Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 67(3) 2013 355-362

Bayesian Prediction in Spatial Small Area Models

Yogita Gharde, Anil Rai and Seema Jaggi
Indian Agricultural Statistics Research Institute, New Delhi

Received 29 April 2011; Revised 01 May 2013; Accepted 19 August 2013

SUMMARY

The small area models make use of explicit linking models based on random area specific effects that account for between areas variation apart from variations explained by auxiliary variables included in the model. The basic small area model considers the random area effects as independent. In practice, it should be more reasonable to assume that the random effects between the neighbouring areas are correlated. In this context, many models have been developed in recent past (Singh *et al.* 2005 and Pratesi and Salvati 2008, Salvati *et al*. 2012). In the present study, a spatial unit level small area model is obtained using Geographically Weighted Regression (GWR) approach. Further, the spatial model is studied under Hierarchical Bayes (HB) framework to improve small area estimates. Small area HB estimates are obtained using Gibbs sampling. The effects of incorporating spatial information in the model through three spatial weighting procedures are compared. Results show that estimates from new spatial model in HB framework are more efficient than the empirical approach.

Keywords: Spatial unit level model, Geographically Weighted Regression, Hierarchical Bayes, Gibbs sampling.

1. INTRODUCTION

Small area estimation is used when the sample sizes within the areas are too small to provide reliable direct survey estimates. In making estimates for small area with adequate level of precision, it is often necessary to use indirect estimators (small area models) that borrow strength from related areas and/or time periods. These models make use of explicit linking models based on random area specific effects that account for between areas variation apart from variations explained by auxiliary variables included in the model. The basic small area model considers the random area specific effects as independent. In case, the variable under study is spatial in nature, it should be more reasonable to assume that the random effects between the neighbouring areas are spatially correlated and thus correlation decays to zero as distance between any two areas in the population increases.

A popular approach of incorporating the spatial information in the model is through Geographically Weighted Regression (GWR) (Brunsdon *et al.* 1996, Fotheringham *et al.* 2002) which assumes that the regression coefficients vary spatially across the geography of interest. GWR approach extends the traditional regression model by allowing local rather than global regression parameters to be estimated by representing non-stationary local phenomena. Salvati *et.al.* (2012) investigated GWR-based small area estimation under the M-quantile modelling approach. In particular, they specify an M-quantile GWR model that is a local model for the M-quantile of the conditional distribution of the outcome variable given the covariates. Here, a new spatial unit level model has been developed using GWR approach in which regression coefficients vary spatially across the area of interest. Empirical Best Linear Unbiased Predictor (EBLUP) of small area mean and Mean Squared Error

Corresponding Author: Yogita Gharde
E-mail address: yogita_iasri@rediffmail.com

(MSE) estimator of variance have been obtained for this spatial model.

By incorporating the spatial effects in this model through Geographically Weighted Regression (GWR) approach, it may possible to improve small area estimates. Due to incorporating spatial effects, number of parameters becomes more. So, MSE estimation of this EBLUP is more complicated. To tackle this complication, Hierarchical Bayes approach is used to find the estimates for the parameters of the model and inferences for different parameters are exact unlike the EBLUP estimation. This motivates us to use HB estimates for spatial data. Datta and Ghosh (1991) introduced a Hierarchical Bayes (HB) approach for prediction in general mixed linear models. Datta *et al.* (1998) considered a superpopulation approach to HB prediction of small area mean vectors using the multivariate nested error regression model of Fuller and Harter (1987). You and Rao (2000) presented multilevel models in a HB framework and estimated small area means through its posterior mean. You and Rao (2003) also developed two-step approach to obtain design-consistent small area estimates by utilizing survey weights. You and Chapman (2006) considered the situation where the sampling error variances are estimated individually by direct estimators. They constructed a full HB model for the direct survey estimators and the sampling error variances. This HB approach automatically takes account of the extra uncertainty of estimating the sampling error variances, especially when the area specific sample sizes are small.

Therefore, in the present study, proposed spatial model is again studied in HB framework and three different weight matrices were used to incorporate the spatial effects in the model. Further, it is always difficult to obtain accurate information about the distribution of the variances. So, sensitivity analysis was done to know the effects caused by the choice of different priors on posterior means and variances. Hereby, in section 2, a spatial unit level small area model is developed using GWR approach and in section 3, model is proposed under HB framework. Section 4 presents results related to statistical properties of the estimator and weighting approaches. In this section, the effects of incorporating spatial information through three spatial weighting approaches (i.e. Neighbourhood criteria, Gaussian-decay and Spherical variogram approach) are compared empirically through spatial simulation study and section 5 shows the results of sensitivity analysis of small area estimates to the choice of values of different parameters. Finally in section 6, conclusion is given.

2. SPATIAL UNIT LEVEL SMALL AREA MODEL

Consider unit level model given by Battese *et al.* (1988)

$$y_{ij} = \boldsymbol{x}_{ij}^{T}\beta + v_i + e_{ij},\ j = 1, 2, \dots n_i;\ i = 1, 2, \dots m \quad (1)$$

where, y_{ij} denote the value of the variable of interest or dependent variable for the j^{th} unit in small area i and let $\boldsymbol{x}_{ij} = (x_{ij1}, x_{ij2}, \dots, x_{ijp})$ denote the vector of values of the p unit level auxiliary variables associated with this unit, n_i is the number of sampling units in small area i and m is the number of small areas. β is a vector of p unknown fixed effects, v_i is the random area effect associated with small area i, assumed to have mean zero and variance σ_v^2, and e_{ij} is random error associated with j^{th} unit of i^{th} small area with mean zero and variance σ_e^2. The error terms are assumed to be mutually independent, within small area as well as between small areas. Further, it is assumed that they are normally distributed. In matrix notation, model (1) is expressed as

$$\boldsymbol{Y_i} = \boldsymbol{X_i}\beta + v_i\mathbf{1}_{n_i} + \boldsymbol{e}_i,\ j = 1, 2, \dots n_i;\ i = 1, 2, \dots, m \quad (2)$$

where,

$\boldsymbol{Y}_i = (y_{i1}, y_{i2}, \dots y_{in_i})$, $\boldsymbol{X_i} = (\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \dots, \boldsymbol{x}_{in_i})$ is $n_i \times p$ matrix and $e_i = (\boldsymbol{e}_{i1}, e_{i2}, \dots, e_{in_i})$. The variance-covariance matrix of $\boldsymbol{Y}_i$ is given as

$$\text{var}(\boldsymbol{Y}_i) = \sigma_e^2\mathbf{I}_{n_i} + \sigma_v^2\mathbf{1}_{n_i}\mathbf{1}_{n_i}^T \quad (3)$$

Here, $\mathbf{1}_{n_i}$ is the unit column vector of length n_i and $\mathbf{I}_{n_i}$ is the identity matrix of order n_i.

Let,

$$v_i = z_i^T\delta_i + \in_i,\ i = 1, 2, \dots, m \quad (4)$$

where, z_i is a vector of q area level covariates of area i, δ_i is q × 1 vector of regression coefficients associated with i^{th} area and $\in_i$ is error term associated with area effect v_i. Further, in matrix notation, this can be written as

$$\boldsymbol{v} = \mathbf{Z}\delta + \in, \quad (5)$$

where, $\boldsymbol{v}$ is $m \times 1$ vector of area effects for all areas, $\mathbf{Z}$ is block diagonal matrix of area level covariates of order $m \times mq$, $\delta = (\delta_1, \delta_2, ...\delta_m)$ is $mq \times 1$ vector of regression coefficients and $\in = (\in_1, \in_2, ... \in_m)$ is vector of error terms. The spatial distance weight matrix for the i^{th} area is given by

$$\mathbf{W}_i = \mathbf{L}_i^T \mathbf{L}_i = \begin{bmatrix} \gamma_{i1} & 0 & . & . & . & 0 \\ 0 & \gamma_{i2} & . & . & . & 0 \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ 0 & 0 & . & . & & \gamma_{im} \end{bmatrix} \quad (6)$$

Here, the matrix $\mathbf{L}_i$ can be defined as the gramian root of the weight matrix $\mathbf{W}_i$ and can be written as

$$\mathbf{L}_i = \begin{bmatrix} \sqrt{\gamma_{i1}} & 0 & . & . & . & 0 \\ 0 & \sqrt{\gamma_{i2}} & . & . & . & 0 \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ 0 & 0 & . & . & & \sqrt{\gamma_{im}} \end{bmatrix} \quad (7)$$

where, γ_{ik} represents the spatial weight given to k^{th} area with respect to i^{th} area. It is expected that as distance between i^{th} and k^{th} area increases, the weight assigned by $\mathbf{W}_i$ *i.e.* γ_{ik} decreases. This means spatial effect decreases with distance and distant areas are likely to be more heterogeneous. Here, three types of weighting approaches are used to incorporate the spatial effects in the model. In first approach, (Neighbourhood criteria method) for i^{th} area, the element γ_{ik}, of the weight matrix $\mathbf{W}_i$ assigned weight as

$$\gamma_{ik} = \begin{cases} r_i^{-1} & \text{if i and k are contiguous} \\ 0, & \text{otherwise.} \end{cases}$$

where, r_i is the total number of areas that share an edge with area i (including area i itself). Further, it is also possible to introduce location effect through γ_{ik} with a continuous function such as Gaussian distance-decay based weighting as

$$\gamma_{ik} = \exp\left(-d_{ik}^2 / 2r_0^2\right), \; r_0 = \max(d_{ik})$$

Here, the value of the weight would decay gradually with distance, to the extent that when $d_{ik} = r_0$ the weighting would be 0.5. The simplest way of denoting the spatial dependence of these areas is the use of authorized variogram function. In this approach, elements of $\mathbf{W}_i$ *i.e.* γ_{ik} are obtained by using following equations.

$$\gamma_{ik} = 1 - r(d_{ik})$$

This also preserves the symmetric nature of weight matrix. Main authorized variogram functions are exponential, spherical and Gaussian variograms. Spherical variogram function is defined as

$$r(d_{ik}) = \begin{cases} \left\{\left(\dfrac{3d_{ik}}{2r_m} - \dfrac{1}{2}\left(\dfrac{d_{ik}}{r_m}\right)^3\right)\right\}, & \text{if } d_{ik} < r_0 \\ 1, & \textit{otherwise} \end{cases}$$

Premultipling $\mathbf{1}^T\mathbf{L}_i$ in equation (5), one can get the following equation

$$\mathbf{1}^T\mathbf{L}_i\boldsymbol{v} = \mathbf{1}^T\mathbf{L}_i\,\mathbf{Z}\gamma + \mathbf{1}^T\mathbf{L}_i\in \quad (8)$$

where, $\mathbf{1}$ denote the vector of one's. Further, above equation can be written as

$$v_i^* = z_i^{*T}\delta_i + \in_i^* \quad (9)$$

where,

$$v_i^* = \sqrt{\gamma_{i1}}v_1 + \sqrt{\gamma_{i2}}v_2 + ... + \sqrt{\gamma_{im}}v_m,$$

$$z_i^{*T}\delta_i = \sqrt{\gamma_{i1}}\mathbf{Z}_1^T\delta_1 + \sqrt{\gamma_{i2}}\mathbf{Z}_2^T\delta_2 + ... + \sqrt{\gamma_{im}}\mathbf{Z}_m^T\boldsymbol{\delta}_m$$

and

$$\in_i^* = \sqrt{\gamma_{i1}}\in_1 + \sqrt{\gamma_{i2}}\in_2 + ... + \sqrt{\gamma_{im}}\in_m$$

Thus, for i^{th} area, the model (2) can be written as

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + v_i^*\mathbf{1}_{n_i} + \boldsymbol{e}_i, \; j = 1, 2, ... n_i; \; i = 1, 2, ..., m \quad (10)$$

Putting the value of v_i^* from (9) above model can be written as

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + z_i^{*T}\delta_i\mathbf{1}_{n_i} + \in_i^*\mathbf{1}_{n_i} + \boldsymbol{e}_i,$$
$$j = 1, 2, ..., n_i; \; i = 1, 2, ..., m \quad (11)$$

For simplicity in further analysis, drop the star sign (*) from above equation and thus model (11) can be written as

$$\boldsymbol{Y}_i = \boldsymbol{x}_i\boldsymbol{\beta} + z_i^T\delta_i\mathbf{1}_{n_i} + \in_i\mathbf{1}_{n_i} + \boldsymbol{e}_i,$$
$$j = 1, 2, ..., n_i; \; i = 1, 2, ..., m \quad (12)$$

Unit level model is given by

$$y_{ij} = \boldsymbol{x}_{ij}^T \beta + z_i^T \delta_i + \in_i + e_{ij},$$
$$j = 1, 2, ..., n_i;\ i = 1, 2, ..., m \quad (13)$$

3. HIERARCHICAL BAYES ESTIMATION

Let us consider the model (13) in hierarchical Bayes (HB) framework and can be written as

$$y_{ij} \mid \beta, \delta_i, \in_i, \sigma_e^2 \sim \mathrm{N}(\boldsymbol{x}_{ij}^T \beta + z_i^T \delta_i + \in_i, \sigma_e^2),$$
$$j = 1, 2, ...\ n_i;\ i = 1, 2, ..., m$$
$$\in_i \mid \sigma_\in^2 \sim \mathrm{N}(0, \sigma_\in^2),\ i = 1, 2, ..., m$$

and marginal prior distributions of the parameter will be

$$\beta \sim \mathrm{N_p}(\mathbf{0}, \mathrm{C}), \delta \sim \mathrm{N_{mq}}(\mathbf{0}, \mathrm{D}),$$

$$\sigma_e^2 \sim \mathrm{IG}(a_0, b_0) \text{ and } \sigma_\in^2 \sim \mathrm{IG}(a_1, b_1)$$

Here, a_0, b_0, a_1, b_1 are known and IG(a_0, b_0) denotes the inverted gamma distribution with shape and scale parameters a_0 and b_0 respectively. Under the HB approach, the posterior mean $E(\bar{Y}_i \mid \{y_{ij}\})$ has been used as point estimate for mean and the posterior variance $V(\bar{Y}_i \mid \{y_{ij}\})$ as a measure of variability. In order to estimate $E(\bar{Y}_i \mid \{y_{ij}\})$ and $V(\bar{Y}_i \mid \{y_{ij}\})$, the Gibbs sampling method given by Gelfand and Smith (1990) has been used. Full conditional distributions for the Gibbs sampler have been obtained as follows

$$(\beta \mid \delta, \in_i, \sigma_\in^2, \sigma_e^2, \boldsymbol{Y}) \sim \mathrm{N_p}\left(\frac{1}{\sigma_e^2}\left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\sigma_e^2} + \boldsymbol{C}^{-1}\right)^{-1} \times \boldsymbol{X}^T(\boldsymbol{Y} - \boldsymbol{UZ}\delta + \boldsymbol{U} \in), \left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\sigma_e^2} + \boldsymbol{C}^{-1}\right)^{-1}\right),$$

$$(\delta \mid \beta, \in_i \sigma_\in^2, \sigma_e^2, \boldsymbol{Y}) \sim \mathrm{N_{mq}}\left(\frac{1}{\sigma_e^2}\left(\frac{\boldsymbol{Z}^T\boldsymbol{U}^T\boldsymbol{UZ}}{\sigma_e^2} + \boldsymbol{D}^{-1}\right)^{-1} \boldsymbol{Z}^T\boldsymbol{U}^T(\boldsymbol{Y} - \boldsymbol{S}), \left(\frac{\boldsymbol{Z}^T\boldsymbol{U}^T\boldsymbol{UZ}}{\sigma_e^2} + \boldsymbol{D}^{-1}\right)^{-1}\right),$$

$$(\in_i \mid \beta, \delta_i, \sigma_\in^2, \sigma_e^2, \boldsymbol{Y}_i) \sim \mathrm{N}\left(\left(n_i + \frac{\sigma_e^2}{\sigma_\in^2}\right)^{-1} \mathbf{1}_{n_i}^T(\boldsymbol{Y}_i - \boldsymbol{Q}_i), \left(\frac{n_i}{\sigma_e^2} + \frac{1}{\sigma_\in^2}\right)^{-1}\right),$$

$$\sigma_e^2 \mid \beta, \delta, \in_i, \sigma_\in^2, \boldsymbol{Y}) \sim \mathrm{IG}[a_0 + \frac{n}{2}, b_0 + \frac{1}{2}[\boldsymbol{Y} - (\boldsymbol{X}\beta + \boldsymbol{UZ}\delta + \boldsymbol{U} \in)]^T[\boldsymbol{Y} - (\boldsymbol{X}\beta + \boldsymbol{UZ}\delta + \boldsymbol{U} \in)]],$$

and $(\sigma_\in^2 \mid \beta, \delta, \in_i, \sigma_e^2, Y) \sim \mathrm{IG}(\left(a_1 + \frac{m}{2}\right), b_1 + \frac{1}{2} \in^T \in)$

where,

$\boldsymbol{Q}_i = \boldsymbol{X}_i\beta + \boldsymbol{Z}_i^T \delta_i \mathbf{1}_{n_i}$,

$\boldsymbol{S} = \boldsymbol{X}\beta + \mathbf{U}\in$, ($j$ = 1, 2, ..., n_i) and (i = 1, 2, ..., m)

$$\mathbf{U} = \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0} & . & . & . & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{n_2} & . & . & . & \mathbf{0} \\ . & . & . & & & . \\ . & . & & & . & . \\ . & . & & & . & . \\ \mathbf{0} & \mathbf{0} & . & . & & \mathbf{1}_{n_m} \end{bmatrix}_{n \times m}$$

$$\mathbf{Z} = \begin{bmatrix} \boldsymbol{Z}_1^{*T} & \mathbf{0} & . & . & . & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Z}_2^{*T} & . & . & . & \mathbf{0} \\ . & . & . & & & . \\ . & . & & & . & . \\ . & . & & & . & . \\ \mathbf{0} & \mathbf{0} & . & . & & \boldsymbol{Z}_m^{*T} \end{bmatrix}_{m \times mq}$$

It is important to note that proper priors have been used for all the unknown parameters to ensure that all the posterior distributions are proper (Hobert and Casella 1996). Here, values for the parameters of the priors (*i.e.*, hyperparameters) were chosen very small to reflect a fairly vague knowledge of the prior distributions. Now, it is straight forward to draw samples from these full conditional distributions as they have closed-form. After a "burn-in" period of B = 5000 iterations, next 5000 samples were considered to obtain $\beta^{(k)}, \delta_i^{(k)}$ (i = 1, 2 , ..., m), $\in_i^{(k)}$ (i = 1, 2, ...m), $\sigma_e^{2(k)}$ and $\sigma_\in^{2(k)}$, k = 1, 2, ...,d}. The prior parameters of inverted gamma distributions were also chosen to be very small.

Conditional on $\{y_{ij}\}, \beta, \delta, \sigma_e^2$ and $\sigma_\in^2$ the posterior distribution of $\in_i$ from the unit level model can be obtained as

$$(\in_i \mid \beta, \delta, \sigma_\in^2, \sigma_e^2, \boldsymbol{Y}) \sim \mathrm{N}(\left(n_i + \frac{\sigma_e^2}{\sigma_\in^2}\right)^{-1} \mathbf{1}_{n_i}^T(\boldsymbol{Y}_i - \boldsymbol{X}_i\beta$$

$$+\boldsymbol{Z}_i^T\delta_i\mathbf{1}_{n_i})),(\frac{n_i}{\sigma_e^2}+\frac{1}{\sigma_\in^2})^{-1})$$

Hence it can be seen that,

$$E(\in_i|\,\beta,\delta,\sigma_\in^2,\sigma_e^2,\boldsymbol{Y}_i) = \hat{\in}_i\,(\beta,\delta,\sigma_\in^2,\sigma_e^2,\boldsymbol{Y}_i)$$

$$= \left(n_i+\frac{\sigma_e^2}{\sigma_\in^2}\right)^{-1}\mathbf{1}_{n_i}^T(\boldsymbol{Y}_i-(\boldsymbol{X}_i\beta+\boldsymbol{Z}_i^T\delta_i\mathbf{1}_{n_i}))$$

Thus the conditional posterior mean of $\boldsymbol{Y}_i$ is

$$E(\bar{Y}_i\,|\,\beta,\delta,\sigma_\in^2,\sigma_e^2,\boldsymbol{Y}_i) = \hat{\bar{\boldsymbol{Y}}}_i(\beta,\delta,\sigma_\in^2,\sigma_e^2)$$
$$= \bar{\boldsymbol{X}}_i^T\hat{\beta}+\boldsymbol{Z}_i^T\delta_i+\hat{\in}_i\,(\beta,\delta,\sigma_\in^2,\sigma_e^2,\boldsymbol{Y}_i)$$

Similarly, the conditional posterior variance of $\bar{Y}_i$ is given by

$$V(\hat{\bar{Y}}_i\,|\,\beta,\delta,\sigma_\in^2,\sigma_e^2,\boldsymbol{Y}_i) = \hat{V}_i(\beta,\delta,\sigma_\in^2,\sigma_e^2)$$
$$= \left(\frac{n_i}{\sigma_e^2}+\frac{1}{\sigma_\in^2}\right)^{-1}$$

Now, using the simulated samples of $\{\beta^{(k)}$, $\delta_i^{(k)}$ (i = 1, 2, ..., m), $\in_i^{(k)}$ (i = 1, 2, ..., m), $\sigma_e^{2(k)}$ and $\sigma_\in^{2(k)}$, k = 1, 2, ..., $d\}$ the unconditional posterior mean and posterior variance of $\bar{Y}_i$ were estimated. The posterior mean $E(\bar{Y}_i\,|\{y_{ij}\})$ is estimated as

$$\hat{\bar{Y}}_i^{HB} = \frac{1}{d}\Sigma_{k=1}^d\hat{\bar{Y}}_i(\beta^{(k)},\delta_i^{(k)},\sigma_\in^{2(k)},\sigma_e^{2(k)}) \quad (14)$$

Similarly, the postierior variance $V(\bar{Y}_i\,|\{y_{ij}\})$ is estimated as

$$\hat{V}^{HB}(\bar{Y}_i) = \frac{1}{d}\Sigma_{k=1}^d\hat{V}_i(\beta^{(k)},\delta_i^{(k)},\sigma_\in^{2(k)},\sigma_e^{2(k)})$$
$$+\frac{1}{d}\Sigma_{k=1}^d\left[\hat{\bar{Y}}_i(\beta^{(k)},\delta_i^{(k)},\sigma_\in^{2(k)},\sigma_e^{2(k)})\right]^2$$
$$-\left[\frac{1}{d}\Sigma_{k=1}^d\hat{\bar{Y}}_i(\beta^{(k)},\delta_i^{(k)},\sigma_\in^{2(k)},\sigma_e^{2(k)})\right]^2 \quad (15)$$

In following section, a comparison of HB estimates with corresponding spatial EBLUP (SEBLUP) estimates is made for spatial unit level small area model through simulation study.

4. SIMULATION STUDY

In this study, spatial population structure has been generated assuming mean value of the dependent variable y fixed for a given area which is located at the centroid of the population in the map. Since, total 15 areas have been considered in the population, therefore, mean value of the other 14 areas for the dependent variable y has been generated assuming spatial pattern based on distance from this area. In this pattern, the values of y depend on distance from the centroid. Further, the mean values of auxiliary variable x have been generated using mean values of y and bivariate normal distribution, keeping the value of correlation coefficient between mean values of y and mean values of x fixed, *i.e.* $\rho = 0.7$. In order to generate the unit level data for each area, the bivariate normal population has been assumed and coefficient of variation has been fixed at 15% for y and x in each of the small areas. Further, values of other area level covariate *i.e*. z, have been generated following the similar approach using mean values of y and fixed correlation coefficient *i.e*. $\rho = 0.8$ between y and z. It may be noted that, number of units in each of the small area are different, ranging from 335 to 430 with total population of 5945 units.

The sample size n_i (i = 1, 2, ... ,15) is random within each area ranges from 2 to 5. In each simulation run it sums up to 50 for the whole population under study. Sampling units were selected with simple random sampling without replacement. The unit level model is given by

$$y_{ij} = \beta_0 + x_{ij1}\beta_1 + z_i\delta_i + \in_i + e_{ij},\ j = 1, 2, ..., n_i;\quad i = 1, 2, ..., m \quad (16)$$

In order to implement the Gibbs sampler based on above model (16), the following priors were assumed for different parameters $\beta_0 \sim N(0, 10^4)$, $\beta_1 \sim N(0, 10^4)$, $\delta \sim N_{15}(0,\mathbf{D})$, $\sigma_\in^2 \sim IG(0.001,0.001)$, $\sigma_e^2 \sim IG(0.001,0.001)$, where $\mathbf{D}$ is the matrix with diagonal elements of the order 10^4 and non-diagonal elements of order 10^3. After a burn-in period of 5000 iterations, next 5000 samples were kept for $\{\beta_0^{(k)},\beta_1^{(k)},\delta^{(k)},\in_i^{(k)}$ (i = 1, 2, ...,m), $\sigma_e^{2(k)}$ and $\sigma_\in^{2(k)}$, k = 1, 2, ..., $d\}$. $\delta_i^{(k)}$ was obtained by multiplying $\delta^{(k)}$ with vector (0, 0, 0, ..., 1, 0, 0, ...,0) 1 in the i^{th} place and 0 everywhere else. Gibbs sampler for the model (16) was implemented using WinBugs software. The WinBugs program constructs the necessary full conditional distributions and carries out the Gibbs sampling. Prior and initial values were generated using this software. Gibbs sampler was first run for a burn-in period of 5000 iterations and then 5000 more iteration were run and kept for analysis and estimation. This was done for all three different weighting approaches *i.e.* neighbourhood criteria, gaussian-decay and spherical variogram approach and posterior mean and posterior variances were obtained. Posterior means are almost same for all three approaches. Following figure shows posterior variances of Hierarchical Bayesian estimates for three different weighting approaches.

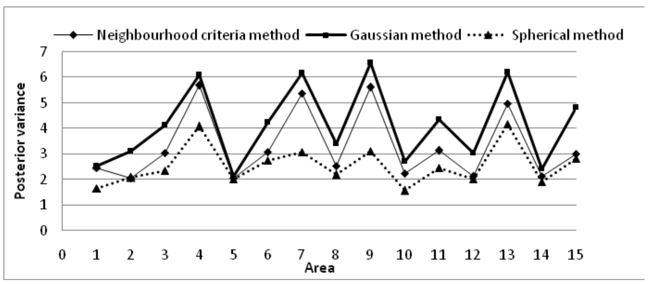

Fig 1. Posterior variances of Hierarchical Bayes estimates for spatial unit level small area model for three different weighting approaches.

Fig. 1 shows that posterior variances are different for all three weighting approaches. Variances are least for all small areas in case of spherical variogram approach as compare to other two methods. It can easily be seen that Neighbourhood criteria method performs well as compare to Gaussian decay function method of incorporating spatial-effects. Table 1 shows the % bias and %gain in efficiency of SHB estimates with respect to SEBLUP estimates of spatial small area models.

where,

$$\text{Percentage bias} = \left[(\hat{\bar{Y}}_i^{SEBLUP} - \hat{\bar{Y}}_i^{SHB}) / \hat{\bar{Y}}_i^{SEBLUP}\right] \times 100$$

Percentage gain in efficiency

$$= \left[(\hat{V}_i^{SEBLUP} - \hat{V}_i^{SHB}) / \hat{V}_i^{SEBLUP}\right] \times 100$$

The Table 1 shows that when sample size is more, % bias is very less and it increases with decrease in the sample size. % bias is almost equal in all the three weighting approaches. It also shows that there is % gain in efficiency in HB approach with respect to EBLUP approach for all the three weighting methods. The % gain in efficiency is highest in case of neighbourhood criteria method as compared to its counterparts.

Table 1. Percentage bias and percentage gain in efficiency of Spatial HB (SHB) estimates with respect to the Spatial-EBLUP (SEBLUP) estimates for three weighting approaches.

Area	Sample size	Neighbourhood criteria method		Gaussian method		Spherical method	
		% bias	% gain in efficiency	% bias	% gain in efficiency	% bias	%gain in efficiency
1	5	–0.30	3.83	–0.30	0.30	–0.29	1.46
2	4	–12.27	32.33	–12.30	5.58	–12.37	22.43
3	3	–6.50	0.58	–6.50	0.76	–6.56	1.10
4	2	–11.70	0.91	–11.69	11.79	–11.69	37.07
5	4	–1.56	10.17	–1.53	26.00	–1.39	38.86
6	3	5.89	14.88	5.90	1.76	5.96	29.34
7	2	–11.22	3.15	–11.20	3.82	–11.19	6.73
8	4	–6.32	34.27	–6.32	7.73	–6.40	25.19
9	2	–2.18	11.76	–2.21	4.37	–2.34	14.49
10	5	–9.48	2.27	–9.50	11.57	–9.55	38.57
11	3	9.37	19.22	9.34	8.06	9.25	48.06
12	4	1.23	26.46	1.23	0.13	1.29	29.24
13	2	–0.43	7.11	–0.45	0.44	–0.50	37.32
14	4	–0.16	1.03	–0.15	19.21	–0.03	43.52
15	3	–8.63	21.13	–8.68	2.14	–8.80	42.30

Sensitivity Analysis: In practice, it is always difficult to obtain accurate information about the distribution of the variances. Here, inverted gamma distribution was assumed on variance components. Now, the interest is to know the effects caused by the choice of different priors. Basically, the sensitivity of posterior means to the choice of different priors on the variance components was evaluated to understand this effect. In order to see the sensitivity of the posterior estimates to the choice of a_i and b_i, different values of these parameters *i.e.* 0.001, 0.01, 0.1, 1 and 10 have been set. Following figures show the posterior variance for different gamma values for three different weighting approaches.

Sensitivity analysis shows that posterior estimates of small area means and variances (Fig 2, 3, 4) are almost same when a_i and b_i are small (< = 0.1). This indicates small area mean estimates and variances are very much stable for small values of a_i and b_i, as there is negligible difference among the estimates. But as a_i and b_i increase, small area means are stable but posterior variances are decreasing rapidly. This indicates that posterior variances decrease as the priors become more informative. You and Rao (2000) also showed that as a_i and b_i increase, the posterior estimates of the variances decrease and there is no difference among the estimates for very small values of parameters. So, if there is strong prior information on a_i and b_i (for example 1 or 10), the posterior estimates of the variance components will be significantly different from non-informative priors. Thus small values of parameters can be considered as non-informative priors.

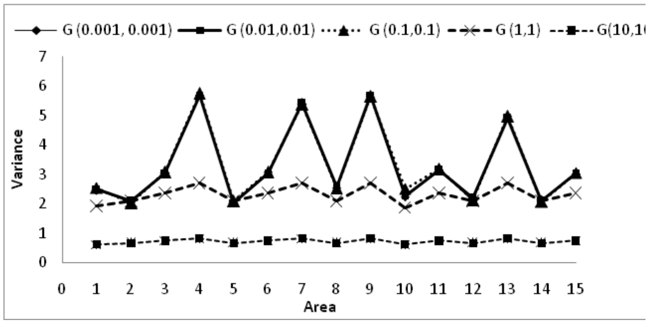

Fig 2. Posterior variances of Hierarchical Bayes estimates for neighbourhood criteria method.

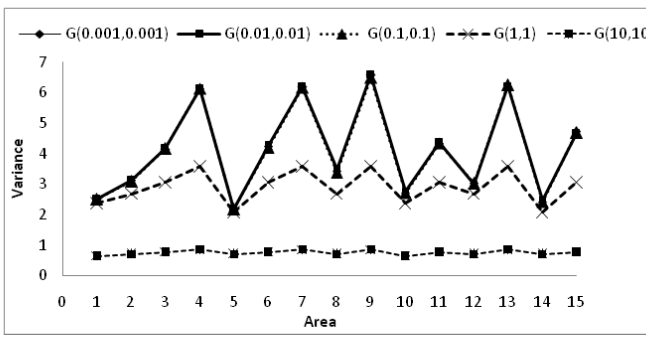

Fig 3. Posterior variances of Hierarchical Bayes estimates for Gaussian method.

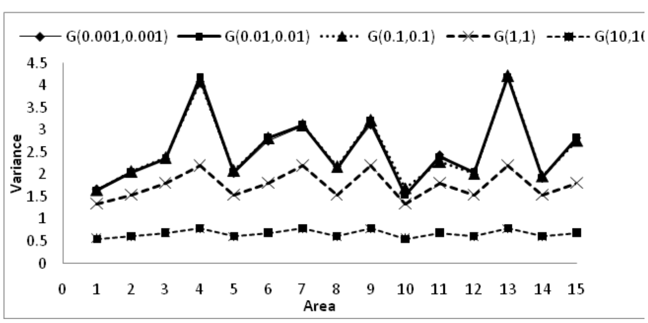

Fig 4. Posterior variances of Hierarchical Bayes estimates for Spherical method.

5. CONCLUSION

In this study, attempt has been made to obtain small area estimates for spatial unit level model under hierarchical Bayes (HB) framework. Three different weighting approaches were used to incorporate the spatial effects in the model in the form of elements in the weight matrix. Study shows that posterior mean for all three weighting approaches are almost same in all small areas but variances are different. Variances are least for all small areas in case of spherical variogram approach as compare to other two methods. It can easily be seen that neighbourhood criteria method performs well as compare to Gaussian-decay function method of incorporating spatial effects.

Further, it shows that when sample size is more, %bias is very less and it increases with decrease in the sample size. %bias is almost equal in all the three weighting approaches. Although, it also shows that there is % gain in efficiency in spatial-HB (SHB) approach with respect to SEBLUP approach for all the three weighting methods but neighbourhood criteria of incorporating weights seems to have more impact in gain in efficiency. Sensitivity analysis shows that posterior estimates of small area means and variances are almost same when parameters of inverted gamma distribution which are assumed for variances, are small (< = 0.1). But as the value of these parameters increase, small area means are stable but posterior variances decrease rapidly. This indicates that posterior variances decrease as the priors become more informative.

ACKNOWLEDGEMENTS

Authors are grateful to the Associate Editor and referees for their valuable comments.

REFERENCES

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *J. Amer. Statist. Assoc.*, **83**, 28–36.

Brunsdon, C., Fotheringham, A.S. and Charlton, M.E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Anal.*, **28**, 281-298.

Brunsdon, C., Fotheringham, S. and Charlton, M. (1998). Geographically weighted regression-modelling spatial non-stationary. *The Statistician*, **47(3),** 431-443.

Datta, G.S. and Ghosh, M. (1991). Bayesian prediction in linear models: applications to small area estimation. *Ann. Statist.,* **19,** 1748-1770.

Datta, G.S., Day, B. and Maiti, T. (1998). Multivariate bayesian small area estimation: An application to survey and satellite data. *Sankhya: The Indian Journal of Statistics,* Special Issue on Bayesian Analysis, **A60**, 344-362.

Fotheringham, A.S. Brunsdon, C., and Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons, Chichester, UK.

Fuller, W.A. and Harter, R.M. (1987). The multivariate components of variance model for small area estimation. In: R. Platek, J.N.K Rao, C.E. Sarndal and M.P. Singh (Eds.), *Small Area Statistics*, Wiley, New York, pp. 103-123.

Gelfand, A.E. and Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.*, **85**, 398-409.

Ghosh, M. and Rao, J.N.K. (1994). Small area estimation: an appraisal (with discussion). *Statistical Science*, **9(1)**, 55-93.

Hobert, J.P. and Cassella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *J. Amer. Statist. Assoc.*, **91**, 1461-1473.

Moura, F. and Holt, D. (1999). Small area estimation using multilevel models. *Survey Methodology*, **25**, 73-80.

Pfeffermann, D. (2002). Small area estimation - New developments and directions. *Intern. Statist. Rev.*, **70**, 125-143.

Pratesi, M. and Salvati, N. (2008). Small area estimation: the EBLUP estimator based on spatially correlated random area effects. *Statist. Methods Appl.*, **17**, 113-141.

Rao, J.N.K. (2003): *Small Area Estimation*. Wiley, London.

Salvati, Nicola, Tzavidis, Nikos, Pratesi Monica and Chambers, Ray (2012). Small area estimation via M-quantile Geographically Weighted Regression. *Test*, **21(1)**, 1-28.

Singh, B., Shukla, G. and Kundu, D. (2005). Spatio-temporal models in small area estimation. *Survey Methodology,* **31**,183-195.

You, Y. and Rao, J.N.K. (2000). Hierarchical bayes estimation of small area means using multi-level models. *Survey Methodology,* **26**, 173-181.

You, Y. and Rao, J.N.K. (2003). Pseudo hierarchical Bayes small area estimation combining unit level models and survey weights. *J. Statist. Plann. Inf.*, **111**, 197-208.

You, Y. and Chapman, B. (2006). Small area estimation using area level models and estimated sampling variances. *Survey Methodology,* **32(1)**, 97-103.