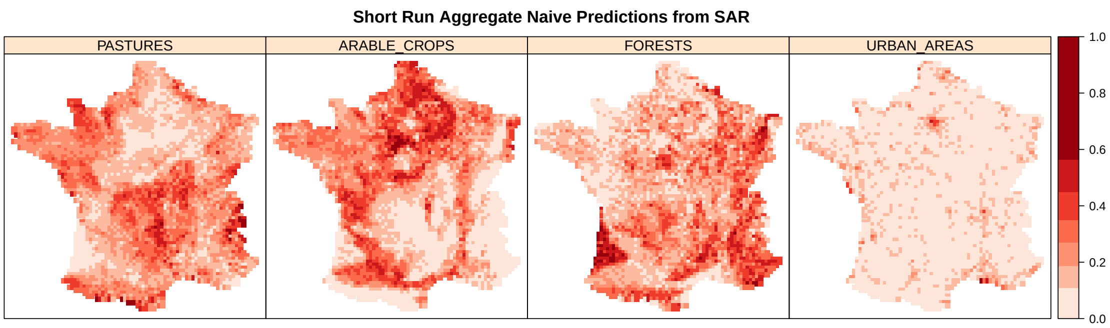

PSTUR03
ARBLE03
FORST03
URBAN03
Long Run Predictions from Individual MNL
PASTURES
ARABLE_CROPS
FORESTS
URBAN_AREAS
Long Run Predictions from Aggregate OLS
PASTURES
ARABLE_CROPS
FORESTS
URBAN_AREAS
Long Run Aggregagte Naive Predictions from SAR
PASTURES
ARABLE_CROPS
FORESTS
URBAN_AREAS
Long Run Aggregate BLUP Predictions from SAR
PASTURES
ARABLE_CROPS
FORESTS
URBAN_AREAS
1.0
0.8
0.6
0.4
0.2
0.0

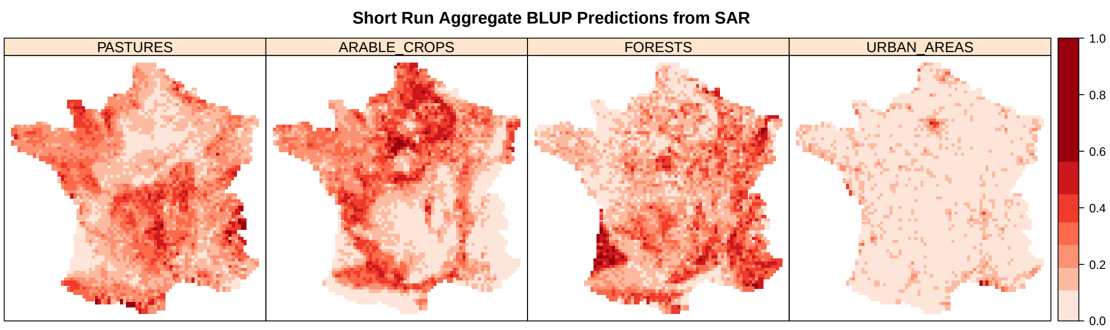
8.4 Observed land use in 2003 and short run predictions











9 Supporting Information (not for publication)

9.1 Estimation results from individual MNL models

Table 8: Individual MNL models on 1993–2003

	Long Run			Short Run		
	arable use	forest use	urban use	arable use	forest use	urban use
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 explanatory variables. Reference= Pastures					

9.2 Estimation results from the models estimated by OLS

Table 9: Linear logit-transformed 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***				
		(0.020)				
FOlog93				0.937***		
				(0.017)		
URlog93						0.847***
						(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 explanatory variables, HC robust standard errors.

9.3 Estimation results from the GAM model

Table 10: GeoAdditive logit-transformed 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***				
		(0.010)				
FOlog93				0.912***		
				(0.006)		
URlog93						0.837***
						(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

Note: *p<0.1; **p<0.05; ***p<0.01.
Reference= Pastures, scaled explanatory variables, bivariate smooth function of coordinates

9.4 Estimation results from the SEM model

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)				
FOlog93				0.920***		
				(0.006)		
URlog93						0.842***
						(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 explanatory variables. Reference= Pastures
The Wald and the LR test are the Wald and the likelihood ratio test for the significance of the spatial error coefficient

9.5 Estimation results from the SXM model

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***				
		(0.011)				
FOlog93				0.897***		
				(0.006)		
URlog93						0.836***
						(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
Note:	*p<0.1; **p<0.05; ***p<0.01					
	scaled explanatory variables. Reference= Pastures					

9.6 Estimation results from the SAR model

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)				
FOlog93				0.890***		
				(0.007)		
URlog93						0.830***
						(0.009)
scale(Arable returns03)	0.297***	0.017	0.069***	−0.005	0.242***	0.034**
	(0.028)	(0.015)	(0.024)	(0.010)	(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)	(0.010)	(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)	(0.010)
scale(Precipitations)	−0.248***	−0.038**	−0.005	−0.040***	−0.063*	−0.014
	(0.030)	(0.017)	(0.010)	(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)	(0.010)
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 explanatory variables. Reference= Pastures
The Wald and the LR test are the Wald and the likelihood ratio test for the significance of the spatial lag coefficient

9.7 Estimation results from the SDM model

Table 14: Spatial Durbin 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***				
		(0.009)				
FOlog93				0.893***		
				(0.006)		
URlog93						0.834***
						(0.008)
scale(Arable returns03)	0.342***	0.079***	−0.119***	−0.055***	0.148**	0.048**
	(0.050)	(0.027)	(0.010)	(0.017)	(0.058)	(0.024)
scale(Pasture returns03)	0.005	0.004	0.042	0.033***	−0.015	−0.029***
	(0.014)	(0.014)	(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)	(0.048)	(0.048)	(0.037)	(0.052)	(0.048)
scale(POP03)	−0.100***	−0.016	−0.011	−0.011	0.115***	0.009
	(0.029)	(0.029)	(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)	(0.070)	(0.120)	(0.076)
scale(Slope)	−0.443***	−0.155	0.603***	0.076***	0.055	0.067
	(0.070)	(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)	(0.070)	(0.070)	(0.070)	(0.063)	(0.063)
scale(Soil depth)	−0.176***	−0.014	0.106	0.014	−0.065	0.001
	(0.065)	(0.065)	(0.065)	(0.065)	(0.067)	(0.002)
scale(Precipitations)	−0.203***	−0.022	0.239***	0.006	−0.129*	−0.068***
	(0.055)	(0.052)	(0.052)	(0.052)	(0.074)	(0.026)
scale(Temperature)	1.086***	0.286	0.376***	0.050	0.399***	0.009
	(0.160)	(0.119)	(0.138)	(0.119)	(0.119)	(0.119)
scale(Humidity)	−0.211	−0.147***	0.301***	0.026	−0.026	0.033
	(0.136)	(0.018)	(0.036)	(0.060)	(0.060)	(0.065)
scale(Radiation)	−0.206	−0.080	−0.541***	−0.019	0.189	0.113
	(0.171)	(0.171)	(0.158)	(0.171)	(0.171)	(0.073)
Constant	−0.242***	−0.109***	−0.058***	0.061	−0.929***	0.013
	(0.023)	(0.019)	(0.017)	(0.019)	(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 explanatory variables. Reference= Pastures
The Wald and the LR test are the Wald and the likelihood ratio test for the significance of the spatial lag coefficient

9.8 Maps at the aggregated scale

Figure 1: Aggregated land use shares in 2003

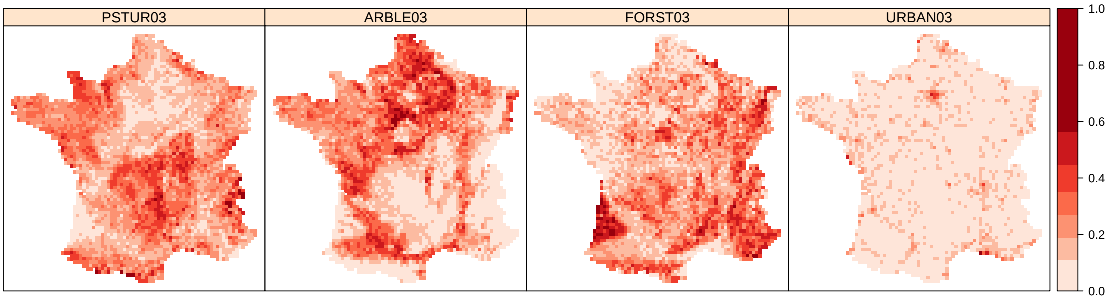


Figure 2: Aggregated land use variations on 1993–2003, in km^2

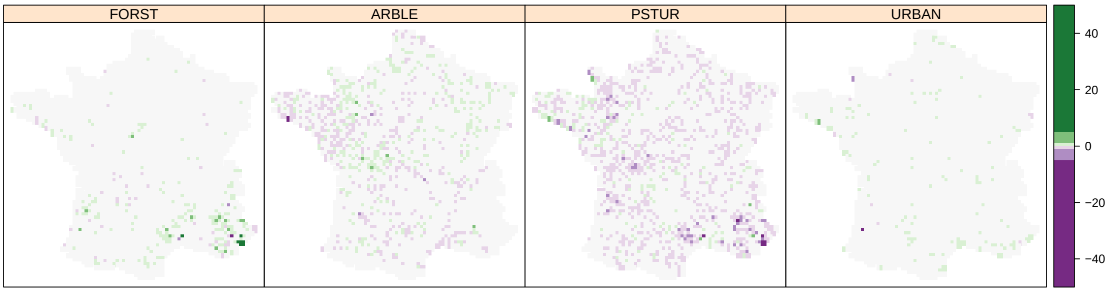


Figure 3: Out of sample 2003 predictions from individual mnl

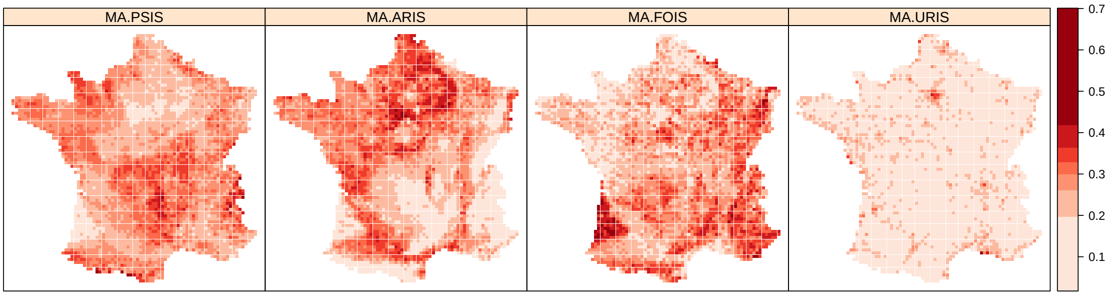


Figure 4: Raw distribution of 1998 aggregated land use shares

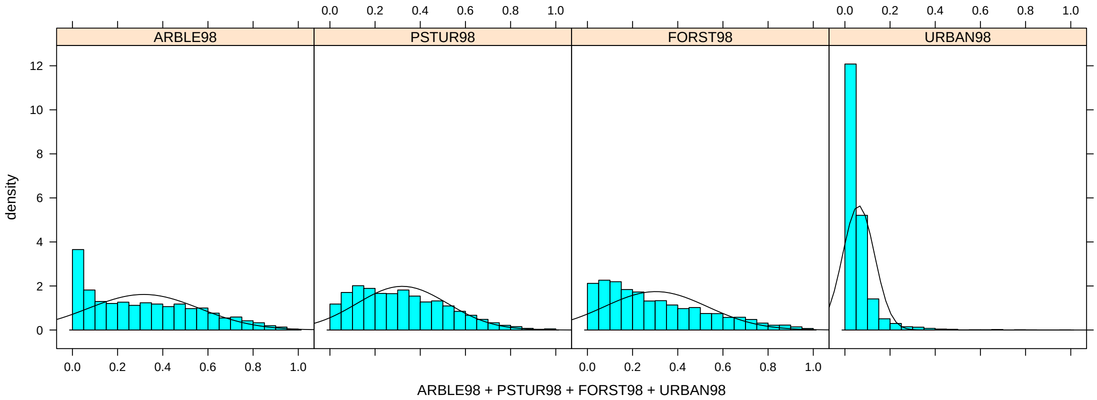


Figure 5: Linearized distribution of 1998 aggregate land use shares