UNITE MIXTE DE RECHERCHE
EN ECONOMIE PUBLIQUE

JOINT RESEARCH UNIT IN
PUBLIC ECONOMICS

2014
CAHIERS DE RECHERCHE
WORKING PAPERS
2014/02

The effects of scale, space and time on the predictive accuracy of land use models

J.S. Ay ; R. Chakir ; J. Le Gallo

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

The effects of scale, space and time on the predictive accuracy of land use models*

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

June 25, 2014

Abstract

The econometric literature about modeling land use choices is highly heterogeneous with respect to the scale of the data, and to the structure of the models in terms of the effects of space and time. This paper proposes a joint evaluation of each of these three elements by estimating a broad spectrum of individual and aggregate, spatial and aspatial, short and long run econometric models on the same detailed French dataset. Considering four land use classes (arable crops, pasture, forest, and urban), all the models are compared in terms of both in- and out-of-sample predictive accuracy. We argue that the aggregate scale allows to model more effectively the effect of space by using spatial econometric models. We show that modeling spatial autocorrelation allow to have very accurate predictions which can even outperform individual models when the appropriate predictors are used. We also found some strong interactions between the effects of scale, space and time which can be of major interest for applied researchers.

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

JEL Classifications: Q15, Q24, R1, C21.

*Corresponding author: J.-S. Ay. Mail: jsay.site@gmail.com, Phone: 331.44.08.86.83.
†CNRS, UMR 7204 CESCO, 55 rue Buffon, 75005 PARIS, France. jsay.site@gmail.com
‡INRA, UMR 210 Économie Publique, 16 rue Claude Bernard, 75005 PARIS, France.
§Université de Franche-Comté, CRESE, 45D, Avenue de l'Observatoire, 25030 BESANÇON, France.

1 Introduction

Land Use Changes (LUC) 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 (Turner et al., 2007; Bateman et al., 2013). Given these impacts and the expected changes in land use in the next decades, prospective analysis requires a thorough understanding of how economics and policies affect land use patterns (Nelson et al., 2008; Lewis, 2010; Wu and Duke, 2014).

The way scale as well as spatial and dynamic structures are accounted for in LUC econometric literature varies greatly.[1] These modeling choices depend on the research question but are also driven by external constraints such as data availability or computation capacity. To the best of our knowledge, their effects on predictive accuracy have never been jointly evaluated. In this context, the main objective of this paper is to compare land use models with respect to their scale (individual vs aggregate), spatial structure (spatial vs aspatial) and time effects (short term vs long term). These comparisons are made at the same aggregate scale, both in terms of in- and out-of-sample predictive accuracy as well as in terms of counterfactual simulations.

With respect to scale, the most distinctive feature of previous works is the use of aggregate or individual data. Until recently, most studies have been based on aggregated data for a region, a country, or other geographic scales with land use shares as the outcome of interest (Plantinga, 1996; Plantinga et al., 1999; Chakir and Le Gallo, 2013). Recent studies increasingly rely on individual data and consider discrete plot-level choices as the outcome of interest (Lewis and Plantinga, 2007; Lubowski et al., 2008; Chakir and Parent, 2009). However, the literature offers mixed evidence about higher predictive accuracy of individual versus aggregated models. In a seminal paper, Grunfeld and Griliches (1960) have examined the relative power of micro and macro models for explaining the variability of the aggregate dependent variable and found that the aggregate equation may explain the aggregate data better. Wu and Adams (2002) examine the issue in the context of predicting land allocation. They show that, even for linear prediction models, deciding to choose micro or macro models to make aggregate predictions cannot generally be resolved by *a priori* reasoning.

With respect to space, the vast majority of past studies assumes spatial independence of land use choices, both at aggregate and individual scales. Recent exceptions include: Chakir and Parent (2009); Sidharthan and Bhat (2012); Chakir and Le Gallo (2013); Ferdous and Bhat (2013); Li et al. (2013). Incorporating spatial dimension into land-use models raises several issues related to econometric estimation, hypothesis testing and prediction (Anselin, 2007; Brady and Irwin, 2011). This is even more challenging in the case of individual multinomial land use models since the introduction of spatial dependence would render discrete choice models analytically intractable, and would require the use of simulation or Bayesian techniques (Fleming and Mae, 2004).[2] Consequently, we consider only spatially independent individual models but we introduce spatial autocorrelation in our aggregate models. The possibility of modeling spatial

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

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

autocorrelation appears as a comparative advantage of some aggregate models.

With respect to time, the most distinctive feature in the literature is the use of short run or long run models. Such modeling choices are typically driven by the research question, but also by constraints on the available data. With cross-section data, only long run models can be estimated (Lichtenberg, 1989; Fezzi and Bateman, 2011) with predictions that have to be interpreted as a steady state equilibrium. With a time dimension in the data, the inertia of LUC can be captured and the predictions can be used to simulate temporal recursive scenarios (Lewis, 2010). The combination of the spatial and the temporal dimensions as considered in this paper, is a very recent literature, currently limited to models with continuous outcomes (Baltagi et al., 2014).

This paper contributes to the previous econometric analyses of LUC in three ways. Firstly, we show how the heterogeneous econometric models of the literature can be grounded on the same microeconomic principles, and can be consistently related to each other even if they are specified with different scales, spatial and temporal structures. In particular, we explicitly introduce a Ricardian framework into LUC models to formalize the relationship between land prices and land returns, consistently relating different scales and time horizons. Secondly, we show how the introduction of recent spatial econometric tools in aggregated land use models enables better predictions than individual, aspatial models with higher numbers of observations. Thirdly, through counterfactual simulations, our wide spectrum of econometric models are found to produce very differentiated results, with strong interactions between scale and time dimensions. In particular, we show that aggregate spatial models perform relatively well in counterfactual long run simulations but not in the short run, which underlies some limits in the use of aggregate data in relation to time.

The paper is organized as follows. In section 2, we present our econometric LUC models. In section 3, we present the in-sample and out-of-sample formulas available to predict LUCs, with some original predictors for spatiotemporal econometric models. In section 4, we present the data and in section 5 the results both in terms of the estimated parameters and the prediction 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 land use choices (Stavins and Jaffe, 1990; Plantinga, 1996; Lubowski et al., 2006), we consider a risk-neutral landowner facing the choice of allocating a parcel of land of uniform quality 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. 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 (Lubowski et al., 2006; Lewis, 2010). As Train (2009) states, the two major implications of the 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 necessary for their 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, Plantinga (1996) 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, $r_{i\ell t}$ 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.[3] 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 $\forall i, \ell, t$:

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

In contrast to the above random utility framework which is shared by most of the literature, the relevant data used to proxy expected one-period returns are heterogeneous in previous studies. It is clear that obtaining precise data about expected returns for each use and each plot of land is challenging.[4] Here, our proposed solution is to match data about land price with the Ricardian formula to proxy the expected returns (Ay et al., 2014). As we will show, this contributes to integrate the scale, space and time dimensions of the different models while keeping them consistent with standard economic theory.

Assuming constant land use, the Ricardian formula states that the observed land price $\overline{r}_{\ell t}$ at t for a land in use ℓ is equal to the net present value of all expected future returns.[5] 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 one-period expected return:

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

This shows that land prices can be used to substitute expected returns in the specification of utility, without ensuring that returns are perfectly observed. Typically, the 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 $\mathrm{E}(r_{i\ell t}) = \mathbf{b}_i^\top \boldsymbol{\gamma}_\ell^B + \overline{r}_{g(i)\ell t} \gamma_\ell^R + \varepsilon_{i\ell t}$ 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 availability, and $\varepsilon_{i\ell t}$ represents the random deviations from the average values due to unobserved non-economic variables. Substituting this approximation of expected net returns in Equation 2 allows us to obtain the following reduced form for utilities:

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

[4] To the best of our knowledge, only studies by Bockstael (1996) and Bell and Bockstael (2000) has used an explicit estimate of urban development profits and Lubowski (2002) constructed county-level profits for crops, pasture, and range.

[5] 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. Moreover, the assumption of land prices inferred from constant land use is consistent with the French data used in our empirical application, see section 4 for details.

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

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 γ_{ℓ}^{D}. 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 γ_{ℓ}^{D}, γ_{ℓ}^{R} and γ_{ℓ}^{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 obtain $\gamma_{\ell}^{R} = 1 - (\kappa_{\ell}/(1 + \tau))$ which allow to identify (up to a constant discount factor) the expected growth rates through the utility parameters associated to land prices.

The stochastic dimension of the framework is only related to the unobserved components $\varepsilon_{i\ell t}$ and their associated densities. McFadden (1974) identifies three standard assumptions about error terms that allow obtaining a multinomial logit model: independence, homoskedasticity and extreme value distribution (i.e., Gumbel). 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 Aggregate models

There is an important literature on econometric aggregate land use models: Lichtenberg (1989), Stavins and Jaffe (1990), Wu and Segerson (1995) and Plantinga (1996), and Miller and Plantinga (1999) are the most significant papers. The underlying microeconomic theory is identical to that in the previous section, but individual choices are aggregated 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 from i to an aggregate scale g considers that a representative landowner chooses land use shares $S_{g\ell t}$ instead. By identification with the individual model, the observed shares of land use ℓ in t for $g = 1, \dots, G$ are then expressed as $(\forall \ell = 1, \dots, L)$:

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

The meanings and dimensions of these variables are the same as in the previous subsection, capital letters represent aggregate values 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 period as explanatory variables. 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.

When estimating an aggregated land use share model, 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 to one. The most common strategy in the literature is to specify the shares as logistic functions (Wu and Segerson, 1995; Chakir and Le Gallo, 2013). This has the advantage of being linearly tractable thanks to the "logit-linear transformation" (Zellner and Lee, 1965). We propose in this paper to compare this strategy with two others proposed in the literature: Fractional logit model (FRA) proposed by Papke and Wooldridge (1993) and Dirichelet (DIR) model proposed by Mullahy (2010). The FRA and DIR models have the advantage of dealing with the bounded nature of land use shares as well as the adding constraint without transformation. However, these two specifications have the drawback of being nonlinear which make them difficult to introduce spatial autocorrelation.

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 aggregate land use share model is approximately:[6]

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

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 (Considine and Mount, 1984), but Chakir and Le Gallo (2013) show that estimating inter-equation correlations doesn't improve the predictive accuracy of the model. Therefore, to simplify the results, we skip this aspect and first estimate the linearized equations by Ordinary Least Squares (OLS).

Space can easily be introduced in these models by including a smoothed function of the geographical coordinates of the grids' centroids in $\mathbf{B}_g$. This leads to semi-parametric Generalized Additive Models (GAM), estimated by penalized likelihood techniques (Hastie and Tibshirani, 1986; Wood, 2004) also used in the literature (Wang et al., 2013). Because in this case, spatial autocorrelation is not modeled explicitly, we do not consider such models as being spatial econometric models but we nevertheless include them in our comparative set.

2.3 Spatial autocorrelation

The spatial econometric literature is extensive (Cliff and Ord, 1981; Anselin, 1988; LeSage and Pace, 2009; Anselin, 2010) and provides a number of ways to deal with spatial autocorrelation. Nevertheless, introducing spatial dependence in discrete choice models is still problematic econometrically, especially with high numbers of observations (Fleming and Mae, 2004; Smirnov, 2010). An important consequence of introducing spatial dimension in discrete choice models is the complex covariance structure due to heteroskedasticity and the necessity to linearize the objective functions (Klier and McMillen, 2008; Li et al., 2013). Moreover, it implies high dimension integrals in order to compute the likelihood function (Anselin, 2002) and relies on complex optimization algorithms, dependent on starting values and tolerances of the fixed-point

[6]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 inconvenient that can be visually evaluated from Appendix A.10. More rigorously, it will be also evaluated by comparing the predictions with those from other models, as we consider this as a necessity of linearized logistic models, often used in the literature.

iterations for the generalized method of moments (Knittel and Metaxoglou, 2012). To avoid such complications associated with spatial autocorrelation in discrete choice models, we focus in this paper on introducing spatial autocorrelation in the aggregate land use models only. Hence, the possibility of modeling spatial autocorrelation is, by definition, a comparative advantage of aggregate models.

In the context of aggregated land use share models, we introduce spatial autocorrelation through three additional terms corresponding to three spatial mechanisms that will be formally tested with model estimation (LeSage and Pace, 2009).[7] Let $\mathbf{W}$ be a $G \times G$ spatial weight matrix, which summarizes the spatial connectivity structure of the observations. Once multiplied to a variable and if it is row-standardized, it returns the weighted average of the values of the neighbors of each observations. We consider the most general spatial econometric model applied to aggregate 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 interactions that lead land use at t to depend upon neighboring land uses at t and $t-1$. 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 temporally 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 and $\eta_{\ell t}$ is an iid error term. The third and last spatial term is 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 $\theta_\ell^D = \boldsymbol{\theta}_\ell^R = \boldsymbol{\theta}_\ell^B = 0$ (Bivand, 2002; Bivand et al., 2013, also called SARAR(1,1) by Kelejian and Prucha, 1999), 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 $\theta_\ell^D = \lambda_\ell = 0$. Based on maximum likelihood, we will estimate all of these spatial specifications in the empirical part of the paper, test the relative importance of the three sources of spatial autocorrelation and compare their implications in terms of predictive abilities.

2.4 Temporal autocorrelation

Because 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. Consequently, we call them short run models with inertia on LUC. Predicting the

[7] For the reason stated above, we only consider spatial autocorrelation in the linearized land use share models and not in the DIR or FRA models.

long run equilibrium (or steady state) of land use is also of interest in applied works, and is of social and political concern. Therefore, we include long run models in our comparative set for predictive accuracy. The long run equivalents of previous models correspond to the limit of the expectation of the endogenous variable when t tends to infinity. We begin the presentation by the aspatial aggregated models.

Noting from Equation 7 that $\overline{\mathbf{R}}_{g(t+1)} = \boldsymbol{\kappa} \times \overline{\mathbf{R}}_{gt}$ allows us to write land use changes between t and $t+1, t+2, \ldots, t+T$ and recursively substitute equations from previous periods, to obtain:

$$\begin{aligned}\widetilde{S}_{g\ell(t+T)} \approx \widetilde{S}_{g\ell t}\left(\beta_\ell^D\right)^T + \left[\boldsymbol{\kappa}^T + \boldsymbol{\kappa}^{T-1}\beta_\ell^D + \cdots + \boldsymbol{\kappa}\left(\beta_\ell^D\right)^{T-1}\right]\overline{\mathbf{R}}_{gt}^\top \boldsymbol{\beta}_\ell^R \\ + \left[1 + \beta_\ell^D + \cdots + \left(\beta_\ell^D\right)^{T-1}\right]\mathbf{B}_g^\top \boldsymbol{\beta}_\ell^B + \xi_{g\ell(t+T)} + \beta_\ell^D \xi_{g\ell(t+T-1)} + \cdots + \left(\beta_\ell^D\right)^{T-1}\xi_{g\ell t}\end{aligned} \tag{9}$$

which describes the net LUC between t and $t+T$. The convergence of the expectation of such land use shares when T tends to infinity is based on the stability conditions that $|\,\beta_\ell^D\,| < 1$, $\ell = 1, \ldots, L$, and the fact that (by definition) $\mathrm{E}(\xi_{g\ell s}) = 0$, $\forall s$. These conditions ensure that the first term of the RHS of Equation 9 vanishes, as the infinite sum of the error terms. Therefore the long run equilibrium does not depend upon the initial value of $\widetilde{S}_{g\ell t}$ and reduce to a cross-sectional relationship. The static long run equation obtained is simply: $\lim_{T\to\infty} \mathrm{E}\big(\widetilde{S}_{g\ell(t+T)}\big) \approx \overline{\mathbf{R}}_g^{*\top} \boldsymbol{\beta}_\ell^{R*} + \mathbf{B}_g^\top \boldsymbol{\beta}_\ell^{B*}$ with the convergence condition about the growth rate of net returns $|\,\kappa_\ell\,| > |\,\beta_\ell^D\,|$, $\ell = 1, \ldots, L$.[8] Considering land use and land price curretly observed as driven by such a long term equilibria, the models can be estimated with classical techniques: both by OLS or GAM with geographical coordinates. Using the property about the sum of a geometric series, one can show that the estimated long run coefficients correspond to $\boldsymbol{\beta}_\ell^{R*} = \boldsymbol{\kappa}\boldsymbol{\beta}_\ell^B/(\boldsymbol{\kappa} - \beta_\ell^D)$ and $\boldsymbol{\beta}_\ell^{B*} = \boldsymbol{\beta}_\ell^B/(1-\beta_\ell^D)$.

Such long run equilibria can also be obtained from spatial econometric models with very similar computations close to LeSage and Pace (2009), reported in details in subsection A.1. Here, we only report the reduced forms that will be estimated and used to predict long run land use equilibrium from spatial econometric models:

$$\begin{aligned}\lim_{T\to\infty} \mathrm{E}\big(\widetilde{S}_{g\ell(t+T)}\big) \approx (I_G - \rho_\ell^*\mathbf{W})^{-1}(\overline{\mathbf{R}}_t^* \boldsymbol{\beta}_\ell^{R*} + \mathbf{W}\overline{\mathbf{R}}_t^* \boldsymbol{\theta}_\ell^{R*}) \\ + (I_G - \rho_\ell^*\mathbf{W})^{-1}\big(\mathbf{B}\boldsymbol{\beta}_\ell^{B*} + \mathbf{W}\mathbf{B}\boldsymbol{\theta}_\ell^{B*}\big) + (I_G - \rho_\ell^*\mathbf{W})^{-1}(I_G - \lambda_\ell^*\mathbf{W})^{-1}\xi_{g\ell}\end{aligned} \tag{10}$$

where I_G is the identity matrix of dimension $(G \times G)$.

Previous equations for $\ell \neq l$ correspond to the reduced form of the SMC model without lagged land use shares at the RHS, but still with spatially autocorrelated errors. It is worth mentioning that such long run equilibrium models present some simultaneous spatial autocorrelation on the dependent variable through the terms $(I_G - \rho_\ell^*\mathbf{W})^{-1}$ that can be factorized.[9] The equivalence with the parameters from short run spatial econometric models can be recovered as $\boldsymbol{\beta}_\ell^{R*} = \boldsymbol{\kappa}\boldsymbol{\beta}_\ell^R/(\boldsymbol{\kappa} - \beta_\ell^D)$, $\boldsymbol{\beta}_\ell^{B*} = \boldsymbol{\beta}_\ell^B/(1-\beta_\ell^D)$, $\boldsymbol{\theta}_\ell^{R*} = \boldsymbol{\kappa}\boldsymbol{\theta}_\ell^R/(\boldsymbol{\kappa} - \beta_\ell^D)$, $\boldsymbol{\theta}_\ell^{B*} = \boldsymbol{\theta}_\ell^B/(1-\beta_\ell^D)$, $\lambda_\ell^* = \lambda_\ell/(1-\beta_\ell^D)$ and $\rho_\ell^* = \rho_\ell + (\theta_\ell^D/(1-\beta_\ell^D))$.

[8] When $T \to \infty$, the net returns also tend to infinity. We consider that the long run equilibrium for land use is reached earlier, when the economic returns are equal to $\overline{\mathbf{R}}_g^*$.

[9] The long run spatial multipliers ρ_ℓ^* can be factorized thanks to two simplifications: ignoring second order spatial relationships and neglecting the attenuation effect of the growth of net returns. We will show that these simplifications will not change the main economic interpretations and predictive accuracies.

3 Performing predictions

Predicting LUC from all the econometric models of our comparative framework needs a careful analysis of the scope of interest (in-sample or out-of-sample predictions, temporal forecasting or not) and the available conditioning information (the endogenous/exogenous variables of other periods, and of neighboring units). In particular, any relevant comparative exercise between aspatial/spatial and short/long run models should be based on a precise description of the conditioning information sets, and the properties (bias, efficiency) of the predictors used.

For both individual and aggregate 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 predicting the value of the dependent variable for the same period 2003 that is not used to estimate the models. This is the case most interesting for policy implications. Moreover, in both cases, in-sample and out-of-sample predictions are performed for short-run and long-run models.

3.1 From aspatial individual models

For the individual MNL models, the direct predictions (without changing exogenous variables) consist, for each plot i, of a fitted probability vector $\widehat{\mathbf{p}}_{it}$ of being in each use at t. Assuming $L = 4$ and that each observation counts for 100 ha (in anticipation of our empirical application), the predicted probabilities can easily be converted into aggregate LUC. For example, consider a plot i which counts for 100 ha of annual crop in period $t-1$ 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 aggregate 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. Predicting from fractional models (FRA and DIR) is more direct as they produce well-shaped aggregate shares. To evaluate the effect of these desirable prediction properties, we also estimate some linear probability models on individual data that do not take account of the discrete nature of land use choices but are less computationally intensive. Within this framework, 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 $\overline{\mathbf{r}}_{g_i(t+1)}$ allow to obtain the vector $\widehat{\mathbf{p}}_{i(t+1)}$ of predictions and compute the aggregate land use shares.

3.2 From aspatial aggregate models

On aggregate models, we compare three categories of predictors, the form of which differ for in-sample and out-of-sample prediction and also for short and long run models.

The first category of predictors are the aspatial predictors. They are based on short-run and long-run models without endogenous or exogenous spatial lag variables and without spatial error terms, i.e. Equation 6 and Equation 7. Their form are described in the first section of Table 1 both for the in-sample and the out-of-sample case. They only involve the knowledge of the current observations for the exogenous 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 in the aspatial models, these predictors are the best linear unbiased. More specifically, we compute four aspatial predictors : (i) aspatial predictor on a simple land use share model estimated by OLS (ii) aspatial predictor on a land use share model including a smoothed function of the geographical coordinates of the grids' centroids (GAM) and (iii) the aggregate shares from fractional models.

3.3 From spatial aggregate models

The issue of prediction in spatial econometric models has gained considerable attention in the last decade. For instance, Baltagi and Li (1999) and Baltagi et al. (2012) derive the Best Linear Unbiased Predictor (BLUP, see Goldberger (1962)) for static spatial panel data models with random effects and Baltagi et al. (2014) derive the BLUP for a dynamic spatial panel data model with random effects. When it comes to comparing the predictive performances of models, Baltagi and Li (2006) 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. Angulo and Trivez (2010) 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, Kholodilin et al. (2008) 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, Schanne et al. (2010) 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.

The first category of spatial predictors is labeled reduced spatial predictors. They are based on the range of spatial models described in Equation 8 when they are rewritten in reduced form:

$$\begin{aligned}\widetilde{S}_{\ell t} \approx (I - \rho_\ell \mathbf{W})^{-1}(\beta_\ell^D \widetilde{S}_{\ell(t-1)} + \theta_\ell^D \mathbf{W} \widetilde{S}_{\ell(t-1)} \\ + \overline{\mathbf{R}}_t \boldsymbol{\beta}_\ell^R + \mathbf{W}\overline{\mathbf{R}}_t \boldsymbol{\theta}_\ell^R + \mathbf{B}\boldsymbol{\beta}_\ell^B + \mathbf{W}\mathbf{B}\boldsymbol{\theta}_\ell^B + \lambda_\ell \mathbf{W}\xi_{\ell t} + \eta_{\ell t})\end{aligned} \tag{11}$$

The form of reduced spatial predictors in the in-sample and out-of-sample cases are described in the second sections 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 Kholodilin et al. (2008) and Schanne et al. (2010). They also correspond to the first unbiased predictor suggested by Kelejian and Prucha (2007).[10] 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 independent variables.

[10]The case figure considered by Kelejian and Prucha (2007) ("leave-one-out" predictors) and also by Pace and LeSage (2008) ("ex-sample predictors") is different as they are concerned with spatial out-of-sample prediction: the case where for a cross-section of observations, part of the observations on the dependent variable is missing and needs to be imputed.

Table 1: The available predictors from 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. The same predictors are also used to perform counterfactual scenarios at subsection 5.4.

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	$\widehat{\widetilde{S}}_{\ell t} = \widehat{\rho}_\ell \mathbf{W}\widetilde{S}_{\ell t} + \widehat{\beta}_\ell^D \widetilde{S}_{\ell(t-1)} + \widehat{\theta}_\ell^D \mathbf{W}\widetilde{S}_{\ell(t-1)} + \mathbf{X}_t\widehat{\boldsymbol{\beta}} + \widehat{\lambda}_\ell \mathbf{W}(\widetilde{S}_{\ell t} - \mathbf{X}_t\widehat{\boldsymbol{\beta}})$	$\widehat{\widetilde{S}}_{\ell(t+1)} = (\widehat{\rho}_\ell + \widehat{\theta}_\ell^D)\mathbf{W}\widetilde{S}_{\ell t} + \widehat{\beta}_\ell^D \widetilde{S}_{\ell t} + \mathbf{X}_t\widehat{\boldsymbol{\beta}} + \widehat{\lambda}_\ell \mathbf{W}(\widetilde{S}_{\ell t} - \mathbf{X}_t\widehat{\boldsymbol{\beta}})$

The third category of predictors 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 sections 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 (Bivand, 2002) 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 inertia over time.

4 Data

4.1 Land use data

Data on land use are extracted from the TERUTI survey (AGRESTE, 2004), which is carried out every year by the statistical services of the French Ministry of Agriculture. It collects data on land use through the whole continental territory of France. It counts 550,903 points localized at the level of the smallest administrative division (commune) observed from 1992 to 2003. The survey uses a systematic area frame sampling with a two-stage sampling design. In the first stage, the total land area of France is divided into 12×12 km grids. For each of the 4,700 grids there are four aerial photographs which cover 3.5 km^2 each. In the second stage, on each photograph, a 6×6 grid determines 36 points (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 aggregate items:[11] 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 2 presents the raw transitions 1993–2003.

Table 2: 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

[11]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.

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

As mentioned in footnote 4, the presence of zeros in the denominator of the logit transformation is a limit of the logit transformation for aggregate modeling that is overridden by adding ϵ both in the numerator and the denominator. As Figure 5 and Figure 6 of the Appendix B.10 show, the logit transformation produces some mass probabilities around the value -7 but the distribution of the outcome is undoubtedly closer to that of a normal distribution than raw land use shares were.

4.2 Explanatory variables

The theoretical literature on land use suggests that the explanatory variables introduced in models include the net return to each land use and the distribution of land quality. As presented in subsection 2.1, we use current land price to proxy the net return to each land use according to the Ricardian formula and the assumption of no anticipation about land-use changes. We use precise data about land prices of arable land and pastures from the French ministery of agriculture that are in accordance with these assumptions. In effect, because such data are collected to reflect the "agricultural productive value" of land, the sales that are judged to be too much driven by land-use change anticipation (by urbanization but also by nature conservation) are dropped from official data by agents that generally know the context well. Forest returns are proxied by the current aggregate production (in value) divided by the current forest acreages, and population densities are used as proxies for the economic returns from urban use. Finally, we include some biophysical attributes: slope, altitude, water holding capacity (WHC), and climate. The following Table 3 displays summary statistics for these variables aggregated at the grid scale but numerous are available at the individual scale.

Table 3: 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

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

5 Results

5.1 Specifications

Our comparative set includes a wide spectrum of econometric LUC models of the literature. We estimate a total of 12 types of models at different scales, and with different space and time structures. Moreover, each specification is estimated two times. The short run models (with time-lagged land use) are estimated on the 1993–1998 and 1993–2003 periods, and the long run models (without time-lagged land use) are estimated on 1998 and 2003 cross-sections. Because the predictions are computed from the values 2003 of explanatory variables in all cases, the first series of (both short run and long run) models is used to perform out-of-sample predictions and the second for in-sample predictions.

To keep the models comparable, we use the same specifications for the effects of explanatory variables.[12] 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 aggregate 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 raw estimation results based on maximum likelihood of different model specifications are provided in the Supporting Information B (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 Dirichlet estimations (close to the fractional FRA), 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 request.

[12]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. We choose to not consider this possibility of more complex regression functions as a comparative advantage of individual models.

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 probability, are not found to be restrictive (homoskedasticity cannot be rejected by a score test, p-value= 0.283). Moreover, these assumptions are associated with the classical restriction of Independence of Irrelevant Alternatives for which Hausman-McFadden specification tests were performed, with mixed evidence. The independence is not rejected for two uses: pasture and urban (p-values are respectively 0.001, 0.005) but is rejected for arable and forest at 5%. This means that the former choices can be dropped from the choice set without significant modification to the model (i.e., they are robust to the IIA restriction), a property that does not apply to the latter two choices. In the literature, use of nested multinomial logit is found not to change the results (Lubowski et al., 2008; Li et al., 2013).

By comparing the significance of coefficients from aspatial models (Table 9 to 12 in the SI B), 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 land uses. On aggregate 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. The spatial smoothed functions estimated by the GAMs are displayed in Figure 7 and Figure 8 of Appendix B.11. 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 aggregate) and whatever the considered land use, 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 price keep their significance for individuals models but are not longer significant in the aggregate models. It is clearly related to the number of observation as this effect of the loss of significance is also present for fractional models. This can be considered as an expression of multicolinearity in short run aggregate models that is not present in the short run individual models because they use more observations (Kennedy, 2003).

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 Florax et al., 2003, Elhorst, 2010 or Le Gallo, 2014 for reviews of these spatial specification searches). Instead, 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 LeSage and Pace (2009), 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. However, as shown in Figure 9 and Figure 10 of SI B.12 that display the Moran scatter plots of regression residuals in the OLS and GAM models, it also allows to decrease or render spatial error autocorrelation insignificant.

Table 7 and Table 8 in the subsection A.2 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 a 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 7 and Table 8), 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 aggregate grid level. The comparisons are reported in the panels A, B, Table 4 for in-sample predictions and in panels A, B, Table 5 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 aggregate models, and (iii) spatial aggregate 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 int terms of predictive accuracy whereas the advantage of using spatial econometrics techniques is clearly presented in Table 1 above.

Looking in more details the results of Table 4 and Table 5, 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 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. Spatial models gains relative to OLS are half of the gains of OLS relatively to the benchmark REF. Thus, the effect is strong. In the same magnitudes, the GAM is in an intermediate position between the spatial and the aspatial models. For the aspatial models (both

Table 4: In-sample Root Mean Square Errors for the different predictors: *The rows marked REF report the benchmark RMSE from constant predictions (i.e., the mean share for each unit). The columns MEAN report the row means of RMSE. OLS is for Ordinary Least Squares, GAM for Generalized Additive Models, FRA for aggregate fractional model, DIR for aggregate Dirichlet model, SEM for spatial error model, SXM for model with spatially-lagged explanatory variables, SAR for spatial autoregressive model, SDM for spatial Durbin model, SAC for the spatial error spatial autoregressive model, 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
MNL	0.155	0.1491	0.1709	0.0618	0.1408
DIR	0.1558	0.1527	0.1735	0.0694	0.1436
Aggregate 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
Aggregate aspatial models					
OLS	0.0405	0.0379	0.0282	0.0182	0.0324
GAM	0.0396	0.0366	0.028	0.0181	0.0317
MNL	0.0516	0.0465	0.043	0.0295	0.0434
DIR	0.0541	0.0498	0.0461	0.0304	0.046
Aggregate 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 5: Out-of-sample Root Mean Square Errors for different predictors: *The rows marked REF report the benchmark RMSE from constant predictions (i.e., the mean share for each unit). The columns MEAN report the row means of RMSE. OLS is for Ordinary Least Squares, GAM for Generalized Additive Models, FRA for aggregate fractional model, DIR for aggregate Dirichlet model, SEM for spatial error model, SXM for model with spatially-lagged explanatory variables, SAR for spatial autoregressive model, SDM for spatial Durbin model, SAC for the spatial error spatial autoregressive model, 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	0.1675	0.1717	0.19	0.0677	0.1567
GAM	0.1342	0.1525	0.1645	0.0656	0.1347
FRA	0.1616	0.1556	0.1771	0.0633	0.1464
DIR	0.1625	0.1567	0.1737	0.0714	0.1468
Aggregate 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
Aggregate aspatial models					
OLS	0.0317	0.0308	0.0202	0.0128	0.0251
GAM	0.0308	0.0309	0.0212	0.0133	0.0251
FRA	0.0478	0.0427	0.0416	0.0293	0.0409
DIR	0.0503	0.0461	0.0459	0.0299	0.0437
Aggregate 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

aggregate and individual) the predictive abilities are rather similar and the individual linear probability model is the worst. Note that the multicollinear models such as SAC and SMC, perform the best, according to a property that multicollinearity does not bias the predictions. Including lagged land uses for short run predictions drastically decreases the RMSE, and the differences between estimation techniques also decrease significantly. The spatial models perform best, but the performance of the GAM model is also quite similar. More importantly, the inclusion of temporal lag implies a loss of relative performance in the models (aggregate and individual) based on discrete outcomes: FRA, DIR, LPB, MNL. These results are confirmed by the maps presented in the appendix A.3 and A.4.

5.4 Counter-factual simulations

Turning to the counter-factual simulations, we use the 1993–1998 models in out-of-sample predictions (i.e., for 2003) to simulate the impacts of a subsidy of 200 euros per ha and per year for pastures. This represents an example of a green policy which proposes incentives for extensive land use in order to improve water quality and preserve biodiversity. We compute the differences in terms of aggregate pasture acreages in 2003 relatively to what is observed. Note that spatial predictions take into account both the direct and indirect effects of payments, so they imply no differences for the interpretation of changes to acreages at the aggregate scale.

In terms of aggregate changes in pasture acreages as a consequence of increasing their economic returns, there is a significant gap between individual and aggregate models when comparing short run and long run models. Predictions are relatively similar from a long term perspective (except for GAM and SEM which appear to strongly underestimate the effects) but in the short run, the aggregate models clearly underestimate the effects. We interpret this result as indicating an indisputable advantage of individual models (amount of information), which show significant effects of economics returns in short run models when lagged land use captures a lot of the effect. In contrast, no specific patterns emerge for the aggregate models in either the short run or long run. OLS seems to be in mid position, between the spatial econometric models that sometimes perform well (SXM in the long run and SAR in the short run) but appears really contrasted with the individual model. Aggregate FRA and DIR perform less badly in relation to predictive accuracy terms, at least in the long run.

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

Table 6: Simulation of 200 euros payments for pasture: acreages variations *The table reports the variations of pasture acreages for 2003 at the national scale on the basis of models estimated on the period 1993–1998. The t stat. are relative to the nullity of the correlation with the individual MNL model. The units are in thousand ha.*

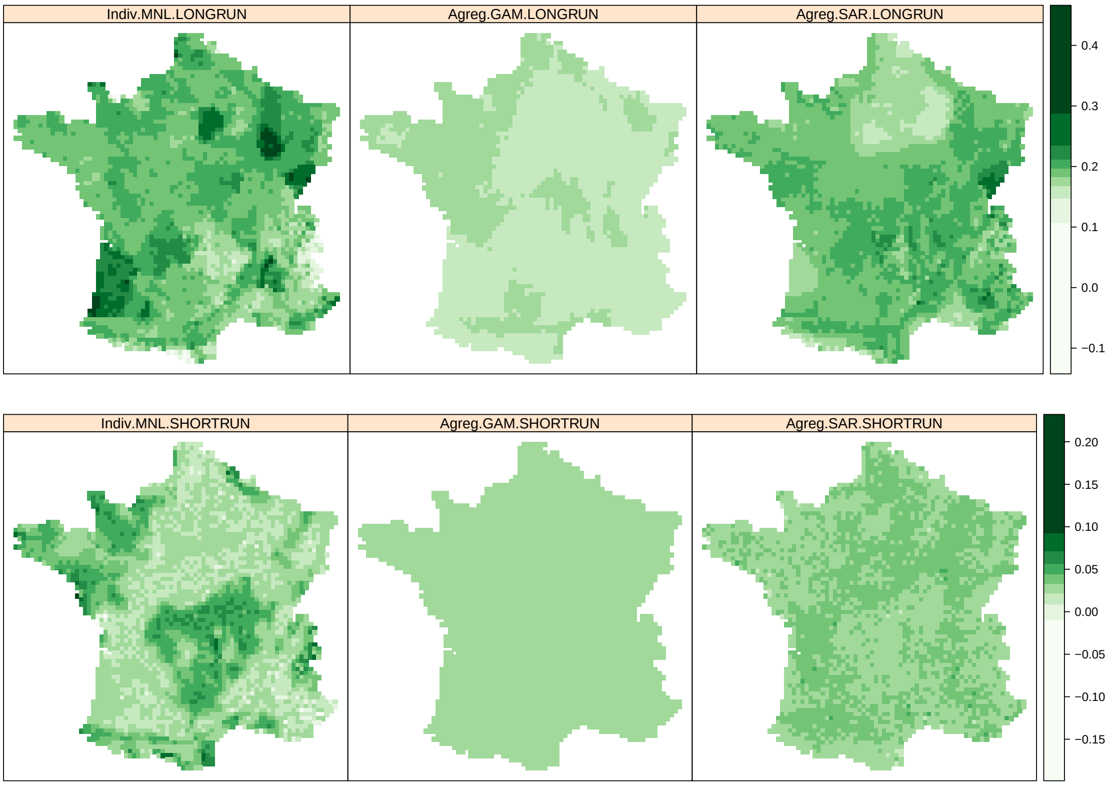
	LONG RUN			SHORT RUN		
	NET EFFECT	COR(MNL)	t stat.	NET EFFECT	COR(MNL)	t stat.
Aggregate predictors						
OLS	+ 714.5	− 0.018	− 1.155	+ 41.82	+ 0.131	+ 8.084
GAM	+ 126.4	− 0.212	− 13.32	+ 30.52	+ 0.243	+ 15.36
FRA	+ 651.3	+ 0.061	+ 3.727	− 3.467	− 0.292	− 18.72
DIR	+ 568.8	+ 0.049	+ 3.020	+ 13.12	+ 0.271	+ 17.30
Spatial Reduced predictors						
SEM	+ 298.6	+ 0.074	+ 4.547	+ 45.05	+ 0.128	+ 7.938
SXM	+ 713.5	+ 0.224	+ 14.08	− 13.12	− 0.028	− 1.775
SAR	+ 667.3	+ 0.300	+ 19.13	+ 91.00	+ 0.012	+ 0.735
SDM	+ 885.7	+ 0.169	+ 10.51	+ 51.13	+ 0.001	+ 0.071
SAC	+ 382.1	+ 0.025	+ 1.558	+ 18.10	+ 0.023	+ 1.400
SMC	+ 285.8	+ 0.016	+ 9.992	− 27.25	− 0.033	− 2.011
Spatial structural predictors						
SEM	+ 181.2	+ 0.026	+ 1.624	+ 46.62	+ 0.157	+ 9.732
SXM	+ 326.4	+ 0.031	+ 1.913	+ 8.484	− 0.022	− 1.380
SAR	+ 325.7	+ 0.054	+ 3.333	+ 19.99	+ 0.061	+ 3.767
SDM	+ 391.9	+ 0.021	+ 1.311	+ 7.408	− 0.021	− 1.299
SAC	+ 27.93	+ 0.125	+ 7.726	+ 17.19	+ 0.032	+ 1.200
SMC	+ 174.9	+ 0.131	+ 8.120	− 23.96	− 0.012	− 0.786
Individual predictors						
LPB	+ 526.7	+ 0.783	+ 35.12	+ 150.2	+ 0.848	+ 45.72
MNL	+ 756.2	1.000		+ 166.2	1.000	

6 Conclusion

It is widely acknowledged that micro-economic behaviors are more accurately analyzed using individual data and models. However, with the recent developement of individual land use models, aggregate models are still appealing as they ease estimation with more refined econometric technics and are less data and computational demanding. The comparative advantage of individual models in term of prediction accuracy has remained an open question that was investigate in this paper.

We have compared the predictive abilities of a wide spectrum of econometric models of land use for different scales (individual vs aggregate), spatial (with vs without spatial autocorrelation) and temporal structures (short term vs long term). More specifically, we have showed how the introduction of a spatial dimension in aggregate models matters for improving their predictions related to aggregate changes in land use. Our results suggest that: (i) introducing spatial autocorrelation in aggregate grid-level models improves their predictive accuracy and even outperforms individual models if appropriate predictors are used, (ii) a specification including lagged land use as explanatory variable in the aggregated as well in the individual models, outperforms any other spec-

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



ification where only economic and biophysical variables are included, (iii) in terms of policy simulation, individual models perform better than aggregate models and cannot be totally substituted.

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, we show how the introduction of spatial autocorrelation in aggregated land use models allow more precise predictions than individual models. However, for policy simulation aggregated models give very different results from individual models. So, individual land use models are needed for simulation for example if the focus is to study the impact of land use changes on greenhouse gas emissions or other local environmental issues such as biodiversity loss or ground-water pollution.

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. However, our results suggest that model that best predicts at aggregated scale is not necessarily the most appropriate for simulation of public policies. The introduction of spatial autocorrelation can significantly improve the prediction of aggregated models however it doesn't allow to make them more precise in terms of policy simulation. Thus, choosing the best land use econometric model will depend on the purpose of the study.

7 Acknowledgements

This research has been founded by the FRB (Fondation de Recherche sur la Biodiversité) and GDF–SUEZ through the MOBILIS project. The authors also acknowledge the financial support from French Agence Nationale de la Recherche through the ModULand project (ANR–11–BSH1–005).

References

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

Angulo, A. and Trivez, F. (2010). The impact of spatial elements on the forecasting of spanish labour series. *Journal of Geographical Systems* 12: 155–174.

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

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

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

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

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

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

Baltagi, B. and Li, D. (1999). Prediction in the spatially autocorrelated error component model. *Econometric Theory* 15: 2: 259.

Baltagi, B. and Li, D. (2006). Prediction in the panel data model with spatial correlation: the case of liquor. *Spatial Economic Analysis* 1: 175–195.

Baltagi, H. B., Fingleton, B. and Pirotte, A. (2014). Estimating and forecasting with a dynamic spatial panel data model. *Oxford Bulletin of Economics and Statistics* forthcoming.

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

Bell, K. P. and Bockstael, N. E. (2000). Applying the generalized-moments estimation approach to spatial problems involving micro-level data. *Review of Economics and Statistics* 82: 72–82.

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

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

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

Brady, M. and Irwin, E. (2011). Accounting for spatial effects in economic models of land use:

Recent developments and challenges ahead. *Environmental & Resource Economics* 48: 487–509.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Kennedy, P. (2003). *A Guide to Econometrics*. Blackwell.

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

Klier, T. and McMillen, D. P. (2008). Clustering of auto supplier plants in the United States. *Journal of Business & Economic Statistics* 26.

Knittel, C. R. and Metaxoglou, K. (2012). Estimation of random-coefficient demand models: Two empiricists' perspective. *Review of Economics and Statistics* .

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

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

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

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

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

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

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

Lubowski, R. N. (2002). 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* .

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

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

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

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

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

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

Pace, R. K. and LeSage, J. P. (2008). Spatial econometric models, prediction. In Shekhar, S. and Xiong, H. (eds), *Encyclopedia of Geographical Information Science*. 10.1007/978-0-387-35973-1: Springer-Verlag.

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

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

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

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

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

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

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

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

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

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

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

Wang, Y., Kockelman, K. M. and Wang, X. C. (2013). Understanding spatial filtering for analysis of land use-transport data. *Journal of Transport Geography* 31: 123–131.

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

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

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

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

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

A Appendix (for publication)

A.1 Long run equilibrium from spatial econometric models

We begin the computations with Equation 8 of the main text, which describes the short run SMC model of logit linearized land use shares:

$$\begin{aligned}\widetilde{S}_{\ell t} \approx\ & \rho_\ell \mathbf{W} \widetilde{S}_{\ell t} + \beta_\ell^D \widetilde{S}_{\ell(t-1)} + \theta_\ell^D \mathbf{W} \widetilde{S}_{\ell(t-1)} \\ & + \overline{\mathbf{R}}_t \boldsymbol{\beta}_\ell^R + \mathbf{W}\overline{\mathbf{R}}_t \boldsymbol{\theta}_\ell^R + \mathbf{B}\boldsymbol{\beta}_\ell^B + \mathbf{W}\mathbf{B}\boldsymbol{\theta}_\ell^B + \lambda_\ell \mathbf{W}\xi_{\ell t} + \eta_{\ell t}\end{aligned} \tag{12}$$

Updating the reduced form of this relationship for $t+1, t+2, \dots, t+T$ and recursively substituting the equations from previous periods, we obtain:

$$\begin{aligned}(I_G - \rho_\ell \mathbf{W})\widetilde{S}_{\ell(t+T)} \approx\ & \left[\kappa^T + \kappa^{T-1}(\beta_\ell^D I_G + \theta_\ell^D \mathbf{W}) + \dots + \kappa(\beta_\ell^D I_G + \theta_\ell^D \mathbf{W})^{T-1}\right](\overline{\mathbf{R}}_t \boldsymbol{\beta}_\ell^R + \mathbf{W}\overline{\mathbf{R}}_t \boldsymbol{\theta}_\ell^R) \\ & + \left[I_G + (\beta_\ell^D I + \theta_\ell^D \mathbf{W}) + \dots + (\beta_\ell^D I_G + \theta_\ell^D \mathbf{W})^{T-1}\right](\mathbf{B}\boldsymbol{\beta}_\ell^B + \mathbf{W}\mathbf{B}\boldsymbol{\theta}_\ell^B + \lambda_\ell \mathbf{W}\xi_{g\ell}) \\ & + \eta_{\ell(t+T)} + (\beta_\ell^D I_G + \theta_\ell^D \mathbf{W})\eta_{\ell(t+T-1)} + \dots + (\beta_\ell^D I_G + \theta_\ell^D \mathbf{W})^{T-1}\eta_{\ell t}.\end{aligned}$$

If the convergence conditions presented in the main text apply, we can simplify some geometric sums as follows:

$$\begin{aligned}(I_G - \rho_\ell \mathbf{W})\widetilde{S}_{\ell(t+T)} \approx\ & \left[I_G - (\beta_\ell^D I_G + \theta_\ell^D \mathbf{W})/\boldsymbol{\kappa}\right]^{-1}(\overline{\mathbf{R}}_t^* \boldsymbol{\beta}_\ell^R + \mathbf{W}\overline{\mathbf{R}}_t^* \boldsymbol{\theta}_\ell^R) \\ & + (I_G - \beta_\ell^D I_G - \theta_\ell^D \mathbf{W})^{-1}(\mathbf{B}\boldsymbol{\beta}_\ell^B + \mathbf{W}\mathbf{B}\boldsymbol{\theta}_\ell^B + \lambda_\ell \mathbf{W}\xi_{g\ell}) \\ & + \eta_{\ell(t+T)} + (\beta_\ell^D I_G + \theta_\ell^D \mathbf{W})\eta_{\ell(t+T-1)} + \dots + (\beta_\ell^D I_G + \theta_\ell^D \mathbf{W})^{T-1}\eta_{\ell t}.\end{aligned}$$

Finally, after rearranging the autoregressive terms and taking the $T \to \infty$ long run expectation, we obtain:

$$\begin{aligned}(I_G - \rho_\ell \mathbf{W})\mathrm{E}_{T\to\infty}(\widetilde{S}_{\ell(t+T)}) \approx\ & \frac{\kappa}{\kappa - \beta_\ell^D}\left(I_G - \frac{\theta^D}{\kappa - \beta_\ell^D}\mathbf{W}\right)^{-1}(\overline{\mathbf{R}}_t^* \boldsymbol{\beta}_\ell^R + \mathbf{W}\overline{\mathbf{R}}_t^* \boldsymbol{\theta}_\ell^R) \\ & + \frac{1}{1 - \beta_\ell^D}\left(I_G - \frac{\theta^D}{1 - \beta_\ell^D}\mathbf{W}\right)^{-1}(\mathbf{B}\boldsymbol{\beta}_\ell^B + \mathbf{W}\mathbf{B}\boldsymbol{\theta}_\ell^B + \lambda_\ell \mathbf{W}\xi_{g\ell})\end{aligned}$$

Equation 10 in the main text and the relationships between short and long run parameters are established with the factorization of the spatial multipliers based of two simplifications mentioned in footnote 7. We assume that $\kappa \approx 1$ and neglect the second-order neighboring effects $\mathbf{W}^2 \approx 0$:

$$\left(I_G - \rho_\ell \mathbf{W} - \frac{\theta_\ell^D}{1 - \beta_\ell^D}\mathbf{W} + \frac{\rho_\ell \theta_\ell^D}{1 - \beta_\ell^D}\mathbf{W}^2\right)^{-1} \approx \left[I_G - \left(\rho_\ell + \frac{\theta_\ell^D}{1 - \beta_\ell^D}\right)\mathbf{W}\right]^{-1}$$

A.2 Spatial coefficient from aggregate models

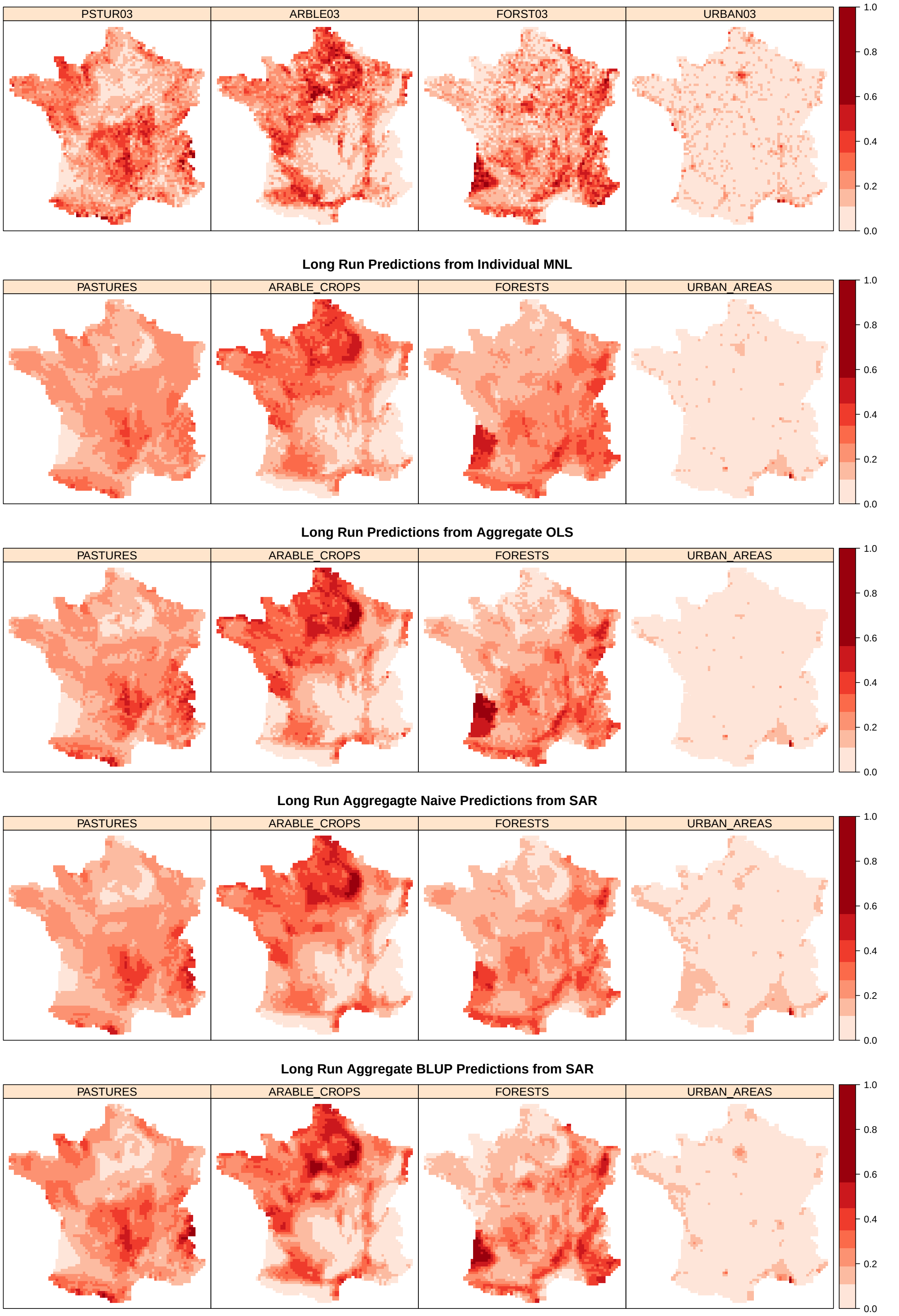
Table 7: Spatial coefficients for long run models

	Spatial Error components: λ_ℓ^*			Spatial Lag components: ρ_ℓ^*		
	AR	FO	UR	AR	FO	UR
SEM	0.6449**	0.7349**	0.4991**			
	(0.0183)	(0.0248)	(0.0217)			
SXM	0.626**	0.7019**	0.4902**			
	(0.0177)	(0.0158)	(0.0216)			
SAR				0.5654**	0.7017**	0.4586**
				(0.0171)	(0.0151)	(0.0209)
SDM				0.6205**	0.6944**	0.4877**
				(0.0174)	(0.0153)	(0.0215)
SAC	0.9093**	0.9306**	0.8166**	−0.6221**	−0.7208**	−0.6248**
	(0.0129)	(0.0086)	(0.0195)	(0.047)	(0.0426)	(0.0502)
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)

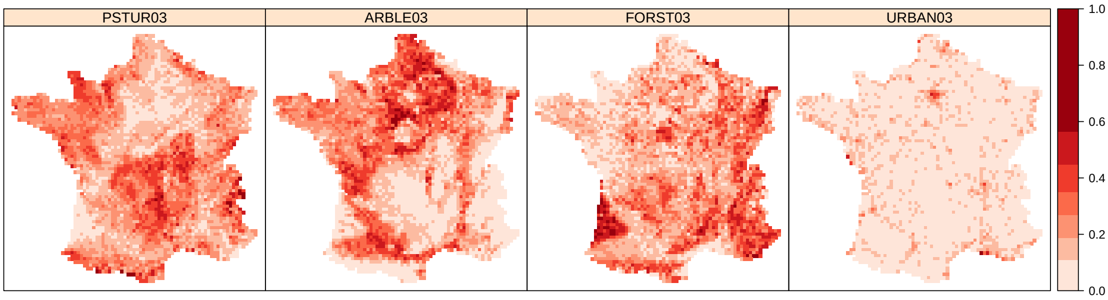
Table 8: Spatial coefficients for short run models

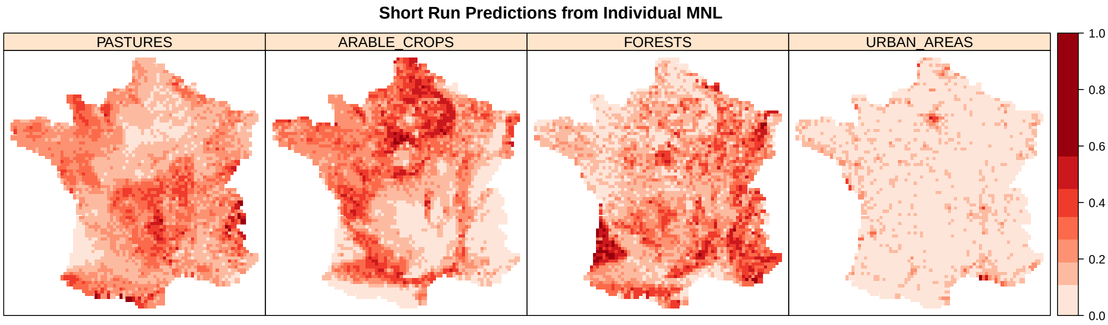
	Spatial Error components: λ_ℓ			Spatial Lag components: ρ_ℓ		
	AR	FO	UR	AR	FO	UR
SEM	0.1134**	0.3004**	0.2246**			
	(0.022)	(0.0295)	(0.0278)			
SXM	0.0473**	0.2404**	0.2**			
	(0.0131)	(0.0282)	(0.0288)			
SAR				0.1335**	0.1256**	0.1122**
				(0.0129)	(0.0087)	(0.0134)
SDM				0.0629**	0.2427**	0.2011**
				(0.0302)	(0.029)	(0.0287)
SAC	−0.1119**	0.1451**	0.1338**	0.1572**	0.1103**	0.0755**
	(0.0334)	(0.0324)	(0.0361)	(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)

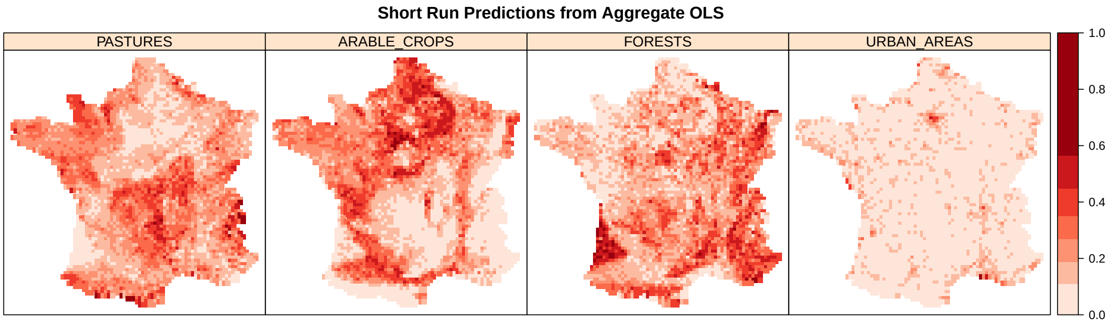
A.3 Observed land use in 2003 and long run predictions

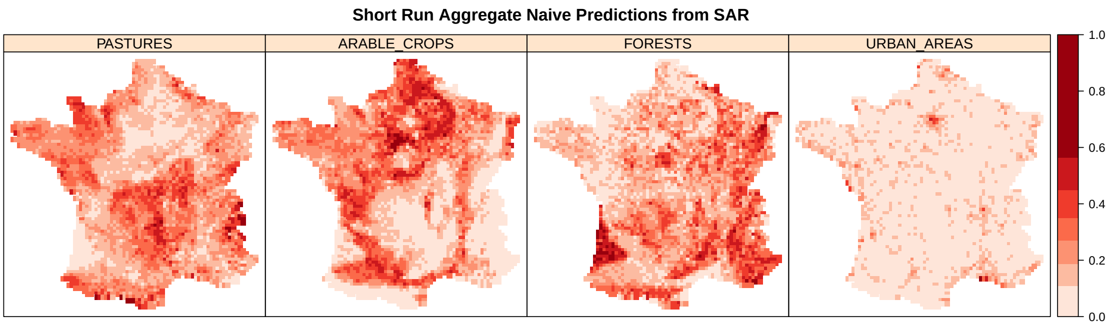


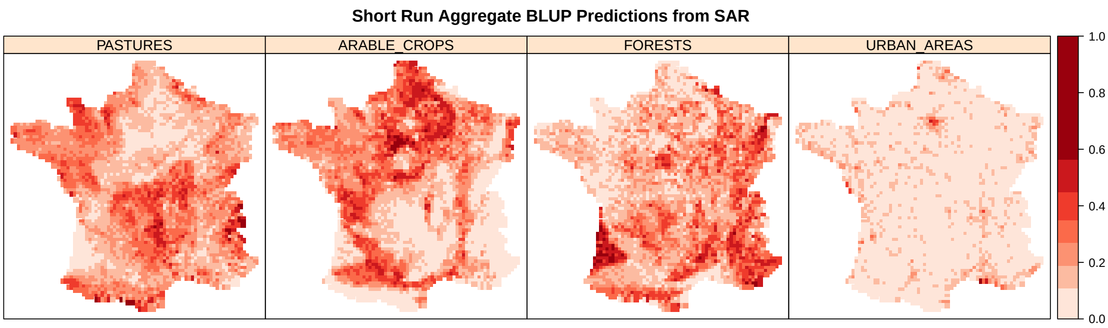
A.4 Observed land use in 2003 and short run predictions











B Supporting Information (not for publication)

B.1 Raw results from individual MNL

Table 9: 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					

B.2 Raw results from OLS

Table 10: 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.

B.3 Raw results fom GAM

Table 11: 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
UBRE	1.932	0.595	1.509	0.208	1.599	0.399

Note: *p<0.1; **p<0.05; ***p<0.01.
Reference= Pastures, scaled explanatory variables, bivariate smooth function of coordinates, see subsection B.11

B.4 Raw results from FRA fractional

Table 12: Aggregate FRA fractional models of land use on 1993–2003

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

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

B.5 Raw results from SEM

Table 13: 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

B.6 Raw results from SXM

Table 14: 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
Wald Test (df = 1)	1,251.796***	12.987***	1,982.371***	72.671***	514.239***	48.084***
LR Test (df = 1)	905.610***	2.284	1,271.983***	65.701***	421.610***	46.372***

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

B.7 Raw results from SAR

Table 15: 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					

B.8 Raw results from SDM

Table 16: Spatial Durban Models of land use on 1993–2003

	Arable Share		Forest Share		Urban Share	
	long run	short run	long run	short run	long run	short run
ARlog93		0.831***				
		(0.009)				
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

B.9 Maps at the aggregate scale

Figure 2: Aggregated land use shares in 2003

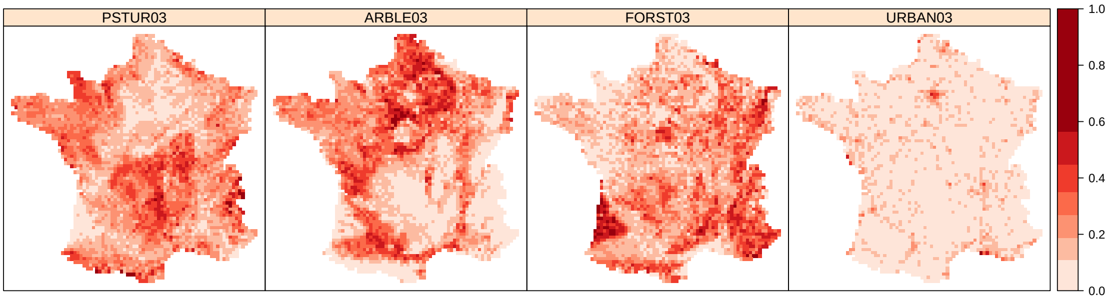


Figure 3: Aggregated land use variations on 1993–2003, in km^2

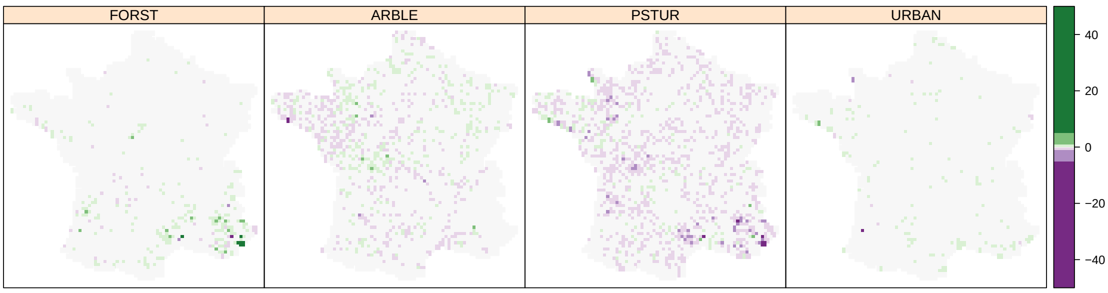
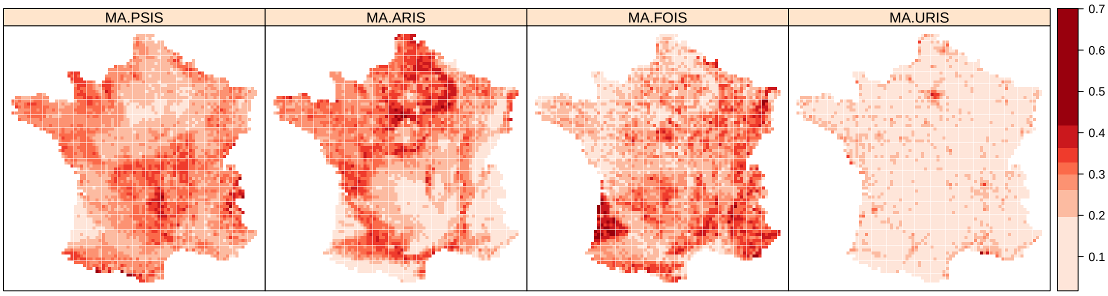


Figure 4: Out of sample 2003 predictions from individual mnl



B.10 Aggregate outcome variables

Figure 5: Raw distribution of 1998 aggregate land use shares

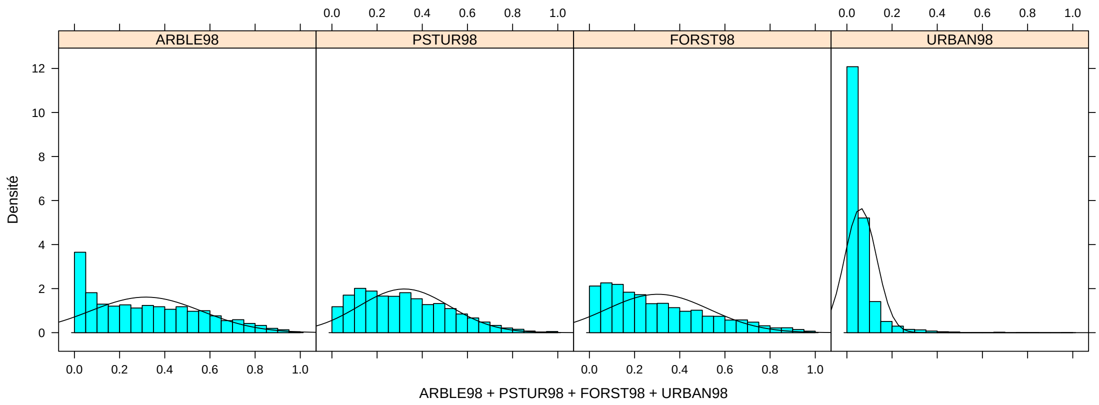
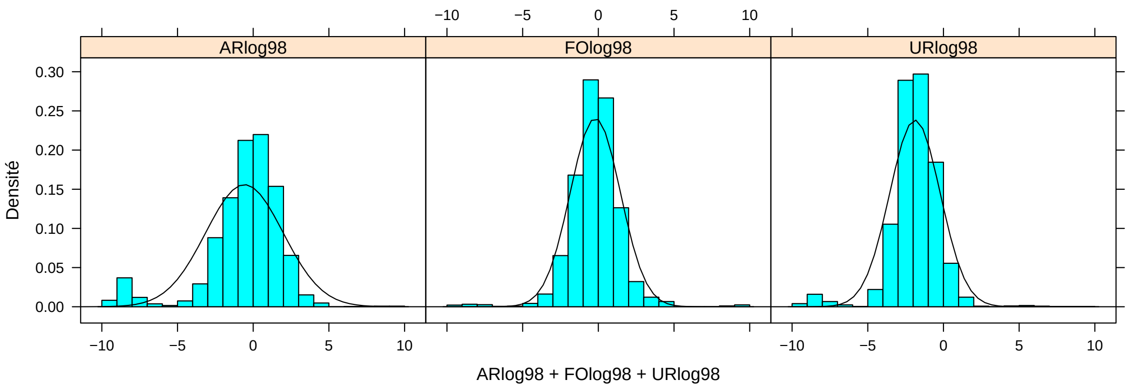


Figure 6: Linearized distribution of 1998 aggregate land use shares



B.11 Spatial Smoothing Functions

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

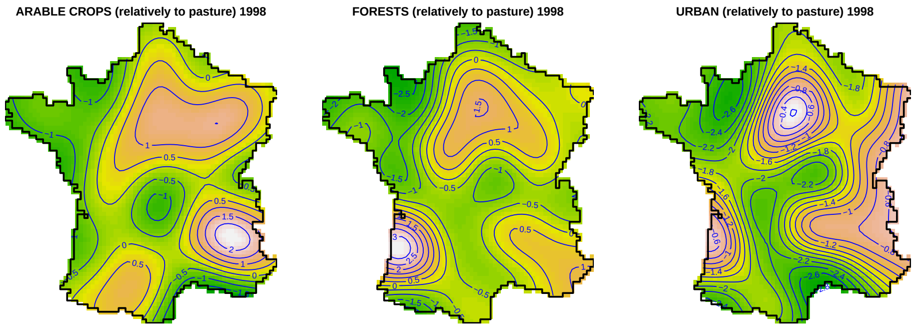
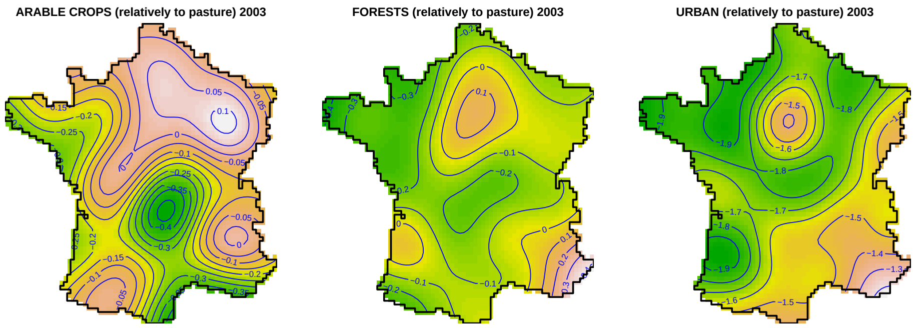


Figure 8: Semi-parametric smoothing functions of geographical coordinates: with temporal lags



B.12 Morans' *I* on residuals

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

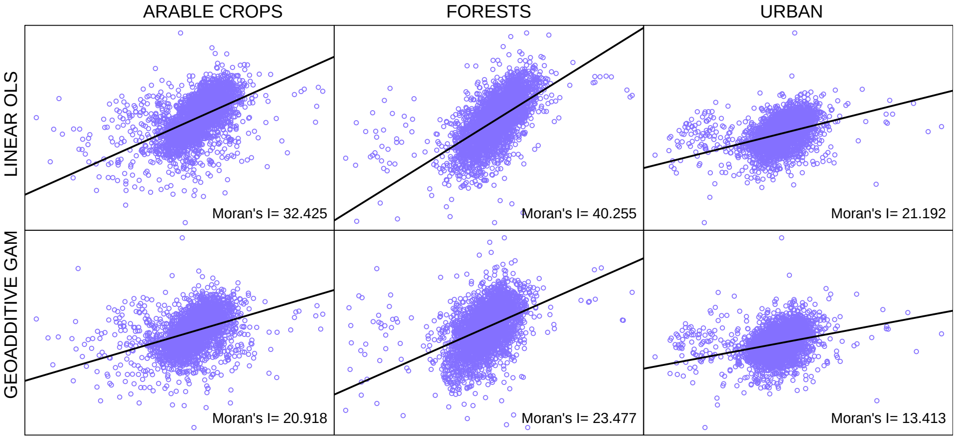


Figure 10: Morans' *I* from OLS and GAM with temporal lags