REVSTAT – Statistical Journal
Volume 10, Number 3, November 2012, 285–308

SMALL AREA ESTIMATION USING A SPATIO-TEMPORAL LINEAR MIXED MODEL

Authors: Luis N. Pereira
– Escola Superior de Gestão, Hotelaria e Turismo,
Centro de Investigação sobre o Espaço e as Organizações,
Universidade do Algarve, Portugal
lmper@ualg.pt

Pedro S. Coelho
– ISEGI – Universidade Nova de Lisboa, Portugal
Faculty of Economics, Ljubljana University, Slovenia
psc@isegi.unl.pt

Received: February 2012 Accepted: April 2012

Abstract:

- In this paper it is proposed a spatio-temporal area level linear mixed model involving spatially correlated and temporally autocorrelated random effects. An empirical best linear unbiased predictor (EBLUP) for small area parameters has been obtained under the proposed model. Using previous research in this area, analytical and bootstrap estimators of the mean squared prediction error (MSPE) of the EBLUP have also been worked out. An extensive simulation study using time-series and cross-sectional data was undertaken to compare the efficiency of the proposed EBLUP estimator over other well-known EBLUP estimators and to study the properties of the proposed estimators of MSPE.

Key-Words:

- *empirical best linear unbiased prediction; linear mixed model; mean squared prediction error estimation; small area estimation; spatial correlation; temporal autocorrelation.*

AMS Subject Classification:

- 62J05, 62F40, 62D05.

1. INTRODUCTION

Large scale repeated sample surveys are usually designed to produce reliable estimates of several characteristics of interest for large subgroups of a population, from which samples are drawn. A subgroup may be a geographical region or a group obtained by cross-classification of demographic factors such as age or gender. However, for effective planning in a wide variety of fields, there is a growing demand to produce similar estimates for smaller subgroups for which adequate sample sizes are not available. In fact, sample sizes are frequently very small or even zero in many subgroups of interest (small areas), resulting in unreliable direct design-based small area estimates. This creates a need to employ indirect estimators that "borrow information" from related small areas and time periods through linking models using recent census or current administrative data, in order to increase the effective sample size and thus precision. Such indirect estimators are often based on explicit Linear Mixed Models (LMM) that provide a link to a related small area through the use of supplementary data. The empirical best linear unbiased prediction (EBLUP) approach is one of the most popular methods for the estimation of small area parameters of interest. This approach uses an appropriate LMM which captures several salient features of the areas and combines information from censuses or administrative records conjointly with the survey data. When time-series and cross-sectional data are available, longitudinal LMM might be useful to take simultaneously advantage of both the possible spatial similarities among small areas and the expected time-series relationships of the data in order to improve the efficiency of the small area estimators. Although there is some research on temporal (e.g. Rao & Yu, 1994; Datta *et al.*, 2002; Saei & Chambers, 2003; Pereira & Coelho, 2010) and on spatial small area estimation using LMM (e.g. Salvati, 2004; Petrucci *et al.*, 2005; Petrucci & Salvati, 2006; Chandra *et al.*, 2007; Pratesi & Salvati, 2008), there is a need of research in the field of small area estimation using LMM with spatio-temporal information. Such kind of estimation might account simultaneously for the spatial dependence and the chronological autocorrelation in order to strengthen the small area estimates. This can be achieved by incorporating in the model both area specific random effects and area-by-time specific random effects. The area specific random effects could then be linked by a spatial process, while the area-by-time random effects could be linked by a temporal process. This approach is definitely more complex than a simple regression method and its success depends on the ability to define a suitable spatial neighbourhood, to specify properly spatial and temporal processes and to estimate additional parameters.

Thus, the main goal of this paper is to propose a simple and intuitive spatio-temporal LMM involving spatially correlated and temporally autocorrelated random area effects, using both time-series and cross-sectional data. The proposed model is an extension of three well-known small area models in the literature.

Under the proposed model, two research questions are addressed. Firstly, we analyse the extent to which the spatial and the temporal relationships in the data justify the introduction of a spatial and a temporal autoregressive parameters in the model. This is carried out via a simulation study which compares the efficiency of the proposed EBLUP estimator against other well-known EBLUP estimators, by taking into account the joint effects of the following components: the sampling variances of the direct estimators of the small area parameters; the variance components of the random effects; and the spatial and temporal autocorrelation parameters. Secondly, we discuss how to measure the uncertainty of the proposed EBLUP. This is carried out via a simulation study which compares the accuracy of the analytical and the bootstrap estimators of the mean squared prediction error (MSPE) introduced in this paper.

Singh *et al.* (2005) proposed the only existing work on small area estimation using spatial-temporal approaches. They proposed a spatio-temporal state space model via Kalman filtering estimation which, like our model, borrows strength from past surveys, neighbour small areas and a set of covariates. However, our model, unlike the model due to Singh *et al.* (2005), makes use of a different estimation method and incorporates independent specific random effects. Note that the model due to Singh *et al.* (2005) considers an interaction between the spatial dependence and the temporal autocorrelation, since the spatial process is stated in the design matrix of random effects.

This paper is organized as follows. In Section 2 it is proposed a spatio-temporal area level LMM. The BLUP and EBLUP of the mixed effects are also provided in this section. Section 3 discusses the measure of uncertainty of the proposed EBLUP. In this section it is proposed both an analytical and a parametric bootstrap method to estimate the MSPE of the EBLUP. The design of the simulation study, as well as its empirical results on the efficiency of the proposed EBLUP and on the properties of the proposed estimators of MSPE, is reported in Section 4. Finally, the paper ends with a conclusion in Section 5, which summarizes the main advantages of the proposed methodology and identifies further issues of research.

2. THE SPATIO-TEMPORAL MODEL

2.1. Proposed model

Let $\boldsymbol{\theta} = \mathrm{col}_{1\leq i\leq m}(\boldsymbol{\theta}_i)$ be a $mT\times 1$ vector of the parameters of inferential interest and assume that $\mathbf{y} = \mathrm{col}_{1\leq i\leq m}(\mathbf{y}_i)$ is its design-unbiased direct survey estimator. Here $\boldsymbol{\theta}_i = \mathrm{col}_{1\leq t\leq T}(\theta_{it})$, $\mathbf{y}_i = \mathrm{col}_{1\leq t\leq T}(y_{it})$ and y_{it} is the direct survey estimator of the parameter of interest for small area i at time t, $\theta_{it}(i = 1, ..., m;$ $t = 1, ..., T)$. Thus the sampling error model is given by:

$$\mathbf{y} = \boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{2.1}$$

where $\boldsymbol{\epsilon} = \mathrm{col}_{1\leq i\leq m;1\leq t\leq T}(\varepsilon_{it})$ is a $mT\times 1$ vector of sampling errors. We assume that $\boldsymbol{\epsilon} \overset{iid}{\sim} N(\mathbf{0}; \mathbf{R})$, where $\mathbf{R} = \mathrm{diag}_{1\leq i\leq m;1\leq t\leq T}(\sigma^2_{it})$ is a $mT\times mT$ matrix with known sampling variances of the direct estimators. We propose the following linking model in which the parameters of inferential interest are related to area-by-time specific auxiliary data through a linear model with random effects:

$$\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{v} + \mathbf{u}_2 \tag{2.2}$$

where $\mathbf{X} = \mathrm{col}_{1\leq i\leq m;1\leq t\leq T}(\mathbf{x}'_{it})$ is a $mT\times p$ design matrix of area-by-time specific auxiliary variables with rows given by $\mathbf{x}'_{it} = (x_{it1}, ..., x_{itp})$ $(1\times p)$, $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)'$ $(p\times 1)$ is a column vector of regression parameters, $\mathbf{v} = \mathrm{col}_{1\leq i\leq m}(\nu_i)$ is a $m\times 1$ vector of random area specific effects and $\mathbf{u}_2 = \mathrm{col}_{1\leq i\leq m;1\leq t\leq T}(u_{2,it})$ is a $mT\times 1$ vector of random area-by-time specific effects. Further, $\mathbf{Z}_1 = \mathbf{I}_m \otimes \mathbf{1}_T$ $(mT\times m)$ where $\mathbf{I}_m$ is an identity matrix of order m and $\mathbf{1}_T$ $(T\times 1)$ is a column vector of ones. We assume that $\mathbf{X}$ has full column rank p and $\mathbf{v}$ is the vector of the second order variation.

In order to take into account for the spatial dependence among small areas we propose the use of a simple spatial model in the random area specific effects. In particular, we propose the use of the simultaneous autoregressive (SAR) process (Anselin, 1992), where the vector $\mathbf{v}$ satisfies:

$$\mathbf{v} = \phi\,\mathbf{W}\mathbf{v} + \mathbf{u}_1 \;\;\Rightarrow\;\; \mathbf{v} = (\mathbf{I}_m - \phi\,\mathbf{W})^{-1}\mathbf{u}_1\ , \tag{2.3}$$

where ϕ is a spatial autoregressive coefficient which defines the strength of the spatial relationship among the random effects associated with neighboring areas and $\mathbf{W} = \{w_{ij}\}$ $(m\times m)$ is a known spatial proximity matrix which indicates whether the small areas are neighbors or not $(i, j = 1, ..., m)$. A simple common way to specify $\mathbf{W}$ is to define $w_{ij} = 1$ if the area i is physically contiguous to area j and $w_{ij} = 0$ otherwise. In this case $\mathbf{W}$ is a contiguity matrix. The most common way to define $\mathbf{W}$ is in the row standardized form, that is, restricting rows to satisfy $\sum_{j=1}^m w_{ij} = 1$, for $i = 1, ..., m$. It is yet possible to create more elaborate weights as functions of the length of common boundary between the small areas (Wall, 2004). Further, $\mathbf{u}_1 = \mathrm{col}_{1\leq i\leq m}(u_{1i})$ is a $m\times 1$ vector of independent error terms satisfying $\mathbf{u}_1 \overset{iid}{\sim} N(\mathbf{0}; \sigma^2_u\mathbf{I}_m)$.

In order to borrow strength across time we propose the use of autocorrelated random effects. In particular, we propose that $u_{2,it}$'s follow a common first-order autoregressive [AR(1)] process for each small area:

$$u_{2,it} = \rho\, u_{2,i,t-1} + \xi_{it}\ , \qquad |\rho| < 1\ , \tag{2.4}$$

where ξ_{it}'s are the error terms satisfying $\xi_{it} \overset{iid}{\sim} N(0;\sigma^2)$ and ρ is a temporal autoregressive coefficient which measures the level of chronological autocorrelation.

Combining models (2.1)–(2.4), the proposed model involving spatially correlated and temporally autocorrelated random area effects may be written in matrix form as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\upsilon} + \boldsymbol{\epsilon}\ , \tag{2.5}$$

where $\mathbf{Z} = [\mathbf{Z}_1\ \mathbf{I}_{mT}]$, $\mathbf{Z}_1 = \mathbf{I}_m \otimes \mathbf{1}_T$ and $\boldsymbol{\upsilon} = [\mathbf{v}'\ \mathbf{u}_2']'$. Further, we assume that error terms $\mathbf{v} = (\mathbf{I}_m - \phi\,\mathbf{W})^{-1}\mathbf{u}_1$, $\mathbf{u}_2$ and $\boldsymbol{\epsilon}$ are mutually independent distributed with $\mathbf{u}_1 \overset{iid}{\sim} N(\mathbf{0};\sigma_u^2\mathbf{I}_m)$, $\mathbf{u}_2 \overset{iid}{\sim} N(\mathbf{0};\sigma^2\mathbf{I}_m \otimes \boldsymbol{\Gamma})$ and $\boldsymbol{\epsilon} \overset{iid}{\sim} N(\mathbf{0};\mathbf{R})$, where $\boldsymbol{\Gamma}(T\times T)$ is a matrix with elements $\rho^{|r-s|}/(1-\rho^2)$, $r,s = 1,...,T$ and $\mathbf{R} = \mathrm{diag}_{1\le i\le m;1\le t\le T}(\sigma_{it}^2)$. We can now see that model (2.5) is a special case of the general LMM with a block diagonal covariance matrix of $\boldsymbol{\upsilon}$, given by $\mathbf{G} = \mathrm{diag}_{1\le k\le 2}(\mathbf{G}_k)$, where $\mathbf{G}_1$ and $\mathbf{G}_2$ are the covariance matrices of $\mathbf{v}$ and $\mathbf{u}_2$, respectively. As showed by Salvati (2004) and by Rao & Yu (1994), these covariance structures are given by $\mathbf{G}_1 = E(\mathbf{v}\mathbf{v}') = \sigma_u^2\big[(\mathbf{I}_m - \phi\mathbf{W})'(\mathbf{I}_m - \phi\mathbf{W})\big]^{-1}$ and $\mathbf{G}_2 = E(\mathbf{u}_2\mathbf{u}_2') = \sigma^2\mathbf{I}_m \otimes \boldsymbol{\Gamma}$, respectively. It follows that the covariance matrix of $\mathbf{y}$ is:

$$\mathbf{V} \;=\; \mathrm{diag}_{1\le i\le m;1\le t\le T}(\sigma_{it}^2) + \mathbf{Z}_1\sigma_u^2\mathbf{B}^{-1}\mathbf{Z}_1' + \sigma^2\mathbf{I}_m \otimes \boldsymbol{\Gamma}\ , \tag{2.6}$$

where $\mathbf{B} = (\mathbf{I}_m - \phi\mathbf{W})'(\mathbf{I}_m - \phi\mathbf{W})$. Note that $\mathbf{V}$ is not a block diagonal covariance structure, like in the context of the well known Fay–Herriot and Rao–Yu models. Finally, note that the temporal model due to Rao & Yu (1994) can be obtained from model (2.5) setting $\phi = 0$, as well as the spatial model due to Salvati (2004) can be obtained from model (2.5) setting $T = 1$, $\rho = 0$ and $\sigma^2 = 0$. The model proposed by Fay & Herriot (1979) is also a particular case of model (2.5) since it can be obtained setting $T = 1$, $\phi = 0$, $\rho = 0$ and $\sigma^2 = 0$. However, if the spatial and temporal autocorrelation parameters are not equal to zero then the model proposed by Singh *et al.* (2005) cannot be obtained from model (2.5) because it is assumed in this model that the error terms are mutually independent.

2.2. The BLUP

The current small area parameter, $\theta_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \nu_i + u_{2,it}$, is a special case of the linear combination $\tau = \mathbf{k}_{it}'\boldsymbol{\beta} + \mathbf{m}_{it}'\boldsymbol{\upsilon}$, where $\mathbf{k}_{it}' = \mathbf{x}_{it}'$ and $\mathbf{m}_{it}' = [\mathbf{m}_{1i}'\ \mathbf{m}_{2it}']$ in which $\mathbf{m}_{1i}' = (0,...,0,1,0,...,0)$ is a $1\times m$ vector with value 1 in the i^{th} position and 0 elsewhere, and $\mathbf{m}_{2it}' = (0,...,0,1,0,...,0)$ is a $1\times mT$ vector with value 1 in the $(it)^{\text{th}}$ position and 0 elsewhere. Noting that model (2.5) is a special case of the general LMM, thus the BLUP estimator of $\tau = \theta_{it}$ can be obtained from Henderson's general results (Henderson, 1975). Assuming first that $\boldsymbol{\psi} = (\sigma^2, \sigma_u^2, \phi, \rho)'$

is fully known, the BLUP of θ_{it} is given by:

$$\widetilde{\theta}_{it} = \widetilde{\theta}^{H}_{it}(\boldsymbol{\psi}) = \mathbf{x}'_{it}\widetilde{\boldsymbol{\beta}} + \mathbf{h}'_{it}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\widetilde{\boldsymbol{\beta}}) \,, \tag{2.7}$$

where $\widetilde{\boldsymbol{\beta}} = \widetilde{\boldsymbol{\beta}}(\boldsymbol{\psi}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$ is the best linear unbiased estimator of $\boldsymbol{\beta}$ and $\mathbf{h}'_{it}$ is a $1{\times}mT$ vector which captures the potential spatial and temporal autocorrelation present in the i^{th} small area. Further, $\mathbf{h}'_{it} = \sigma_u^2\boldsymbol{\varsigma}'_i \otimes \mathbf{1}'_T + \sigma^2\boldsymbol{\zeta}'_{it}$ where $\boldsymbol{\varsigma}'_i = \{\varsigma_{ii'}\}$ is the t^{th} row of the $\mathbf{B}^{-1}$ matrix and $\boldsymbol{\zeta}'_{it}$ is a $1{\times}mT$ vector with m T-dimensional blocks, with the t^{th} row of the $\boldsymbol{\Gamma}$ matrix, $\boldsymbol{\gamma}_t$, in the i^{th} block and null vectors, $\mathbf{0}_{1\times T}$, elsewhere, $i, i' = 1, ..., m$; $t = 1, ..., T$. This estimator can be classified as a combined estimator, since it can be decomposed into two components: a synthetic estimator, $\mathbf{x}'_{it}\widetilde{\boldsymbol{\beta}}$, and a correction factor, $\mathbf{h}'_{it}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\widetilde{\boldsymbol{\beta}})$. We can say that the weights in $\mathbf{h}'_{it}\mathbf{V}^{-1}$ allow a correction of the synthetic part of the estimator (2.7) through the regression residuals from the small area that is the target of inference at t^{th} time point and from this area at previous time periods, but also from the regression residuals from other small areas at the t^{th} time point which are spatially correlated with the target small area. From expression (2.7) it is also possible to observe that when a particular small area is not represented in the sample of period t, it is still possible to estimate the correction factor through the spatial and/or temporal autocorrelation, if there are data collected for small area i in at least one of the previous samples and/or for one related small area at time period t. This is certainly a very appealing characteristic of the estimator (2.7): it is possible to avoid the reduction of the proposed estimator to a pure synthetic estimator, even when the sample size of period t in the i^{th} small area is null.

2.3. The two-stage estimator

The BLUP estimator depends on the parameters of the vector $\boldsymbol{\psi} = (\sigma^2, \sigma_u^2, \phi, \rho)'$, but in practice they are unknown and have to be estimated from the data. As far as the estimation of $\boldsymbol{\psi}$ is concerned, a number of methods have been proposed in the literature, such as the maximum likelihood method (Fisher, 1922; Hartley & Rao, 1967), the restricted maximum likelihood method (Thompson, 1962; Patterson & Thompson, 1971) and the analysis of variance (ANOVA) method (Henderson, 1953), among others. While the likelihood-based methods assume the normality of the error terms, the ANOVA method is free of this kind of assumptions. So, in the present work we have decided to estimate the variance components through an extension of Henderson method 3 (Henderson, 1953) to the model (2.5) with spatial correlated errors through a SAR process, ν_i, temporal autocorrelated errors through an AR(1) process, $u_{2,it}$, and independent sampling errors, ε_{it}. Furthermore, we have assumed that the autoregressive coefficients are known, due to difficulties on getting efficient and admissible estimators for ρ, as was reported by Fuller (1987) and by Rao & Yu (1994). Thus, from this point forward we define the vector of variance components as $\boldsymbol{\psi} = (\sigma^2, \sigma_u^2)'$.

We first obtain an unbiased estimator of σ^2. For this purpose, model (2.5) is transformed in order to eliminate the vector of random area specific effects $\mathbf{v}$. First transform $\mathbf{y}_i$ to $\mathbf{z}_i = \mathbf{P}\mathbf{y}_i$ such that the covariance matrix $V(\mathbf{P}\mathbf{u}_{2i}) = \sigma^2 \mathbf{I}_T$. In this Prais–Winsten transformation we use the decomposition $\boldsymbol{\Gamma} = \mathbf{P}^{-1}(\mathbf{P}^{-1})'$, where $\mathbf{P}(T \times T)$ has the following form: $p_{1,1} = (1-\rho^2)^{1/2}$, $p_{t,t'} = 1, \forall\, t = t'$ for $t, t' = 2, ..., T$, $p_{t+1,t} = -\rho$ for $t = 1, ..., T-1$ and remaining $p_{t,t'} = 0$ (Judge *et al.*, 1985). Pre-multiplying model (2.5) by $\mathrm{diag}_{1\le i\le m}(\mathbf{P})$, the transformed model is given by:

$$\mathbf{z}^* = \mathbf{H}^* \boldsymbol{\beta} + \mathrm{diag}_{1\le i\le m}(\mathbf{f})\, \mathbf{v} + \mathrm{col}_{1\le i\le m}(\mathbf{P}\mathbf{u}_{2i}) + \mathrm{col}_{1\le i\le m}(\mathbf{P}\boldsymbol{\epsilon}_i)\ , \tag{2.8}$$

where $\mathbf{z}^* = \mathrm{col}_{1\le i\le m}(\mathbf{z}_i)$, $\mathbf{H}^* = \mathrm{col}_{1\le i\le m}(\mathbf{H}_i)$, $\mathbf{z}_i = \mathbf{P}\mathbf{y}_i$, $\mathbf{H}_i = \mathbf{P}\mathbf{X}_i$ and $\mathbf{f} = \mathbf{P}\mathbf{1}_T = \mathrm{col}_{1\le i\le m}(f_t)$, with $f_1 = (1-\rho^2)^{1/2}$ and $f_t = 1-\rho$ for $2 \le t \le T$. Next we transform $\mathbf{z}_i$ to $\mathbf{z}_i^{(1)} = (\mathbf{I}_T - \mathbf{D})\,\mathbf{z}_i$, where $\mathbf{D} = (\mathbf{f}\mathbf{f}')/c$ is a $T \times T$ matrix with $c = \mathbf{f}'\mathbf{f} = (1-\rho)\big[T - (T-2)\rho\big]$. Pre-multiplying now model (2.8) by $\mathbf{D}^* = \mathrm{diag}_{1\le i\le m}(\mathbf{I}_T - \mathbf{D})$ and noting that $(\mathbf{I}_T - \mathbf{D})\mathbf{f} = \mathbf{0}_{T\times 1}$, then the transformed model reduces to:

$$\mathbf{z}^{(1)} = \mathbf{H}^{(1)} \boldsymbol{\beta} + \mathbf{e}^{(1)}\ , \tag{2.9}$$

where $\mathbf{z}^{(1)} = \mathrm{col}_{1\le i\le m}\big(\mathbf{z}_i^{(1)}\big)$, $\mathbf{H}^{(1)} = \mathrm{col}_{1\le i\le m}\big(\mathbf{H}_i^{(1)}\big)$, $\mathbf{H}_i^{(1)} = (\mathbf{I}_T - \mathbf{D})\mathbf{P}\mathbf{X}_i$ and $\mathbf{e}^{(1)} = \mathrm{col}_{1\le i\le m}\big[(\mathbf{I}_T - \mathbf{D})\,\mathbf{P}(\mathbf{u}_{2i} + \boldsymbol{\epsilon}_i)\big]$. Further, we can see at this moment that $E\big(\mathbf{e}^{(1)}\big) = \mathbf{0}_{mT\times 1}$ and $V\big(\mathbf{e}^{(1)}\big) = \mathrm{diag}_{1\le i\le m}\big[(\mathbf{I}_T - \mathbf{D})\,(\sigma^2 \mathbf{I}_T + \mathbf{P}\mathbf{R}_i\mathbf{P}')\,(\mathbf{I}_T - \mathbf{D})'\big]$ do not involve σ_u^2, thus we can estimate σ^2 through the reduced model (2.9) using the residual sum of squares. Let $\widehat{\mathbf{e}}^{(1)\prime}\,\widehat{\mathbf{e}}^{(1)}$ be the residual sum of squares obtained by regressing $\mathbf{z}^{(1)}$ on $\mathbf{H}^{(1)}$ using ordinary least squares (OLS). An unbiased estimator of σ^2 is given by:

$$\widetilde{\sigma}^2 = \left(\widehat{\mathbf{e}}^{(1)\prime}\widehat{\mathbf{e}}^{(1)} - \mathrm{tr}\left\{\left[\mathbf{D}^* - \mathbf{H}^{(1)}\big(\mathbf{H}^{(1)\prime}\mathbf{H}^{(1)}\big)^{-}\mathbf{H}^{(1)\prime}\right]\mathbf{R}^{(1)}\right\}\right)\left[m(T-1) - r(\mathbf{H}^{(1)})\right]^{-1}, \tag{2.10}$$

where $\mathbf{R}^{(1)} = \mathrm{diag}_{1\le i\le m}(\mathbf{P}\mathbf{R}_i\mathbf{P})$ and $\mathbf{A}^-$ is the Moore–Penrose generalized inverse of $\mathbf{A}$. Although this estimator has been deducted in the context of the spatio-temporal proposed model, it is equal to the estimator proposed by Rao & Yu (1994) in the context of their temporal model. This happens due to the fact that we have transformed model (2.5) in order to eliminate the vector of random area specific effects $\mathbf{v}$, which accounts for the spatial dependence among small areas. Rao & Yu (1994) showed that estimator (2.10) is unbiased and asymptotically consistent.

Turning to the estimation of σ_u^2, we transform $\mathbf{z}_i$ to $z_i^{(2)} = c^{-1/2}\,\mathbf{f}'\mathbf{z}_i$ such that $u_{2i}^{(2)} = c^{-1/2}\,\mathbf{f}'\mathbf{P}\mathbf{u}_{2i}$ has mean 0 and variance σ^2. Pre-multiplying model (2.8) by $\mathrm{diag}_{1\le i\le m}(c^{-1/2}\,\mathbf{f}')$ and noting that $c^{-1/2}\,\mathbf{f}'\mathbf{f} = c^{1/2}$, we obtain the following transformed model:

$$\mathbf{z}^{(2)} = \mathbf{H}^{(2)}\boldsymbol{\beta} + \mathbf{e}^{(2)} , \tag{2.11}$$

where $\mathbf{z}^{(2)} = \mathrm{col}_{1\le i\le m}(c^{-1/2}\mathbf{f}'\mathbf{z}_i)$, $\mathbf{H}^{(2)} = \mathrm{col}_{1\le i\le m}(c^{-1/2}\mathbf{f}'\mathbf{H}_i)$ and $\mathbf{e}^{(2)} = c^{1/2}\mathbf{v} + \mathbf{u}_2^{(2)} + \boldsymbol{\epsilon}^{(2)}$, in which $\mathbf{u}_2^{(2)} = \mathrm{col}_{1\le i\le m}(c^{-1/2}\mathbf{f}'\mathbf{P}\mathbf{u}_{2i})$ and $\boldsymbol{\epsilon}^{(2)} = \mathrm{col}_{1\le i\le m}(c^{-1/2}\mathbf{f}'\mathbf{P}\boldsymbol{\epsilon}_i)$. The error term of model (2.11) has $E(\mathbf{e}^{(2)}) = \mathbf{0}_{m\times 1}$ and $V(\mathbf{e}^{(2)}) = c\sigma_u^2\mathbf{B}^{-1} + \sigma^2\mathbf{I}_m + \mathbf{R}^{(2)}$, where $\mathbf{R}^{(2)} = \mathrm{diag}_{1\le i\le m}(c^{-1}\mathbf{f}'\mathbf{P}\mathbf{R}_i\mathbf{P}'\mathbf{f})$. Let $\widehat{\mathbf{e}}^{(2)\prime}\widehat{\mathbf{e}}^{(2)}$ be the residual sum of squares obtained by regressing $\mathbf{z}^{(2)}$ on $\mathbf{H}^{(2)}$ using OLS. An unbiased estimator of σ_u^2 is given by:

$$\widetilde{\sigma}_u^2 = \left\{\widehat{\mathbf{e}}^{(2)\prime}\widehat{\mathbf{e}}^{(2)} - \mathrm{tr}\big(\mathbf{P}_{H^{(2)}}\mathbf{R}^{(2)}\big) - \widetilde{\sigma}^2\Big[m - r\big(\mathbf{H}^{(2)}\big)\Big]\right\} \Big/ \Big[c \times \mathrm{tr}\big(\mathbf{P}_{H^{(2)}}\mathbf{B}^{-1}\big)\Big] , \tag{2.12}$$

where $\mathbf{P}_{H^{(2)}} = \mathbf{I}_m - \mathbf{H}^{(2)}\big(\mathbf{H}^{(2)\prime}\mathbf{H}^{(2)}\big)^{-}\mathbf{H}^{(2)\prime}$ and $\widetilde{\sigma}^2$ is given by (2.10). The unbiasedness of $\widetilde{\sigma}_u^2$ follows by noting that $E(\widehat{\mathbf{e}}^{(2)\prime}\widehat{\mathbf{e}}^{(2)}) = c\sigma_u^2\,\mathrm{tr}(\mathbf{P}_{H^{(2)}}\mathbf{B}^{-1}) + \sigma^2\big[m - r(\mathbf{H}^{(2)})\big] + \mathrm{tr}(\mathbf{P}_{H^{(2)}}\mathbf{R}^{(2)})$. Since $\widetilde{\sigma}^2$ and $\widetilde{\sigma}_u^2$ can take negative values, we truncate them at zero and use $\widehat{\sigma}^2 = \max\{0, \widetilde{\sigma}^2\}$ and $\widehat{\sigma}_u^2 = \max\{0, \widetilde{\sigma}_u^2\}$. The truncated estimators are no longer unbiased but they are still asymptotically consistent.

A two-stage estimator of θ_{it} can now be obtained from (2.7) by replacing $\boldsymbol{\psi} = (\sigma^2, \sigma_u^2)'$ for $\widehat{\boldsymbol{\psi}} = (\widehat{\sigma}^2, \widehat{\sigma}_u^2)'$:

$$\widehat{\theta}_{it} = \widehat{\theta}_{it}^H(\widehat{\boldsymbol{\psi}}) = \mathbf{x}_{it}'\widehat{\boldsymbol{\beta}} + \big(\widehat{\sigma}_u^2\boldsymbol{\varsigma}_i' \otimes \mathbf{1}_T' + \widehat{\sigma}^2\boldsymbol{\zeta}_{it}'\big)\,\widehat{\mathbf{V}}^{-1}\big(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\big) . \tag{2.13}$$

This estimator is not a genuine EBLUP because we have assumed known autoregressive coefficients. An adequate name for estimator (2.13) would be partial EBLUP, because just the variance components were replaced by their estimators. However, for simplicity it is called EBLUP along this manuscript. Following Kackar & Harville (1984) and Rao & Yu (1994), we may note that estimator (2.13) is unbiased as estimators of variance components, (2.10) and (2.12), are even functions of $\mathbf{y}$ and translation-invariant.

3. ESTIMATION OF THE MSPE OF THE TWO-STAGE ESTIMATOR

Under the normality of the random effects and random errors, following the Kackar & Harville (1984) identity and using the Henderson's general result (Henderson, 1975), the MSE of an EBLUP can be decomposed as:

$$MSE\big[\widehat{\theta}_{it}(\widehat{\boldsymbol{\psi}})\big] = g_{1it}(\boldsymbol{\psi}) + g_{2it}(\boldsymbol{\psi}) + E\big[\widehat{\theta}_{it}(\widehat{\boldsymbol{\psi}}) - \widetilde{\theta}_{it}(\boldsymbol{\psi})\big]^2 , \tag{3.1}$$

where E means the expectation with respect to model (2.5), $g_{1it}(\boldsymbol{\psi})$ represents the uncertainty of the EBLUP due to the estimation of the random effects and is of order $o(1)$, $g_{2it}(\boldsymbol{\psi})$ is due to the estimation of the fixed effects and is of order $o(m^{-1})$, and the last term measures the uncertainty due to the estimation of the variance components. The first two terms can be analytically evaluated, due to the linearity of the EBLUP in the data vector $\mathbf{y}$, from the following closed formulas,

$$g_{1it}(\boldsymbol{\psi}) = \sigma_u^2 \varsigma_{ii} + \frac{\sigma^2}{1-\rho^2} - \big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)' \mathbf{V}^{-1} \big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big) \tag{3.2}$$

and

$$\begin{aligned} g_{2it}(\boldsymbol{\psi}) &= \Big[\mathbf{x}_{it} - \mathbf{X}'\mathbf{V}^{-1}\big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big]' (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &\quad \times \Big[\mathbf{x}_{it} - \mathbf{X}'\mathbf{V}^{-1}\big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big] . \end{aligned} \tag{3.3}$$

However, the third term on the right of equation (3.1) does not have a closed-form expression, due to the non-linearity of the EBLUP in the data vector $\mathbf{y}$, and therefore an approximation is needed.

3.1. Analytical approximation of the MSPE estimator

Following Kackar & Harville (1984) Taylor series approximation and then the lines of Prasad & Rao (1990), we propose the following analytical approximation of the third term of (3.1):

$$E\Big[\widehat{\theta}_{it}(\widehat{\boldsymbol{\psi}}) - \widetilde{\theta}_{it}(\boldsymbol{\psi})\Big]^2 \approx \operatorname{tr}\Big[\mathbf{L}_{it}(\boldsymbol{\psi}) \mathbf{V}(\boldsymbol{\psi}) \mathbf{L}'_{it}(\boldsymbol{\psi}) \overline{\mathbf{V}}(\widehat{\boldsymbol{\psi}})\Big] = g_{3it}(\boldsymbol{\psi}) , \tag{3.4}$$

where $\mathbf{L}_{it}(\boldsymbol{\psi}) = \frac{\partial \mathbf{b}'_{it}(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}}$, $\mathbf{V}(\boldsymbol{\psi})$ is given by (2.6) and $\overline{\mathbf{V}}(\widehat{\boldsymbol{\psi}})$ denotes the asymptotic covariance matrix of $\widehat{\boldsymbol{\psi}}$. Using $\mathbf{b}'_{it}(\boldsymbol{\psi}) = \big(\sigma_u^2 \boldsymbol{\varsigma}'_i \otimes \mathbf{1}'_T + \sigma^2 \boldsymbol{\zeta}'_{it}\big)\mathbf{V}^{-1}$ it follows that $\mathbf{L}_{it}(\boldsymbol{\psi}) = \big(\frac{\partial \mathbf{b}_{it}}{\partial \sigma^2}, \frac{\partial \mathbf{b}_{it}}{\partial \sigma_u^2}\big)'$ is a $2 \times mT$ matrix with two blocks given by:

$$\frac{\partial \mathbf{b}'_{it}}{\partial \sigma^2} = \Big[\boldsymbol{\zeta}'_{it} - \big(\sigma_u^2 \boldsymbol{\varsigma}'_i \otimes \mathbf{1}'_T + \sigma^2 \boldsymbol{\zeta}'_{it}\big) \mathbf{V}^{-1} \big(\mathbf{I}_m \otimes \boldsymbol{\Gamma}\big)\Big] \mathbf{V}^{-1} \tag{3.5}$$

and

$$\frac{\partial \mathbf{b}'_{it}}{\partial \sigma_u^2} = \Big[\boldsymbol{\varsigma}'_i \otimes \mathbf{1}'_T - \big(\sigma_u^2 \boldsymbol{\varsigma}'_i \otimes \mathbf{1}'_T + \sigma^2 \boldsymbol{\zeta}'_{it}\big) \mathbf{V}^{-1} \big(\mathbf{Z}_1 \mathbf{B}^{-1} \mathbf{Z}'_1\big)\Big] \mathbf{V}^{-1} . \tag{3.6}$$

Let $\mathbf{A}_{it}(\boldsymbol{\psi}) = \mathbf{L}_{it}(\boldsymbol{\psi}) \mathbf{V}(\boldsymbol{\psi}) \mathbf{L}'_{it}(\boldsymbol{\psi}) = \{a_{kl}\}$, thus it follows from previous results that it is a 2×2 symmetric matrix with elements:

$$\begin{aligned} a_{11} &= \Big[\boldsymbol{\zeta}_{it} - (\mathbf{I}_m \otimes \boldsymbol{\Gamma}) \mathbf{V}^{-1} \big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big]' \mathbf{V}^{-1} \\ &\quad \times \Big[\boldsymbol{\zeta}_{it} - (\mathbf{I}_m \otimes \boldsymbol{\Gamma}) \mathbf{V}^{-1} \big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big] , \end{aligned} \tag{3.7}$$

$$\begin{aligned} a_{22} &= \Big[\boldsymbol{\varsigma}_i \otimes \mathbf{1}_T - (\mathbf{Z}_1 \mathbf{B}^{-1} \mathbf{Z}'_1) \mathbf{V}^{-1} \big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big]' \mathbf{V}^{-1} \\ &\quad \times \Big[\boldsymbol{\varsigma}_i \otimes \mathbf{1}_T - (\mathbf{Z}_1 \mathbf{B}^{-1} \mathbf{Z}'_1) \mathbf{V}^{-1} \big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big] \end{aligned} \tag{3.8}$$

and

$$a_{12} = a_{21} = \Big[\boldsymbol{\zeta}_{it} - (\mathbf{I}_m \otimes \boldsymbol{\Gamma})\,\mathbf{V}^{-1}\big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big]'\mathbf{V}^{-1} \times \Big[\boldsymbol{\varsigma}_i \otimes \mathbf{1}_T - (\mathbf{Z}_1 \mathbf{B}^{-1} \mathbf{Z}_1')\,\mathbf{V}^{-1}\big(\sigma_u^2 \boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \sigma^2 \boldsymbol{\zeta}_{it}\big)\Big]. \tag{3.9}$$

It remains to obtain the elements of $\overline{\mathbf{V}}(\widehat{\boldsymbol{\psi}})$ in expression (3.4). Following the lines of Rao & Yu (1994), we propose the evaluation of those elements using a lemma on the covariance of two quadratic forms of normally distributed variables (Jiang, 2007, p. 238). For this purpose we have to write $\widetilde{\sigma}^2$ and $\widetilde{\sigma}_u^2$ as quadratic forms of normally distributed variables. These estimators can be written as:

$$\widetilde{\sigma}^2 = k_1\, \mathbf{a}'\mathbf{C}_1\mathbf{a} + k_2 \tag{3.10}$$

and

$$\widetilde{\sigma}_u^2 = k_3\, \mathbf{a}'\mathbf{C}_2\mathbf{a} + k_4\, \mathbf{a}'\mathbf{C}_1\mathbf{a} + k_5\ , \tag{3.11}$$

where $k_1 = \big[m(T-1) - r(\mathbf{H}^{(1)})\big]^{-1}$, $k_2 = -\operatorname{tr}\big\{\big[\mathbf{D}^* - \mathbf{H}^{(1)}(\mathbf{H}^{(1)\prime}\mathbf{H}^{(1)})^- \mathbf{H}^{(1)\prime}\big]\,\mathbf{R}^{(1)}\big\}$ $\big[m(T-1) - r(\mathbf{H}^{(1)})\big]^{-1}$, $k_3 = \big[c \times \operatorname{tr}(\mathbf{P}_{H^{(2)}}\mathbf{B}^{-1})\big]^{-1}$, $k_4 = -k_1\big[m - r(\mathbf{H}^{(2)})\big]\big[c \times \operatorname{tr}(\mathbf{P}_{H^{(2)}}\mathbf{B}^{-1})\big]^{-1}$ and $k_5 = \big\{-k_2\big[m - r(\mathbf{H}^{(2)})\big] - \operatorname{tr}(\mathbf{P}_{H^{(2)}}\mathbf{R}^{(2)})\big\}\,\big[c \times \operatorname{tr}(\mathbf{P}_{H^{(2)}}\mathbf{B}^{-1})\big]^{-1}$ are constants. Furthermore $\mathbf{a} = \mathbf{Z}_1\mathbf{v} + \mathbf{u}_2 + \boldsymbol{\epsilon} \sim N(\mathbf{0}; \mathbf{V})$, $\mathbf{C}_1 = \mathbf{C}^{(1)\prime}\big[\mathbf{I}_{mT} - \mathbf{C}^{(1)}\mathbf{X}\,(\mathbf{X}'\mathbf{C}^{(1)\prime}\mathbf{C}^{(1)}\mathbf{X})^- \mathbf{X}'\mathbf{C}^{(1)\prime}\big]\mathbf{C}^{(1)}$ and $\mathbf{C}_2 = \mathbf{C}^{(2)\prime}\big[\mathbf{I}_m - \mathbf{C}^{(2)}\mathbf{X}\,(\mathbf{X}'\mathbf{C}^{(2)\prime}\mathbf{C}^{(2)}\mathbf{X})^- \mathbf{X}'\mathbf{C}^{(2)\prime}\big]\mathbf{C}^{(2)}$ are symmetric matrices with $\mathbf{C}^{(1)} = \operatorname{diag}_{1\le i\le m}\big[(\mathbf{I}_T - \mathbf{D})\mathbf{P}\big]$ and $\mathbf{C}^{(2)} = \operatorname{diag}_{1\le i\le m}(c^{-1/2}\mathbf{f}'\mathbf{P})$, respectively. Let $\overline{\mathbf{V}}(\widehat{\boldsymbol{\psi}}) \equiv \mathbf{D} = \{d_{kl}\}$, thus it follows from previous results that it is a 2×2 symmetric matrix with elements:

$$d_{11} = V(\widetilde{\sigma}^2) = 2\,k_1^2 \operatorname{tr}(\mathbf{C}_1\mathbf{V}\mathbf{C}_1\mathbf{V})\ , \tag{3.12}$$

$$d_{22} = V(\widetilde{\sigma}_u^2) = 2\,k_3^2 \operatorname{tr}(\mathbf{C}_2\mathbf{V}\mathbf{C}_2\mathbf{V}) + 4\,k_3k_4 \operatorname{tr}(\mathbf{C}_1\mathbf{V}\mathbf{C}_2\mathbf{V}) + 2\,k_4^2 \operatorname{tr}(\mathbf{C}_1\mathbf{V}\mathbf{C}_1\mathbf{V}) \tag{3.13}$$

and

$$d_{12} = d_{21} = \operatorname{Cov}(\widetilde{\sigma}^2; \widetilde{\sigma}_u^2) = 2\,k_1k_3 \operatorname{tr}(\mathbf{C}_1\mathbf{V}\mathbf{C}_2\mathbf{V}) + 2\,k_1k_4 \operatorname{tr}(\mathbf{C}_1\mathbf{V}\mathbf{C}_1\mathbf{V})\,. \tag{3.14}$$

Following Prasad & Rao (1990), it can be assumed that $E\big[g_{1it}(\widehat{\boldsymbol{\psi}}) + g_{2it}(\widehat{\boldsymbol{\psi}})\big] = g_{1it}(\boldsymbol{\psi}) + g_{2it}(\boldsymbol{\psi}) - g_{3it}(\boldsymbol{\psi})$. Thus it follows that a bias corrected analytical estimator of MSPE of two-stage estimator is given by:

$$mspe^A\big[\widehat{\theta}_{it}(\widehat{\boldsymbol{\psi}})\big] = g_{1it}(\widehat{\boldsymbol{\psi}}) + g_{2it}(\widehat{\boldsymbol{\psi}}) + 2\,g_{3it}(\widehat{\boldsymbol{\psi}})\ . \tag{3.15}$$

3.2. Bootstrap approximation of the MSPE estimator

In this section we introduce an alternative way of approximating the MSPE of the EBLUP by a simple bootstrap procedure using similar ideas to Butar &

Lahiri (2003). Hereafter we describe a bootstrap procedure designed for estimating the MSPE under the spatio-temporal small area model (2.5), and using ANOVA estimates of variance components introduced in Section 2.3. In general, the parametric bootstrap method consists of generating parametrically a large number of area bootstrap samples from the model fitted to the original data, re-estimating the model parameters for each bootstrap sample and then estimating the separate components of the MSPE. Assuming that ρ and ϕ are known, the parametric bootstrap procedure follows the next steps:

1. Calculate estimates of the variance components, $\widehat{\sigma}^2$ and $\widehat{\sigma}_u^2$, from the initial data, $\mathbf{y}$, using the method of moments. Then fit model (2.5) in order to estimate the fixed effects, $\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}(\mathbf{y};\widehat{\boldsymbol{\psi}})$ with $\widehat{\boldsymbol{\psi}} = (\widehat{\sigma}^2, \widehat{\sigma}_u^2)'$.
2. Compute the EBLUP estimates of θ_{it}, $\widehat{\theta}_{it}(\widehat{\boldsymbol{\psi}})$, and the first two terms of MSPE, $g_{1it}(\widehat{\boldsymbol{\psi}})$ and $g_{2it}(\widehat{\boldsymbol{\psi}})$.
3. Generate m independent copies of a variable $\mathbf{u}_1^*$, with $\mathbf{u}_1^* \sim N(\mathbf{0};\widehat{\sigma}_u^2\mathbf{I}_m)$. From this values, construct the random vector $\mathbf{v}^* = (\mathbf{I}_m - \phi\mathbf{W})^{-1}\mathbf{u}_1^*$, assuming that ϕ is known.
4. Generate mT independent copies of a variable $\boldsymbol{\xi}^*$, with $\boldsymbol{\xi}^* \sim N(\mathbf{0};\widehat{\sigma}^2\mathbf{I}_{mT})$, independently of the generation of $\mathbf{u}_1^*$. From this values, construct the random vector $\mathbf{u}_2^*$, assuming that ρ is known.
5. Generate mT independent copies of a variable $\boldsymbol{\epsilon}^*$, with $\boldsymbol{\epsilon}^* \sim N(\mathbf{0};\mathbf{R})$, independently of the generation of $\mathbf{u}_1^*$ and $\boldsymbol{\xi}^*$.
6. Construct the bootstrap data $\mathbf{y}^* = \mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{Z}\boldsymbol{\upsilon}^* + \boldsymbol{\epsilon}^*$ where $\boldsymbol{\upsilon}^* = [\mathbf{v}^{*\prime}\mathbf{u}_2^{*\prime}]'$.
7. Calculate bootstrap estimates of the variance components, $\widehat{\sigma}^{2*}$ and $\widehat{\sigma}_u^{2*}$, from the bootstrap data, $\mathbf{y}^*$, and then fit model (2.5) in order to obtain bootstrap estimates of the fixed effects, $\widehat{\boldsymbol{\beta}}^* = \widehat{\boldsymbol{\beta}}(\mathbf{y};\widehat{\boldsymbol{\psi}}^*)$ with $\widehat{\boldsymbol{\psi}}^* = (\widehat{\sigma}^{2*}, \widehat{\sigma}_u^{2*})'$.
8. Calculate bootstrap estimates of the EBLUP, as well as estimates of the two components of the MSPE of BLUP, using bootstrap estimates of the variance components, $\widehat{\boldsymbol{\psi}}^*$:

$$\widehat{\theta}_{it}^* = \widehat{\theta}_{it}^*(\mathbf{y};\widehat{\boldsymbol{\beta}}^*;\widehat{\boldsymbol{\psi}}^*) = \mathbf{x}_{it}'\widehat{\boldsymbol{\beta}}^* + \left(\widehat{\sigma}_u^{2*}\boldsymbol{\varsigma}_i' \otimes \mathbf{1}_T' + \widehat{\sigma}^{2*}\boldsymbol{\zeta}_{it}'\right)\left[\widehat{\mathbf{V}}(\widehat{\boldsymbol{\psi}}^*)\right]^{-1}\left(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}^*\right),$$

$$\begin{aligned} g_{1it}^* = g_{1it}^*(\mathbf{y};\widehat{\boldsymbol{\beta}}^*;\widehat{\boldsymbol{\psi}}^*) &= \widehat{\sigma}_u^{2*}\varsigma_{ii} + \frac{\widehat{\sigma}^{2*}}{1-\rho^2} \\ &\quad - \left(\widehat{\sigma}_u^{2*}\boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \widehat{\sigma}^{2*}\boldsymbol{\zeta}_{it}\right)'\left(\widehat{\mathbf{V}}^*\right)^{-1}\left(\widehat{\sigma}_u^{2*}\boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \widehat{\sigma}^{2*}\boldsymbol{\zeta}_{it}\right), \end{aligned}$$

$$\begin{aligned} g_{2it}^* = g_{2it}^*(\mathbf{y};\widehat{\boldsymbol{\beta}}^*;\widehat{\boldsymbol{\psi}}^*) &= \left[\mathbf{x}_{it} - \mathbf{X}'(\widehat{\mathbf{V}}^*)^{-1}\left(\widehat{\sigma}_u^{2*}\boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \widehat{\sigma}^{2*}\boldsymbol{\zeta}_{it}\right)\right]' \\ &\quad \times \left[\mathbf{X}'(\widehat{\mathbf{V}}^*)^{-1}\mathbf{X}\right]^{-1}\left[\mathbf{x}_{it} - \mathbf{X}'(\widehat{\mathbf{V}}^*)^{-1}\left(\widehat{\sigma}_u^{2*}\boldsymbol{\varsigma}_i \otimes \mathbf{1}_T + \widehat{\sigma}^{2*}\boldsymbol{\zeta}_{it}\right)\right]. \end{aligned}$$

9. Repeat steps 3–8 B times. Let $\widehat{\sigma}^{2*(b)}$ and $\widehat{\sigma}_u^{2*(b)}$ be the bootstrap estimates of variance components, $\widehat{\boldsymbol{\psi}}^{*(b)} = (\widehat{\sigma}^{2*(b)}, \widehat{\sigma}_u^{2*(b)})'$; and $\widehat{\boldsymbol{\beta}}^{*(b)}$, $\widehat{\theta}_{it}^{*(b)}$, $g_{1it}^{*(b)}$ and $g_{2it}^{*(b)}$ the bootstrap estimates of $\boldsymbol{\beta}$, θ_{it}, g_{1it} and g_{2it}, respectively, all of them obtained in the b^{th} bootstrap replication, $b = 1, ..., B$.

10. Calculate a bootstrap estimate of g_{3it} using the following Monte Carlo approximation:

$$g_{3it}^* = B^{-1} \sum_{b=1}^{B} \big(\widehat{\theta}_{it}^{*(b)} - \widehat{\theta}_{it}\big)^2 .$$

Since it is known that the quantity $g_{1it}(\widehat{\boldsymbol{\psi}}) + g_{2it}(\widehat{\boldsymbol{\psi}})$ is a biased estimator of $g_{1it}(\boldsymbol{\psi}) + g_{2it}(\boldsymbol{\psi})$ (Prasad & Rao, 1990), thus following the lines of Butar & Lahiri (2003), a bias corrected bootstrap estimator of MSPE of the two-stage estimator can be defined as:

$$\text{(3.16)} \qquad mspe^B \big[\widehat{\theta}_{it}(\widehat{\boldsymbol{\psi}})\big] = 2\big[g_{1it}(\widehat{\boldsymbol{\psi}}) + g_{2it}(\widehat{\boldsymbol{\psi}})\big] - B^{-1} \sum_{b=1}^{B} \big[g_{1it}^{*(b)} + g_{2it}^{*(b)}\big] + g_{3it}^* \;.$$

4. A MONTE CARLO SIMULATION STUDY

In order to assess the merits of our spatio-temporal estimator, in this section we present a simulation study designed for comparing the efficiency of the proposed EBLUP estimator (ST) against other well-known EBLUP estimators, such as the Fay–Herriot (FH), the Salvati (NS) and the Rao–Yu (RY) estimators. The first is one of the paradigms in small area estimation exploring neither spatial nor chronological similarities; the Salvati estimator is a well-known small-area estimator that explores spatial similarities in data; and the last one is a reference for small area estimation with chronological correlation. This simulation study also aims to study the accuracy of the proposed estimators of MSPE of the EBLUP.

In the simulated experiments we assume that the proposed EBLUP estimator performs better than the others when the spatio-temporal model provides a good fit. Thus we have decided to generate an artificial population of $\mathbf{y}$-values using model (2.5). We have considered model (2.5) with $p = 2$, that is, a constant and one explanatory variable, $\mathbf{x}_{it} = (1, x_{it})'$. The mT values of x_{it} were generated from a uniform distribution in the interval [0,1]. The true model coefficients were $\boldsymbol{\beta} = (1, 2)'$, the random area specific effects variance were $\sigma_u^2 \in \{0.5; 1.0\}$, the random area-by-time specific effects variance were $\sigma^2 \in \{0.25; 0.50; 1.00\}$, the temporal autoregressive coefficients were $\rho \in \{0.2; 0.4; 0.8\}$ and the spatial autoregressive coefficients were $\phi \in \{0.25; 0.50; 0.75\}$. The row standardized $\mathbf{W}$ matrix was kept fixed in all simulations and it corresponded to contiguous NUTSIII in one European country (NUTSIII is a geocode standard for referencing the subdivisions of European countries for statistical purposes). We have selected $T = 7$ and $m = 28$ divided into four uniform groups. Further, we have considered three different σ_{it}^2 patterns (see Table 1). Note that sampling variances are more dispersed in pattern (C) than in pattern (A).

Table 1: Values of the sampling variances for the simulation study.

i	**1–7**	**8–14**	**15–21**	**22–28**
Pattern (A)	1.0	1.0	1.0	1.0
Pattern (B)	0.6	0.8	1.1	1.5
Pattern (C)	0.3	0.6	1.0	2.0

In addition, we have assumed known ρ and ϕ, although the bootstrap method can also accommodate the situation of unknown coefficients. The random effects and errors were generated independently from a Normal distribution with zero mean. Finally, the vector of mT values of the interest variable, $\mathbf{y}$, was generated from the cross-sectional and time-series stationary small area model (2.5).

4.1. The efficiency of the EBLUP

The simulation study designed for comparing the efficiency of the proposed EBLUP estimator over other well-known EBLUP estimators has followed the algorithm described below:

1. Generate $L = 1{,}000$ samples of initial data, $\mathbf{y}^{(l)} = (y_{11}^{(l)}, ..., y_{it}^{(l)}, ..., y_{mT}^{(l)})$, as described above, $l = 1, ..., L$.
2. Fit the Fay–Herriot (FH), the Salvati (NS) and the Rao–Yu (RY) models, as well as the spatio-temporal (ST) model, to the initial data, $\mathbf{y}^{(l)}$, using method of moments estimates of those variance components, for each $l = 1, ..., L$.
3. Compute the EBLUP estimates under each of those models, $\widehat{\theta}_{it}^{a}$, $a \in$ {FH,NS, RY,ST}.
4. Calculate the Monte Carlo approximation of the relative efficiency (RE) of the proposed EBLUP over the other three EBLUP estimators as follows: $RE_{it} = \frac{MSE(\widehat{\theta}_{it}^{a})}{MSE(\widehat{\theta}_{it}^{ST})} \times 100$, where $MSE(\widehat{\theta}_{it}^{a}) = L^{-1} \sum_{l=1}^{L} \big(\widehat{\theta}_{it}^{a(l)} - y_{it}^{a(l)}\big)^2$ is the empirical MSE of each EBLUP, $a \in$ {FH,NS,RY,ST}. That measure of efficiency is calculated at area-by-time level. To summarize results, we have computed an average global measure over the mT small areas: $ARE = \frac{1}{mT} \sum_{i=1}^{m} \sum_{t=1}^{T} RE_{it}$.

These results are shown in Tables 2 to 4 for the sampling variance patterns (A), (B) and (C), respectively.

Table 2: Average relative efficiency of the proposed EBLUP over other EBLUP estimators — σ^2_{it} pattern (A).

σ^2	σ^2_u	$\phi = 0.25$			$\phi = 0.50$			$\phi = 0.75$		
		ρ=0.2	ρ=0.4	ρ=0.8	ρ=0.2	ρ=0.4	ρ=0.8	ρ=0.2	ρ=0.4	ρ=0.8
FH estimator										
0.25	0.50	130	133	151	136	138	156	156	159	177
	1.00	161	165	185	172	176	196	213	217	237
0.50	0.50	126	132	168	130	136	173	149	155	192
	1.00	153	160	199	162	169	209	197	205	246
1.00	0.50	120	129	186	124	133	190	138	147	206
	1.00	140	151	211	148	158	219	175	186	249
NS estimator										
0.25	0.5	130	133	150	135	138	156	153	156	176
	1.0	161	165	185	171	175	195	209	213	235
0.50	0.5	126	132	168	130	136	172	146	152	191
	1.0	153	160	199	161	169	209	194	202	245
1.00	0.5	120	129	186	123	132	190	135	145	206
	1.0	140	150	211	147	157	219	172	183	249
RY estimator										
0.25	0.5	101	101	101	105	105	104	122	121	119
	1.0	102	102	102	109	109	108	136	135	132
0.50	0.5	101	101	101	105	104	104	119	118	117
	1.0	102	102	102	108	108	108	132	131	129
1.00	0.5	101	101	101	104	104	105	116	116	117
	1.0	101	101	102	107	107	108	128	128	127

To begin with, it is interesting to note from Tables 2 to 4 that there is not much difference in efficiency among the three sampling variance patterns, which agrees with the results obtained by Pratesi & Salvati (2009) under a model with spatially correlated random effects. In addition, it should be emphasized that the gain in efficiency of the proposed EBLUP over other estimators slightly increases with the dispersion of the sampling variances. Note that the efficiency of the proposed estimator over the Fay–Herriot and the Salvati estimators increases mainly for higher values of the temporal autocorrelation coefficient, while the efficiency of that estimator over the Rao–Yu estimator tends to increase for higher values of the spatial and temporal autocorrelation coefficients.

Comparing the ARE, among the three sampling variance patterns (see Tables 2 to 4), we observe that substantial gains in efficiency are achieved when it is used the proposed EBLUP over other EBLUP estimators, especially over the

Table 3: Average relative efficiency of the proposed EBLUP over other EBLUP estimators — σ_{it}^2 pattern (B).

σ^2	σ_u^2	$\phi=0.25$			$\phi=0.50$			$\phi=0.75$		
		ρ=0.2	ρ=0.4	ρ=0.8	ρ=0.2	ρ=0.4	ρ=0.8	ρ=0.2	ρ=0.4	ρ=0.8
FH estimator										
0.25	0.5	130	133	151	135	138	157	157	161	180
	1.0	163	167	189	175	179	201	219	224	247
0.50	0.5	124	131	171	130	136	177	149	156	199
	1.0	154	162	206	164	172	217	201	210	258
1.00	0.5	119	129	191	123	132	195	137	147	211
	1.0	140	151	217	148	159	226	176	188	258
NS estimator										
0.25	0.5	129	133	151	134	138	157	154	158	179
	1.0	163	167	189	174	178	201	216	221	245
0.50	0.5	124	131	171	129	136	177	146	154	198
	1.0	153	161	206	163	171	217	198	207	257
1.00	0.5	119	128	191	122	132	195	135	146	211
	1.0	140	151	217	147	158	226	173	186	257
RY estimator										
0.25	0.5	101	101	101	106	106	105	124	124	121
	1.0	102	102	102	111	110	110	140	139	136
0.50	0.5	101	101	101	106	106	105	122	121	119
	1.0	102	102	102	110	109	109	135	135	132
1.00	0.5	101	101	101	105	105	106	118	118	118
	1.0	102	102	102	108	108	109	130	130	129

Fay–Herriot estimator and over the Salvati estimator. On average the gains in efficiency over these estimators are about 69% for pattern (A), 71% for pattern (B) and 78% for pattern (C). There are also some gains in efficiency of the proposed EBLUP over the Rao–Yu estimator, although they are smaller than the ones observed against the other two estimators (overall average gain in efficiency is equal to 11%, 12% and 13%, respectively, for patterns (A), (B) and (C). Note that the ARE values of the proposed EBLUP over the Rao–Yu estimator are negligible for $\phi = 0.25$ and $\phi = 0.50$. However, for $\phi = 0.75$ the proposed EBLUP performs clearly better than the Rao–Yu estimator in terms of efficiency. Let us also observe from Tables 2 to 4 that gains in efficiency tend to increase with σ_u^2, ρ and ϕ, i.e., the higher values of these three parameters the stronger gains in efficiency of the proposed estimator over the others. Nevertheless, there is a slight decrease on ARE values with the increase of σ^2. It is also worth noting that the proposed EBLUP shows significant gains in efficiency over the NS estimator even for small ρ and that these gains increase with the increase of ϕ. This means that the introduction of the chronological autocorrelation in small area estimation models has a better effect in the efficiency of the estimators than the

Table 4: Average relative efficiency of the proposed EBLUP over other EBLUP estimators — σ^2_{it} pattern (C).

σ^2	σ^2_u	$\phi=0.25$			$\phi=0.50$			$\phi=0.75$		
		ρ=0.2	ρ=0.4	ρ=0.8	ρ=0.2	ρ=0.4	ρ=0.8	ρ=0.2	ρ=0.4	ρ=0.8
FH estimator										
0.25	0.5	123	127	152	129	134	160	156	161	191
	1.0	165	172	204	181	188	222	238	247	285
0.50	0.5	118	127	183	124	134	191	147	158	220
	1.0	154	166	230	166	179	246	211	227	302
1.00	0.5	113	125	203	118	130	209	134	147	229
	1.0	138	151	236	146	161	247	177	194	286
NS estimator										
0.25	0.5	123	127	152	129	133	160	153	159	190
	1.0	165	172	204	180	187	221	235	244	283
0.50	0.5	117	127	183	123	133	191	144	156	219
	1.0	154	166	230	165	179	246	208	224	301
1.00	0.5	113	125	203	117	129	209	131	145	229
	1.0	138	151	236	146	160	247	174	191	286
RY estimator										
0.25	0.5	98	98	98	105	105	104	129	129	127
	1.0	100	100	100	111	111	111	150	149	145
0.50	0.5	99	98	99	105	105	105	127	127	123
	1.0	100	100	100	110	110	110	144	143	138
1.00	0.5	99	99	100	105	105	106	123	122	121
	1.0	100	100	101	109	109	109	135	135	132

introduction of the spatial correlation. All these results confirm the superiority of the proposed spatio-temporal model in comparison to other well-known small area models when data exhibits both spatial and chronological correlations with the considered structure. Finally, it is important to note that our findings are consistent with the simulation results obtained by Singh *et al.* (2005) when it is assumed $\phi = 0.75$ and $\rho = 0.50$. Therefore, the spatial dependence and the temporal autocorrelation should be exploited to strengthen the small area estimates, no matter it is assumed or not an interaction between the autoregressive coefficients.

4.2. The accuracy of the MSPE of the EBLUP

The simulation study designed for studying the accuracy of the proposed estimators of the MSPE of the EBLUP has followed the algorithm described below:

1. Generate $L = 1{,}000$ samples of initial data, $\mathbf{y}^{(l)} = (y_{11}^{(l)}, ..., y_{it}^{(l)}, ..., y_{mT}^{(l)})$, as described above, $l = 1, ..., L$.

2. From the initial data, $\mathbf{y}^{(l)}$, calculate estimates of the variance components, $\sigma^{2(l)}$ and $\sigma_u^{2(l)}$, using the method of moments, and then fit model (2.5) in order to estimate the fixed effects $\widehat{\boldsymbol{\beta}}^{(l)} = \widehat{\boldsymbol{\beta}}(\mathbf{y}^{(l)}; \widehat{\boldsymbol{\psi}}^{(l)})$, where $\widehat{\boldsymbol{\psi}}^{(l)} = (\widehat{\sigma}^{2(l)}, \widehat{\sigma}_u^{2(l)}, \rho, \phi)'$, for each $l = 1, ..., L$.

3. Compute the EBLUP estimates, $\widehat{\theta}_{it}^{(l)}(\widehat{\boldsymbol{\psi}}^{(l)})$ and their analytic MSPE estimates (A-MSPE), $mspe^{A(l)}(\widehat{\theta}_{it}^{(l)})$, for each $l = 1, ..., L$.

4. Generate B bootstrap data sets as described in Section 3.2 from estimates $\sigma^{2(l)}$ and $\sigma_u^{2(l)}$, and then compute the bootstrap MSPE estimates: $mspe^{B(l)}\ (\widehat{\theta}_{it}^{(l)})$ (B-MSPE), for each $l = 1, ..., L$.

5. Compute the empirical values of MSPE for each i^{th} small-area at t^{th} time point, $MSPE_{it}$, which are the benchmark values, with $R = 5,000$ independent data sets in order to ensure high precision.

6. Calculate the Monte Carlo approximations of each MSPE estimative of the EBLUP ($mspe$), their relative bias (RB) and relative MSE (RMSE) as follows: $mspe_{it}^{b} = L^{-1} \sum_{l=1}^{L} mspe_{it}^{b(l)}$, $RB_{it} = L^{-1} \sum_{l=1}^{L} \frac{mspe_{it}^{b(l)} - MSPE_{it}}{MSPE_{it}} \times 100$ and $RMSE_{it} = L^{-1} \sum_{l=1}^{L} \frac{(mspe_{it}^{b(l)} - MSPE_{it})^2}{MSPE_{it}} \times 100$, where $b \in \{A, B\}$ denotes the different MSPE estimators. These measures are calculated at area-by-time level. To summarize results, we have produced three global measures over the mT small areas: the percentage of areas where the relative bias is negative (RBN), the average of absolute relative bias (AARB), $AARB = \frac{1}{mT} \sum_{i=1}^{m} \sum_{t=1}^{T} |RB_{it}|$ and the average relative MSE (ARMSE), $ARMSE = \frac{1}{mT} \sum_{i=1}^{m} \sum_{t=1}^{T} RMSE_{it}$.

Tables 5 to 7 report the percent RBN, percent AARB and percent ARMSE of the analytical and the bootstrap MSPE estimators of the EBLUP, for $\rho = 0.2$, $\rho = 0.4$ and $\rho = 0.8$, respectively. The variance components patterns have a significant effect on the performance of both MSPE estimators, unlike the spatial autoregressive coefficient. For that reason, we have decided to report results only for $\phi = 0.25$ and $\phi = 0.50$. Note that Molina *et al.* (2009) showed, under a model with spatially correlated random effects, that the level of spatial dependence does not significantly affect the performance of the analytical and parametric bootstrap MSPE estimators.

The results in Table 5 suggest that the bootstrap MSPE estimator can compete with the analytical MSPE estimator in terms of bias and accuracy when $\rho = 0.2$. From Table 5 we can see that the analytical estimator presents smaller percent AARB and ARMSE (although of the same order of magnitude) than the resampling-based estimator when $\sigma^2 = 0.25$ and $\sigma^2 = 1.00$. On the other hand, the results for those measures reveal that the bootstrap estimator is somewhat better than the analytical estimator when $\sigma^2 = 0.50$. In Table 5 we can also see that both MSPE estimators tend to underestimate the true MSPE of the EBLUP for the majority of small areas, but this underestimation decreases as long as the variance, σ^2, increases. Furthermore, note that the analytical estimator has slight more negative bias than the other estimator.

Table 5: RBN, ARB and ARMSE of MSPE estimators, $\rho = 0.2$.

σ^2	σ_u^2	A-MSPE	B-MSPE	A-MSPE	B-MSPE
		$\phi = 0.25$		$\phi = 0.50$	
RBN (%)					
0.25	0.5	100.000	100.000	100.000	100.000
	1.0	100.000	100.000	100.000	100.000
0.50	0.5	99.908	99.867	99.908	99.878
	1.0	99.755	99.571	99.755	99.582
1.00	0.5	13.520	12.133	13.959	12.480
	1.0	9.153	8.051	9.296	8.245
AARB (%)					
0.25	0.5	58.674	59.872	58.753	59.995
	1.0	57.213	58.409	56.528	57.788
0.50	0.5	34.770	34.063	34.856	34.169
	1.0	31.991	31.246	32.121	31.421
1.00	0.5	17.287	18.090	17.294	18.124
	1.0	20.449	21.504	20.303	21.313
ARMSE (%)					
0.25	0.5	26.138	27.592	26.261	27.757
	1.0	24.899	26.317	23.984	25.434
0.50	0.5	8.637	8.501	8.696	8.568
	1.0	7.374	7.250	7.442	7.332
1.00	0.5	1.925	2.087	1.933	2.103
	1.0	2.537	2.775	2.516	2.741

Table 6 shows that the analytical MSPE estimator is slightly better than the bootstrap MSPE estimator when $\rho = 0.4$. In fact, the results reported in this table indicate again that there is not much difference on bias and accuracy between these MSPE estimators. Although the analytical estimator is always the best in terms on precision (it has the lowest percent ARMSE), the bootstrap estimator is somewhat better than the analytic estimator in terms of bias (according to AARB measure) when $\sigma^2 = 0.50$. The systematic underestimation of the true MSPE of the EBLUP is also revealed by Table 6.

The results reported in Table 7 suggest that the analytical MSPE estimator is the winner when $\rho = 0.8$. From this table we can see that gains on bias and accuracy are reached when it is used the analytical estimator instead of the

Table 6: RBN, ARB and ARMSE of MSPE estimators, $\rho = 0.4$.

σ^2	σ_u^2	A-MSPE	B-MSPE	A-MSPE	B-MSPE
		$\phi = 0.25$		$\phi = 0.50$	
RBN (%)					
0.25	0.5	100.000	100.000	100.000	100.000
	1.0	100.000	100.000	100.000	100.000
0.50	0.5	99.755	99.592	99.786	99.561
	1.0	99.429	99.010	99.408	99.051
1.00	0.5	26.316	22.939	26.347	22.908
	1.0	21.173	17.602	21.531	17.898
AARB (%)					
0.25	0.5	58.363	60.733	58.417	60.808
	1.0	56.305	58.634	56.385	58.757
0.50	0.5	36.561	36.026	36.595	36.037
	1.0	34.199	33.548	34.293	33.679
1.00	0.5	15.019	15.818	15.107	15.959
	1.0	16.875	18.092	16.853	18.052
ARMSE (%)					
0.25	0.5	26.036	28.725	26.128	28.852
	1.0	23.948	26.509	24.058	26.661
0.50	0.5	9.741	9.829	9.779	9.864
	1.0	8.592	8.658	8.649	8.727
1.00	0.5	1.645	1.812	1.665	1.846
	1.0	2.025	2.294	2.024	2.288

bootstrap estimator in measuring the uncertainty of the EBLUP, especially when $\sigma^2 = 0.25$. Furthermore, it should be noted that the underestimation of the true MSPE of the EBLUP is more pronounced when $\sigma^2 = 1.00$.

Comparing the results reported in Tables 5 to 7, among different levels of spatial and temporal autocorrelation, we can observe that both estimators show similar bias and accuracy in most cases, since they are very close on AARB and ARMSE measures. We can also observe for both estimators that: (i) the spatial autoregressive coefficient does not have a significant impact on the performance of the MSPE estimators; (ii) the performance of both MSPE estimators, in what concerns the AARB and the ARMSE, tends to improve for higher values of the variance components (mainly σ^2); (iii) the number of estimates with positive bias tends to increase for higher values of the variance components; (iv) the analytical estimator shows consistently negative bias in more small areas than the bootstrap estimator; and (v) the gain of accuracy of the analytical estimator over the bootstrap one tends to increase with the strength of chronological autocorrelation. Finally, our results reported in Tables 5 to 7 suggest that both MSPE estimators perform well, especially for higher values of the variance components

Table 7: RBN, ARB and ARMSE of MSPE estimators, $\rho = 0.8$.

σ^2	σ_u^2	A-MSPE	B-MSPE	A-MSPE	B-MSPE
		$\phi = 0.25$		$\phi = 0.50$	
RBN (%)					
0.25	0.5	99.990	99.898	99.980	99.878
	1.0	99.990	99.878	99.980	99.837
0.50	0.5	96.378	94.061	96.051	93.969
	1.0	96.163	93.469	95.990	93.184
1.00	0.5	59.765	54.153	59.571	54.122
	1.0	59.469	53.673	59.306	53.561
AARB (%)					
0.25	0.5	50.048	60.057	49.776	60.065
	1.0	49.143	58.869	49.062	58.844
0.50	0.5	36.395	38.984	36.257	38.966
	1.0	35.730	38.231	35.662	38.218
1.00	0.5	21.665	21.954	21.789	21.964
	1.0	21.666	22.110	21.773	22.173
ARMSE (%)					
0.25	0.5	19.944	29.258	19.774	29.329
	1.0	19.173	28.053	19.131	28.100
0.50	0.5	10.696	12.913	10.636	12.927
	1.0	10.350	12.460	10.320	12.467
1.00	0.5	3.502	3.883	3.555	3.889
	1.0	3.494	3.943	3.539	3.969

(for example, for $\sigma^2 = 1.00$ and $\sigma_u^2 = 0.5$ or $\sigma_u^2 = 1.0$) and for lower values of the temporal autoregressive coefficient, and can be in most cases used to adequately access the accuracy of the proposed EBLUP estimator. Analogous findings were reached by Pereira & Coelho (2010) when comparing the performance of MSPE estimators under a cross-sectional and time-series stationary model.

5. CONCLUSIONS

In this work we have studied the problem of "borrowing information" from related small areas and time periods in order to strengthen the estimators of the small area parameters of interest. In particular, we have proposed a spatio-temporal area level LMM involving spatially correlated and temporally autocorrelated random area effects, using both time-series and cross-sectional data. Under this model, we first obtained a partial EBLUP estimator. We then proposed two estimators of the MSPE of that EBLUP.

In the simulation study we have studied the efficiency of the proposed EBLUP estimator over other well-known EBLUP estimators and we have studied the accuracy of the proposed estimators of MSPE. Our empirical results based on simulated data have shown that the proposed EBLUP estimator can lead to remarkable efficiency gains. This is especially true over both the sectional FH and the spatial NS estimators and when the autocorrelation coefficients are high. It should also be noted that our results have shown mild gains in efficiency from the inclusion of a spatial structure into the Rao–Yu cross-sectional and time-series model, i.e. there are gains in efficiency of the proposed EBLUP over the temporal RY estimator.

Under several simulated scenarios for the variance components and autocorrelation coefficients, our empirical results have revealed that both MSPE estimators perform well. Furthermore, our results indicate that the analytical MSPE estimator performs slightly better than the bootstrap one on bias and precision, although its superiority is not uniform. In particular, it is to be noticed that this gain tend to be more conspicuous when chronological correlation is strong.

We believe that the proposed methodology can provide a useful tool for practitioners working with spatially correlated and temporally autocorrelated data in the context of small area estimation problems. Nonetheless, some of the issues mentioned in this paper require further theoretical work and/or more extensive simulation studies. For example, we have assumed known and positive autoregressive coefficients, but in practice these parameters are unknown and could be negative. A further issue relates to on what happens when random area effects do not follow approximately a SAR process or when random area-by-time effects do not follow an AR(1) process? Our expectation is that whenever spatial and chronological correlations exist, this can be a viable approach due to its simplicity of implementation and the fact of allowing to incorporate all the available information in the estimation process. Nevertheless, additional work is still needed to understand the properties of the partial EBLUP estimator and particularly of its MSPE approximations when spatial or chronological correlation processes significantly departure from the considered ones.

ACKNOWLEDGMENTS

The authors thank the Fundação para a Ciência e a Tecnologia (FCT).

REFERENCES

[1] Anselin, L. (1992). *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Boston.

[2] Butar, F.B. and Lahiri, P. (2003). On measures of uncertainty of empirical Bayes small area estimators, *Journal of Statistical Planning and Inference*, **112**, 1–2, 63–76.

[3] Chandra, H.; Salvati, N. and Chambers, R. (2007). Small area estimation for spatially correlated populations - a comparison of direct and indirect model-based methods, *Statistics in Transition – new series*, **8**, 2, 331–350.

[4] Datta, G.S.; Lahiri, P. and Maiti, T. (2002). Empirical Bayes estimation of median income of four-person families by state using time series and cross-sectional data, *Journal of Statistical Planning and Inference*, **102**, 1, 83–97.

[5] Fay, R.E. and Herriot, R.A. (1979). Estimates of income for small places: An application of James-Stein procedures to census data, *Journal of the American Statistical Association*, **74**, 366, 269–277.

[6] Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics, *Philosophical Transactions of the Royal Society A*, **222**, 309–368.

[7] Fuller, W.A. (1987). *Measurement Error Models*, John Wiley & Sons, New York.

[8] Hartley, H.O. and Rao, J.N.K. (1967). Maximum-likelihood estimation for the mixed analysis of variance model, *Biometrika*, **54**, 1–2, 93–108.

[9] Henderson, C.R. (1953). Estimation of variance and covariance components, *Biometrics*, **9**, 2, 226–252.

[10] Henderson, C.R. (1975). Best linear unbiased estimation and prediction under a selection model, *Biometrics*, **31**, 2, 423–447.

[11] Jiang, J. (2007). *Linear and Generalized Linear Mixed Models and Their Applications*, Springer-Verlag, New York.

[12] Judge, G.G.; Griffiths, W.E.; Hill, R.C.; Lütkepohl, H. and Lee, T. (1985). *The Theory and Practice of Econometrics, second ed.*, Wiley, New York.

[13] Kackar, R.N. and Harville, D.A. (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models, *Journal of the American Statistical Association*, **79**, 388, 853–862.

[14] Molina, I.; Salvati, N. and Pratesi, M. (2009). Bootstrap for estimating the MSE of the Spatial EBLUP, *Computational Statistics*, **24**, 3, 441–458.

[15] Patterson, H.D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal, *Biometrika*, **58**, 3, 545–554.

[16] Pereira, L.N. and Coelho, P.S. (2010). Assessing different uncertainty measures of EBLUP: a resampling-based approach, *Journal of Statistical Computation and Simulation*, **80**, 7, 713–727.

[17] Petrucci, A.; Pratesi, M. and Salvati, N. (2005). Geographic information in small area estimation: small area models and spatially correlated random area effects, *Statistics in Transition*, **7**, 3, 609–623.

[18] PETRUCCI, A. and SALVATI, N. (2006). Small area estimation for spatial correlation in watershed erosion assessment, *Journal of Agricultural, Biological and Environmental Statistics*, **11**, 2, 169–182.

[19] PRASAD, N.G.N. and RAO, J.N.K. (1990). The estimation of the mean squared error of small-area estimators, *Journal of the American Statistical Association*, **85**, 409, 163–171.

[20] PRATESI, M. and SALVATI, N. (2008). Small area estimation: the EBLUP estimator based on spatially correlated random area effects, *Statistical Methods & Applications*, **17**, 1, 113–141.

[21] PRATESI, M. and SALVATI, N. (2009). Small area estimation in the presence of correlated random area effects, *Journal of Official Statistics*, **25**, 1, 37–53.

[22] RAO, J.N.K. and YU, M. (1994). Small-area estimation by combining time-series and cross-sectional data, *The Canadian Journal of Statistics*, **22**, 4, 511–528.

[23] SAEI, A. and CHAMBERS, R. (2003). *Small area estimation under linear and generalized linear mixed models with time and area effects*, Working Paper M03/15, Southampton Statistical Sciences Research Institute, Southampton. Available from `http://eprints.soton.ac.uk/8165/`.

[24] SALVATI, N. (2004). *Small area estimation by spatial models: the spatial Empirical Best Linear Unbiased Prediction (Spatial EBLUP)*, Working Paper No 2004/3, Dipartimento di Statistica "G. Parenti", Firenze. Available from `http://eprints.soton.ac.uk/8165/`.

[25] SINGH, B.B.; SHUKLA, G.K. and KUNDU, D. (2005). Spatio-temporal models in small area estimation, *Survey Methodology*, **31**, 2, 183–195.

[26] THOMPSON, W.A. (1962). The problem of negative estimates of variance components, *The Annals of Mathematical Statistics*, **33**, 1, 273–289.

[27] WALL, M.M. (2004). A close look at the spatial structure implied by the CAR and SAR models, *Journal of Statistical Planning and Inference*, **121**, 2, 311–324.