

Spatiality in small area estimation: A new structure with a simulation study

Yadollah Mehrabi,¹ Amir Kavousi,² Ahmad-Reza Baghestani,^{3,4}
Mojtaba Soltani-Kermanshahi⁵

¹Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences; ²Workplace Health Promotion Research Center and Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences; ³Physiotherapy Research Centre, Shahid Beheshti University of Medical Sciences; ⁴Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran; ⁵Social Determinants of Health Research Center, School of Medicine, Semnan University of Medical Sciences, Semnan, Iran

Abstract

In numerous practical applications, data from neighbouring small areas present spatial correlation. More recently, an extension of the Fay–Herriot model through the Simultaneously Autoregressive (SAR) process has been considered. The Conditional Autoregressive (CAR) structure is also a popular choice. The reasons of using these structures are theoretical properties, computational advantages and relative ease of interpretation. However, the assumption of the non-singularity of matrix (I_{m-pw}) is a problem. We introduce here a novel structure of the covariance matrix when approaching spatiality in small area estimation (SAE) comparing that with the commonly used SAR process. As an example, we present synthetic data on grape production with spatial correlation for 274 municipalities in the region of Tuscany as base data simulating data at each area and comparing the results. The SAR

process had the smallest Root Average Mean Square Error (RAMSE) for all conditions. The RAMSE also generally decreased with increasing sample size. In addition, the RAMSE values did not show a specific behaviour but only spatially correlation coefficient changes led to a stronger decrease of RAMSE values than the SAR model when our new structure was applied. The new approach presented here is more flexible than the SAR process without severe increasing RAMSE values.

Introduction

Sample surveys have for a long time been employed as a cost-effective remedy for data gathering. Such require proper statistics to effectively assess entire populations but need also to produce reliable outcomes for subpopulations, mostly entitled domains or areas, which are said to be “small” if the domain-specific sample size is not large sufficient to produce a direct estimate with acceptable precision. Because of the increasing demand for reliable estimates for these areas, Small Area Estimation (SAE) techniques are exposed to growing popularity in survey sampling (Rao, 2003; Rao and Wu, 2009; Soltani-Kermanshahi *et al.*, 2017a; 2017b).

A direct estimate of a small area parameter(s) depends entirely on the samples from that area which is usually not reliable. However, much of the popularity and usefulness of SAE techniques can be attributed to model-based techniques, now widely used and with a growing application during the past three decades. These approaches take strength from neighbouring areas and use proper techniques that connect direct estimates from the small areas. When an area-level summary of covariate variables exists, we can usefully utilize these models. A fundamental area-level model was first proposed by Fay and Herriot (1979) and followed up by others (Rao, 2003; Waller and Gotway 2004; Rao and Wu, 2009; Soltani-Kermanshahi *et al.*, 2017a; 2017b).

In numerous practical applications, data from neighbouring small areas present spatial correlations. In these instances, unless enough auxiliary variables are accessible, between-area correlations should in some way be illustrated in the covariance structure of the model. Nevertheless, the establishment of a structure of dependence between small areas requires a serious conceptual difference with regard to the traditional structure of independent small areas, where the entire covariance matrix is block-diagonal (Prasad and Rao, 1990; Molina *et al.*, 2009; Gharde *et al.*, 2012). Cressie (1991) already employed a model with spatially correlated random effects in the context of SAE. More recently, an extension

Correspondence: Mojtaba Soltani-Kermanshahi, Department of Epidemiology and Biostatistics, School of Medicine, Semnan University of Medical Sciences, Semnan, Iran.
Tel.: +98.9125285381.
E-mail: msoltani@farabi.tums.ac.ir

Key words: Spatiality; small area estimation; simultaneously autoregressive; exponential structure; simulation.

Acknowledgments: The authors are thankful to SCI, Statistical Research and Training Center and Mr. Yuosef Bashiri for providing data for this project. We also thank Shahid Beheshti University of Medical Sciences for all spiritual supports.

Conflict of interests: The authors declare no potential conflict of interests.

Received for publication: 26 February 2020.
Accepted for publication: 7 September 2020.

©Copyright: the Author(s), 2020
Licensee PAGEPress, Italy
Geospatial Health 2020; 15:872
doi:10.4081/gh.2020.872

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (CC BY-NC 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

of the Fay–Herriot model through the simultaneously autoregressive (SAR) process has been considered (Singh *et al.*, 2005; Petrucci and Salvati, 2006; Pratesi and Salvati, 2008). The model has the following form:

$$Y = X\beta + v + e \quad (\text{Eq. 1.1})$$

where Y is the vector of design-unbiased direct estimator available for each of the small areas (from 1 to m); X the vector of area-level auxiliary covariates; e the vector of independent sampling errors; and v the result of a SAR process with the unknown autoregression parameter ρ (with range -1 to 1) and proximity matrix W (Anselin, 1988; Cressie, 1993). Then v can be expressed as:

$$v = (I_m - \rho W)^{-1} u \quad (\text{Eq. 1.2})$$

Here, $u = (u_1, \dots, u_m)^T$ and has mean 0 and covariance matrix $\sigma_u^2 I_m$, where I_m denotes the $m \times m$ identity matrix and σ_u^2 is an unknown parameter (Anselin 1988; Molina *et al.*, 2009).

The Conditional Auto-Regressive (CAR) structure is also a popular choice and, in some papers, more common than the CAR. The SAR structure is known as a subset of CAR structure (Arab *et al.*, 2017). The reasons of using these structures are due to their theoretical properties, computational advantages and ease of interpretations (Cressie, 1993; Yasui and Lele, 1997; Arab *et al.*, 2017; Li *et al.*, 2009).

Regrettably, the application of traditional covariance-based spatial statistical models is in many cases either inappropriate or computationally inefficient. Moreover, conventional methods are often incapable of allowing the researcher to quantify uncertainties corresponding to the model parameters since the space parameters of most complex spatial and spatio-temporal models are very large (Arab *et al.*, 2017). A problem of the SAR structure is the assumption of non-singularity of the matrix $(I_m - \rho W)$. Singularity of that matrix leads not to real solutions for all of the equations, so we cannot use the entire situation of that matrix (Molina *et al.*, 2009). Another problem of the SAR model is the estimation of ρ . Since, one reason for using SAE techniques is the small sample size, then we have a low precision in estimation of ρ .

A widely adopted choice for this correlation function is the Mat'ern function, *i.e.*

$$[(2\eta\sqrt{\nu} d)^\nu K_\nu(2\eta\sqrt{\nu} d)] / [2^{(\nu-1)} \Gamma(\nu)] \quad (\text{Eq. 1.3})$$

where η measures the correlation decay with the distance; ν a smoothness parameter; $\Gamma(\cdot)$ the conventional Gamma function; and $K_\nu(\cdot)$ the modified Bessel function of the second kind of order ν . Special cases of the Mat'ern function include the exponential model:

$$\rho(d) = \exp(-d) \quad (\text{Eq. 1.4})$$

when the smoothness parameter $\nu = 0.5$, the 'decay parameter' $\eta = 1$ and d is the distance between two areas (Waller and Gotway, 2004; Li *et al.*, 2009).

In this paper, we used geographic information such as latitude and longitude to increase the precision of ρ estimation. We introduce a novel structure in covariance matrix by using spatiality in SAE and compare it with the commonly used SAR process. In the first part of this research, we applied a new perspective of structure in the covariance matrix of v in Eq.1.1 by using an exponential

model in the Mat'ern function to predict household food expenditures (HFE) in urban areas of Iran. In the second part of the study, we conducted a simulation study based on grape production data in Tuscany, Italy to compare the effects of two methods of spatiality on the precision of estimations.

Statistical approach

Suppose that we have M areas in the population and only m areas are sampled. The SAR model with spatial dependence in SAE has the linear form (Eq. 1.1). Then the co-variance of v has the following form as described by Molina *et al.* (2009):

$$G(\omega) = \sigma_u^2 [(I_m - \rho W)^T (I_m - \rho W)]^{-1} \quad (\text{Eq. 2.1})$$

where the vector of variance components is denoted by ω . Alternatively, we used a new structure in the v covariance matrix (Eq. 1.1) by using an exponential model in the Mat'ern function as follows:

$$u = (u_1, \dots, u_m) \sim \text{MvN}(0, V), \quad e_i \sim (0, \sigma_i^2), \quad v_{ij} = \sigma_u^2 \rho(d_{ij}), i, j = 1, 2, \dots, m \quad (\text{Eq. 2.2})$$

where V is a variance-covariance matrix with elements v_{ij} and correlation function $\rho(\cdot)$, an isotropic correlation function that decays as the Euclidean distance $d_{ij} = \|s_i - s_j\|$ between two individuals increases. In the exponential model $\rho(\cdot)$ has the form describe by Eq. 1.4. In this study, we used Restricted Maximum Likelihood (REML) and Spatial Empirical Best Linear Unbiased Prediction (SEBLUP) to estimate σ_u^2 and β 's, respectively. $P < 0.05$ was considered as significant.

Real data analysis

The data used were collected by the Statistical Center of Iran (SCI) in 2013. The sample size was optimum at the province level but not at the district level (SCI, 2013). The Iranian Rural and Urban Household Expenditures and Income Survey (IRUHEIS) are carried out annually by SCI. The sampling design was three-stage with stratification. In the first stage, the census areas were determined and selected; in the second, the urban blocks and rural settlements were selected and at the third, the sample households were selected. We analyzed urban data of IRUHEIS 2013, collected between 21 April 2013 and 20 April 2014.

In 2013, a total of 387 districts were selected by IRUHEIS out of 429 districts of Iran's urban areas. Of the 18,876 households that participated in IRUHEIS 2013, we analysed complete data consisting of 18,850 households. The highest and lowest sample sizes belonged to Tehran (998 samples) and Narmashir (5 samples) districts, respectively. The HFE includes all payments made to purchase the needed nutritious and essential food items including meat, dairy products, cereal and bean, bread and flour, biscuits and cakes, oil and butter, fruits and vegetables, nuts, sweets and sugar, additives and dressings, as well as cigarettes and tobacco (SCI, 2013). We also took into account some socio-economic factors that may potentially affect the HFE. These variables have been measured at the area level rather than the individual one. The data were extracted from reports disseminated by the Iran Ministry of Interior and also by SCI.

We considered district-level variables emanating from Iran's census of 2011 that included the Average Number of Households (ANH); the average number of rooms of each household; sex



ratio; the Proportion of Male Households (PMH); the proportion of the active population employed; the proportion of population of the following age groups: >65 years, 25-64 years, 15-24 years and <15 years; and the Proportion of Higher Education (PHE), (SCI, 2011). We also considered the Gross Domestic Product (GDP); the proportion of households that had joined a charity organization; the distance from province capital; the per capita income for municipalities; and the migration rate. These data had been produced by the Ministry of Interior in a project conducted for identification of less developed regions of the country. We also used geographic information such as latitude and longitude of each area capital to calculate exponential correlations (Iranian Noojum, 2018).

In order to reach the highest correlation with dependent variables, an appropriate transformation of independent variables (power for continuous variables and logarithm for proportion variables) was used. We also utilized the Variance Inflation Factors (VIF) to assess the collinearity among independent variables, with VIF>5 or coefficients of multiple determination with respect to other independent variables more than 0.8, which indicates serious multicollinearity for the predictor (Besley *et al.*, 1980). The forward selection method was used to establish the final model. In the forward method, independent variables (ordered by their correlation with dependent variable) are inserted into the model serially where only significant variables remain.

According to VIF values, the variable proportion of population at 25 to 64 years was omitted (Supplementary file S1).

Simulation with grape production data

To investigate the effects of spatiality on precision of estimation a simulation study was carried out using the new structure proposed in this paper. We used synthetic data based on grape production in the region of Tuscany, Italy with a spatial correlation for 274 municipalities. That dataset used included a frame with 274 observations of the following variables: i) the direct estimators of the mean agrarian surface area used for production of grape (in hectares) for each Tuscany municipality (grapehect); ii) the agrarian surface area (in hectares) used for production (area); iii) the average number of working days in the reference year 2000 (workdays); and iv) the sampling variance of the direct estimators for each Tuscany municipality (var). We also added longitudinal and latitude for each area to that data. Population data come from the Italian Agricultural Census for the region of Tuscany in the year 2000 (Pratesi *et al.*, 2013; Molina and Marhuenda, 2015). We generated data for each of the 274 areas by the following approach: i) Multivariate normal distribution with mean and co-variance of the

correspondent real data was used, *i.e.*, for the new structure σ_u^2 was 165 and d_{ij} calculated from the following equation:

$$d_{ij} = \sqrt{(l_i - l_j)^2 + (t_i - t_j)^2} \tag{Eq. 4.1}$$

where i and j are the latitude and longitude of areas. This gives a σ_u^2 of 101 and the autoregression parameter ρ 0.2 for the SAR structure. ii) The sample size at each area was closely mimicked the real data to make the simulation as realistic as possible; iii) Data generation was repeated 1000 times at each of 274 small areas and mean and variance of every area were computed; iv) To achieve small area estimations, we took advantage of the auxiliary variables included in the continues variables of area (in hectares) and workdays from Tuscany in the year 2000, as they were found to be significant in real data analysis. v) We fitted the SAR and exponential structure spatial model to each of 1000 generated datasets and saved the estimates for each area. vi) To assess the precision of SAEs, we employed Root Average Mean Square Errors (RAMSEs) for small area estimates:

$$RAMSE = \sqrt{\sum_{i=1}^{274} MSE_i / 274} \tag{Eq. 4.2}$$

Table 1. Results of spatial (exponential) model on households' food expenditures.

Variables	Beta	SE	P-value
ANH	4.59x10 ³	1.98x10 ³	0.021
PMH	5.45x10 ⁴	9.36x10 ³	<0.001
PHE	6.20x10 ³	1.48x10 ³	<0.001
$\sigma_v^2=1.31 \times 10^8$		AIC=13619	

ANH: Average Number of Household; PMH: proportion of Male Households (headed by a male); PHE: Proportion of Higher Education; SE=Standard Error; AIC=Akaike Information Criterion.

Table 2. Descriptive statistics of variables in the Tuscany data.

Variable	Minimum	Maximum	Mean	SD
Grapehect	0.25	343.00	69.51	49.45
Area	35.81	2890	639.20	571.89
Workdays	22.15	497.70	149.60	67.49

Table 3. Root Average Mean Square Error of small area estimates with area sample size and spatiality correlation coefficient changes in two types of spatiality structures by simulation.

Sample Size	Method	Spatiality correlation coefficient			
		20% reduction	Real	20% increase	
20% reduction	SAR	219.42	219.19	219.30	
	Exponential	222.19	222.32	222.21	
	Real	SAR	218.79	218.87	218.87
		Exponential	222.11	221.76	221.70
20% increase	SAR	218.67	218.27	218.51	
	Exponential	221.59	221.52	221.67	

SAR: Simultaneously Auto-Regressive process.

where has the following form:

$$MSE_i = \frac{1}{1000} \sum_{q=1}^{1000} (\hat{g}_{i,q} - \hat{g}_{i,M})^2 \quad (\text{Eq. 4.3})$$

where $\hat{g}_{(i,M)}$ is the estimate of g_i based on real data with the two models (M) and $\hat{g}_{(i,q)}$ shows the estimated g_i at run q (Soltani-Kermanshahi *et al.*, 2017a; 2017b). For better assessment, we used sample size and correlation sensitive analysis using the real sample size and correlation as well as 20% above and below of it. All calculations used utilized the R package SAE (Molina and Marhuenda, 2015) and the codes given in Supplementary files S2 and S3).

Simulation results

The Tuscany data had 274 areas and the descriptive statistics of the three variables used are shown in Table 2. The results of RAMSE with sample size and spatiality correlation coefficient sensitivity are shown in Table 3 and Figures 2 and 3. Accordingly, SAR process approximately had smallest RAMSE for all conditions. Also, with increasing sample size, RAMSE roughly decreased. The increase in accuracy powerfully related to areas sample size. In addition, with increasing spatiality correlation coefficient, RAMSEs have not specific behavior but in most times, they decreased, approximately.

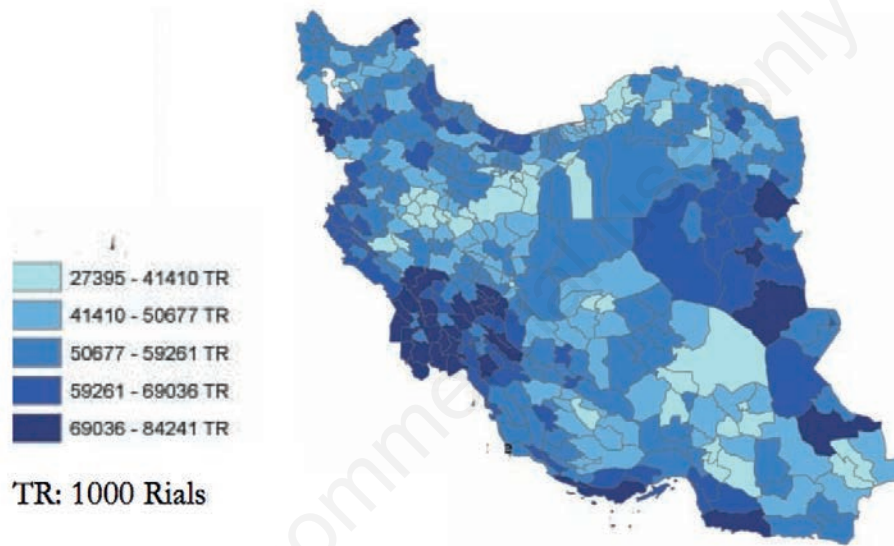
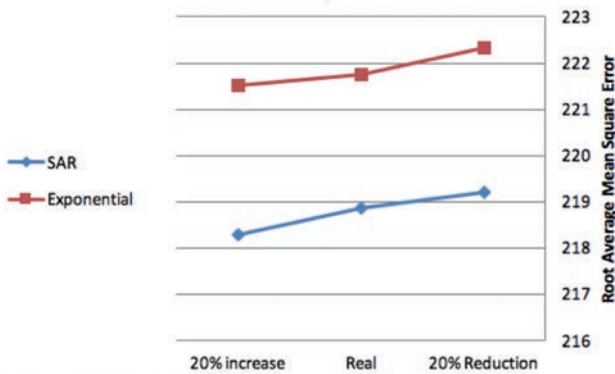
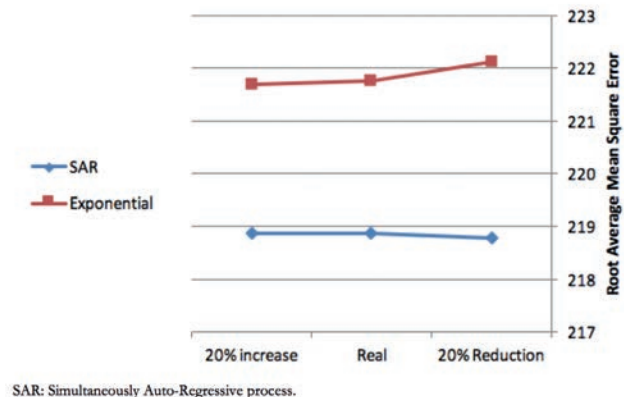


Figure 1. The zoning of Annual Household Food Expenditures at the district-level in Iran estimated with spatial (exponential) model.



SAR: Simultaneously Auto-Regressive process.

Figure 2. Root Average Mean Square Error of small area estimates with area sample size changes in two types of spatiality structures by simulation.



SAR: Simultaneously Auto-Regressive process.

Figure 3. Root Average Mean Square Error of small area estimates with spatiality correlation coefficient changes in two types of spatiality structures by simulation.

Discussion and conclusions

This paper proposes a new structure in the covariance matrix of v (see Eq. 1.1) by using an exponential model in Mat'ern function (Cressie, 1993; Li *et al.*, 2009) after considering a Fay–Herriot model with correlated random area effects according to the SAR process. Our study shows three significant variables, ANH, PMH and PHE, had a significant effect on HFE (Table 1), the zoning of which showed that the border districts, especially those on the western border of Iran, had a higher HFE than in other parts of the country (Figure 1). Although we had expected the highest HFE in the districts of the capital, our findings showed the opposite. One reason for this could be the way Iran has developed economically.

From a purely statistical point of view, the position of the small area is related to modelling of its parameters and further improvement in the EBLUP estimator can be achieved by including probable spatial interaction among random area effects as discussed previously (Petrucci and Salvati, 2004; Pratesi and Salvati, 2009). Indeed, the inclusion of ancillary variables to obtain the spatial effects may be beneficial even when the strength of the spatial link is weak (Petrucci *et al.*, 2005).

The simulation study with sensitive analysis used to compare the RAMSE values of the SAR and the proposed new structure showed that that sample size changes have a higher impact on RAMSE than spatiality correlation coefficient changes. One potential problem with SAE is small subdomain samples. The expected increased precision with increased sample sizes (Soltani-Kermanshahi *et al.*, 2017b) was corroborated by our observation of decreasing RAMSE values with bigger samples. In our simulations process, there were no areas without sample data. For such areas, for which the values of the covariates at the area-level are available from any other data source, possible estimators are $\hat{Y}_i = X_i \beta_i$ (Molina *et al.*, 2009). In addition, our results showed that only spatially correlation coefficient changes lead to a stronger decrease of RAMSE values than SAR model when our new structure was applied.

The SAR process had the smallest RAMSE values for all conditions in relation to the new process. Even if SAR has appealing theoretical properties, computational advantages and provides ease of interpretation (Cressie, 1993; Yasui and Lele, 1997; Li *et al.*, 2009), it assumes non-singularity of matrix $(I_m - \rho W)$, which can be avoided by the new process. In general, the RAMSE values were increased about 1.5% without the assumption of non-singularity of matrix $(I_m - \rho W)$ by the new process. The new process is also more flexible than the SAR process without severe increasing in RAMSEs. With high spatial correlations in the study variables, spatial EBLUP with correlated random area effects following a SAR process, works better (Petrucci *et al.*, 2005; Pratesi and Salvati, 2009). However, in our work the RAMSE values increased in certain situations with increasing spatiality correlation coefficients.

The biggest difference of our study related to similar studies was the insufficient number of households in some districts, but the degree of weakness decreased by using the SAE approach. At least we could find the admissible model to predict the HFE in urban areas of Iran. Furthermore, one of the basic limitations in SAE methodology is the inclusion of highly correlated variables with dependent variables. In this study, we tried to find those variables and gain access to them.

References

- Anselin L, 1988. Spatial econometrics: methods and models. Kluwer Academic Publishers, Dordrecht; Boston.
- Arab A, Hooten MB, Wikle CK, 2017. Hierarchical spatial models. In: Encyclopedia of GIS. Springer Publishers, Boston. doi:10.1007/978-3-319-17885-1_564
- Besley DA, Kuh E, Welsch RE, 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. John Wiley and Sons., New York
- Cressie N, 1991. Small-area prediction of undercount using the general linear model. In Proceedings of statistics symposium 90: measurement and improvement of data quality, pp. 93-105. Ottawa: Statistics Canada.
- Cressie N, 1993. Statistics for spatial data: Wiley series in probability and statistics. Wiley-Interscience NY 15, pp.105-209.
- Fay III RE, Herriot RA, 1979. Estimates of income for small places: an application of James-Stein procedures to census data. J Am Stat Assoc 74: 269-77.
- Gharde Y, Rai A, Chandra H, 2012. Hierarchical Bayes small area estimation approach for spatial data. J Indian Soc Agric Stat 66:259-68.
- Iranian Noojum, 2018. [Longitude and latitude of the Iran's districts.] Available from: <http://www.noojum.com/other/astronomy-tools/187-online-tools/6266-longitude-latitude.html> Accessed on: October 2018. [Website in Persian].
- Li Y, Tang H, Lin X, 2009. Spatial linear mixed models with covariate measurement errors. Stat Sin 19:1077.
- Molina I, Marhuenda Y, 2015. SAE: An R package for small area estimation. R J 7:81-98.
- Molina I, Salvati N, Pratesi M, 2009. Bootstrap for estimating the MSE of the Spatial EBLUP. Comput Stat 24:441-58.
- Petrucci A, Salvati N, 2004. Small area estimation using spatial information. The rathbun lake watershed case study. Dipartimento di Statistica "G. Parenti" viale morgagani, pp.59-50134. Available from: https://www.academia.edu/1376417/Small_Area_Estimation_Using_Spatial_Information_The_Rathbun_Lake_Watershed_Case_Study Accessed on: September 2020.
- Petrucci A, Salvati N, 2006. Small area estimation for spatial correlation in watershed erosion assessment. J agric Biol Environ Stat 11:169-82.
- Petrucci A, Pratesi M and Salvati N, 2005. Geographic information in small area estimation: small area models and spatially correlated random area effects. Stat Transition 7:609-23.
- Prasad NN, Rao JN, 1990. The estimation of the mean squared error of small-area estimators. J Am Stat Assoc 85:163-71.
- Pratesi M, Salvati N, 2008. Small area estimation: the EBLUP estimator based on spatially correlated random area effects. Stat Meth Appl 17:113-41.
- Pratesi M, Salvati N, 2009. Small area estimation in the presence of correlated random area effects. J Off Stat 25:37.
- Pratesi M, Giusti C, Marchetti S, 2013. Small area estimation of poverty indicators. In: Survey data collection and integration. Springer, Berlin, Heidelberg.
- Rao JNK, 2003 Small area estimation. John Wiley & Sons.
- Rao JNK, Wu W, 2009. Sample Surveys: Inference and Analysis In: Handbook of statistics 29B. Available from: <https://www.sciencedirect.com/handbook/handbook-of-statistics/vol/29/part/PB>
- Singh BB, Shukla GK, Kundu D, 2005. Spatio-temporal models in



- small area estimation. *Surv Methodol* 31:183.
- Soltani-Kermanshahi M, Mehrabi Y, Kavousi A, Baghestani AR, Mohammadi-Nasrabadi F, Aliakbari-Saba R, 2017a. Households' food expenditures in urban areas of Iran: An application of small area estimation. *Epidemiol Biostat Public Health* 14. doi:10.2427/12704
- Soltani-Kermanshahi M, Mehrabi Y, Kavousi A, Baghestani AR, Mohammadi-Nasrabadi F, 2017b. Effects of sampling variance estimation methods on precision of small area estimation. *Jp J Biostat* 14:93-106.
- Statistical Center of Iran (SCI), 2011. [National Census of Population and Housing.] Available from: https://nnt.sci.org.ir/_LAYOUTS/15/PsaLogins/sciLogin.aspx?ReturnUrl=%2fsites%2fnnt%2f_layouts%2f15%2fAuthenticate.aspx%3fSource%3d%252Fsites%252Fnnt%252FSitePages%252Freport%255F90%252Fgozaresh%255Fostani%255Fshahrestani%252Easpx&Source=%2Fsites%2Fnnt%2FSitePages%2Freport%5F90%2Fgozaresh%5Fostani%5Fshahrestani%2Easpx Accessed on: October 2018. [Website in Persian].
- Statistical Center of Iran (SCI), 2013. [Results of Urban Households' Expenditures and Income Survey.] Available from: http://amar.org.ir/Portals/0/Files/fulltext/1392/n_ahdkhsh_92.pdf Accessed on: October 2018. [Website in Persian].
- Waller LA, Gotway CA, 2004. *Applied spatial statistics for public health data*. John Wiley & Sons. doi:10.1002/0471662682
- Yasui Y, Lele S, 1997. A regression method for spatial disease rates: an estimating function approach. *J Am Stat Assoc* 92:21-32.

Non-commercial use only