WORKING PAPERS

N° 1098

June 2021

"Spatial simultaneous autoregressive models for compositional data: Application to land use"

Christine Thomas-Agnan, Thibault Laurent, Anne Ruiz-Gazen, T.H.A Nguyen, Raja Chakir and Anna Lungarska

Toulouse School of Economics

Spatial simultaneous autoregressive models for compositional data: Application to land use

Christine Thomas-Agnan [1], Thibault Laurent [2], Anne Ruiz-Gazen [1], Nguyen Thi Huong An [1,3], Raja Chakir [4] & Anna Lungarska [4]

[1] *Toulouse School of Economics, University of Toulouse Capitole, France*
[2] *Toulouse School of Economics, CNRS, University of Toulouse Capitole, France*
[3] *Danang University of Architecture, Vietnam*
[4] *Université Paris-Saclay, INRAE, AgroParisTech, Economie Publique, 78850, Thiverval-Grignon, France.*

Abstract. Econometric land use models study determinants of land-use-shares of different classes:"agriculture", "forest", "urban" and "other" for example. Land-use-shares have a compositional nature as well as an important spatial dimension. We compare two compositional regression models with a spatial autoregressive nature in the framework of land use. We study the impact of the choice of coordinate space and prove that a choice of coordinate representation does not have any impact on the parameters in the simplex as long as we do not impose further restrictions. We discuss parameters interpretation taking into account the non linear structure as well as the spatial dimension. In order to assess the explanatory variables impact, we compute and interpret the semi-elasticities of the shares with respect to the explanatory variables and the spatial impact summary measures.

Keywords. spatial error regression models, spatial lag regression models, land-use-share model, simplicial regression, semi-elasticities, compositional data.

1 Introduction

Land use and land use changes are among the main human pressures on the environment in terms of biodiversity loss, carbon cycle and water quality [Foley et al., 2005, Pielke, 2005, Lal, 2004, Verburg et al., 2013]. Given the environmental impacts and societal issues associated to land use, numerous empirical land use models have been developed among different disciplines. The aim of these models is to support future land-use planning and environmental impact assessments of land-use change. All these models have mainly two different research focuses: explaining the underlying process behind land use or predicting spatial land use patterns or dynamics. The distinction between these two approaches is often not easy [Shmueli, 2010]. In summary, the purpose of econometric models, mostly developed by economists, is to test theoretical results and identify the economic processes behind observed land use changes and patterns. This leads to the search for parsimonious models, i.e. models with limited number of explanatory variables suggested by the theoretical model. Statisticians, geographers and other researchers

outside of economics typically have focused on prediction models of land use [Munroe and Müller, 2007, Veldkamp and Lambin, 2001]. These models are useful for prediction of land use patterns but they provide little insight into the underlying economic and other processes that generate these patterns. There is clearly no land use model which is preferable in absolute terms. The choice of the best model is largely dependent on the research question, data availability and calculation costs.

In terms of methodology, econometric land use studies can be classified into different groups depending on the type of data used (aggregated data vs. micro-geographic data), the type of model (spatially explicit vs. aspatial model) and the land use categories under consideration (rural vs. urban, agriculture vs. forest, and developed vs. non-developed).

Among land use models, we restrict attention to econometric formulations based on simultaneous spatial autoregressive models (SAR) because they are common in the econometric land use literature (see for instance [LeSage and Pace, 2009]) and easy to fit. There are alternative approaches based on conditional autoregressive models such as [Pirzamanbein et al., 2018] and [Leininger et al., 2013] but they require MCMC techniques, see [Nguyen et al., 2019] for further details.

Our starting point is the application to land use in [Chakir and Lungarska, 2017] and [Lungarska and Chakir, 2018] which is based on a spatial error model (SEM) in alr coordinates. The objective, using [Nguyen et al., 2019], is to propose and compare alternative models based on alternative coordinate systems (ilr instead of alr) and alternative model formulations in the same family (LAG instead of SEM as we will explain later on). Using the same dataset, we estimate both models and question model specification and interpretation. The two approaches differ in several important dimensions. First of all, the first approach specifies the model in an alr coordinate space whereas the second one uses an ilr coordinate space, see e.g. [Pawlowsky-Glahn et al., 2015] for the definition of alr and ilr transformations. Secondly the first approach uses a spatial error model formulation for each coordinate separately while the second approach uses a joint multivariate spatial lag model. Finally, the first approach performs a separate maximum likelihood for each coordinate whereas the second one uses the simultaneous Spatial two stage least squares method (acronym S2SLS) which will be presented in Section 2. This confrontation raises many questions: in particular is it possible to write a simplex formulation of the first model ? and since parameters of the models are linked to a particular transformation, what are the links between parameters associated to different transformations ? For models using ilr coordinates, the link between parameters in the simplex space and in coordinate space is well known ([Chen et al., 2017]). We extend these formulas to models using alr coordinates.

Section 3.2 is devoted to the interpretation of model parameters. The first approach uses plots of the fitted shares as a function of each explanatory variable. [Morais and Thomas-Agnan, 2020] prove that semi-elascities are natural tools for the interpretation of model parameters in simplicial regression models with a compositional dependent variable because they are derived from simplicial derivatives and because they describe variations in the simplex rather than in coordinate space. However their framework does not encompass the case of spatial models. We thus

derive an approximation formula for semi-elasticities in the present models which is illustrated later in Section 4.3. In Section 4, after presenting the data set and paying attention to the treatment of missing values according to compositional data analysis (CoDa) principles, we analyze it with the different models and propose a comparison.

2 Multivariate SAR models

In this section, after recalling some classical facts about compositional data analysis, we recall the models from [Chakir and Lungarska, 2017] and [Nguyen et al., 2019] with a common notation and refer to the first one as to the UniSEM model and to the second one as to the MultiLAG model.

2.1 Notations

In the Compositional Data analysis (CoDa) literature, a D-composition $\mathbf{u}$ is a vector with positive components conveying relative information which can be represented by an element of the so-called simplex space $\mathbf{S}^D$ defined by [Aitchison, 1986]:

$$\mathbf{S}^D = \left\{ \mathbf{u} = (u_1, \ldots, u_D)^T : u_m > 0, m = 1, ..., D; \sum_{m=1}^{D} u_m = 1 \right\},$$

where ^{T} is the transpose operator. $\mathbf{S}^D$ can be equipped with a vector structure and the compositional/Aitchison inner product.

The analysis of compositional data makes use of some classical log-ratio transformations which are the additive log-ratio (alr), the centered log-ratio (clr) and the isometric log-ratio (ilr) transformations. In this paper, we will use alr and ilr transformations.
For the reference level D, let us recall that the alr transformation $\mathbf{u}^*$ of a vector $\mathbf{u}$ in $\mathbf{S}^D$ is defined by:

$$\mathbf{u}^* = \mathrm{alr}_D(\mathbf{u}) = (\ln(u_1/u_D), \ldots, \ln(u_{D-1}/u_D))$$

and its inverse transformation alr_D^{-1} is given by:

$$\mathbf{u} = \mathrm{alr}_D^{-1}(\mathbf{u}^*) = C(\exp(u_1^*), \ldots, \exp(u_{D-1}^*), 1)$$

where $C(.)$ denotes the closure operation defined by

$$C(\mathbf{w}) = \left(\frac{w_1}{\sum_{m=1}^{D} w_m}, \cdots, \frac{w_D}{\sum_{m=1}^{D} w_m} \right),$$

for any vector $\mathbf{w} \in \mathbb{R}_+^D$. A similar definition can be given changing the reference level D into any other level $m = 1, \ldots, D-1$.

A given additive log-ratio transformation for reference level $m = 1, \ldots, D$ can also be expressed using a $(D-1) \times D$ matrix $\mathbf{F}_m$ by $\mathrm{alr}_m(\mathbf{u}) = \mathbf{F}_m \ln(\mathbf{u})$ for a vector $\mathbf{u} \in \mathbf{S}^D$. For example, in the case $m = D$, $\mathbf{F}_D = [\mathbf{I}_{D-1} \; -\mathbf{j}_{D-1}]$ where $\mathbf{j}_{D-1}$ denotes the $(D-1)$-dimensional column vector of ones and $\mathbf{I}_{D-1}$ is the $(D-1) \times (D-1)$ identity matrix.

Similarly, in order to define an ilr transformation, let $\mathbf{V}$ be a $D \times (D-1)$ contrast matrix associated to a given orthonormal basis $(\mathbf{e}_1, \cdots, \mathbf{e}_{D-1})$ of $\mathbf{S}^D$ by $\mathbf{V} = \mathrm{clr}(\mathbf{e}_1, \cdots, \mathbf{e}_{D-1})$, where clr is understood columnwise and where $g(\mathbf{u}) = (\prod_{i=1}^D u_i)^{1/D}$ is the geometric mean of the components of $\mathbf{u} \in \mathbf{S}^D$:

$$\mathrm{clr}(\mathbf{u}) = \left(\ln \frac{u_m}{g(\mathbf{u})} \right)_{m=1,\ldots,D}.$$

For each such matrix $\mathbf{V}$, an isometric log-ratio transformation (ilr) is then defined by:

$$\mathbf{u}^* = \mathrm{ilr}_V(\mathbf{u}) = \mathbf{V}^T \ln(\mathbf{u})$$

where the logarithm of $\mathbf{u} \in \mathbf{S}^D$ is understood componentwise. Note that we use the same star notation throughout the paper to indicate either an ilr or an alr transformed vector. It is to emphasize the similarity of treatment and the meaning will be clear from the context. The inverse transformation is given by:

$$\mathbf{u} = \mathrm{ilr}_V^{-1}(\mathbf{u}^*) = C(\exp(\mathbf{V}\mathbf{u}^*)).$$

Moreover, the compositional product of a matrix by a composition vector is denoted by $\boxdot$. It is defined for a $D \times D$ matrix $\mathbf{B} = (b_{lm})$, $l, m = 1, \ldots D$ such that $\mathbf{B}\mathbf{j}_D = \mathbf{0}_D$ and $\mathbf{B}^T\mathbf{j}_D = \mathbf{0}_D$ (with $\mathbf{0}_D$ a D-dimensional vector of zeros) and for a vector $\mathbf{u} \in \mathbf{S}^D$ by:

$$\mathbf{B} \boxdot \mathbf{u} = C\left(\prod_{m=1}^D u_m^{b_{1m}}, \cdots, \prod_{m=1}^D u_m^{b_{Dm}} \right)^T.$$

As detailed in [Pawlowsky-Glahn et al., 2015], the previous expression comes from the usual product matrix definition in the coordinate space for an ilr_V transformation: $\mathbf{B} \boxdot \mathbf{u} = \mathrm{ilr}_V^{-1}(\mathbf{B}^* \mathrm{ilr}_V(\mathbf{u}))$ where $\mathbf{B}^* = \mathbf{V}^T\mathbf{B}\mathbf{V}$. We also have

$$\mathbf{B} \boxdot \mathbf{u} = \mathrm{ilr}_V^{-1}\left(\mathbf{V}^T\mathbf{B}\ln(\mathbf{u})\right).$$

Note that this matrix product in the simplex does not depend on the particular choice of contrast matrix $\mathbf{V}$. This product also verifies the following property (see the proof in Appendix 6.1):

$$\mathbf{B} \boxdot \mathbf{u} = \mathrm{alr}_m^{-1}(\mathbf{F}_m\mathbf{B}\ln(\mathbf{u})).$$

In what follows, the data we consider is made of samples of composition vectors and is stored in a $n \times D$ matrix $\mathbf{Y} = (\mathbf{Y}_{il})$, $i = 1, \ldots, n,\ l = 1, \ldots, D$. We propose to generalize the matrix product to a $n \times D$ matrix $\mathbf{Y}$ of composition vectors multiplied by a $D \times D$ matrix $\mathbf{B}$ (such that $\mathbf{B}\mathbf{j}_D = \mathbf{0}_D$ and $\mathbf{B}^T\mathbf{j}_D = \mathbf{0}_D$) by applying the usual matrix product to each row of $\mathbf{Y}$. We have

$$\mathbf{Y} \boxdot \mathbf{B} = \mathrm{ilr}_V^{-1}\left(\ln(\mathbf{Y})\mathbf{B}\mathbf{V}\right) = \mathrm{alr}_m^{-1}\left(\ln(\mathbf{Y})\mathbf{B}\mathbf{F}_m^T\right). \tag{1}$$

For a $n \times D$ matrix of compositions $\mathbf{Y}$, and a $n \times n$ weight matrix $\mathbf{W}$, [Nguyen et al., 2019] define the following operation $\triangle$ as a map from the cartesian product of simplex spaces $(\mathbf{S}^D)^n$ to itself specified by

$$\mathbf{W}\triangle\mathbf{Y} = \mathrm{ilr}_V^{-1}(\mathbf{W}\mathrm{ilr}_V(\mathbf{Y})) = \mathrm{ilr}_V^{-1}(\mathbf{W}\ln(\mathbf{Y})\mathbf{V}). \tag{2}$$

The operation $\mathbf{W}\triangle\mathbf{Y}$ defines a $n \times D$ matrix of compositions whose components are weighted geometric means of the $\mathbf{Y}$ values weighted by the $\mathbf{W}$ weights.

2.2 The MultiLAG model

In [Nguyen et al., 2019], the authors propose to define the MultiLAG model for a simplex-valued dependent variable. In the present paper, we consider a simpler version of this model which involves in each ilr coordinate equation spatial lags of the other coordinates but not the other coordinates themselves.

Let us consider a sample of size n. Let $\mathbf{Y}$ be a $n \times D$ matrix of dependent compositional vectors and $\mathbf{X}$ be a $n \times K$ matrix of classical (non-compositional) explanatory variables. An extension to models including compositional explanatory variables would be straightforward with some adjustments for the interpretation section. Let $\mathbf{W}$ be a $n \times n$ spatial weight matrix specifying the neighborhood structure. For n spatial locations, the elements w_{ij} of the matrix $\mathbf{W}$ are measures of proximity between locations i and j. The model can be written in the simplex as follows

$$\mathbf{Y} = (\mathbf{W}\triangle\mathbf{Y}) \boxdot \mathbf{R} \oplus \mathbf{X} \odot \boldsymbol{\beta} \oplus \boldsymbol{\epsilon}, \tag{3}$$

where $\mathbf{R}$ is a $D \times D$ matrix of parameters such that $\mathbf{R}\mathbf{j}_D = \mathbf{0}_D$ and $\mathbf{R}^T\mathbf{j}_D = \mathbf{0}_D$, $\boldsymbol{\beta}$ is a $K \times D$ matrix of parameters. $\boldsymbol{\epsilon}$ is a $n \times D$ matrix of compositional errors satisfying the following conditions. Denoting by $\boldsymbol{\epsilon}_{.l}$ the columns of $\boldsymbol{\epsilon}$ and by $\boldsymbol{\epsilon}_{i.}$ its rows, we assume that $\mathbb{E}(\boldsymbol{\epsilon}^*_{i.}{\boldsymbol{\epsilon}^*_{j.}}^T) = \boldsymbol{\Sigma}^*$ if individuals i and j are equal and $\mathbf{0}$ if they are different, where $\boldsymbol{\Sigma}^*$ is a $(D-1) \times (D-1)$ covariance matrix. Note that the sum in the simplex sense ($\oplus$) over individuals of the rows of the matrix $\boldsymbol{\epsilon}$ is the neutral element of the simplex.

Following [Nguyen et al., 2019], the model can be written in ilr coordinate space with a coordinate-wise formulation:

$$\mathbf{Y}^*_{.l} = \sum_{m \in S_l^{\mathbf{WY}^*}} \mathbf{R}^*_{ml} \mathbf{WY}^*_{.m} + \mathbf{X}_{S_l^{\mathbf{X}}} \boldsymbol{\beta}^*_{S_l^{\mathbf{X}}} + \boldsymbol{\epsilon}^*_{.l},\ l = 1, \ldots, D-1. \tag{4}$$

where $\mathbf{R}^*$ and $\boldsymbol{\beta}^*$ are the parameters of the model in coordinate space and where $\mathbf{Y}^*_{.l}$ and $\boldsymbol{\epsilon}^*_{.l}$ denote respectively $\text{ilr}_V(\mathbf{Y}_{.l})$ and $\text{ilr}_V(\boldsymbol{\epsilon}_{.l})$ for any given contrast matrix $\mathbf{V}$. We allow for using a different set of explanatory variables in each equation. For this reason, $S_l^{\mathbf{Y}^*},\ S_l^{\mathbf{X}},\ S_l^{\mathbf{WY}^*}$ denote the sets of indices of the variables which appear in the l^{th} equation for $\mathbf{Y}^*,\ \mathbf{X},\ \mathbf{WY}^*$ respectively.

It is important to note that $\mathbf{R}^*$ and $\boldsymbol{\beta}^*$ depend upon the specific choice of $\mathbf{V}$ despite the fact that we do not indicate it in the notation for simplification purposes. The relation between $\mathbf{R}^*$ and $\mathbf{R}$, and between $\boldsymbol{\beta}^*$ and $\boldsymbol{\beta}$ will be explored in Section 3. This model can be easily estimated via the S2SLS method in a multivariate fashion. Two stage least squares also called instrumental variables regression is popular in econometrics when an explanatory variable of interest is correlated with the error term and it has been adapted to the spatial models by [Kelejian and Prucha, 1998]. Ignoring the dependence between coordinates and performing a univariate estimation coordinate by coordinate would result in the same parameters estimates but incorrect standard errors estimates.

In view of Section 4, we need another formulation of the MultiLAG spatial model in ilr coordinate space with a matrix formulation as in [Kelejian and Prucha, 2004] since [Nguyen et al., 2019] only give a coordinatewise formulation. Let $\mathbf{Y}^*$ be the $n \times (D-1)$ matrix whose columns are the ilr coordinates $\mathbf{Y}^*_{.l}$. The matrix formulation of model (4) is given by

$$\mathbf{Y}^* = \mathbf{WY}^*\mathbf{R}^* + \left[\mathbf{X}_{S_1^{\mathbf{X}}} \boldsymbol{\beta}^*_{S_1^{\mathbf{X}}} \cdots \mathbf{X}_{S_{D-1}^{\mathbf{X}}} \boldsymbol{\beta}^*_{S_{D-1}^{\mathbf{X}}}\right] + \boldsymbol{\epsilon}^*, \tag{5}$$

where $[A_1 \ldots A_k]$ is a notation for a block matrix with blocks A_1 through A_k.
In order to be able to compute fitted values and marginal effects, we also need a formulation of the so-called reduced form of the model: it is an equivalent formulation of the model where the dependent variable appears only on the left hand side of the equation. It is possible to do so using a vectorized formulation. For a matrix $\mathbf{A}$, let $\text{vec}_c\mathbf{A}$ be the column vectorization obtained by stacking the columns of $\mathbf{A}$.

Theorem 0.1 *The reduced form of model (4) expressed in vectorized formulation is given by*

$$vec_c\mathbf{Y}^* = \left(\mathbf{I}_{n(D-1)} - ((\mathbf{R}^*)^T \otimes \mathbf{W})\right)^{-1} \left[\mathbf{X}_{S_1^{\mathbf{X}}} \boldsymbol{\beta}^*_{S_1^{\mathbf{X}}} \cdots \mathbf{X}_{S_{D-1}^{\mathbf{X}}} \boldsymbol{\beta}^*_{S_{D-1}^{\mathbf{X}}}\right]^T + vec_c\boldsymbol{\epsilon}^* \tag{6}$$

where $\otimes$ denotes the Kronecker product of matrices.

The matrix $\mathbf{A}(\mathbf{W}) = (\mathbf{I}_{n(D-1)} - ((\mathbf{R}^*)^T \otimes \mathbf{W}))^{-1}$ is called the filter matrix. The proofs of equation (5) and Theorem 0.1 are in the appendix.

2.3 The UniSEM model

Even though [Chakir and Lungarska, 2017] do not use the compositional data analysis formalism, their model can be formulated in a similar fashion as that of [Nguyen et al., 2019]. Indeed if we use the same star notation for denoting now an alr transformation instead of an ilr one, it is easy to see that the specification (3) in [Chakir and Lungarska, 2017] can be reformulated as

$$\begin{cases} \mathbf{Y}^*_{.l} & = \mathbf{X}_{S_l^{\mathbf{X}}} \boldsymbol{\beta}^*_{S_l^{\mathbf{X}}} + \boldsymbol{\epsilon}^*_{.l} \\ \boldsymbol{\epsilon}^*_{.l} & = \sum_{m \in S_l^{\mathbf{W}\mathbf{Y}^*}} \mathbf{R}^*_{ml} \mathbf{W} \boldsymbol{\epsilon}^*_{.m} + \boldsymbol{\nu}^*_{.l} \end{cases} \tag{7}$$

with the restriction that the $\mathbf{R}^*$ matrix is diagonal and that the alr coordinates are independent (i.e. the errors $\boldsymbol{\nu}^*_{.l}$ are independent). Therefore the reduced form of the model can be derived as in Theorem 0.1. It is also easy to see that we can reformulate this model in the simplex as

$$\begin{cases} \mathbf{Y} & = \mathbf{X} \odot \boldsymbol{\beta} \oplus \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} & = (\mathbf{W} \triangle \boldsymbol{\epsilon}) \boxdot \mathbf{R} \oplus \boldsymbol{\nu} \end{cases}. \tag{8}$$

The motivation for returning to the simplex is twofold. In this work, we need to compare models which are using different alr and ilr transformations and correspond to different coordinate spaces: only the simplex parameters are comparable. Moreover [Morais and Thomas-Agnan, 2020] advocate the interpretation of parameters in the simplex as being more informative. Model 8 will be called the MultiSEM model when no restriction is made on the $\mathbf{R}^*$ matrix and UniSEM with the diagonality restriction as in [Chakir and Lungarska, 2017].

2.4 Summary of models

We just recalled the definition of the MultiLAG and the UniSEM models which are solutions for defining a compositional dependent variable regression model with spatial dependence. The MultiLAG takes into account the possible dependence between coordinates. The submodel of MultiLAG obtained by constraining the coordinates to be independent as in the UniSEM will be called the UniLAG. For the same type of data, an ordinary regression model for compositional dependent variable without spatial dependence could also be considered and we will later refer to it as the OLM model for ordinary linear model (see [Filzmoser et al., 2018]). Fitting the UniLAG or the UniSEM just requires code for spatial univariate regression models as in the R package *spdep* [Bivand and Wong, 2018]. Fitting the MultiLAG with the S2SLS method just requires code for multivariate regression such as the R function *lm* but we keep the implementation of MultiSEM for future work because of the lack of specific code for this case.

3 Model specification and interpretation

3.1 Influence of the transformation

The impact of the coordinate system on the model is crucial for compositional data: indeed a model formulation involving some restrictions on the parameters which are dependent upon a particular choice of coordinate system is inappropriate since it means that the model does not have an intrinsic simplex formulation. [Chakir and Lungarska, 2017] specify the model in coordinate space after choosing a particular alr transformation. We have shown in Section 2 that this model can also be formulated in the simplex. [Nguyen et al., 2019] first specify the model in coordinate space with a particular ilr transformation, but then show that the model can also be specified in the simplex. If the formulation of a model is done in the simplex, it is independent of any choice of transformation. A natural question is then to establish the relationships between parameters in the simplex and parameters in coordinate space.

[Nguyen et al., 2019] show that the relationship between the simplex parameters of (3) $\mathbf{R}$ and $\boldsymbol{\beta}$ and the coordinate parameters of (4) $\mathbf{R}^*$ and $\boldsymbol{\beta}^*$ is then

$$\begin{cases} \mathbf{R} = \mathbf{V}\mathbf{R}^*\mathbf{V}^T \\ \boldsymbol{\beta} = \mathrm{ilr}_V{}^{-1}(\boldsymbol{\beta}^*) \end{cases} \tag{9}$$

which can also be written as $\mathbf{V}^T\mathbf{R} = \mathbf{R}^*\mathbf{V}^T$.

Consider now the UniSEM model specified by (8). Let us derive the correspondence between the parameters in alr space and in the simplex. The following theorem establishes the equivalent of (9) for alr. Our statements and proofs below concern the alr_D transformations but are easily generalized to any alr.

Theorem 0.2 *If* $\mathbf{R}$ *and* $\boldsymbol{\beta}$ *are the simplex parameters of model (8) and if* $alr_D(\mathbf{R}) = \mathbf{R}^*$ *and* $alr_D(\boldsymbol{\beta}) = \boldsymbol{\beta}^*$*, then*

$$\begin{cases} \mathbf{R} = \mathbf{K}_D\mathbf{R}^*\mathbf{F}_D \\ \boldsymbol{\beta} = alr_D{}^{-1}(\boldsymbol{\beta}^*) \end{cases} \tag{10}$$

where $\mathbf{F}_D$ *is the matrix associated to the alr transformation* alr_D *and* $\mathbf{K}_D$ *is defined by*

$$\mathbf{K}_D = \begin{bmatrix} \mathbf{I}_{D-1} - \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T/D \\ -\mathbf{j}_{D-1}^T/D \end{bmatrix}.$$

Proof. Let us start with model (3) and apply the alr_D transformation to both sides of the equation. Using equation (1), it is easy to see that

$$\mathrm{alr}_D(\mathbf{R} \boxdot \mathbf{W}\triangle\mathbf{Y} \oplus \mathbf{X} \odot \boldsymbol{\beta} \oplus \boldsymbol{\epsilon}) = \mathbf{F}_D\mathbf{R}\ln(\mathbf{W}\triangle\mathbf{Y}) + \mathbf{X}\mathrm{alr}_D(\boldsymbol{\beta}) + \mathrm{alr}_D(\boldsymbol{\epsilon}) \tag{11}$$

Comparing to (11) the right hand side of (4) where star would mean the alr_D transformation, we get that $\mathrm{alr}_D(\boldsymbol{\beta})$ corresponds to $\boldsymbol{\beta}^*$(which proves the second part of

(10)) and that $\mathbf{F}_D\mathbf{R} = \mathbf{R}^*\mathbf{F}_D$. Then lemma 0.1 in Appendix 6.3 allows to show that $\mathbf{F}_D\mathbf{R} = \mathbf{R}^*\mathbf{F}_D$ is equivalent to the first equation of (10).

Note that these relationships between simplex-space parameters and coordinate-space parameters induce the same relationships between their estimated counterparts for the MultiLAG model when estimated by S2SLS (see [Nguyen et al., 2019]). It also holds for the UniSEM estimated by maximum likelihood since maximum likelihood is preserved by a one to one transformation. Finally, we see that if we specify the model in the simplex, the choice of a particular transformation for estimation does not change the final result in the simplex. However a restriction like $\mathbf{R}^*$ is diagonal, used in [Chakir and Lungarska, 2017], in a given coordinate space (ilr or alr) does not imply that the same is true in another coordinate space. In this sense one can say that it is not a simplex assumption.

Moreover another issue with the choice of coordinate system is the relationships between the mean and variance parameters associated to each system. For the case of two ilr transformations, the result can be found in [Pawlowsky-Glahn et al., 2015]. Let us derive the same type of relationship for the case of two alr transformations and for the case of an alr and an ilr transformation.

Theorem 0.3 *If we assume that* $alr_D(\mathbf{X}) \sim \mathcal{N}_{\mathcal{S}_D}(\boldsymbol{\mu}_{alr_D}, \boldsymbol{\Sigma}_{alr_D})$, *and if* $\mathbf{V}$ *is a contrast matrix, then the means and variances of* $ilr_V(\mathbf{X})$ *and* $alr_j(\mathbf{X})$ *are related as follows to those of* $alr_D(\mathbf{X})$

$$\boldsymbol{\mu}_{ilr_\mathbf{V}} = \mathbf{V}^T\mathbf{K}_D\boldsymbol{\mu}_{alr_D} \tag{12}$$
$$\boldsymbol{\Sigma}_{ilr_\mathbf{V}} = \mathbf{V}^T\mathbf{K}_D\boldsymbol{\Sigma}_{alr_D}(\mathbf{V}^T\mathbf{K}_D)^T \tag{13}$$
$$\boldsymbol{\mu}_{alr_j} = \mathbf{F}_j\mathbf{K}_D\boldsymbol{\mu}_{alr_D} \tag{14}$$
$$\boldsymbol{\Sigma}_{alr_j} = \mathbf{F}_j\mathbf{K}_D\boldsymbol{\Sigma}_{alr_D}(\mathbf{F}_j\mathbf{K}_D)^T \tag{15}$$

Proof. We recall that $\mathrm{alr}_j(\mathbf{X}) = \mathbf{F}_j \ln(\mathbf{X})$ for $j = 1, \ldots, D$ and $\mathrm{ilr}_V(\mathbf{X}) = \mathbf{V}^T \ln(\mathbf{X})$. Let us prove that

$$\mathrm{ilr}_V(\mathbf{X}) = \mathbf{V}^T\mathbf{K}_D\mathrm{alr}_D(\mathbf{X}) \tag{16}$$
$$\mathrm{alr}_j(X) = \mathbf{F}_j\mathbf{K}_D\mathrm{alr}_D(\mathbf{X}) \tag{17}$$

It is easy to prove that $\mathbf{F}_D\mathbf{K}_D = \mathbf{V}^T\mathbf{V} = \mathbf{I}_{D-1}$ and $\mathbf{K}_D\mathbf{F}_D = \mathbf{K}_j\mathbf{F}_j = \mathbf{V}\mathbf{V}^T$. Thus,

$$\begin{aligned}\mathrm{ilr}_V(\mathbf{X}) &= \mathbf{V}^T\mathbf{V}\mathrm{ilr}_V(\mathbf{X}) \\ &= \mathbf{V}^T\mathbf{V}\mathbf{V}^T \ln(\mathbf{X}) \\ &= \mathbf{V}^T\mathbf{K}_D\mathbf{F}_D \ln(\mathbf{X}) \\ &= \mathbf{V}^T\mathbf{K}_D\mathrm{alr}_D(\mathbf{X})\end{aligned}$$

and

$$\begin{aligned}\mathrm{alr}_j(\mathbf{X}) &= \mathbf{F}_j\mathbf{K}_j\mathrm{alr}_j(\mathbf{X}) \\ &= \mathbf{F}_j\mathbf{K}_j\mathbf{F}_j\ln(\mathbf{X}) \\ &= \mathbf{F}_j\mathbf{K}_D\mathbf{F}_D\ln(\mathbf{X}) \\ &= \mathbf{F}_j\mathbf{K}_D\mathrm{alr}_D(\mathbf{X})\end{aligned}$$

Equations (12) and (13) (resp. (14) and (15)) are then directly derived from (16) (resp. (17)).

This theorem implies that if Σ_{alr_D} is diagonal, then Σ_{ilr_V} is not necessarily diagonal and Σ_{alr_j} neither implying that this restriction is not simplex compatible.

3.2 Model interpretation

Concerning the impact of explanatory variables, we first focus on looking at how the fitted shares in the simplex depend upon the values of a particular explanatory variable. For the UniSEM model, the question is quite simple: the fitted values in coordinate space are straightforward to compute and their inverse alr transformation yield the fitted shares.
For the MultiLAG model, thanks to Theorem 0.1, it is first easy to derive the fitted values of $\mathbf{Y}^*$ in coordinate space by substituting the parameters by their estimates and we get

$$\mathrm{vec}_c\,\hat{\mathbf{Y}}^* = \left(\mathbf{I}_{n(D-1)} - ((\hat{\mathbf{R}}^*)^T \otimes \mathbf{W})\right)^{-1} \begin{bmatrix} \mathbf{X}_{S_1}\hat{\boldsymbol{\beta}}_1^* \\ . \\ \mathbf{X}_{S_{D-1}}\hat{\boldsymbol{\beta}}_{D-1}^* \end{bmatrix} \tag{18}$$

Applying the inverse ilr transformation to (18), we can compute the fitted shares in the simplex and use them to illustrate the impact of the covariates by a graph as follows. Since this investigation is relative to a single variable at a time, we can drop the variable index in this paragraph to simplify the derivation. With this notation, for a given covariate $\mathbf{X}$, x_j denotes the value of $\mathbf{X}$ at location j.

For a given spatial unit and a given covariate, we imagine changing the value of that particular covariate at that particular location holding everything else constant. We denote by $\mathbf{Y}_{im}(x_j + \delta)$ the value of the component m of the share vector $\mathbf{Y}$ for location i obtained by increasing by δ the value of $\mathbf{X}$ at location j (note that we only indicate the location that changes). We then compute the new fitted shares $\hat{\mathbf{Y}}_{im}(x_j + \delta)$ for a grid of points $x_j + \delta$ in the range of the covariate of interest. Finally we draw a scatterplot of the fitted shares as a function of the grid points. Some Figures in Section 4 illustrate this approach.

A further step in this investigation implies computing the semi-elasticities in the simplex as in [Morais and Thomas-Agnan, 2020]. This comprises two issues. The first issue is the computation of spatial impacts in coordinate space and the second one the evaluation of semi-elasticities in the simplex.

For the first issue, indeed parameters in a simultaneous spatial autoregressive model cannot be interpreted as in ordinary linear models due to the non-linear filter matrix in the expected value of the dependent variable (see [LeSage and Pace, 2009]). We obtain the impacts in coordinate space thanks to the reduced form of the model (6). For a given explanatory variable $\mathbf{X}$ and a given component m, the impact on $\mathbf{Y}^*_{im}$ of changing $\mathbf{X}$ at location j is defined by $\frac{\partial \mathbb{E}\mathbf{Y}^*_{im}}{\partial x_j}$. If $a_{im:jm}(\mathbf{W})$ denote the generic element of $A(\mathbf{W})$, for location i, j, component m, we have

$$\frac{\partial \mathbb{E}\mathbf{Y}^*_{im}}{\partial x_j} = \boldsymbol{\beta}^*_m a_{im:jm}. \tag{19}$$

If a variable $\mathbf{X}$ does not appear in the equation of coordinate m, then set the corresponding parameter to zero (this will not happen in our application). The estimated impacts are then obtained by substituting the parameters by their estimates.
For the second issue, the formulas in [Morais and Thomas-Agnan, 2020] cannot be applied directly due to the fact that the UniSEM and MultiLAG models involve this non-linear filter. Extensions of [Morais and Thomas-Agnan, 2020] adapted to the presence of a spatial filter can be found in [Thomas-Agnan et al., 2020]. In the present work, we propose an approximation of these semi-elasticities as follows. [Morais and Thomas-Agnan, 2020] show that the effect of increasing a classical explanatory variable $\mathbf{X}$ by an additive amount δ on unit j has a multiplicative impact on the components of the dependent vector $(\mathbf{Y}_{i1} \dots \mathbf{Y}_{iD})$ (for statistical unit i):

$$\mathbb{E}\mathbf{Y}_{im}(x_j + \delta) \sim \mathbb{E}\mathbf{Y}_{im}(x_j) \left(1 + \mathrm{se}_{ij;m}\right), \tag{20}$$

where

$$\mathrm{se}_{ij;m} = \frac{\partial \ln \mathbb{E}\mathbf{Y}_{im}}{\partial x_j}(x_j)$$

is classically called semi-elasticity of the m^{th} component $\mathbf{Y}_{im}$ ($m = 1, \dots, D$ and $i, j = 1, \dots, n$) with respect to $\mathbf{X}$. [Morais and Thomas-Agnan, 2020] show that in the non spatial model these semi-elasticities depend on the share vector at the point of interest but not on the location at which we increase the explanatory variable. Formula (20) yields a way to estimate these approximate semi-elasticities using a finite difference approach for a small value of δ and using fitted shares $\hat{\mathbf{Y}}_{im}(x_j + \delta)$. For each data point we obtain a full structure of size $n^2(D-1)$ of semi-elasticities. To summarize the spatial impacts, [LeSage and Pace, 2009] propose summary measures called the direct, indirect and total impacts. Here, one first needs to re-arrange the semi-elasticities vectors thus computed at the individual level into D matrices of size $n \times n$ (one matrix of semi-elasticities per component). For each component m, the average of the diagonal terms will represent the average direct impact (acronym ADI) $\mathrm{ADI}_m = \sum_{i=1}^n \mathrm{se}_{ii;m}/n$. The average indirect impact (acronym AII) should be the average of the sum of the extra-diagonal terms. In order to alleviate its evaluation, we propose to restrict the sum to neighboring points as follows:

$$\mathrm{AII}_m = \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}\mathrm{se}_{ij;m} \sim \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i,\, j\in N(i)}\mathrm{se}_{ij;m}, \tag{21}$$

where $N(i)$ is the set of first order neighbors of a given location i. Because this averaging will hide some variability, we propose to introduce the local contributions to the direct impact and indirect impact respectively as $\mathrm{se}_{ii;m}$ and $\sum_{j\neq i,\, j\in N(i)}\mathrm{se}_{ij;m}$ (see Section 4 for their illustration).

4 Data and results

4.1 Land use data

The dataset and descriptive statistics of the variables used in this study are presented in Table 1. More detail can be found in the supplementary material available online (see `http://www.thibault.laurent.free.fr/code/land_use_spat_coda/`). Land use data are collected from the Corine Land Cover (CLC) databases for France at the scale of 100m × 100m (1 ha) for the year 2000 (except for the Paris area). Land uses are grouped into four categories: "agriculture", "forest", "urban" and "other" where "agriculture" is aggregated from crops and pastures. We calculate the land use shares for each grid cell of 8 km × 8 km. Land use shares are expressed as the sum of the land use classes in hectares divided by the surface of the grid cell. We using the simple zero replacement strategy suggested by [Aitchison, 1986] and fully described in [Martín-Fernández et al., 2003]. We choose to use a weight matrix **W** based on first order queen contiguity: i.e. two polygons are neighbors if they share a common side or a common vertex.

According to the economic literature, standard drivers of land use allocation, consistent with the von Thünen and Ricardian conceptual models, are the measures of the net return to each use and biophysical suitability. However, net returns are not easily observable for all land uses and one solution is to use proxies. For example, due to data limitations, a measure of the average net return to urban land is difficult to construct. An alternative is to use a variable that proxies for these returns. One measure frequently used in land use share models is population density, which indicates the pressures for urbanization. [Chakir and Lungarska, 2017] compared three possible proxies of agricultural land rent: agricultural land price, farm income and the land shadow price. These comparisons are made based on various criteria including: consistency with the theoretical hypothesis, prediction quality and specification tests. They conclude that shadow price proxy provides stable and intuitive results. In this paper, we consider the same explanatory variables as in [Lungarska and Chakir, 2018] but we replace the spatial Durbin error model by the spatial error model because the lagged explanatory variables of the spatial Durbin error model resulted in collinearity issues with the S2SLS method.

Following [Pawlowsky-Glahn and Egozcue, 2011], to guide our choice of ilr transformation (contrast matrix), we look at the biplot associated with the first two principal components (see the supplementary material available online).

The first two principal axes explain 94% of the variability in the data. On the first one, the land use "other" is opposed to the rest of the land uses while on the second one, the land use "forest" is opposed to "urban". We thus propose the following contrast matrix: the first ilr coordinate opposes the land use "other" to the geometric mean of the other land uses while the second opposes "agricultural" to the geometric mean of "forest" and "urban" and the third "forest" to the "urban" land use.

	Description	mean	St.dev.	min	q.25%	q.50%	q.75%	max
Land use shares								
agriculture	Share of crops and pastures	0.60	0.29	0	0.39	0.67	0.85	1.00
forest	Share of forest	0.26	0.22	0	0.07	0.20	0.41	1.00
other	Share of other uses	0.09	0.17	0	0.00	0.01	0.07	1.00
urban	Share of urban	0.05	0.09	0	0.01	0.02	0.05	0.99
Explanatory variables								
Shadow Price	Agricultural land shadow price (AROPAj, spatialized) (k€/ha)	0.55	0.22	0.00	0.40	0.48	0.68	1.11
Forest revenue	Forest revenues (FFSM++, regional) (€/ha)	137.68	66.51	28.93	85.72	127.75	187.99	308.04
Pop. Income	Population revenues (INSEE, commune) (k€/year/household)	12.31	3.24	0.00	10.38	11.97	13.98	41.80
Pop. density	Population density (INSEE, "carottage") (household/ha)	5.43	2.27	2.75	4.48	4.90	5.58	58.72
slope	Terrain average slope (GTOPO30) (%)	4.33	6.15	0.00	1.14	2.10	4.23	47.72
Texture	Soil's texture classes				1	2	3	4
	Number of cells				1242	4820	3120	579

Table 1: Descriptive statistics for land use shares and for the explanatory variables.

4.2 Model results

In this section, we fit the UniSEM model and the MultiLAG model, but also consider two other models based on the same variables introduced in Section 2.4: the ordinary linear model (OLM) and the UniLAG model which can both be fit separately for each equation by maximum likelihood.

Table 2 shows the results of the OLM and the MultiLAG regression models, the most simple and the most complex of our models. Results for the other models can be found in the supplementary material. Most explanatory variables are significant for the three ilr coordinates. The Moran test (see for instance [Cliff and Ord, 1981]) on the OLM residuals is highly significant motivating the need for a spatial model. Most of the spatial autocorrelation parameters of $\mathbf{R}^*$ are significant too. However it is not possible to directly compare the parameters of the OLM with those of the

MultiLAG due to the spatial filter (see [LeSage and Pace, 2009]) and the parameters of the MultiLAG with those of the UniSEM due to the fact that they use different coordinate spaces. Hence the need to go back to the simplex.

Table 2: Results of the OLM vs. MultiLAG

	OLM in the ILR space			MultiLAG in the ILR space		
	ilr_1	ilr_2	ilr_3	ilr_1	ilr_2	ilr_3
Intercept	-0.369*	3.582***	-6.353***	1.498***	2.499***	-4.657***
	(0.168)	(0.089)	(0.13)	(0.318)	(0.18)	(0.271)
shadow price	0.137	0.515***	0.755***	-0.117	0.09	0.214*
	(0.138)	(0.073)	(0.107)	(0.121)	(0.069)	(0.103)
forest revenues	-0.008***	0	0.001***	-0.001**	0	0.001***
	(0)	(0)	(0)	(0)	(0)	(0)
population density	0.121***	-0.138***	0.279***	0.018	-0.119***	0.223***
	(0.013)	(0.007)	(0.01)	(0.013)	(0.007)	(0.011)
population income	-0.137***	-0.117***	0.191***	-0.096***	-0.074***	0.158***
	(0.009)	(0.005)	(0.007)	(0.013)	(0.007)	(0.011)
slope	0.203***	-0.097***	-0.124***	0.072***	-0.029***	-0.081***
	(0.005)	(0.003)	(0.004)	(0.008)	(0.005)	(0.007)
texture 2	-0.737***	0.815***	0.743***	-0.442***	0.242***	0.344***
	(0.086)	(0.045)	(0.066)	(0.085)	(0.048)	(0.073)
texture 3	-1.562***	0.944***	0.691***	-0.71***	0.347***	0.369***
	(0.092)	(0.049)	(0.071)	(0.092)	(0.052)	(0.078)
texture 4	-1.81***	1.324***	0.405***	-0.846***	0.666***	0.122
	(0.136)	(0.072)	(0.105)	(0.123)	(0.07)	(0.105)
$\mathbf{R}^*_{1.}$	-	-	-	0.784***	0.052*	0.14***
				(0.037)	(0.021)	(0.032)
$\mathbf{R}^*_{2.}$	-	-	-	-0.031	0.677***	0.275***
				(0.066)	(0.037)	(0.056)
$\mathbf{R}^*_{3.}$	-	-	-	0.229***	0.178***	0.44***
				(0.044)	(0.025)	(0.038)
$\mathbf{\Sigma}^*_{1.}$	7.216	-0.233	-0.544	4.461	0.158	-0.668
$\mathbf{\Sigma}^*_{2.}$	-0.233	2.002	-0.075	0.158	1.434	-0.481
$\mathbf{\Sigma}^*_{3.}$	-0.544	-0.075	4.312	-0.668	-0.481	3.238
Nb. Obs.	9760	9760	9760	9760	9760	9760
Moran's I test	76.02***	64.29***	70.14***	-	-	-
LMlag	5834***	4008***	4448***	-	-	-

Note: *p<0.1; **p<0.05; ***p<0.01

Following [Glahn et al., 2012], we propose to compare the fit of these models using a mean square error measure built with the Aitchison geometry (norm denoted by $\| \, . \, \|_A$):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \| \, \hat{\mathbf{Y}}_i \ominus \mathbf{Y}_i \, \|_A^2 = \frac{1}{n} \sum_{i=1}^{n} \| \, \text{ilr}_V \left(\hat{\mathbf{Y}}_i \right) - \text{ilr}_V \left(\mathbf{Y}_i \right) \, \|^2, \tag{22}$$

for any contrast matrix $\mathbf{V}$. Note that since the goal is to evaluate a prediction ability of the model, instead of using a fitted value understood as an estimate of $\mathbb{E}(\mathbf{Y}_i)$, we rather use a prediction of $\mathbf{Y}_i$. There are several ways to compute predicted values in spatial models, therefore we indicate which formula we use among best prediction (BP), trend signal (TS) and trend corrected (TC) (see [Goulard et al., 2017] for detail). We use univariate predictors due to the lack of code to evaluate multivariate versions. For the MultiLAG model, because the ilr is isometric with respect to the Aitchison geometry, it is enough to compute the MSE in coordinate space. For the UniSEM model, this is not true anymore since the alr is not isometric. For this reason, we first compute the fitted response in the simplex and then compute the MSE in an ilr coordinate space (and it does not depend upon which particular one). Table 3 summarizes the MSE of all considered models. The MSE of UniSEM, UniLAG (using TS prediction) and UniLAG (using BP prediction) are very close to each other. The UniSEM model reaches the smallest MSE for the second and third ilr. Results also show that the total MSE of the UniLAG with the BP prediction method (see [Goulard et al., 2017]) is the smallest among all models.

Table 3: MSE for all models

Coordinates	**Univariate**					**Multivariate**	
	OLM	SEM	LAG(TS)	LAG(TC)	LAG(BP)	LAG(TC)	LAG(TS)
ilr1	7.21	4.52	4.49	6.99	4.36	9.13	4.55
ilr2	2.00	1.38	1.43	1.99	1.43	2.58	1.53
ilr3	4.31	2.91	3.01	4.40	2.97	4.49	3.57
Sum	13.52	8.81	8.93	13.38	**8.76**	16.20	9.65

4.3 Interpretation

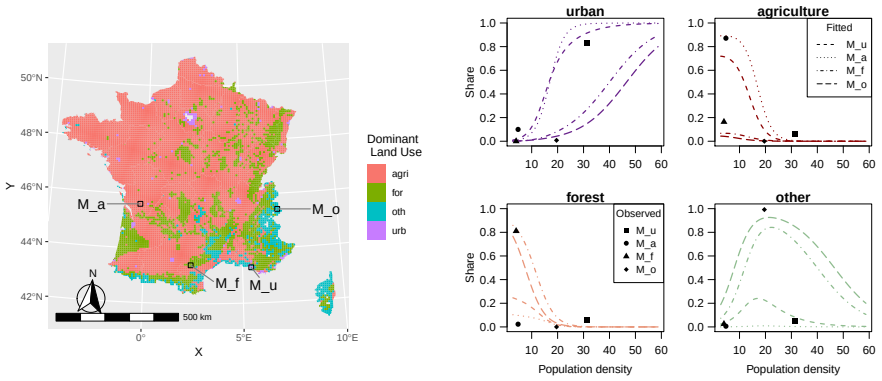
In this section, we try to better understand the impact of population density changes on land use using the techniques proposed in Section 3.2. We first interpret the fitted shares and the semi-elasticities thanks to very simple but informative scatterplots. Then we illustrate the interpretation of the direct and indirect impacts in the MultiLAG model using maps.

4.3.1 Fitted shares and semi-elasticities

We first select four particular grid cells, shown on Figure 1a, based on the fact that they are typical of a dominant share (more than 80%) of each particular land use (M_u for "urban", M_a for "agriculture", M_f for "forest" and M_o for "other"). We then draw the fitted shares scatterplot presented in Section 3.2 for a a sequence of 100 equally spaced values from 2.75 (minimum) to 58.72 (maximum) of the variable

population density while the other variables are fixed at the observed values for the four selected points. Figure 1b shows the fitted values obtained with the OLM model.

Each selected location is represented by a different dash type for the fitted shares and a different symbol for the observed shares. Population density shows a positive impact on "urban" use and a negative impact on both "agriculture" and "forest" uses. The effect on "other" use is first increasing and then decreasing. One can emphasize the variability of these curves according to the selected locations: "urban" use is gaining ground rapidly at the expense of "agriculture" for M_u and M_a. For M_f and M_o, it seems that the decline of the "forest" use is first beneficial to the "other" use before switching to the "urban" use.



(a) Four selected points with a respective dominance of "agriculture", "urban", "forest" and "other".

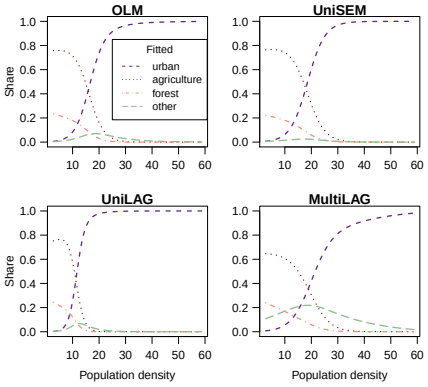
(b) Fitted shares curves according to population density for the four selected points.

Fig. 1: Selected points with a dominant land usetext and their fitted shares curves according to population density for the OLM model.

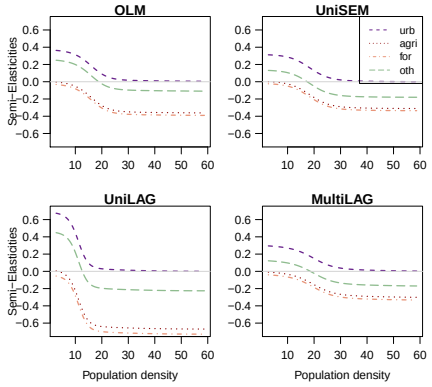
Figure 2a shows the average fitted shares according to population density for the OLM, UniSEM, UniLAG and MultiLAG models. At the exception of the UniLAG model, we see that the three other models, OLM, UniSEM and MultiLAG, give relatively comparable results for the fitted "urban" and "agricultural" uses: in particular these two curves intersect around the point where density is equal to 20 household per hectare whereas the intersection occurs at 10 for the UniLAG. Only the MultiLAG model shows a less flat shape of "other" uses. Generally, an increase in population density results in an increase of "urban" areas at the expense of "agricultural" land. This phenomenon has been observed in France since many years in rural peri-urban areas with the sprawl of cities and road infrastructure [Chakir and Madignier, 2006].

Figure 2b present the average semi-elasticities with respect to population density for the four models. They are computed with a value of $\delta = 10^{-10}$ for the approximation (20). The average semi-elasticities have the same shape for the UniSEM and MultiLAG models. The semi-elasticities are always positive for "urban" land use and negative for "forest" and "other" land uses. This means that an increase

in population density results in an increase of the "urban" and "other" use at the expense of "forest" and "agricultural" uses. The semi-elasticity of "other" land use remains positive until a population density around 20 households per hectare and becomes negative after that value. This indicates that in areas with low urban density, increasing population density would increase "urban" areas and "other" land uses at the expense of agriculture and forestry. For example, taking the selected point represented by a circle (agriculture dominant area) with a fitted share vector equal to $(0.006, 0.899, 0.092, 0.003)$ (in the order "urban", "agriculture", "forest" and "other") with the MultiLAG model and a vector of semi-elasticities equal to $(0.305, 0.001, -0.029, 0.131)$, an increase of one household per hectare, all other variables remaining constant, results approximately in an increase of the "urban" share by 30% and a decrease of forestry by 3% leaving the other two shares almost constant. In areas with high urban density (above 20 households per hectare), "urban" use would gain ground on all other uses including "other" land use. Note that the "other" land use may correspond to a temporary land use awaiting conversion to "urban" use either from "forest" or "agricultural" lands.



(a) Fitted shares according to population density for OLM, SEM, UniLAG and MultiLAG models.



(b) Average semi-elasticities with respect to population density for OLM, SEM, UniLAG and MultiLAG models.

Fig. 2: Fitted shares and average semi-elasticities for population density

4.3.2 Direct and indirect impacts

Figure 3 is an illustration of the local contributions to the direct and indirect impacts caused by an increase of one unit of population density at the cell M_u for the MultiLAG model. M_u is part of the city of Marseille, the second largest city in France. Its neighborhood contains mainly cells with a dominance of "urban" or "other" uses (see Figure 3a). Figures 3b (resp. 3c) represent the change of "urban"

(resp. “other”) use on M_u and its neighborhood. The local direct impact corresponds to the changes observed at M_u itself whereas the local indirect impact correspond to the changes observed at all other cells. We observe that an increase of one unit of population density has a positive direct impact on the “urban” use at the expense of “other” use. On the contrary, we observe a positive indirect impact on the “other” use whereas the indirect impact on the “urban” use is almost zero. Moreover, the indirect impact is all the more strong as the location is closer to the point of interest (see Fig. (3c)) .

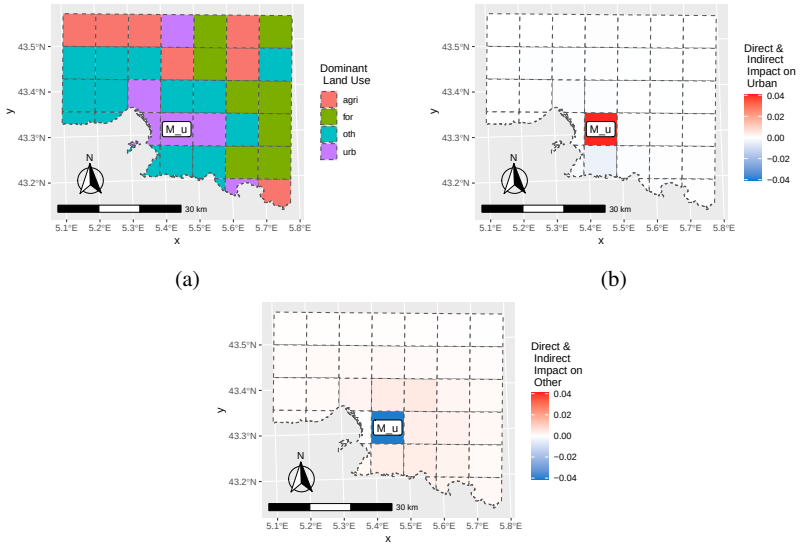
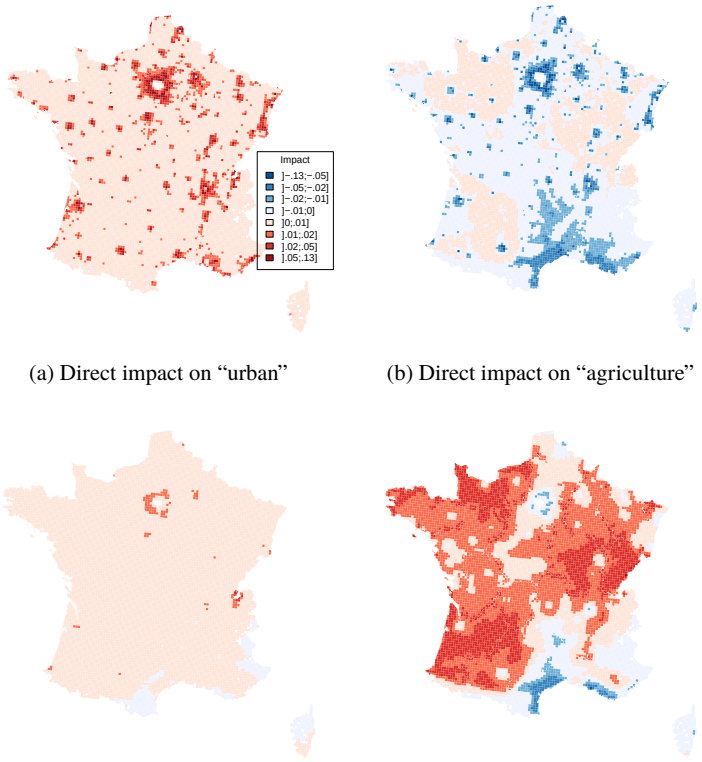


Fig. 3: Region of Marseille: (a) Dominant land uses; (b) Local impact on “urban” share; (c) Local impact on “other” share

The maps in Figure 4 summarize direct and indirect impacts of population density on “urban” and “agricultural” land use shares. They show that an increase in population density has a larger increase impact on “urban” land in already urbanized areas and it happens at the expense of “agricultural” land located in peri-urban areas.

The maps in Figure 5 show direct and indirect impacts of population density on “forest” and “other” land use shares. These maps show that an increase in population density results in an increase of “other” land uses at the expense of “forest” mainly in less urbanized areas. This confirms the fact that “other” land use could be a transition use from “forest” to “urban” land use when the population density increases.



(a) Direct impact on “urban”

(b) Direct impact on “agriculture”

(c) Indirect impact on “urban”

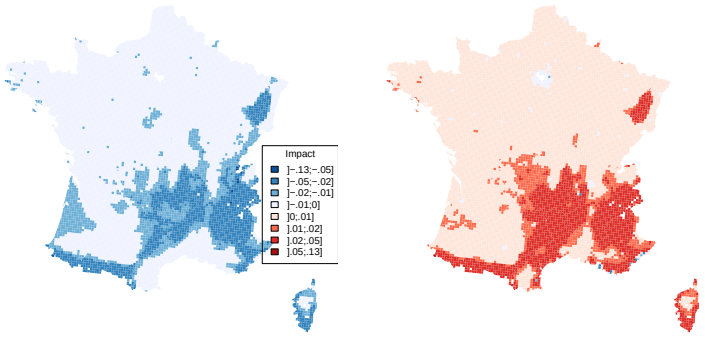
(d) Indirect impact on “agriculture”

Fig. 4: Direct and Indirect impact of population density on “urban” and “agricultural” shares

5 Conclusion

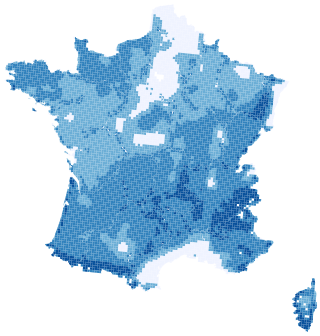
Through this application to land use, we have explored the differences between two families of models which we call the UniSEM and the MultiLAG. We have proved that a choice of coordinate representation does not have any impact on the parameters in the simplex as long as we do not impose further restrictions. Indeed we have seen that a restriction like the matrix of spatial autocorrelation parameters in coordinate space being diagonal is not a simplex assumption: it is dependent on a particular choice of coordinate space. Therefore a multivariate approach without any restrictions on the parameters is preferable. Finally we could have considered a MultiSEM model but did not do so due to the lack of code to fit such a multivariate model and this can be considered as a perspective for future work.

We have demonstrated that approximations of the semi-elasticities can be easily computed and interpreted in the simplex. The application results yield interesting insights for understanding the drivers of land use. Several directions of improve-

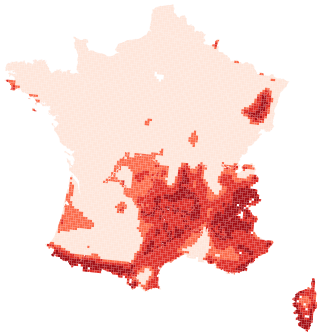


(a) Direct impact on “forest”

(b) Direct impact on “other”



(c) Indirect impact on “forest”



(d) Indirect impact on “other”

Fig. 5: Direct and indirect impact of population density on “forest” and “other” shares

ments could be pursued. First of all, one could derive formulas for the statistical significance of the semi-elasticities. The prediction formulas which have been used in the MSE evaluation are univariate: using multivariate counterparts (which haven’t been derived yet for these models to our knowledge) would be preferable.

Acknowledgements The authors are grateful to Vera Pawlowsky-Glahn for her valuable contributions to the CODA field. They also thank two anonymous referees and the editors for their helpful comments. Thibault Laurent, Anne Ruiz-Gazen and Christine Thomas-Agnan acknowledge funding from ANR under grant ANR-17-EURE-0010 (Investissements d’Avenir program). Raja Chakir and Anna Lungaska acknowledge the support of the Agence Nationale de la Recherche as part of the "Investments d’Avenir" Programme within STIMUL (Scenarios Towards integrating multi-scale land use tools) flagship project (LabEx BASC; ANR-11- LABX-0034) and Cland Institut de convergence (ANR-16-CONV-0003)

6 Appendix

6.1 Matrix product expression using alr transformation

Let us consider the alr_D transformation $\text{alr}_D(\mathbf{u}) = \mathbf{F}_D \ln(\mathbf{u})$ for a vector $\mathbf{u} \in \mathbf{S}^D$ and with $\mathbf{F}_D = [\mathbf{I}_{D-1} \ -\mathbf{j}_{D-1}]$. The result can be generalized to any alr_m transformation.

$$\text{alr}_D\left(\mathbf{B} \boxdot \mathbf{u}\right) = \mathbf{F}_D \ln\left(\mathbf{B} \boxdot \mathbf{u}\right) = \mathbf{F}_D \ln\left(\text{ilr}_V^{-1}\left(\mathbf{B}_V^* \text{ilr}_V\left(\mathbf{u}\right)\right)\right)$$

where $\mathbf{V}\mathbf{B}_V^* = \mathbf{B}\mathbf{V}$. Since $\mathbf{V}\mathbf{V}^T = \mathbf{I}_D - \mathbf{j}_D\mathbf{j}_D^T/D$ and $\mathbf{F}_D\mathbf{j}_D = \mathbf{0}_D$, we have

$$\text{alr}_D\left(\mathbf{B} \boxdot \mathbf{u}\right) = \mathbf{F}_D \ln\left(C\left(\exp(\mathbf{V}\mathbf{B}_V^* \text{ilr}_V\left(\mathbf{u}\right)\right)\right) = \mathbf{F}_D\mathbf{V}\mathbf{B}_V^* \text{ilr}_V\left(\mathbf{u}\right) = \mathbf{F}_D\mathbf{B}\ln(\mathbf{u}).$$

6.2 Writing the model in reduced form in coordinate space

Recognizing that the matrix form of the term $\sum_{m \in S_l^{\mathbf{WY}^*}} \mathbf{R}_{ml}^* \mathbf{W}\mathbf{Y}_{.m}^*$ is $\mathbf{W}\mathbf{Y}^*\mathbf{R}^*$, we get easily that the matrix formulation of the model can be written

$$\mathbf{Y}^* = \mathbf{W}\mathbf{Y}^*\mathbf{R}^* + \left[\mathbf{X}_{S_1^{\mathbf{X}}}\boldsymbol{\beta}^*_{S_1^{\mathbf{X}}} \cdots \mathbf{X}_{S_{D-1}^{\mathbf{X}}}\boldsymbol{\beta}^*_{S_{D-1}^{\mathbf{X}}}\right] + \boldsymbol{\epsilon}^*,$$

We are going to use the following property: if we have three matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ such that the number of columns of $\mathbf{A}$ is equal to the number of rows of $\mathbf{B}$ and the number of columns of $\mathbf{B}$ is equal to the number of rows of $\mathbf{C}$, then

$$\text{vec}_c(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}_c(\mathbf{B}). \tag{23}$$

Using (23), we get

$$\text{vec}_c(\mathbf{W}\mathbf{Y}^*\mathbf{R}^*) = \left((\mathbf{R}^*)^T \otimes \mathbf{W}\right) \text{vec}_c(\mathbf{Y}^*).$$

Therefore the c-vectorization of the whole model is

$$\text{vec}_c\mathbf{Y}^* = \left((\mathbf{R}^*)^T \otimes \mathbf{W}\right) \text{vec}_c(\mathbf{Y}^*) + \begin{bmatrix} \mathbf{X}_{\mathbf{S_1}}\boldsymbol{\beta}_1^* \\ . \\ \mathbf{X}_{\mathbf{S_{D-1}}}\boldsymbol{\beta}_{D-1}^* \end{bmatrix} + \text{vec}_c\boldsymbol{\epsilon}^*$$

and the reduced form of the model in vectorized form is

$$\text{vec}_c\mathbf{Y}^* = (\mathbf{I}_{n(D-1)} - ((\mathbf{R}^*)^T \otimes \mathbf{W}))^{-1} \begin{bmatrix} \mathbf{X}_{\mathbf{S_1}}\boldsymbol{\beta}_1^* \\ . \\ \mathbf{X}_{\mathbf{S_{D-1}}}\boldsymbol{\beta}_{D-1}^* \end{bmatrix} + \text{vec}_c\boldsymbol{\epsilon}^*.$$

6.3 Lemma used in the proof of Theorem 0.1

Lemma 0.1 *Let* $\mathbf{j}_D$ *(resp.* $\mathbf{0}_D$*) denote the* D*-dimensional column vector of ones (resp. zeros) and* $\mathbf{I}_D$ *be the* $D \times D$ *identity matrix. Let* $\mathbf{K} = \begin{bmatrix} \mathbf{I}_{D-1} - \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T/D \\ -\mathbf{j}_{D-1}^T/D \end{bmatrix}$ *a* $D \times (D-1)$ *matrix and* $\mathbf{F} = [\mathbf{I}_{D-1} \; -\mathbf{j}_{D-1}]$ *a* $(D-1) \times D$ *matrix. Let* $\mathbf{B}$ *be a* $D \times D$ *matrix such that* $\mathbf{B}\mathbf{j}_D = \mathbf{0}_D$ *and* $\mathbf{B}^T\mathbf{j}_D = \mathbf{0}_D$ *and* $\mathbf{B}^*$ *be a* $(D-1) \times (D-1)$ *matrix. We have that* $\mathbf{FB} = \mathbf{B}^*\mathbf{F}$ *is equivalent to* $\mathbf{B} = \mathbf{K}\mathbf{B}^*\mathbf{F}$.

We have

$$\mathbf{K} = \begin{bmatrix} 1-\frac{1}{D} & -\frac{1}{D} & -\frac{1}{D} & \dots & -\frac{1}{D} \\ -\frac{1}{D} & 1-\frac{1}{D} & -\frac{1}{D} & \dots & -\frac{1}{D} \\ -\frac{1}{D} & -\frac{1}{D} & 1-\frac{1}{D} & \dots & -\frac{1}{D} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{D} & -\frac{1}{D} & \frac{1}{D} & \dots & 1-\frac{1}{D} \\ -\frac{1}{D} & -\frac{1}{D} & -\frac{1}{D} & \dots & -\frac{1}{D} \end{bmatrix} \quad \text{and} \quad \mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & -1 \\ 0 & 1 & 0 & \dots & 0 & -1 \\ 0 & 0 & 1 & \dots & 0 & -1 \\ \vdots & \vdots & & \ddots & 0 & -1 \\ 0 & 0 & 0 & & 1 & -1 \end{bmatrix}$$

Let $\mathbf{B}$ be a $D \times D$ matrix such that $\mathbf{B}\mathbf{j}_D = \mathbf{0}_D$ and $\mathbf{B}^T\mathbf{j}_D = \mathbf{0}_D$ and $\mathbf{B}^*$ be a $(D-1) \times (D-1)$ matrix. First, let us prove that if $\mathbf{B} = \mathbf{K}\mathbf{B}^*\mathbf{F}$ then $\mathbf{FB} = \mathbf{B}^*\mathbf{F}$. It is easy to show that $\mathbf{FK} = \mathbf{I}_{D-1}$. Thus, if $\mathbf{B} = \mathbf{K}\mathbf{B}^*\mathbf{F}$,

$$\mathbf{FB} = \mathbf{FK}\mathbf{B}^*\mathbf{F} = \mathbf{I}_{D-1}\mathbf{B}^*\mathbf{F} = \mathbf{B}^*\mathbf{F}.$$

Let us now prove the converse. We have $\mathbf{B}\mathbf{j}_D = \mathbf{B}^T\mathbf{j}_D = \mathbf{0}_D$ and we assume $\mathbf{FB} = \mathbf{B}^*\mathbf{F}$. We write $\mathbf{B}$ in blocks as follows:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{b}_2 \\ \mathbf{b}_3^T & b_4 \end{bmatrix}$$

where $\mathbf{B}_1$ is a $D-1 \times D-1$ matrix, $\mathbf{b}_2$ and $\mathbf{b}_3^T \in \mathbf{R}^{D-1}$ and $\mathbf{b}_4 \in \mathbf{R}$. Using the fact that $\mathbf{B}\mathbf{j}_D = \mathbf{B}^T\mathbf{j}_D = \mathbf{0}_D$, we get

$$\begin{cases} \mathbf{b}_2 = -\mathbf{B}_1\mathbf{j}_{D-1} \\ \mathbf{b}_3 = -\mathbf{B}_1^T\mathbf{j}_{D-1} \\ \mathbf{b}_4 = \mathbf{j}_{D-1}^T\mathbf{B}_1\mathbf{j}_{D-1} \end{cases}. \tag{24}$$

To find $\mathbf{B}$, we only need to find $\mathbf{B}_1$. Using (24), we write the $(D-1) \times D$ matrix $\mathbf{FB}$ as a function of $\mathbf{B}_1$:

$$\mathbf{FB} = [(\mathbf{I}_{D-1} + \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T)\mathbf{B}_1 \quad -(\mathbf{I}_{D-1} + \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T)\mathbf{B}_1\mathbf{j}_{D-1}].$$

Furthermore, $\mathbf{B}^*\mathbf{F} = [\mathbf{B}^* \quad -\mathbf{B}^*\mathbf{j}_{D-1}]$. So, $\mathbf{FB} = \mathbf{B}^*\mathbf{F}$ implies

$$\left(\mathbf{I}_{D-1} + \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T\right)\mathbf{B}_1 = \mathbf{B}^*$$

The inverse matrix of $\left(\mathbf{I}_{D-1} + \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T\right)$ is $\left(\mathbf{I}_{D-1} + \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T/D\right)$. Thus,

$$\mathbf{B}_1 = (\mathbf{I}_{D-1} + \mathbf{j}_{D-1}\mathbf{j}_{D-1}^T/D)\mathbf{B}^*. \tag{25}$$

Using (24) and (25), it is now easy to check that $\mathbf{B} = \mathbf{K}\mathbf{B}^*\mathbf{F}$.

References

[Aitchison, 1986] Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall, London. Reprinted in 2003 with additional material by the Blackburn Press.

[Bivand and Wong, 2018] Bivand, R. S. and Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, 27(3):716–748.

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

[Chakir and Madignier, 2006] Chakir, R. and Madignier, A.-C. (2006). Analyse des changements d'occupation des sols en France entre 1992 et 2003. *Economie Rurale*, 296.

[Chen et al., 2017] Chen, J., Zhang, X., and Li, S. (2017). Multiple linear regression with compositional response and covariates. *Journal of Applied Statistics*, 44(12):2270–2285.

[Cliff and Ord, 1981] Cliff, A. D. and Ord, J. K. (1981). *Spatial processes*. Pion, London.

[Filzmoser et al., 2018] Filzmoser, P., Hron, K., and Templ, M. (2018). *Applied Compositional Data Analysis. With Worked Examples in R*. Springer International Publishing. Springer Nature Switzerland AG, Cham, Switzerland. Springer Series in Statistics.

[Foley et al., 2005] Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., et al. (2005). Global consequences of land use. *Science*, 309(5734):570–574.

[Glahn et al., 2012] Glahn, V., Hron, K., et al. (2012). Simplicial regression. the normal model. *Journal of Applied Probability and Statistics*, 6:87108.

[Goulard et al., 2017] Goulard, M., Laurent, T., and Thomas-Agnan, C. (2017). About predictions in spatial autoregressive models: Optimal and almost optimal strategies. *Spatial Economic Analysis*, 12(2-3):304–325.

[Kelejian and Prucha, 1998] Kelejian, H. H. and Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1):99–121.

[Kelejian and Prucha, 2004] Kelejian, H. H. and Prucha, I. R. (2004). Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of Econometrics*, 118(1-2):27–50.

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

[Leininger et al., 2013] Leininger, T. J., Gelfand, A. E., Allen, J. M., and Silander, J. A. (2013). Spatial regression modeling for compositional data with many zeros. *Journal of Agricultural, Biological, and Environmental Statistics*, 18(3):314–334.

[LeSage and Pace, 2009] LeSage, J. and Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

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

[Martín-Fernández et al., 2003] Martín-Fernández, J. A., Barceló-Vidal, C., and Pawlowsky-Glahn, V. (2003). Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Mathematical Geology*, 35(3):253–278.

[Morais and Thomas-Agnan, 2020] Morais, J. and Thomas-Agnan, C. (2020). Covariates impacts in compositional models and simplicial derivatives. *Austrian Journal of Statistics*. To appear.

[Munroe and Müller, 2007] Munroe, D. K. and Müller, D. (2007). Issues in spatially explicit statistical land-use/cover change (lucc) models: Examples from western Honduras and the Central Highlands of Vietnam. *Land use policy*, 24(3):521–530.

[Nguyen et al., 2019] Nguyen, T. H. A., Thomas-Agnan, C., Laurent, T., and Ruiz-Gazen, A. (2019). A simultaneous spatial autoregressive model for compositional data. *Spatial Economic Analysis*.

[Pawlowsky-Glahn and Egozcue, 2011] Pawlowsky-Glahn, V. and Egozcue, J. J. (2011). Exploring compositional data with the CoDa-dendrogram. *Austrian Journal of Statistics*, 40(1&2):103–113.

[Pawlowsky-Glahn et al., 2015] Pawlowsky-Glahn, V., Egozcue, J. J., and Tolosana-Delgado, R. (2015). *Modeling and analysis of compositional data*. John Wiley & Sons.

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

[Pirzamanbein et al., 2018] Pirzamanbein, B., Lindström, J., Poska, A., and Gaillard, M.-J. (2018). Modelling spatial compositional data: Reconstructions of past land cover and uncertainties. *Spatial Statistics*, 24:14–31.

[Shmueli, 2010] Shmueli, G. (2010). To explain or to predict? *Statistical science*, pages 289–310.

[Thomas-Agnan et al., 2020] Thomas-Agnan, C., Laurent, T., and Ruiz-Gazen, A. (2020). Covariates impacts in spatial autoregressive models for compositional data. *TSE working paper 20-1162*.

[Veldkamp and Lambin, 2001] Veldkamp, A. and Lambin, E. F. (2001). Predicting land-use change. *Agriculture, ecosystems & environment*, 85(1):1–6.

[Verburg et al., 2013] Verburg, P. H., Erb, K.-H., Mertz, O., and Espindola, G. (2013). Land System Science: Between global challenges and local realities. *Current Opinion in Environmental Sustainability*, 5(5):433–437.