Estadística Española • Volumen 54, número 177 / 2012, pp. 89-111

SAR models with nonparametric spatial trends. A P-spline approach

José-María Montero
Universidad de Castilla-La Mancha(*)

Román Mínguez
Universidad de Castilla-La Mancha(**)

María Durbán
Universidad Carlos III de Madrid(***)

Abstract

This article proposes a model which includes a global spatial trend in an SAR specification to take into account both large-scale spatial dependencies and local spatial autocorrelation. We use penalized splines to estimate the model, as they can be represented as mixed models. As a result we can (i) deal with complex non-linear trends, which are very common in spatial phenomena, (ii) estimate short-range spatial correlation together with the large-scale spatial trend, (iii) decompose the systematic spatial variation into these two components and (iv) estimate the smoothing parameters together with the other parameters of the model. We call this model the P-spline-SAR model. Based on the simulation of 2000 datasets generated by a P-spline-SAR (using both linear and non-linear and non-separable global spatial trends), we conclude that (i) the P-spline-SAR model provides much better estimates of both the global spatial trend and also the spatial autocorrelation term than the pure P-spline or SAR specifications, irrespective of whether the true trend is linear or non-linear; (ii) the estimations of the observed values yielded by the P-Spline-SAR model are equally as accurate as those provided by the best competing alternative. We also empirically illustrate how well the P-spline-SAR model performs using the augmented Harrison and Rubinfeld (1978) hedonic pricing data for Boston SMSA.

Keywords: P-splines, SAR, autocorrelation, global spatial trend.

JEL classification: C14, C15, C21

AMS classification: 91B72, 93E14, 65D07, 62M30, 62G08

(*) Jose.MLorenzo@uclm.es

(**) Roman.Minguez@uclm.es

(***) maria.durban@est-econ.uc3m.es

Modelos SAR con tendencia espacial no paramétrica. Un enfoque P–spline

Resumen

Este artículo propone un modelo que incluye una tendencia espacial global en una especificación SAR para estimar tanto las dependencias espaciales a gran escala como la autocorrelación espacial local. Para la estimación del modelo se utilizan splines con penalización, debido a que pueden ser representados como modelos mixtos. De esta manera, (i) pueden considerarse complejas tendencias no lineales, muy habituales en los fenómenos espaciales, (ii) se puede estimar tanto la tendencia global como la correlación espacial a pequeña escala, (iii) se puede descomponer la variación espacial sistemática en los dos componentes anteriores y (iv) se pueden estimar los parámetros de suavizado junto con los demás parámetros del modelo. A este modelo se le denomina P-spline-SAR. A partir de la simulación de 2000 conjuntos de datos generados por un P-spline-SAR (utilizando tanto una tendencia lineal como una tendencia no lineal y no separable), se concluye que (i) el modelo P-spline-SAR proporciona estimaciones, tanto de la tendencia como del parámetro de autocorrelación espacial, mucho mejores que las de las especificaciones puras P-spline y SAR, independientemente de si la verdadera tendencia es o no lineal; (ii) el modelo P-spline-SAR siempre proporciona estimaciones de la variable observada tan buenas como las del mejor de los modelos competidores. El buen comportamiento del modelo P-spline-SAR también se ilustra empíricamente a partir de la base de datos (aumentada) de precio de la vivienda de Harrison y Rubinfeld (1978) para Boston SMSA.

Palabras clave: P-splines, SAR, autocorrelación, tendencia espacial global.

Clasificación JEL C14, C15, C21

Clasificacion AMS: 91B72, 93E14, 65D07, 62M30, 62G08

1. Introduction

The analysis of spatial (and more recently spatio-temporal) data is currently of great interest to statistical modeling, especially from the econometrics and geostatistical perspectives. Problems related to meteorology, environmental pollution, ecology, epidemiology and economics, among other scientific disciplines, demand the use of statistical models for spatial data.

Here we focus on the econometrics perspective and more specifically on the well known Spatial Autoregressive (SAR) model. SAR models (also cited in the literature as mixed regressive-spatial autoregressive models and spatial lag models) account for spatial autocorrelation expressly by incorporating a spatial term into the standard regression model (see Anselin, 1988, 2007, LeSage and Pace, 2009). Essentially an SAR model expresses the notion that the value of a variable at a given location is related to the values of the same variable measured at nearby locations, reflecting some kind of interaction. However, it may not be the best option when a global spatial trend exists and such a spatial trend is complex, non-linear and this non-linear structure is unknown.

The large-scale variation is crucial when searching for global spatial patterns, as they help to both understand the spatial structure of our datasets and also discover the main features of the phenomenon under study. As stated in Crujeiras and Keilegom (2010), linear trend specifications are not realistic for the trend of most real phenomena that evolve in space and non-linear trends that are present in many examples involving spatially dependent data. For instance, in soil science Snepvangers et al. (2003) describe a sigmoid growth curve for modeling the relationship between irrigation and soil water content; in environmental science, Haas (1996) considers a non-linear trend model for sulphate deposition; and in meteorology, Verkatram (1988) derives a non-linear model for acid deposition, based on a differential equation system. As the case of unknown complex non-linear spatial trends is common in phenomena which evolve in space, in this paper we deal with data generated by combining a global spatial trend and spatial autocorrelation derived from an SAR process. Our approach, rather than requiring an a priori specification of the trend model form, lets the data suggest the form of the model, which results in a highly flexible technique and penalizes over-parameterization. More specifically, the specification we propose comprises a global spatial trend (which can be both complex and non-linear and could include covariates), a spatial signal that captures the spatially autocorrelated deviations from the global spatial trend and a spatially non-autocorrelated noise term.

The estimation of a trend in the spatial context is not a new problem and several approaches have been proposed in recent years. Here, we will consider the use of penalized splines (Eilers and Marx, 1996) to smooth spatial data (Lee and Durbán, 2009). These methods are well established for smoothing data in one or more dimensions (Currie et al., 2006) and have been used in different applications. Penalized splines (or P-splines) are low-rank smoothers that use a basis for regression and impose a penalty so that adjacent coefficients vary smoothly. They have become very popular due to their representation as mixed models (Currie and Durbán, 2002). This makes it possible to include non-linear trends in many models. In particular, P-splines are very attractive in the context of spatial models, since short-range spatial correlation can be estimated together with the large-scale spatial trend. They have been studied in smooth-CAR models (Lee and Durbán, 2009) and here we propose their use in SAR models. The paper is organized as follows. Section 2 introduces P-splines and their mixed model representation in one and two dimensions. Section 3 extends the SAR model by incorporating a two-dimensional P-spline to account for the large-scale trend. A simulation study is carried out in section 4 in order to compare several competing models. Section 5 includes an empirical case study and we conclude with a discussion in section 6.

2. Splines with Penalties: P-splines

Suppose that the vector **y** depends smoothly on a vector **x** (covariate). Then, a smooth model for the data would be given by:

$$\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\varepsilon} = \mathbf{B}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}),$$

where $\boldsymbol{\theta}$ is a vector of coefficients and $\mathbf{B} \equiv \mathbf{B}(\mathbf{x})$ is a regression basis constructed from the vector **x**. The basis can be chosen in different ways. Here we will use a B-splines basis. The P-splines approach (Eilers and Marx, 1996) modifies the likelihood function by adding a penalty term for the adjacent coefficients to control the smoothness of the fit. The penalty matrix is $\mathbf{P} = \lambda\mathbf{D}'\mathbf{D}$, where **D** is a difference matrix (in general, differences of order two are used), and λ is a smoothing parameter. Then, the coefficients $\boldsymbol{\theta}$ are chosen to minimize:

$$\mathrm{S}(\boldsymbol{\theta}) = (\mathbf{y} - \mathbf{B}\boldsymbol{\theta})'(\mathbf{y} - \mathbf{B}\boldsymbol{\theta}) + \boldsymbol{\theta}'\mathbf{P}\boldsymbol{\theta}.$$

For a given λ, the solution to the penalized sum of squares is:

$$\hat{\boldsymbol{\theta}} = (\mathbf{B}'\mathbf{B} + \mathbf{P})^{-1}\mathbf{y}.$$

2.1 P-splines as mixed models

The connection between non-parametric regression and mixed models goes back to the early nineties (Speed, 1991). In the late nineties and the beginning of the last decade, the topic of smoothing with mixed models was developed in the framework of smoothing splines (Brumback and Rice, 1998, Verbyla et al., 1999, among others) and in the context of P-splines with truncated lines as a basis (Wand, 2003). However, the representation of P-splines as mixed models using B-splines as a basis was not studied until recently (Currie and Durbán, 2002) The use of B-splines has the advantage of resulting in a more stable basis for the mixed model.

Given a standard smooth model:

$$\mathbf{y} = \mathbf{B}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}), \qquad [1]$$

where **B** is the B-spline regression matrix, the aim is to look for a new basis that allows the representation of [1] as a mixed model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \quad \boldsymbol{\alpha} \sim N(\mathbf{0}, \mathbf{G}) \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}), \qquad [2]$$

where **X** is the matrix of the fixed effects, **Z** is the matrix of random effects and **G** is a diagonal matrix which depends on the random effects variance σ_α^2. As usual, σ^2 is the variance of the disturbances.

One possible way to find the new basis is to use the singular value decomposition (SVD) of the penalty $\mathbf{D'D}$ to partition it into a null penalty (which corresponds to the fixed part of the model) and a diagonal penalty (for the random part), i.e., $\mathbf{D'D} = \mathbf{U\Sigma U'}$. Now, $\mathbf{D'D}$ has rank *c-q* (where *q* is the order of the penalty and *c* the number of columns of **B,** in general *q=2*), and so, $\mathbf{\Sigma}$ is a diagonal matrix with *q* eigenvalues equal to zero. Then, we define[1]:

$$\mathbf{X} = [\mathbf{1} : \mathbf{x}] \quad \mathbf{Z} = \mathbf{B}\mathbf{U}_{\mathbf{s}}, \qquad [3]$$

where $\mathbf{U_s}$ are the eigenvectors corresponding to the non-zero eigenvalues of $\mathbf{\Sigma}$. Now, we can immediately establish the connection with a mixed model like [2] where $\mathbf{G} = \sigma^2(\lambda\widetilde{\Sigma})^{-1}$ is the covariance matrix of the random effects and $\widetilde{\Sigma}$ is the diagonal matrix of the non-zero eigenvalues of Σ. The smoothing parameter is $\lambda = \sigma^2 / \sigma_\alpha^2$, and so the degree of smoothness can be estimated using standard mixed models methodology. In the context of mixed models, the usual method for estimating the variance parameters is the residual maximum likelihood (REML) presented in Patterson and Thompson (1971),

$$\ell_R(\sigma_\alpha^2, \sigma_\varepsilon^2) = -\frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\log|\mathbf{X'V^{-1}X}| - \frac{1}{2}\mathbf{y}'(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X'V^{-1}X})^{-1}\mathbf{X'V^{-1}})\mathbf{y}, \qquad [4]$$

where $\mathbf{V} = \mathbf{ZGZ}' + \sigma^2\mathbf{I}$. The vector of parameters $\boldsymbol{\beta}$ and the vector of random effects $\boldsymbol{\alpha}$ are estimated as:

$$\boldsymbol{\beta} = (\mathbf{X'V^{-1}X})^{-1}\mathbf{X'V^{-1}y} \qquad [5]$$

$$\boldsymbol{\alpha} = \mathbf{GZ'V^{-1}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \qquad [6]$$

$$\mathbf{V}^{-1} = \frac{\mathbf{1}}{\sigma^2}(\mathbf{I} - \mathbf{Z}(\mathbf{Z'Z} + \sigma^2\mathbf{G}^{-1})^{-1}\mathbf{Z}'). \qquad [7]$$

2.2 Spatial smoothing with P-splines

For simplicity, let us assume a normally distributed response variable *y*, depending on two spatial coordinates x_1 and x_2 corresponding to geographic longitude and latitude respectively. A smooth model for the data would be given by:

$$\mathbf{y} = f(\mathbf{x}_1, \mathbf{x}_2) + \boldsymbol{\varepsilon} = \mathbf{B}\boldsymbol{\theta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}), \qquad [8]$$

where $\boldsymbol{\theta}$ is a vector of coefficients and $\mathbf{B} \equiv \mathbf{B}(\mathbf{x}_1, \mathbf{x}_2)$ is a regression basis constructed from the covariates (x_1, x_2).

[1] The matrix **X** could be augmented with covariates, as usual in regression models

In the case of two-dimensional smoothing (as is the case of spatial data), the choice of the basis and penalty is even more important and the differences between the approaches are significant. Some authors, like Ruppert et al. (2003) and Kammann and Wand (2003), proposed the use of radial basis functions and mention the connection with other smoothers like thin plate splines (see Wahba, 1990, and Green and Yandell, 1985, for details on thin plate splines). However, these bases have the limitation of being isotropic smoothers; in addition, the selection of knots to construct the basis is not trivial and relies on reduced knots or low-rank kriging approximations based on *space filling algorithms*. Other authors, such as Lang and Brezger (2004), Wood (2006a), Currie et al. (2006) and Lee and Durbán (2009), propose the use of tensor product of B-spline basis functions with equally spaced knots. This is the approach we take here.

In the case of scattered spatial data (see Eilers et al., 2006, for details, the basis is constructed from the tensor product of marginal B-spline bases, defined as the Box-Product or "row-wise". Kronecker product of the individual bases, which is denoted by $\Box$:

$$\mathbf{B}=\mathbf{B}_2 \Box \mathbf{B}_1 = (\mathbf{B}_2 \otimes \mathbf{1}'_{c_1}) \odot (\mathbf{1}'_{c_2} \otimes \mathbf{B}_1), \tag{9}$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are the B-spline bases along the longitude (x_1) and latitude (x_2) coordinates.

The P-spline approach for two-dimensional data also differs from other methods in the penalty applied to the coefficients. Currie et al. (2004, 2006) proposed a penalty matrix based on the penalties associated with each of the marginal bases. More specifically, the bidimensional case is,

$$\mathbf{P} = \lambda_2 \mathbf{D}'_2 \mathbf{D}_2 \otimes \mathbf{I}_{c_1} + \lambda_1 \mathbf{I}_{c_2} \otimes \mathbf{D}'_1 \mathbf{D}_1, \tag{10}$$

where λ_1 and λ_2 are smoothing parameters which tune the smoothness in each direction separately, so that this penalty allows for *anisotropy*. The formulation as a mixed model is again based on a reparameterization of the model. We need to find a *one-to-one* transformation from $\mathbf{B\theta}$ to $\mathbf{X\beta}+\mathbf{Z\alpha}$, so that the composed matrix of $\mathbf{X}$ and $\mathbf{Z}$ has full column rank. This transformation is not unique. We follow a similar approach to that of Lee and Durbán (2009) and Lee (2010), and use the SVD of $\mathbf{P}$ in Eq. [10] (as a function of the SVD of the individual penalties $\mathbf{D}'_1\mathbf{D}_1$ and $\mathbf{D}'_2\mathbf{D}_2$). Then, the mixed model matrices are:

$$\mathbf{X} = \left(\mathbf{1}_n : \mathbf{1}_n \Box \mathbf{x}_1 : \mathbf{x}_2 \Box \mathbf{1}_n : \mathbf{x}_2 \Box \mathbf{x}_1\right) \tag{11}$$

$$\mathbf{Z} = \left(\mathbf{Z}_2 \Box \mathbf{1}_n : \mathbf{Z}_2 \Box \mathbf{x}_1 : \mathbf{1}_n \Box \mathbf{Z}_1 : \mathbf{x}_2 \Box \mathbf{Z}_1 : \mathbf{Z}_2 \Box \mathbf{Z}_1\right). \tag{12}$$

Matrices $\mathbf{Z}_i$, i=1,2, are the matrices for the random effects (defined in [3]) associated with each of the marginal bases. The covariance matrix of random effects now becomes:

$$\mathbf{G}=\sigma^2\begin{pmatrix}\lambda_2\tilde{\boldsymbol{\Sigma}}_2\otimes\mathbf{I}_2 & & \\ & \lambda_1\mathbf{I}_2\otimes\tilde{\boldsymbol{\Sigma}}_1 & \\ & & \lambda_1\mathbf{I}_{c_2-2}\otimes\tilde{\boldsymbol{\Sigma}}_1+\lambda_2\tilde{\boldsymbol{\Sigma}}_2\otimes\mathbf{I}_{c_1-2}\end{pmatrix}^{-1}. \quad [13]$$

The estimation of fixed and random effects is carried out as in the one-dimensional case.

3. P-spline-SAR model

We define the P-spline-SAR (PS-SAR) model for the bi-dimensional case as:

$$\mathbf{y}=\mathbf{B}\boldsymbol{\theta}+\rho\mathbf{W}_n\mathbf{y}+\boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon}\sim N(\mathbf{0},\sigma^2\mathbf{I}), \quad [14]$$

where $\mathbf{B}\equiv\mathbf{B}(\mathbf{x}_1,\mathbf{x}_2)$ is the bi-dimensional regression basis, $\mathbf{W}_n$ is a row-standardized spatial weights matrix which takes into account the n closest neighbours, $\mathbf{W}_n\,\mathbf{y}$ captures the spatial lags of the dependent variable and ρ is a spatial parameter that measures the existing spatial dependence of **y.** $\boldsymbol{\theta}$ has length c_1 x c_2 which indicates that the model is clearly over-parameterized. This is the reason why the coefficients are penalized in the form $\boldsymbol{\theta}'\mathbf{P}\boldsymbol{\theta}$, with **P** defined in Eq. [10].

The model [14] can also be presented as a linear mixed model:

$$\mathbf{A}\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\boldsymbol{\alpha}+\boldsymbol{\varepsilon} \quad \boldsymbol{\alpha}\sim N(\mathbf{0},\mathbf{G}) \quad \boldsymbol{\varepsilon}\sim N(\mathbf{0},\sigma^2\mathbf{I}), \quad [15]$$

where:

$$\mathbf{A}=\mathbf{I}_n-\rho\mathbf{W}_n, \quad [16]$$

and the other matrices are defined in equations [11] to [13].

It is important to note that in the data generation process (DGP) given by:

$$\begin{aligned}\mathbf{y}&=\mathbf{A}^{-1}\mathbf{B}(\mathbf{x}_1,\mathbf{x}_2)+\mathbf{A}^{-1}\boldsymbol{\varepsilon}=\\ &=\mathbf{A}^{-1}\left(\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\boldsymbol{\alpha}\right)+\mathbf{u} \quad \boldsymbol{\alpha}\sim N(\mathbf{0},\mathbf{G}) \quad \mathbf{u}\sim N(\mathbf{0},\sigma^2\mathbf{A}^{-1}\mathbf{A}^{-1'}),\end{aligned} \quad [17]$$

the variable **y** will show spatial dependencies in the systematic part, given by $\mathbf{A}^{-1}\mathbf{B}$ $(\mathbf{x}_1, \mathbf{x}_2)$, as well as in the vector of errors, **u**, although originally the vector of errors $\boldsymbol{\varepsilon}$ did not show any spatial dependence. Moreover, if the spatial trend, $f(\mathbf{x}_1, \mathbf{x}_2)$, is captured by the regression basis, $\mathbf{B}$ $(\mathbf{x}_1, \mathbf{x}_2)$, this model is able to factorize the systematic variation as the product of the trend by the dense matrix $\mathbf{A}^{-1}$. It is worth noting that neither the pure P-spline model (PS) nor the SAR specification provide such a factorization in the case where, as usual, the true spatial trend is unknown. In addition, it should be taken into account that the terms representing both types of spatial dependencies share the matrix $\mathbf{A}^{-1}$ and, as a consequence, interact.

The parameter vector $\left(\lambda_1 = \frac{\sigma^2}{\sigma^2_{\alpha_1}}, \lambda_2 = \frac{\sigma^2}{\sigma^2_{\alpha_2}}, \sigma^2, \rho\right)$ can be estimated by modifying the REML function as follows:

$$\ell_R(\lambda_1, \lambda_2, \sigma^2, \rho) = -\frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\log|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}| + \log|\mathbf{A}| \\ -\frac{1}{2}\mathbf{y}'\mathbf{A}'(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1})\mathbf{A}\mathbf{y}. \quad [18]$$

The modified REML function above can be maximized using numerical algorithms. As usual, the vector of fixed effects, $\boldsymbol{\beta}$, and the vector of random effects, $\boldsymbol{\alpha}$, are estimated as:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\hat{\mathbf{A}}\mathbf{y} \quad [19]$$

$$\hat{\boldsymbol{\alpha}} = \hat{\mathbf{G}}\mathbf{Z}'\hat{\mathbf{V}}^{-1}(\hat{\mathbf{A}}\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}), \quad [20]$$

where the inverted covariance matrix $\mathbf{V}^{-1}$ is the same as in [7].

4. Simulation study

In this Section, the performance of PS-SAR models is compared to the performance of PS and pure SAR models. For this purpose, we simulated 2000 datasets generated by a PS-SAR model with two types of spatial trends: non-linear and non-separable (case 1) and linear (case 2). Then, we estimated the three foregoing models. These two types of spatial trends were chosen as opposite cases, so that, predictably, the DGP could be close to a pure PS process (case 1) or to a pure SAR process (case 2). This way, the behavior of the proposed PS-SAR strategy can be checked in the two more extreme cases.

More specifically, the data generating process[2] (DGP) is given by:

$$\mathbf{y} = f(\mathbf{x}_1, \mathbf{x}_2) + \rho\mathbf{W}_n\mathbf{y} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}), \quad [21]$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ were generated from a uniform distribution in (0,1). Two types of trend generating processes (TGP) have been considered:

– Case 1: Non-linear trend. In this case, the trend, $f(\mathbf{x}_1, \mathbf{x}_2)$, is computed as in [15], that is:

[2] All computations were made using R software. More specifically we used the packages spline, spdep, mgcv and scatterplot3d, see R Development Core Team, 2011, Bivand et al., 2011, Wood, 2006b, Ligges and Mächler, 2003, respectively. The codes are available upon request.

$$f(\mathbf{x}_1,\mathbf{x}_2) = 10\pi\sigma_{x_1}\sigma_{x_2}\Big\{1.2\exp\Big(-(\mathbf{x}_1-0.2)^2/\sigma_{x_1}^2-(\mathbf{x}_2-0.3)^2/\sigma_{x_2}^2\Big)+ \\ +0.8\exp\Big(-(\mathbf{x}_1-0.7)^2/\sigma_{x_1}^2-(\mathbf{x}_2-0.8)^2/\sigma_{x_2}^2\Big)\Big\}, \quad [22]$$

with $\sigma_{x_1} = 0.3$ and $\sigma_{x_2} = 0.4$.

Case 2: Linear trend. In this case, the trend is given by an x_1x_2-plane, that is:

$$f(\mathbf{x}_1,\mathbf{x}_2) = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2, \quad [23]$$

with $\beta_0 = 2.5$, $\beta_1 = -0.5$ and $\beta_2 = -0.1$.

The set of parameter values chosen for the simulation procedure is $\rho = (0,0.25,0.5,0.75)$ and $\sigma = (0.5,1)$. This way, we consider both the pure PS case ($\rho = 0$) and the cases where positive spatial autoregressive dependence is high (values of ρ equal to or above 0.5, as in many real situations).

The simulation algorithm is composed of the following five steps:

1. Generate the random vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ (length n=200). Compute the trend values, which remain fixed in all the simulations. Generate the contiguity matrix $\mathbf{W}_n$ on the basis of the simulated coordinates. The elements of $\mathbf{W}_n$ are 1 for the five closest neighbors and 0 otherwise. As usual, $\mathbf{W}_n$ is used in row-standardized form.
2. Choose a couple of values for both ρ and σ and generate n values of $\boldsymbol{\varepsilon} \sim N(\mathbf{0},\sigma^2\mathbf{I})$. Then, compute the values of vector $\mathbf{y}$ as in Eq. [21].
3. Make a REML estimation of the model parameters for PS and PS-SAR models. In both cases, the matrix of regressors $\mathbf{B}(\mathbf{x}_1, \mathbf{x}_2)$ is built with B-spline basis matrices, the number of knots being set at 10. As usual in this type of models, second-order penalties are used. Finally, estimate an SAR model (ML estimation) including $\mathbf{x}_1$ and $\mathbf{x}_2$ as independent regressors (that is, as if a linear trend were known for both coordinates). In both the SAR and PS-SAR models the contiguity matrix is the one specified in step (1).

   Compare the estimates of the trend values obtained with the three models to the simulated trend values, $f(\mathbf{x}_1,\mathbf{x}_2)$. Also compare the estimates of the observed (simulated) values with the values obtained from the DGP.
4. Return to step (2) and repeat the process m=250 times for each possible combination of σ and ρ values.
5. Compute, as follows, the mean squared error (MSE) for both the estimated trend values and the estimated values of the response variable.

$$\mathrm{MSE}_{\mathrm{trend}} = \sum \left[f\left(\mathbf{x}_1, \mathbf{x}_2\right) - \widehat{f\left(\mathbf{x}_1, \mathbf{x}_2\right)} \right]^2 / \mathrm{n} \qquad [24]$$

where

$$\widehat{f\left(\mathbf{x}_1, \mathbf{x}_2\right)} = \begin{cases} \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}\hat{\boldsymbol{\alpha}} \text{ in PS and PS} - \text{SAR} \\ \hat{\beta}_0 + \hat{\beta}_1 \mathbf{x}_1 + \hat{\beta}_2 \mathbf{x}_2 \text{ in SAR} \end{cases}$$

$$\mathrm{MSE}_{\mathrm{resp.var.}} = \sum \left[\hat{\mathbf{A}}\mathbf{y} - \widehat{f\left(\mathbf{x}_1, \mathbf{x}_2\right)} \right]^2 / \mathrm{n} \qquad [25]$$

where

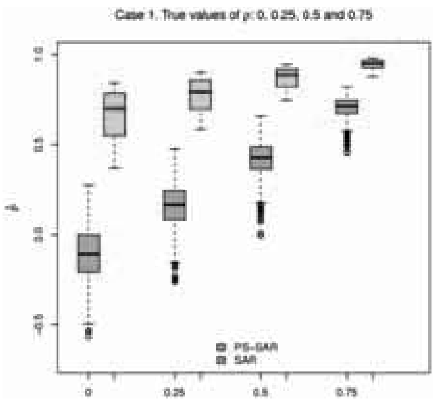
$$\hat{\mathbf{A}} = \begin{cases} \mathbf{I}_\mathrm{n} \text{ in PS} \\ \mathbf{I}_\mathrm{n} - \hat{\rho}\mathbf{W}_\mathrm{n} \text{ in PS} - \text{SAR and SAR} \end{cases}$$

Tables 1 to 4 (in the appendix) and Figures 1 to 3 summarize the simulation results. Next, we highlight the main findings regarding the variance of the error term, the spatial dependence parameter, the estimation of the trend and the estimation of the observed (simulated) values of the dependent variable.

Figure 1

Box-Plot of the estimates of ρ.

Case 1: The DGP includes a non-linear trend. Case 2: The DGP includes a linear trend.



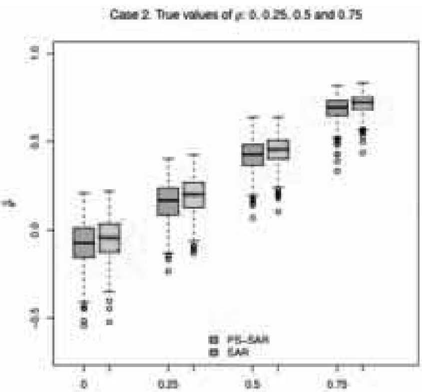


Figure 2

Box-Plot of the MSE when estimating the trend with PS-SAR and SAR models.
Case 1: The DGP includes a non-linear trend. Case 2: The DGP includes a linear trend.

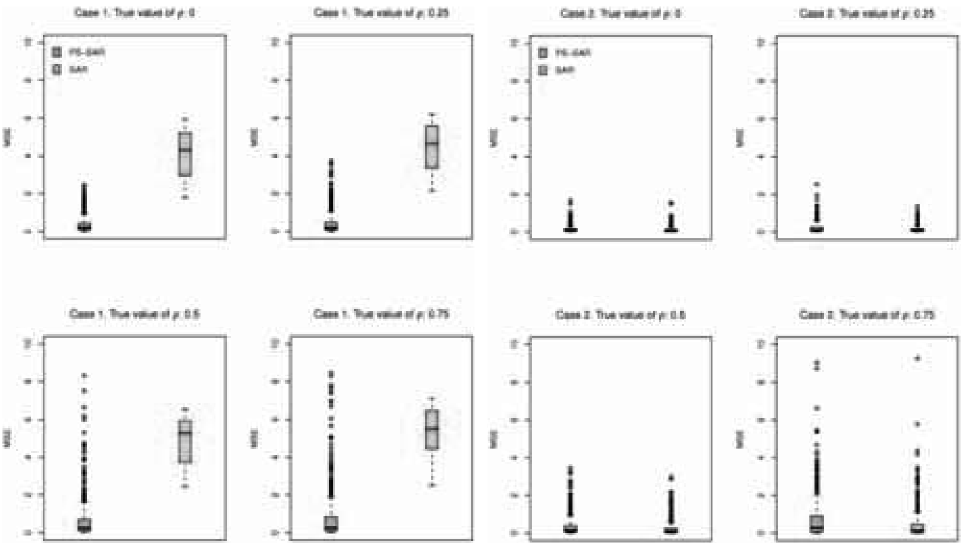
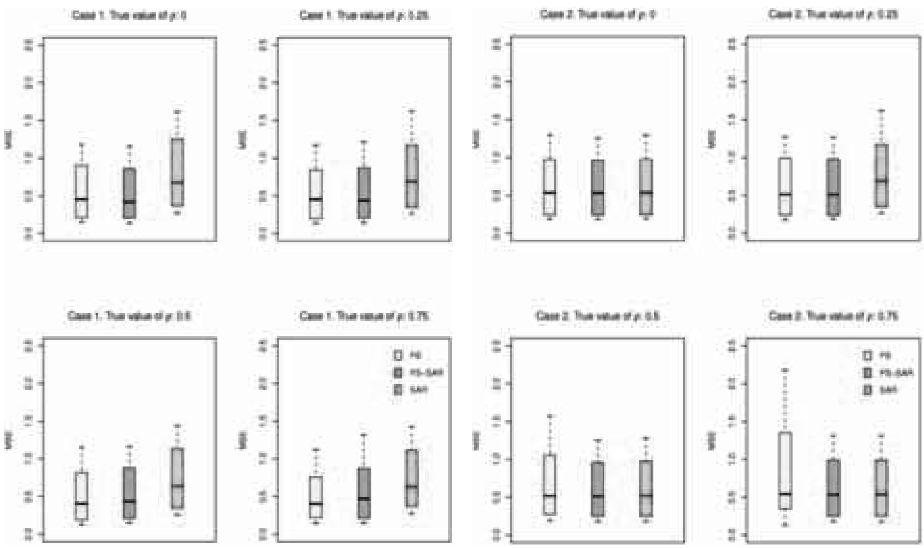


Figure 3

Box-Plot of the MSE when estimating the response variable with PS-SAR, PS and SAR models.
Case 1: The DGP includes a non-linear trend. Case 2: The DGP includes a linear trend.

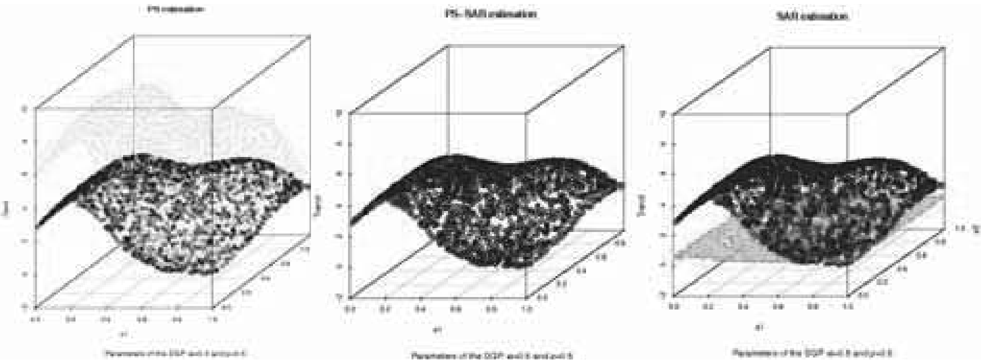


When we focus on the spatial dependence parameter, ρ, we have to distinguish case 1 from case 2. In case 1 (a non-linear trend), Table 1 and Figure 1 show that the estimate provided by PS-SAR tends to be slightly below the true value of the parameter, whereas the estimate provided by the SAR model tends to be close to unity, irrespective of the true value of ρ. More specifically, the larger the true value of ρ, the closer the SAR estimate is to unity. This upward bias is especially pronounced when σ takes low values, indicating that the SAR model is not a proper model to separate global spatial dependence (represented by $f(\mathbf{x}_1,\mathbf{x}_2)$) and local spatial dependence (given by ρ). However, the PS-SAR model overcomes this drawback. In case 2 (a linear trend), Figure 1 indicates that, as expected, the SAR model does not present a systematic positive bias in the estimation of ρ. This is due to the fact that, unlike in case 1, now the trend is well specified because the SAR model includes a linear trend. The estimates provided by the PS-SAR model are similar to the estimates obtained in case of a non-linear trend, which indicates that the PS-SAR estimates of local spatial dependence are the proper estimates, irrespective of the form of the true trend.

The performance of the three competing models when estimating the global trend clearly depends on whether such a trend is linear or non-linear. In case 1, Table 2 and Figure 2 show that the MSE of the PS estimates increases as σ increases; and for values of above 0.5, the PS model can be considered useless for trend estimation (PS estimates are not included in Figure 2 for the sake of a clear visualization). By contrast, the PS-SAR estimates (which also lead to an increasing MSE as the spatial dependence parameter increases) are certainly much more accurate than those obtained by PS. The only exception is when $\rho=0$ because in such a case the DGP is precisely a PS. As can be observed in Table 2, the gap between PS and PS-SAR estimates becomes considerably larger as the true value of ρ increases. As for SAR trend estimates, their MSE always exceeds the MSE obtained with a PS-SAR estimation; however, SAR-MSE only exceeds PS-MSE in the case of values of ρ below 0.5. The reason is that the true trend is non-linear; thus, the PS and PS-SAR models provide better trend estimates than the SAR model. In the case of a high value of ρ, the worst trend estimates are provided by PS. As mentioned previously, when spatial autocorrelation is high, this model is not able to factorize the spatial variation into a global trend and local spatial dependencies; instead, PS trend estimation is a mixture of both large and small scales. In effect, in Figure 4 it can be observed that the PS model overestimates the trend when the parameter of spatial dependence takes relatively high values. This result is quite logical because, as in section 3, the systematic variation in this model is given by $\mathbf{A}^{-1}f(\mathbf{x}_1,\mathbf{x}_2)$ and the PS model is unable to factorize both terms.

Figure 4

True and estimated trends. Case 1: Non-linear trend.
Black: true trend; Grey: estimation.



As expected, the performance of the SAR model improves significantly in the case of a linear global trend (case 2). Figure 2 reveals that the SAR estimation of the trend is substantially better in case of a global linear trend than when the global trend is non-linear. However, the estimates provided by the PS model are as poor as in case 1, especially for high values of ρ. However, the really important finding is that the PS-SAR trend estimates are similar to the SAR estimates even when the DGP coincides with the SAR model. In other words, even in the extreme case of a linear trend, the PS-SAR model remains a good strategy to accurately estimate the global spatial trend.

Finally, regarding the estimation of the observed values in the case of a non-linear global trend, Figure 3 and Table 2, PS and PS-SAR provide similar estimates and are more accurate than SAR estimates. This is a logical result, since both PS and PS-SAR models approximate the DGP better than the SAR specification in the case of a non-linear trend. However, it is worth highlighting that, while the PS estimates are in reality the result of aggregating both the true values of the trend and the local spatial autocorrelation component, the PS-SAR is able to factorize both sources of dependence correctly. In summary, in the case of a non-linear trend, although the PS and PS-SAR estimates of the observed values are similar, only the PS-SAR model provides an adequate decomposition of the global and local variation.

In the case of a linear global trend, the performance of the SAR model improves significantly and there appears to be no significant difference in the estimates provided by the three competing models (Figure 3). The only exception takes place in case of high values of ρ. As can be seen in Figure 3 and Table 4, for ρ=0.75 the PS estimates are significantly worse than the SAR and PS-SAR estimates.

As for the variance of the error term, both the PS model and the PS-SAR strategy estimate the σ parameter properly (albeit with a downward bias), see Tables 1 and 3. In contrast, the SAR model over-estimates σ. This is not an unexpected result, as the SAR model does not properly capture the simulated trend and assigns part of it to the error term of the model.

In summary, from the simulation results it could be concluded that the PS-SAR model is a robust alternative for specifying the global spatial trend. The performance of the model is certainly very acceptable both in case of a simple linear trend and when the true global trend is complex and non-linear. However, the goodness-of-fit of the other two alternative pure models –PS and SAR– clearly depends on the DGP, performing well when the DGP accommodates the model, but very poorly otherwise. The good performance of the PS-SAR model is corroborated by the fact that, irrespective of the DGP, this mixed strategy is always able to correctly decompose the spatial dependencies into a global trend component and a local spatial autocorrelation component.

The great advantage of the PS-SAR model for trend estimation diminishes considerably when it comes to estimating the observed values. But, in any case, the robustness of the PS-SAR strategy is noteworthy, as it always provides estimates of similar quality to that of the best of the two alternative specifications in each trend context.

5. Empirical Case

This section illustrates the performance of the PS-SAR model empirically using the classical Harrison and Rubinfeld (1978) hedonic pricing data for Boston SMSA augmented and corrected by Pace and Gilley (1997). The augmented and corrected database can be found in the spdep library. It includes 506 observations (one observation per census tract) of the dependent variable (median value of owner-occupied homes in 1,000s of dollars) and 16 independent variables: levels of nitrogen oxides, particulate concentrations, average number of rooms, proportion of structures built before 1940, black population proportion, lower status population proportion, crime rate, proportion of area zoned with large lots, proportion of nonretail business areas, property tax rate, pupil-teacher ratio, location contiguous to the Charles River, weighted distances to the employment centers, index of accessibility, and latitude and longitude of the census tract where the house is located.

This database has been widely used to estimate the willingness to pay for clean air, to examine robust estimation, normality of residuals, nonparametric and semi-parametric estimation and other interesting methodological issues. We use it to compare the performance, in terms of MSE, of 8 competing models: 1) The regression model including the 14 original independent variables considered in Harrison and Rubinfeld (1978) (OLS-NST); 2) the regression model with a linear spatial trend (OLS-LST), which includes latitude and longitude among the independent set of variables; 3) the regression model with a quadratic trend (OLS-QST), which was estimated in Pace and Gilley (1997); 4) the SAR model without a trend (SAR-NST); 5) the SAR model with a linear trend (SAR-LST); 6) the SAR model with quadratic trend (SAR-QST); 7) the PS model; and 8) the PS-SAR model.

Table 5 reports the MSE resulting from the estimation of the competing models. Figure 5 depicts the spatial trends estimated by the SAR-LST, SAR-QST and PS-SAR models.

Table 5

MSE in the estimation of the response variable.

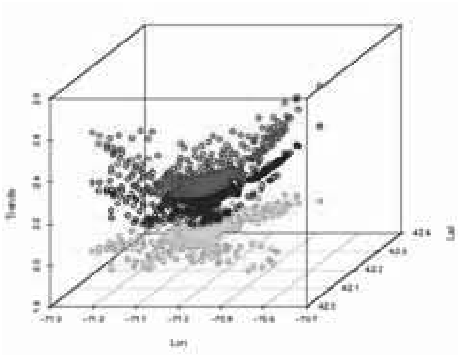
Model	*MSE*	$\hat{\rho}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
OLS-NST	0.03420			
OLS-LST	0.03384			
OLS-QST	0.03290			
SAR-NST	0.02228	0.479		
SAR-LST	0.02207	0.486		
SAR-QST	0.02160	0.480		
PS-SAR	0.02143	0.474	147,355.6	18,802.9
PS	0.03252		14,172.8	95,544,573.0

MSE is computed as in equation [25].

Figure 5

Estimated trends.

Black: PS-SAR estimation; Dark grey: SAR-QST estimation; Light grey: SAR-LST estimation.



The following general conclusions may be drawn from Table 5:

1. The spatial models (SAR and PS-SAR) fit the data better than the non-spatial strategies. This result confirms the finding of Pace and Gilley (1997).

2. The models including a spatial trend (OLS-LST, OLS-QST, SAR-LST, SAR-QST and PS-SAR) provide better estimates than their counterparts of the same name but without a spatial trend. More specifically, OLS-QST and SAR-QST are the OLS and SAR specifications that provide the best fit.

3. The models including a non-parametric spatial trend are preferred to the best models with a parametric trend. In other words, PS is preferred to OLS-QST and PS-SAR is preferred to SAR-QST. The last finding is not an unsurprising result: Figure 5 shows that the spatial trend estimated with PS-SAR lies between the trends estimated with SAR-LST and SAR-QST, but is much closer to the latter.

4. As a result of the foregoing comparisons, PS-SAR is the specification that provides the lowest MSE.

5. Although the MSE obtained with SAR-QST is close to that estimated with PS-SAR, the great advantage of PS-SAR is that it does not need to impose a priori any specific form of spatial trend.

Thus, the PS-SAR model is applied with great success to the augmented and corrected Harrison and Rubinfeld (1978) hedonic pricing data. This specification is able to reduce the MSE by: 35% compared to the best OLS model (OLS-SQT), 4% compared to the pure SAR, 3% compared to SAR-LST and roughly 1% compared to SAR-QST. Finally, the value of ρ in PS-SAR (0.474) is similar to that estimated in the different SAR specifications (0.479 in pure SAR, 0.486 in SAR-LST and 0.480 in SAR-QST). This indicates a certain robustness when estimating short range correlation, whichever the spatial model.

6. Conclusion

SAR models have become very popular specifications for expressing the notion that the value of a variable at a given location is related to the values of the same variable measured at nearby locations, reflecting some kind of interaction. The way SAR models account for large-scale spatial variations is by including the coordinates in the set of covariates, allowing for simplistic polynomial structures or, when the form of the spatial trend is known, more complex forms.

However, simplistic trends are not realistic specifications for the trends of most of the real phenomena that evolve in space. Complex non-linear trends with unknown structure are present in many examples involving spatially dependent data. This is the main reason why we propose the PS-SAR specification, which rather than requiring an a priori specification of the trend model form, lets the data suggest such a form. The proposed PS-SAR model results in a highly flexible technique, penalizes overparameterization and estimates the penalty degree from the data. Our mixed strategy makes it possible to factorize the systematic spatial variation into a global spatial trend component and a local spatial autocorrelation term.

The performance of PS-SAR models when it comes to estimating the global spatial trend, observed values and local correlation has been compared to the performance of PS and pure SAR models. For this purpose, we have simulated 2000 datasets generated by a PS-SAR model with two extreme types of spatial trends: (i) non-linear and non-separable and (ii) linear.

It can be concluded from the simulation results that the PS-SAR model is a robust alternative for specifying the global spatial trend. It performs very well both in the case of a simple linear trend and when the true global trend is complex and non-linear. However, the goodness-of-fit of the other two alternative pure models –PS and SAR– clearly depends on the DGP, performing remarkably well when the DGP accommodates the model and very poorly otherwise. As a consequence, if one of the objectives of the analysis is to decompose the systematic part of the model into a global spatial trend and a component reflecting the existing local spatial autocorrelation, the PS-SAR specification can be considered a good choice.

However, the great advantage of the PS-SAR over the PS approach when estimating the global trend diminishes considerably when it comes to estimating the observed values. In any case, the PS-SAR specification always provides estimates of similar quality to those of the best alternative model, regardless of the form of the trend.

The advantage of the PS-SAR over OLS models and SAR specifications has also been confirmed in an empirical case study. We used the classical Harrison and Rubinfeld (1978) hedonic pricing data for Boston SMSA augmented and corrected by Pace and Gilley (1997). The results obtained show once again that PS-SAR, which does not need an a priori specification of the spatial trend, is able to obtain a better fit to the data than the pure SAR, SAR-LST and SAR-QST models.

There are several possible avenues for future research. Some include empirical applications of PS-SAR, extending the model to non-Gaussian responses, studying of PS-SAR consistency against different misspecifications of the $\mathbf{W}_n$ matrix, comparing between the PS-SAR and geostatistical alternatives and incorporating the temporal dimension into the model.

Acknowledgments: The authors would like to thank Dae-Jin Lee for kindly donating the transfer of R code. We would also like to thank the anonymous reviewers for their useful and valuable comments, which have undoubtely improved the original version of the manuscript. Any remaining errors are our responsability. The work by Jose-María Montero has been partially funded by the Junta de Comunidades de Castilla-La Mancha, under ERDF research project POII10-0250-6975. The work by Maria Durbán and Román Mínguez was supported by the Spanish Ministry of Science and Innovation (project MTM 2008-02901, for M. Durbán, and project MTM 2011-28285-C02-02 in both cases).

References

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

ANSELIN, L. (2007). «Palgrave Handbook of Econometrics», vol. 1, chap. *Spatial Econometrics. Palgrave MacMillan, Great Britain*, 901-969.

BIVAND, R., WITH CONTRIBUTIONS BY MLCAH ALTMAN, ANSELIN, L., ASSUÇÁO, R., BERKE, O., BERNAT, A., BLANCHET, G., BLANKMEYER, E., CARVALHO, M., CHRISTENSEN, B., CHUN, Y., DORMANN, C., DRAY, S., HALBERSMA, R., KRAINSKI, E., LEGENDRE, P., LEWIN-KOH, N., LI, H., MA, J., MILLO, G., MUELLER, W., ONO, H., PERES-NETO, P., PIRAS, G., REDER, M., TIEFELSDORF, M., and YU, D. (2011). «spdep: Spatial dependence: weighting schemes, statistics and models». R package version 0.5-37, URL «http://cran.r-project.org/package=spdep».

BRUMBACK, B. and RICE, J. (1998). «Smoothing spline models for the analysis of nested and crossed samples of curves». *Journal of the American Statistical Association,* 93, 961-994.

CRUJEIRAS, R. and KEILEGOM, I. V. (2010). «Least squares estimation of nonlinear spatial trends». *Computational Statistics and Data Analysis,* 54, 452-465.

CURRIE, I., DURBÁN, M. and EILERS, P. (2004). «Smoothing and forecasting mortality rates». *Statistical Modelling,* 4, 279-298.

CURRIE, I., DURBAN, M. and EILERS, P. (2006). «Generalized linear array models with applications to multidimensional smoothing». *Journal of the Royal Statististical Society, Series B,* 68, 1-22.

CURRIE, I. D. and DURBÁN, M. (2002). «Flexible smoothing with P-splines: A unified approach». *Statistical Modelling,* 2, 333-349.

EILERS, P., CURRIE, I. and DURBÁN, M. (2006). «Fast and compact smoothing on large multidimensional grids». *Computational Statistics and Data Analysis* 50, 61-76.

EILERS, P. and MARX, B. (1996). «Flexible smoothing with *B*-splines and penalties». *Statistical Science,* 11, 89-121.

GREEN, P. and YANDELL, B. (1985). «Semi-parametric generalized linear models». In *Generalized Linear Models,* vol. 32 of *Lecture Notes in Statistcs.* Springer, Berlin, 44-55.

HAAS, T. (1996). «Multivariate spatial prediction in the presence of non-linear trend and covariance non-stationarity». *Environmetrics*, 7, 145-165.

HARRISON, D. and RUBINFELD, D. (1978). «Hedonic Housing Prices and the Demand for Clean Air». *Journal of Environmental Economics and Management,* 5, 81-102.

KAMMANN, E. and WAND, M. (2003). «Geoadditive models». *Journal of the Royal Statistical Society C. Applied Statistics,* 52, 1-18.

LANG, S. and BREZGER, A. (2004). «Bayesian p-splines». *Journal of Computational and Graphical Statistics,* 13, 183-212.

LEE, D.-J. (2010). «Smoothing mixed models for spatial and spatio-temporal data». *Ph.D. thesis, Universidad Carlos III, Leganés.*

LEE, D.-J. and DURBÁN, M. (2009). «Smooth-CAR mixed models for spatial count data». *Computational Statistics and Data Analysis,* 53, 2968-2977.

LESAGE, J. and PACE, R. K. (2009). «Introduction to Spatial Econometrics». *Chapman-Hall, Boca Raton (USA).*

LIGGES, U. and MACHLER, M. (2003). «Scatterplot3d - an r package for visualizing multivariate data». *Journal of Statistical Software,* 8 1-20. URL «http://www.j statsoft.org».

PACE, R. and GILLEY, O. (1997). «Using the Spatial Configuration of the Data to Improve Estimation». *Journal of Real Estate Finance and Economics,* 14, 333-340.

PATTERSON, H. and THOMPSON, R. (1971). «Recovery of inter-block information when block sizes are unequal». *Biometrika,* 58, 545-554.

R DEVELOPMENT CORE TEAM (2011). *R:* «A Language and Environment for Statistical Computing». *R Foundation for Statistical Computing, Vienna, Austria.* URL «http: //www. R-project.org/».

RUPPERT, D., WAND, M. and CARROLL, R. (2003). «Semiparametric Regression». *Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.*

SNEPVANGERS, J., HEUVELINK, G. and HUISMAN, J. (2003). «Soil water content interpolation using spatio-temporal kriging with external drift». *Geoderma,* 112, 253-271.

SPEED, T. (1991). «Comment on "BLUP is a good thing: The estimation of random effects"», by Robinson, G.K. *Statistical Science,* 6, 15-51.

VERBYLA, A., CULLIS, B., KENWARD, M. and WELHAM, S. (1999). «The analysis of designed experiments and longitudinal data using smoothing splines». *Applied Statistics,* 48, 269-312.

VERKATRAM, A. (1988). «On the use of kriging in the spatial analysis of acid precipitation data». *Atmospheric Environment,* 22, 1963-1975.

WAHBA, G. (1990). «Spline Models for Observational Data». *Society for Industrial and Applied Mathematics, Philadelphia.*

WAND, M. (2003). «Smoothing and mixed models». *Computational Statistics,* 18, 223-249.

WOOD, S. (2006a). *mgcv 1.3. R package.* URL «http: //cran.r-project.org/package=mgcv».

WOOD, S. N. (2006b). «Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models». *Biometrics,* 62, 1025-1036.

Appendix of tables

Table1

Summary of σ and ρ estimates. Case 1: Non-linear trend.

True param.		measure	σ estim.			ρ estim.	
σ	ρ		PS	PS-SAR	SAR	PS-SAR	SAR
0.5	0	average	0.492	0.484	0.604	-0.096	0.784
		std. dev.	0.027	0.030	0.031	0.150	0.022
0.5	0.25	average	0.481	0.488	0.587	0.173	0.860
		std. dev.	0.027	0.028	0.029	0.123	0.015
0.5	0.5	average	0.482	0.490	0.586	0.436	0.921
		std. dev.	0.030	0.028	0.028	0.093	0.009
0.5	0.75	average	0.542	0.492	0.600	0.721	0.969
		std. dev.	0.036	0.026	0.027	0.046	0.005
1	0	average	0.993	0.976	1.119	-0.119	0.546
		std. dev.	0.054	0.060	0.057	0.164	0.053
1	0.25	average	0.971	0.979	1.087	0.143	0.695
		std. dev.	0.052	0.055	0.054	0.140	0.035
1	0.5	average	0.968	0.987	1.071	0.399	0.820
		std. dev.	0.057	0.054	0.051	0.113	0.023
1	0.75	average	0.980	0.987	1.057	0.687	0.925
		std. dev.	0.066	0.054	0.053	0.071	0.012

DGP: $\mathbf{y} = f(\mathbf{x}_1, \mathbf{x}_2) + \rho \mathbf{W}_n \mathbf{y} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ where $\mathbf{x}_1$, $\mathbf{x}_2$ have been generated from a $U(0,1)$ and $f(\mathbf{x}_1, \mathbf{x}_2)$ represents the spatial non-linear trend given in [22]. 250 simulations have been generated for each combination of σ and ρ values.

Table 2

Summary of MSE when estimating the trend and the response variable. Case 1: Non-linear trend.

True param.		measure	MSE (Trend)			MSE (Response Variable)		
σ	ρ		PS	PS-SAR	SAR	PS	PS-SAR	SAR
0.5	0	average	0.026	0.276	5.233	0.214	0.205	0.366
		std. dev.	0.007	0.347	0.264	0.024	0.027	0.038
0.5	0.25	average	0.921	0.326	5.587	0.199	0.208	0.346
		std. dev.	0.089	0.496	0.245	0.023	0.025	0.035
0.5	0.5	average	8.073	0.441	5.964	0.192	0.210	0.345
		std. dev.	0.402	0.916	0.230	0.026	0.026	0.033
0.5	0.75	average	72.112	0.404	6.441	0.222	0.212	0.361
		std. dev.	2.762	0.744	0.256	0.034	0.023	0.033
1	0	average	0.087	0.429	2.918	0.904	0.867	1.255
		std. dev.	0.028	0.458	0.430	0.104	0.115	0.128
1	0.25	average	0.991	0.521	3.328	0.846	0.872	1.185
		std. dev.	0.187	0.663	0.418	0.095	0.105	0.117
1	0.5	average	8.297	0.848	3.756	0.813	0.885	1.149
		std. dev.	0.789	1.144	0.464	0.103	0.104	0.109
1	0.75	average	72.631	1.216	4.390	0.765	0.885	1.120
		std. dev.	5.282	1.832	0.515	0.117	0.103	0.113

DGP: $\mathbf{y} = f(\mathbf{x}_1, \mathbf{x}_2) + \rho \mathbf{W}_n \mathbf{y} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N\left(\mathbf{0}, \sigma^2 \mathbf{I}_n\right)$ where $\mathbf{x}_1$, $\mathbf{x}_2$ have been generated from a $U(0,1)$ and $f(\mathbf{x}_1, \mathbf{x}_2)$ represents the spatial non-linear trend given in [22]. 250 simulations have been generated for each combination of σ and ρ values. Then the MSE has been computed from the estimation of the trend and the response variable.

Table 3

Summary of σ and ρ estimates. Case 2: Linear trend.

True param.		measure	σ estim.			ρ estim.	
σ	ρ		PS	PS-SAR	SAR	PS-SAR	SAR
0.5	0	average	0.499	0.497	0.495	-0.071	-0.042
		std. dev.	0.025	0.026	0.025	0.135	0.125
0.5	0.25	average	0.500	0.496	0.494	0.155	0.191
		std. dev.	0.027	0.026	0.026	0.108	0.101
0.5	0.5	average	0.530	0.501	0.497	0.420	0.447
		std. dev.	0.033	0.026	0.026	0.093	0.088
0.5	0.75	average	0.593	0.504	0.500	0.690	0.715
		std. dev.	0.080	0.024	0.024	0.067	0.060
1	0	average	0.997	0.992	0.988	-0.086	-0.054
		std. dev.	0.053	0.053	0.053	0.124	0.120
1	0.25	average	1.008	0.999	0.995	0.156	0.194
		std. dev.	0.055	0.053	0.052	0.114	0.106
1	0.5	average	1.046	0.994	0.989	0.416	0.451
		std. dev.	0.068	0.053	0.053	0.091	0.086
1	0.75	average	1.178	1.009	1.001	0.681	0.710
		std. dev.	0.146	0.057	0.056	0.069	0.059

DGP: $\mathbf{y} = f(\mathbf{x}_1, \mathbf{x}_2) + \rho \mathbf{W}_n \mathbf{y} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ where $\mathbf{x}_1$, $\mathbf{x}_2$ have been generated from a $U(0,1)$ and $f(\mathbf{x}_1, \mathbf{x}_2)$ represents the spatial non-linear trend given in [23]. 250 simulations have been generated for each combination of σ and ρ values.

Table 4

Summary of MSE when estimating the trend and the response variable.
Case 2: Linear trend.

True param.		measure	MSE (Trend)			MSE (Response Variable)		
σ	ρ		PS	PS-SAR	SAR	PS	PS-SAR	SAR
0.5	0	average	0.005	0.127	0.094	0.244	0.242	0.245
		std. dev.	0.004	0.181	0.121	0.025	0.026	0.025
0.5	0.25	average	0.604	0.203	0.134	0.245	0.241	0.245
		std. dev.	0.073	0.278	0.190	0.026	0.026	0.025
0.5	0.5	average	5.330	0.328	0.229	0.274	0.246	0.248
		std. dev.	0.322	0.474	0.363	0.035	0.026	0.026
0.5	0.75	average	47.762	0.692	0.415	0.335	0.249	0.251
		std. dev.	2.078	1.264	0.870	0.099	0.024	0.024
1	0	average	0.024	0.146	0.108	0.974	0.964	0.980
		std. dev.	0.015	0.217	0.184	0.103	0.103	0.105
1	0.25	average	0.649	0.238	0.149	0.993	0.978	0.994
		std. dev.	0.146	0.338	0.224	0.110	0.105	0.104
1	0.5	average	5.363	0.355	0.220	1.060	0.968	0.980
		std. dev.	0.660	0.524	0.391	0.147	0.105	0.104
1	0.75	average	47.589	0.834	0.449	1.315	0.996	1.005
		std. dev.	4.492	1.479	0.733	0.363	0.113	0.114

DGP: $\mathbf{y} = f(\mathbf{x}_1, \mathbf{x}_2) + \rho \mathbf{W}_n \mathbf{y} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ where $\mathbf{x}_1$, $\mathbf{x}_2$ have been generated from a $U(0,1)$ and $f(\mathbf{x}_1, \mathbf{x}_2)$ represents the spatial non-linear trend given in [23]. 250 simulations have been generated for each combination of σ and ρ values. Then the MSE has been computed from the estimation of the trend and the response variable.