

Hierarchical Bayes Small Area Estimation under a Unit Level Model with Applications in Agriculture

Nageena Nazir

Division of Agri-statistics, SKUAST-K, Shalimar, India
nazir.nageena@gmail.com

Shakeel Ahmad Mir

Division of Agri-statistics, SKUAST-K, Shalimar, India
samir@gmail.com

M Iqbal Jeelani

Division of Agri-statistics, SKUAST-K, Shalimar, India
iqbaljeelani@gmail.com

Abstract

To studied Bayesian aspect of small area estimation using Unit level model. In this paper we proposed and evaluated new prior distribution for the ratio $\sigma_v^2 / \sigma_e^2 (= \lambda)$ of variance components in unit level model rather than uniform prior. To approximate the posterior moments of small area means, Laplace approximation method is applied. This choice of prior avoids the extreme skewness, usually present in the posterior distribution of variance components. This property leads to more accurate Laplace approximation. We apply the proposed model to the analysis of horticultural data and results from the model are compared with frequentist approach and with Bayesian model of uniform prior in terms of average relative bias, average squared relative bias and average absolute bias. The numerical results obtained highlighted the superiority of using the proposed prior over the uniform prior. Thus Bayes estimators (with new prior) of small area means have good frequentist properties such as MSE and ARB as compared to other traditional methods viz., Direct, Synthetic and Composite estimators.

Keywords: Small Area Estimation, Unit Level Model, Hierarchical Bayes.

1 Introduction

Model-based small area estimation methods have been widely used in practice due to the increasing demand for precise estimates for local regions and various small areas. It is now generally accepted that the indirect estimates should be based on explicit models that provide links to related areas through the use of supplementary data such as census counts or administrative records; see, for example, (Jiang and Lahiri 2006) and (Rao 2003) for more discussion on model-based small area methods. Also, (Adam et. Al., 2013) summarise the main methodological approaches to SAE and their linkages. [Jiango et. Al., 2013] investigate two new approaches: one relying on the work of Chambers, and the second using the concept of conditional bias to measure the influence of units in the population. [Chambers et. al., 2014] proposed two different analytical mean-squared error estimators for the ensuing bias-corrected outlier robust estimators. [Rao et al., 2013] relaxed the assumption of linear regression for the fixed part of the model and replace it by a weaker assumption of a semi-parametric regression. The model-based estimates are obtained to improve the direct design-based estimates in terms of precision and reliability, *i.e.*, smaller coefficients of variation (CVs). There are two broad classifications for small area models: area level models and unit level models. The basic

unit level model is based on unit level auxiliary variables. In this paper we focus on Unit level models that are related to the unit level values of response through a nested error linear regression model, under the assumption that the nested error and the model error are independent of each other and normally distributed with common mean zero and common or different variances. The nested error unit level regression model was first used to model county crop areas in USA (Battese *et al.*, 1988), they have used the normally distributed common errors variance assumption and revealed that based on the fitting-of-constants method the estimates of errors variances are slightly different from each other. Techniques for validating their model on the basis of unit level auxiliary variables are also considered. This type of model is appropriate for continuous value response variables. Various extensions of this type of model have been proposed to handle binary responses, two-stage sampling within areas, multivariate responses and others (Rao, 2003, 2003a).

The objective of this paper is to consider new improved prior on hyperparameters of variance component $\lambda = \sigma_v^2 / \sigma_e^2$ of Unit Level model and illustrate the usefulness of this models through an application to horticultural survey data. The paper is organized as follows. In section 2, we first study Unit level models including EBLUP estimators of unit level model. Then in section 3 we propose Bayesian formulation of urea level model with new prior for variance component $\lambda = \sigma_v^2 / \sigma_e^2$ and obtain HB inference for small area parameters through the MCMC method using Laplace approximation. In section 4, we apply the proposed model to the analysis of small area data from the Horticultural Survey. We compare the performance the proposed model i.e hierarchical Bayes estimate with the proposed prior (Proposed *HB*) with the hierarchical Bayes estimate with uniform prior (HB(WINBUG)) and EBLUP estimates to investigate the effects of incorporating new prior on the variance component $\lambda = \sigma_v^2 / \sigma_e^2$. Bayesian model comparison and model fit analysis are also provided. Finally in section 5, we offer some concluding remarks.

2 Unit level model

The basic unit level model is based on unit level auxiliary variables. This type of model can be represented by the following mathematical equation

$$y_{ij} = x_{ij}' + v_i + e_{ij}; j = 1, \dots, N_i, i = 1, \dots, m, \quad (2.1)$$

where y_{ij} is the value of the study variable for the j^{th} unit belonging to the i^{th} area, x_{ij} is the unit level $p \times 1$ vector of known covariates, β is a $p \times 1$ vector of unknown regression coefficients, v_i is the random small area effect, and e_{ij} is the error term. We assume $v_i \sim N(0, \sigma_v^2)$ and $e_{ij} \sim N(0, \sigma_e^2)$

2.1 EBLUP estimator of unit level model

For the basic unit level model defined in section (2) the BLUP estimator of μ_i is given as:

$$\tilde{\mu}_i = X_i' \tilde{\beta} + \gamma_i (\bar{y}_{ia} - \bar{x}_{ia}' \tilde{\beta}) \quad (2.1.1)$$

$$\text{and } \gamma_i = \frac{\sigma_v^2}{(\sigma_v^2 + \sigma_e^2 / a_i)} \quad (2.1.2)$$

$$\text{where } \bar{y}_{ia} = \sum a_{ij} y_{ij} / a_i, \bar{x}_{ia} = \sum a_{ij} x_{ij} / a_i, \tilde{\beta} = \left[\sum_i X_i' V_i^{-1} X_i \right]^{-1} \left[\sum_i X_i' V_i^{-1} y_i \right] \quad (2.1.3)$$

where

$$X_i' V_i^{-1} X_i = A_i = \sigma_e^{-2} \left(\sum a_{ij} x_{ij} x_{ij}' - \gamma_i a_i \bar{x}_{ia} \bar{x}_{ia}' \right) \quad X_i' V_i^{-1} y_i = \sigma_e^{-2} \left(\sum a_{ij} x_{ij} y_{ij} - \gamma_i a_i \bar{x}_{ia} \bar{y}_{ia}' \right) \quad (2.1.4)$$

The BLUP estimator (2.1.1) depends on the variance ratio σ_v^2 / σ_e^2 which is unknown in practice. Replacing σ_v^2 and σ_e^2 by estimators $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$ we obtain an EBLUP estimator (Jiang & Lahiri 2006).

$$\hat{\mu}_i = X_i' \hat{\beta} + \hat{\gamma}_i (\bar{y}_{ia} - \bar{x}_{ia}' \hat{\beta}) \quad (2.1.5)$$

Where $\hat{\gamma}_i$ and $\hat{\beta}$ are the values of γ_i and β when σ_v^2 and σ_e^2 are replaced by $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$. The MSE of the BLUP estimator $\tilde{\mu}_i$ can be obtained from the general results of MSE of mixed effects model or by direct calculation. It is given by

$$MSE(\tilde{\mu}_i) = E(\tilde{\mu}_i - \mu_i)^2 = g_{1i}(\sigma_v^2, \sigma_e^2) + g_{2i}(\sigma_v^2, \sigma_e^2) \quad (2.1.6)$$

where

$$g_{1i}(\sigma_v^2, \sigma_e^2) = \gamma_i(\sigma_e^2 / a_i) \quad \text{and} \quad g_{2i}(\sigma_v^2, \sigma_e^2) = (\bar{X}_i - \gamma_i \bar{x}_{ia})' \left(\sum_i A_i \right)^{-1} (\bar{X}_i - \gamma_i \bar{x}_{ia}) \quad (2.1.7)$$

The second order MSE approximation i.e EBLUP estimate of $\tilde{\mu}_i$ is

$$MSE(\hat{\mu}_i) = g_{1i}(\sigma_v^2, \sigma_e^2) + g_{2i}(\sigma_v^2, \sigma_e^2) + g_{3i}(\sigma_v^2, \sigma_e^2) \quad (2.1.8)$$

Where $g_{3i}(\sigma_v^2, \sigma_e^2) = a_i^{-2} (\sigma_v^2 + \sigma_e^2 / a_i)^{-3} h(\sigma_v^2, \sigma_e^2)$

$$h(\sigma_v^2, \sigma_e^2) = \sigma_e^4 \bar{V}_{vv}(\delta) + \sigma_v^4 \bar{V}_{ee}(\delta) - 2\sigma_e^2 \sigma_v^2 \bar{V}_{ve}(\delta) \quad \text{where } \delta = (\sigma_v^2, \sigma_e^2)',$$

Now regarding the estimation of $MSE(\hat{\mu}_i)$ the estimator of MSE is given by

$$mse(\hat{\mu}_i) = g_{1i}(\hat{\sigma}_v^2, \hat{\sigma}_e^2) + g_{2i}(\hat{\sigma}_v^2, \hat{\sigma}_e^2) + 2g_{3i}(\hat{\sigma}_v^2, \hat{\sigma}_e^2) \quad (2.1.9)$$

3 Bayesian formulation of unit level model

Our objective is to study the Bayesian approach in a unit level model for estimating the finite population means $\bar{Y}_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}; i=1, \dots, m$ for small areas. We propose a new prior for the variance component $\lambda (= \sigma_v^2 / \sigma_e^2)$ and use laplace method to approximate the posterior moments involved in the HB approach.

Suppose a sample s_i of size n_i is drawn from the N_i units of the i^{th} area following some specified sampling design; $i = 1, \dots, m$ and $\sum_{i=1}^m n_i = n$. Following Battese *et al.* (1988) and Rao (2003), we assume that the sample values also follow the model (3.4), that is, we do not address the situation involving informative sampling. The assumption that the sample values also obey the model (3.5.1) holds true for simple random sampling from each area. For a proof of this absence of selection bias (Rao, 2003).

We can also write the small area mean \hat{Y}_i as

$$\hat{Y}_i = \frac{1}{N_i} \left\{ \sum_{j \in s_i} y_{ij} + \sum_{j \notin s_i} y_{ij} \right\} = \frac{1}{N_i} \{ n_i \bar{y}_{is} + (N_i - n_i) \bar{y}_{ins} \} = f_i \bar{y}_{is} + (1 - f_i) \bar{y}_{ins} \quad (3.1)$$

where $f_i = n_i / N_i$, $1 - f_i$ is the finite population correction, y_{is} and y_{ins} are the means of the sampled and non-sampled units, respectively, for the i^{th} area. Now applying a hierarchical Bayesian approach to estimate the small area means (3.1) for finite populations. A hierarchical Bayesian version of the Unit level model is given by

- Level 1: $y_{ij} | \beta, v_i, \sigma_e^2 \sim N(x'_{ij} \beta + v_i, \sigma_e^2); j = 1, \dots, N_i, i = 1, \dots, m$
- Level 2: $v_i | \sigma_v^2 \sim N(0, \sigma_v^2)$

Our aim is to find the Bayesian estimator of small area means of unit level model and its measure of uncertainty, which are given by $E(\bar{Y} | y_i)$ and $V(\bar{Y} | y)$ respectively, where y is the vector of sampled values of y . Whether we are in Bayesian or frequentist paradigm depends on whether we assume a prior on the hyperparameters $(\beta, \sigma_v^2, \sigma_e^2)$ at the third level. To obtain the Bayesian summary statistics for Y_i , we proceed as follows (assuming that the sample values follow the same model above, as discussed in Section 2). First we write the likelihood using level 1 and level 2 of the hierarchical Bayesian model as:

$$L(y | \beta, v, \sigma_e^2, \sigma_v^2) \alpha(\sigma_e^2)^{-\frac{n}{2}} (\sigma_v^2)^{-\frac{m}{2}} \exp \left[-\frac{1}{2} \left\{ \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - x'_{ij} \beta - v_i)^2}{\sigma_e^2} + \frac{\sum_{i=1}^m v_i^2}{\sigma_v^2} \right\} \right] \quad (3.2)$$

Now, following Ghosh and Lahiri (1989), Datta and Ghosh (1991), we specify prior on σ_e^2 and λ at level 3 as: $\pi(\beta, \sigma_e^2, \lambda) \propto \pi(\sigma_e^2) \pi(\lambda)$ where $\lambda = \sigma_v^2 / \sigma_e^2$. This completes our hierarchical Bayesian model. Combining the likelihood (3.2) and the prior at level 3, we write the joint distribution of $y, \beta, v = (v_1, \dots, v_m), \sigma_e^2, \lambda$ as

$$f(y, \beta, v, \sigma_e^2, \lambda) \propto \alpha(\sigma_e^2)^{-\frac{n}{2}} \lambda^{-\frac{m}{2}} (\sigma_e^2)^{-\frac{m}{2}} = \exp \left[-\frac{1}{2} \left\{ \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - x'_{ij} \beta - v_i)^2}{\sigma_e^2} + \frac{\sum_{i=1}^m v_i^2}{\lambda \sigma_e^2} \right\} \right] p(\lambda) \quad (3.3)$$

In (3. 3), we consider a flat prior (uniform prior) on σ_e^2 i.e., we consider $\pi(\sigma_e^2) \propto 1, \sigma_e^2 \in [0, \infty]$.

Typically, enough data will be available to estimate σ_e^2 precisely. Any reasonable non-informative prior usually works well for σ_e^2 (Gelman, 2006).

After collecting terms involving v_i from the exponent of (3.7.3), we can say that:

$$v_i | y, \beta, \sigma_e^2, \lambda \sim N \left[(1 - B_i)(\bar{y}_{is} - \bar{x}_{is}'\beta), \frac{\sigma_e^2}{n_i}(1 - B_i) \right] \quad (3.4)$$

where $B_i = 1/(1 + n_i\lambda)$. Integrating out v_i from (3. 3), using matrix notation, we write the joint distribution (without the normalizing constant) of $y, \beta, \sigma_e^2, \lambda$ as

$$\begin{aligned} f(y, \beta, \sigma_e^2, \lambda) &\propto \prod_{i=1}^m \{(1 + n_i\lambda)^{-1/2}\} (\sigma_e^2)^{-n/2} p(\lambda) \\ &\exp \left[-\frac{1}{2\sigma_e^2} \sum \left\{ y_i' y_i + \beta' X_i' X_i \beta - 2y_i' X_i \beta - \frac{\lambda}{1 + \lambda n_i} (y_i' J_i y_i + \beta' X_i' J_i X_i \beta - 2\beta' X_i' J_i y_i) \right\} \right] \\ &= \prod_{i=1}^m \{(1 + n_i\lambda)^{-1/2}\} (\sigma_e^2)^{-n/2} p(\lambda) \exp \left[-\frac{1}{2\sigma_e^2} \{ (\beta - \hat{\beta}_\lambda)' \Sigma_\lambda^{-1} (\beta - \hat{\beta}_\lambda) + T_\lambda \} \right] \end{aligned} \quad (3.5)$$

where

$$T_\lambda = \sum_{i=1}^m y_i' \Sigma_i y_i - \left(\sum_{i=1}^m X_i' \Sigma_i y_i \right)' \left(\sum_{i=1}^m X_i' \Sigma_i X_i \right)^{-1} \left(\sum_{i=1}^m X_i' \Sigma_i y_i \right)$$

and from expression (2.1.4), it follows that

$$\beta | y, \sigma_e^2, \lambda \sim N_p \left[\hat{\beta}_\lambda, \sigma_e^2 \Sigma_\lambda \right] \text{ where } \beta = \left(\sum_{i=1}^m X_i' \Sigma_i X_i \right)^{-1} \left(\sum_{i=1}^m X_i' \Sigma_i y_i \right), \Sigma_\lambda = \left(\sum_{i=1}^m X_i' \Sigma_i X_i \right)^{-1} \quad (3.6)$$

The subscript λ in both $\hat{\beta}_\lambda$ and Σ_λ indicates the dependence of the terms on the ratio of variance component $\lambda = \sigma_v^2 / \sigma_e^2$ alone. But one should note that $V(\beta | y, \sigma_e^2, \lambda) = \sigma_e^2 \Sigma_\lambda$ depends on both λ and σ_e^2 . Integrating out β from (3.5), we find

$$f(y, \sigma_e^2, \lambda) = \prod \left\{ (1 + n_i\lambda)^{-1/2} \right\} (\sigma_e^2)^{-(n/2 - p/2)} |\Sigma_\lambda|^{1/2} \exp \left[-\frac{1}{2\sigma_e^2} T_\lambda \right] p(\lambda) \quad (3.7)$$

From the equation (3.7), it follows that

$$\sigma_e^2 | y_i, \lambda \sim IG \left(\frac{n - p - 2}{2}, \frac{1}{2} T_\lambda \right) \quad (3.8)$$

where $IG(a, b)$ is the inverse gamma distribution with shape parameter “a” and scale parameter “b”. A random variable Z has an inverse Gamma distribution $IG(a, b)$, if its pdf

is given by $f(z) \propto z^{-(a+1)} \exp[-b/z]; z > 0$. Using the property of inverse gamma distribution, we express the conditional posterior mean and variance of σ_e^2 given λ as

$$E(\sigma_e^2 | y, \lambda) = \frac{T_\lambda}{n - p - 4} \quad (3.9)$$

$$V(\sigma_e^2 | y, \lambda) = \frac{2T_\lambda^2}{(n - p - 4)^2 (n - p - 6)} \quad (3.10)$$

Integrating out σ_e^2 from (3.7.7), we can write the posterior distribution of a single variance component λ as

$$f(\lambda | y) = \prod_{i=1}^m \left\{ (1 + n_i \lambda)^{-1/2} \right\} |\Sigma_\lambda|^{1/2} T_\lambda^{-(n-p-2)/2} p(\lambda) \quad (3.11)$$

To present the results in a unified way, we restate (3. 11) using matrix notation as

$$f(\lambda | y) \propto |\Sigma|^{1/2} |X' \Sigma X|^{-1/2} (Y' Q_\lambda Y)^{-(n-p-2)/2} \quad (3.12)$$

where

$$\Sigma = \bigoplus_{i=1}^m \Sigma_i, X = \text{col}_{1 \leq i \leq m} \text{col}_{1 \leq j \leq n_i} x_{ij}, Y = \text{col}_{1 \leq i \leq m} \text{col}_{1 \leq j \leq n_i} y_{ij}, T_\lambda = Y' Q_\lambda Y, Q_\lambda = \Sigma - \Sigma X (X' \Sigma X)^{-1} X' \Sigma$$

This formulation is presented for the ease of writing codes in simulation and data analysis.

3.1 Posterior moments of finite population mean when λ known

We derive $E(\bar{Y}_i | y)$ and $V(\bar{Y}_i | y)$ in two steps. First, we find analytical expressions for these quantities when λ is known. Then using the Laplace approximation, we obtain the final results. Using the iterative expectation and variance technique and the results given in (3..4), (3. 6), and (3. 9), it is not very difficult to get

$$E(\bar{Y}_i | y, \lambda) = (1 - fpc_i B_i) \bar{y}_{is} + fpc_i (\bar{X}_{ins}' - (1 - B_i) \bar{x}_{is}) \hat{\beta}_\lambda = g_i(\lambda), \text{ say} \quad (3.1.1)$$

$$V(\bar{Y}_i | y, \lambda) = \frac{T_\lambda}{n - p - 4} \left[\frac{fpc_i}{N_i} + \frac{(fpc_i)^2 (1 - B_i)}{n_i} + (A_i - D_i)' \Sigma_\lambda (A_i - D_i) \right] = h_i(\lambda), \text{ say}, \quad (3.1.2)$$

where $A_i = fpc_i \bar{X}_{ins}$, $D_i = fpc_i B_i \bar{x}_{is}$, $A_i - D_i = fpc_i \{ \bar{X}_{ins} - (1 - B_i) \bar{x}_{is} \}$. $\hat{\beta}_\lambda$ and Σ_A are defined in (3. 6). Note that Σ_A is different from Σ in). Now let's consider some special case.

Common mean model, Balanced case, and $f_i \approx 0$

Here, besides the common mean assumption, we assume that $n_i = k, \forall_i, n = mk$ and hence $B_i = B = 1/(1 + k\lambda)$. Then the posterior distribution of λ is given by:

$$f(\lambda | y) \propto \frac{(1 + k\lambda)^{-(m-1)/2}}{SSW + B.SSB} \pi(\lambda)$$

where

$$SSW = \sum_i \sum_j (y_{ij} - \bar{y}_{is})^2 \text{ and } SSB = k \sum_i (\bar{y}_{is} - \hat{\mu})^2; \hat{\mu} = \frac{1}{m} \sum_i \bar{y}_{is}$$

are the usual definition of within sum of square and between sum of square used in the balanced one-way ANOVA. After some modification and simplification of (3.7.13) and (3.7.14), we find

$$E(\bar{Y}_i | y, \lambda) = (1 - B)\bar{y}_{is} + B\hat{\mu} \text{ and } V(\bar{Y}_i | y, \lambda) = \left(\lambda B + \frac{B}{n} \right) \left\{ \frac{SSW + B.SSB}{n - 5} \right\}$$

3.2 Choice of prior on λ

To obtain the posterior mean and variance of \bar{Y}_i from (3.1.1) and (3.1.2), we need to perform one-dimensional integral with respect to the posterior distribution of λ . For that we need to assume a prior distribution on λ . The uniform prior on λ is non-informative and yields a posterior distribution of λ for which the mode is identical to the residual maximum likelihood (REML) estimator of λ , following the arguments discussed above. The posterior mode of λ with uniform prior or equivalently, the REML estimator of λ can be zero for a particular application. In many practical applications, the maximum likelihood (ML) or restricted maximum likelihood (REML) estimates of hyperparameters occur at the boundary point. we match the posterior distribution of λ given in (3.11) to the adjusted profile likelihood function of λ to obtain an appropriate prior distribution of λ . This prior leads to a posterior distribution of λ for which the mode is always positive and results in estimators of the shrinkage factor and small area mean that have good frequentist properties. By profile likelihood, we mean the likelihood of λ that does not account for the loss of degrees of freedom due to the estimation of regression coefficient β . This is given by

$$L_p \propto \prod_{i=1}^m \{ (1 + n_i \lambda)^{-1/2} \} T_{\lambda}^{-(n-2)/2} \quad (3.2.1)$$

Note that, to obtain the profile likelihood (3.2.1), we consider β as the only nuisance parameter. We plug in the ML estimate of β in (3.5) [without the term $\pi(\lambda)$] to obtain the joint profile likelihood of λ and σ_e^2 . Then integrating out σ_e^2 , we derive the marginal profile likelihood of λ . This leads to the prior

$$\pi(\lambda) \propto \lambda \left| \sum_{i=1}^m (X_i' \Sigma_i X_i) \right|^{1/2} T_{\lambda}^{-p/2}; \lambda > 0 \quad (3.2.2)$$

To simplify the prior for some special cases, we proceed as follows:

$$\begin{aligned} \Sigma_i &= I_i - \frac{\lambda}{1 + \lambda n_i} J_i = (I_i - \bar{J}_i) + \bar{J}_i - \frac{\lambda}{1 + \lambda n_i} 1_i 1_i' = C_i + 1_i \left(\frac{1}{n_i} - \frac{\lambda}{1 + \lambda n_i} \right) 1_i' \\ &= C_i + \frac{1}{1 + \lambda n_i} \bar{J}_i \end{aligned}$$

where $\bar{J}_i = \frac{1}{n_i} J_i, C_i = I_i - \bar{J}_i$. Thus, we can write

$$X_i' \Sigma_i X_i = X_i' C_i X_i + \frac{1}{1 + \lambda n_i} X_i' \bar{J}_i X_i.$$

For the common mean model, balanced case (i.e. $p = 1, n_i = k \quad \forall i$),

$$X_i' C_i X_i = 0, \quad \frac{1}{1 + \lambda n_i} X_i' \bar{J}_i X_i = \frac{k}{1 + \lambda k} \quad \text{and} \quad T_\lambda = SSW + \frac{SSB}{1 + k\lambda}$$

where SSW and SSB is defined earlier. Hence, simplifying it further we can say

$$\pi(\lambda) \propto \frac{\lambda}{\sqrt{SSB + SSW(1 + \lambda k)}}.$$

3.3 Laplace approximation

The posterior moments of \bar{Y}_i , using the prior in (3.2.2) are not in a closed-form, but can be obtained either by numerical integration or by the Monte Carlo Markov Chain (MCMC) method (Spiegelhalter et al., 1997). We have written a program, using the R2JAGS package in R [R Development Core Team, 2008], that allows the hierarchical Bayes analysis for our new prior using the MCMC method. But the slow computation speed of the MCMC method does not permit its evaluation by repeated use in simulation. Thus, for convenient implementation and evaluation of our hierarchical Bayes method, we approximate the posterior moments of \bar{Y}_i using Laplace's method. This method has been applied by many authors in the context of Bayesian analysis. See [Butar and Lahiri 2002], [Kass Staffey 1989], [Tierney and Kadane 1986] [Tierney et al., 1989]. Following [Kass Staffey 1989] the first order approximation to the posterior mean and variance of θ_i under the area level model described above in section-2 is given by The first order Laplace approximation of the posterior moments of \bar{Y}_i can be given by

$$E(\bar{Y}_i | y) = g_i(\hat{\lambda}) + O(1/m) \quad (3.3.1)$$

$$V(\bar{Y}_i | y) = h_i(\hat{\lambda}) + \left\{ g_i'(\hat{\lambda}) \right\}^2 \frac{1}{i_0} + O(1/m^2) \quad (3.3.2)$$

Below we derive the analytical expressions of the posterior moments of \bar{Y}_i using equation (3.3.1) and (3.3.2) under some special cases.

Common mean model, Balanced case, and $f_i \approx 0$

$$E(\bar{Y}_i | y) = (1 - \hat{B})\bar{y}_{is} + \hat{B}\hat{\mu} + O(1/m)$$

$$V(\bar{Y}_i | y) = \left(\hat{\lambda}\hat{B} + \frac{\hat{B}}{n} \right) \left\{ \frac{SSW + \hat{B}.SSB}{n - 5} \right\} + (\bar{y}_{is} - \mu)^2 V(B | y) + O(1/m^2)$$

Where $V(B | y) = \left\{ \frac{\hat{B}(1 - \hat{B})}{\hat{\lambda}} \right\}^2 1/i_0$, \hat{B} is B evaluated at $\hat{\lambda}$

4 Numerical study

This section compares the performance of the approximate hierarchical Bayesian (HB) approach using our new/proposed prior on λ with several other existing choices. The HB approach is approximate as it is implemented through Laplace approximation. Moreover, we consider one classical approach in estimating small area means that uses the REML method to estimate the variance component λ . Here we study the frequentist properties of the resulting Bayes procedure by a Monte Carlo simulation study under the assumption of a balanced set-up, common mean (μ) model with $f_i \approx 0$. In the tables, we denote the hierarchical Bayes methods resulting from new proposed prior for variance component λ by Proposed HB. For inference about the small area means θ_i we compare our proposed Bayesian estimator with empirical best linear unbiased predictors (EBLUP) and Hierarchical Bayes (HB(WINBUG)) estimators of λ .

Data collected through pilot survey conducted by the Division of Agri-Statistics on estimation of area and yield of apple in District Baramulla has been used for the purpose of our proposed small area estimation. The district Baramulla comprises of 12 blocks viz., Zanigeer, Boniyar, Tangmarg, Wagoora, Sopore, Baramulla, Uri, Pattan, Rohama, Singphora, Rafiabad and Kunzer. Each block consists of different number of villages. A fixed number of five villages were selected at random from each block by simple random sampling. The data set was named apple-1 for analysis and modeling in R/SAS software's. It has 61 rows and 7 columns. The columns names are Blocks, N, n, Yield, Area, Trees, Actual Yield for names of blocks, total number of villages in each block, number of villages selected from each block, yield of apple from each selected village in metric tons, area under apple orchards and total number of apple trees in each of the selected village, actual yield obtained as per census records. This data set has been used for unit level estimation.

4.1 Comparison of different estimators

The performance of different estimators is examined from the accuracy of the point estimates. This is considered through the relative bias and absolute relative bias of different estimators. The different estimators are compared according to four different criteria recommended by the panel on small area estimates of population and income set up by the United States committee on National Statistics (1978), [Datta et al., 2002] [Ghosh et al., 1996],[Datta et al., 2012],[Pfeiffermann 2013.]. We compare different estimators on the basis of average relative bias, average squared relative bias, average absolute bias and average squared deviation. Suppose act_i denotes the true value of the variable for the i th small area, and est_i is any estimate of act_i $i = 1, 2, \dots, m$. Then average relative bias

$$ARB = \frac{1}{m} \sum_{i=1}^m \left| \frac{est_i - act_i}{act_i} \right|$$

average squared relative bias

$$ASRB = \frac{1}{m} \sum_{i=1}^m \left(\frac{est_i - act_i}{act_i} \right)^2$$

average absolute bias

$$AAB = \frac{1}{m} \sum_{i=1}^m |est_i - act_i|$$

average squared deviation

$$ASD = \frac{1}{m} \sum_{i=1}^m (est_i - act_i)^2$$

Now using the above four criteria on the apple data set already discussed above. A Comparison of HB estimates using proposed prior with EBLUP estimates and HB estimates with uniform prior is made using above discussed four different criteria and is reported in Table-1.

Table-1: Comparison of estimators using four different criteria

Criteria Estimates	ARB	ASRB	ABS	ASD
EBLUP	0.0873	0.0122	126.45	30351.32
HB(WINBUG)	0.0561	0.0076	89.64	22587.31
Proposed HB	0.0390	0.0057	65.57	17169.36

It is clear from the Table-1 that the proposed prior for λ performed significantly better than the uniform gamma prior and EBLUP estimates in terms of all the four criterion. This motivates us to argue that our proposed prior is superior even in the cases where the standard hierarchical Bayes estimate with uniform prior seems more appropriate. The value of % ARB is 3.9 per cent in HB (will proposed prior) as compared to 5.61 and 8.73 per cent in HB (with uniform prior) and EBLUP respectively.

Now An Empirical comparison of proposed prior for λ with REML and WINBUG estimates for all the small areas separately using Percent Absolute Relative Error(ARE) and Absolute Error(AE) has been carried out and is depicted in the following Table 2.

Table -2: Empirical comparison of estimators

Small area	EBLUP		WINBUG		NEW	
	ARB	AB	ARB	AB	ARB	AB
1.	3.27	98.48	0.61	18.39	0.28	8.44
2.	3.73	29.09	1.40	9.79	0.11	.80
3.	4.37	25.25	2.33	13.44	0.67	3.88
4.	15.18	132.77	1.73	15.13	0.50	4.41
5.	4.58	168.42	3.15	116.08	2.23	82.16
6.	12.05	85.95	7.01	49.97	4.55	32.46
7.	5.96	37.07	2.89	18.0	1.03	6.42
8.	23.99	406.94	22.93	389.04	20.08	340.62
9.	19.10	342.94	16.97	304.67	15.93	286.02
10.	3.27	56.95	1.81	31.61	0.12	2.09
11.	4.35	104.78	3.81	91.72	0.62	15.02
12.	4.87	31.85	2.74	17.91	0.70	4.63

From the Table-2 it is evident that HB model with proposed prior exhibits smaller errors and a lower incidence of extreme error than either of the HB (with uniform prior) and EBLUP estimates.

Fig.1 shows the comparison of the values of percent relative bias and absolute bias for EBLUP, hierarchical Bayes and proposed Heirarchical Bayes (HB_{LL})

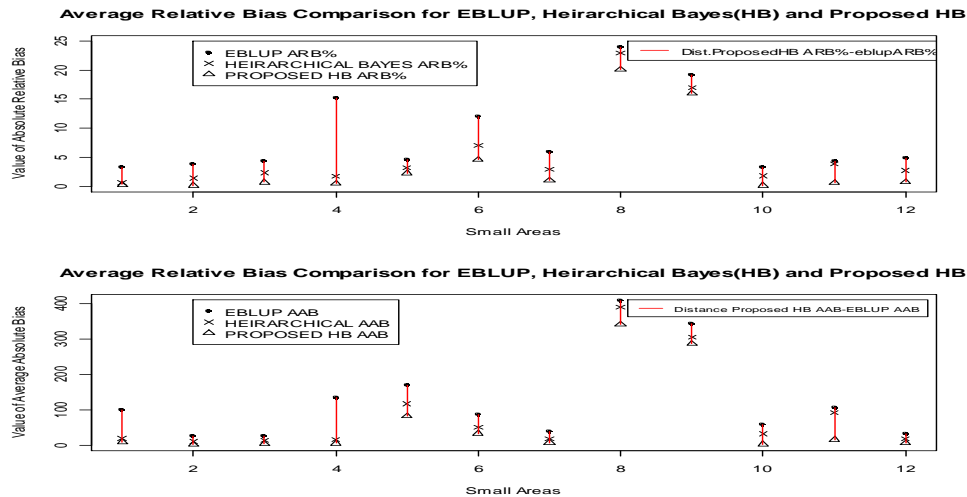


Fig. 1: Comparison of percent average relative bias and absolute bias for EBLUP, HB (will uniform prior) and HB will proposed prior on λ

Fig.1 plots the values of % ARB and AB of EBLUP, hierarchical Bayes and proposed Heirarchical Bayes against small areas. significant disparity is observed among the three estimators. Heirarchical Bayes with the new/proposed prior performed better by

providing the lowest value of both % ARB and AB for each of the small areas compared to EBLUP and hierarchical Bayes (with uniform prior).

Mean Square Error (MSE) of Estimators of Variance components for 12 small areas are reported below in Table 3

Table 3: MSE of Estimators of Variance components

Estimators Small Areas	EBLUP	HB (uniform prior)	HB (proposed prior)
1	6522.55	6197.14	5408.61
2	728.05	670.19	619.88
3	188.79	165.85	138.82
4	2118.14	158.14	114.81
5	10004.3	7853.11	7031.26
6	70.55	59.55	35.19
7	124.68	110.61	90.82
8	1284.61	985.80	920.61
9	551.10	81.08	265.87
10	772.38	750.38	690.18
11	1005.73	905.81	801.52
12	139.17	121.25	103.51

Table-3 reports the different estimates of A for each of the 12 small areas. From the table it is clear that in term of MSE the performance of proposed prior is better than rest of two techniques.

Table -4: EBLUP and HB estimates and associated Standard Errors

Estimators Small Areas	EBLUP		HB (uniform prior)		HB (proposed prior)		Actual value
	Estimate	S.E	Estimate	S.E	Estimate	S.E	Yield
1	3105.28	79.51	3025.19	76.79	3015.24	71.47	3006.8
2	672.85	25.06	698.15	23.87	698.14	22.58	698.94
3	551.29	12.60	563.10	11.34	572.66	10.46	576.54
4	541.67	10.86	859.31	9.75	870.03	9.09	874.44
5	3508.18	100.02	3560.52	89.58	3594.44	83.85	3676.6
6	798.75	7.71	762.77	5.48	745.26	5.01	712.8
7	658.25	12.02	639.18	10.51	627.6	9.78	621.18
8	2103.08	35.84	2085.18	34.37	2036.76	31.31	1696.14
9	1451.80	9.22	1490.07	8.37	1508.72	8.11	1794.74
10	1797.81	27.79	1772.47	26.45	1742.95	25.49	1740.86
11	2508.18	32.49	2495.12	31.39	2418.42	28.46	2403.4
12	621.35	12.17	635.29	11.58	648.57	10.40	653.2

Table-4 reports the EBLUP, HB (uniform prior) and proposed prior estimates and their associated standard errors for all the 12 small areas separately. It is clear from the table that the new/proposed prior for λ provides better estimates than EBLUP and HB (uniform prior). In terms of standard error also new/proposed prior provides better estimates than the EBLUP and HB (uniform prior).

Fig. 2 plots the point estimates of θ_i for EBLUP, HB (with Uniform prior), HB (with new/proposed prior) against the small areas and also provides a comparison of these values with the actual value of yield obtained in each of the small area.

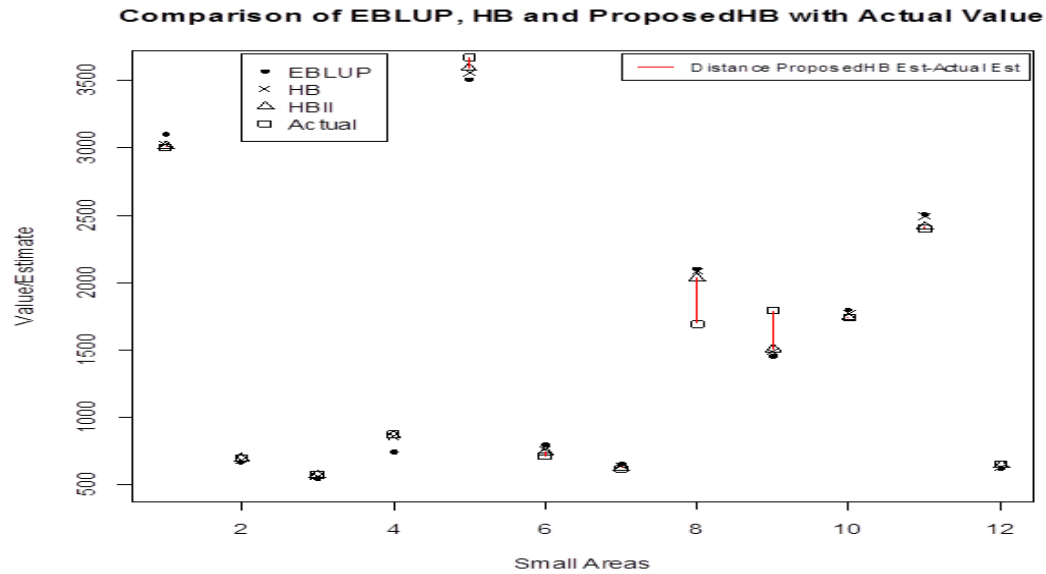


Fig. 2: EBLUP, HB(with Uniform prior), HB(with new/proposed prior) estimates compared to the true means

Fig.2 displays EBLUP, HB(with Uniform prior), HB(with new/proposed prior) estimates and their deviation from the actual mean. Here we can see that the values of θ_i obtained by HB(with new/proposed prior) are more closer to actual values as compared to HB and EBLUP. Thus for the plot also we conclude that among the three techniques discussed the HB(with new/proposed prior) is the best technique for obtaining the estimates.

Table 5: Relative contribution of the three terms to the posterior variance of θ_i using new/proposed prior on λ

Small Areas	$T_1^1(g_{i1})$	$T_2^2(g_{i2})$	$T_3^3(g_{i3})$
1.	4850.13	254.58	151.85
2.	615.81	0.7500	1.40
3.	137.51	0.0064	0.1045
4.	112.99	0.0061	0.0570
5.	6028.26	463.71	259.64
6.	23.89	0.0033	0.0247
7.	88.80	0.0076	0.0868
8.	900.51	4.23	6.48
9.	255.81	1.51	2.35
10.	680.20	2.55	3.49
11.	788.51	3.99	4.41
12.	100.71	0.0050	0.0855

¹Uncertainty in the model in estimating θ_i ²Uncertainty in estimating coefficient β ³Uncertainty in model in estimating the variance component A.

Table-5 reports the relative contribution of the three terms to the posterior variance of θ_i obtained by using the new/proposed prior. The three columns T_1 , T_2 , T_3 exhibits the relative contribution of term1, term 2, term 3 respectively. From the values in the table we conclude that the contribution of the term which accounts for the uncertainty in estimating the variance component is substantially small relative to the first term which accounts for the uncertainty in the model in estimating the small area means.

5 Conclusion

In this paper Bayesian implementation of Unit level model is carried out and new prior is proposed on the variance component λ . Besides being simple, this prior has two main advantages. It removes the possibility of yielding zero estimates for the variance component; the popular choice of uniform prior on λ suffers from this drawback if posterior mode is considered as an estimator. This prior also enjoys good small sample frequentist properties; real agricultural study results justify this conclusion. Also, in order to have closed form expressions of the posterior mean and variance of the true small area mean, Laplace approximation to ratio of integrals, following [Kass and Staffey 1989] is being used. To illustrate the method numerically a real data set on apple has been used and the results showed that the Bayes estimators (with new prior) of small area means have good frequentist properties such as MSE and ARB as compared to other methods viz., EBLUP and HB(WINBUG) estimators.

References

1. Jiang, J. and Lahiri, P. (2006). Estimation of finite population domain means: A model-assisted empirical best prediction approach. *Journal of the American Statistical Association* 101: 301-311.
2. Rao, J.N.K. (2003). *Small Area Estimation*, New Jersey, John Wiley & Sons, Inc.
3. Adam, W., Aitke, G., Anderson, B. and William, (2013). Evaluation and improvements in small area estimation methodologies.
4. Jiango, V.D., Haziza, D. and Duchasne, P. (2013). Controlling the bias of robust small area estimators. *NATSEM* 9: 23-30.
5. Chambers, R., Chandra, H., Salvali, N. and Tzaidis, N. (2014). Outlier robust small area estimation. *Journal of the Royal Statistical Society: Series B* 76(1): 47-69.
6. Rao, J.N.K., Sinha, S.K. and Dumitrescu, L. (2013). Robust small area estimation under semi-parametric mixed models. *Canadian Journal of Statistics* 9999: 1-16.
7. iang, J. and Lahiri, P. (2006). Estimation of finite population domain means: A model-assisted empirical best prediction approach. *Journal of the American Statistical Association* 101: 301-311.
8. Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.
9. Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1997). BUGS: Bayesian inference using Gibbs sampling, Version 0.6. *Biostatistics Unit. Cambridge: MRC*.
10. R Development Core Team, (2008). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
11. Butar, F.B. and Lahiri, P. (2002). Empirical Bayes estimation of several population means and variances under random sampling variances model. *Journal of Statistics Planning Inference* 102: 59-69.
12. Kass, R.E. and Staffey, D. (1989). Approximate Bayesian inference in conditional independent hierarchical models (parametrical empirical Bayes model). *Journal of American Statistical Association* 84: 717- 726.
13. Tierney, L. and Kadane, J.B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of American Statistical Association* 81: 82-86.
14. Tierney, L., Kass, R.E. and Kadane, J.B. (1989). Fully exponential Laplace approximations to expectations and variances of non-positive functions. *Journal American Statistical Association* 84: 710-716.
15. Datta, G.S., Lahiri, P. and Maiti, T. (2002). Estimation of median income of four person families by state using timeseries and cross sectional data. *Journal of Statistical Planning and Influence* 102: 83-97.

16. Ghosh, M., Nangia, N. and Kim, D. (1996). Estimation of Median Income of Four-person Families: A Bayesian Time Series Approach. *Journal of the American Statistical Association* 91: 1423-1431.
17. Datta, G.S., Bell, W.R. and Ghosh, M. (2012). Benchmarking small area estimators. *Journal of Statistical Sciences* 28(1): 1-134.
18. Pfeiffermann, D. (2013). New important development in small area estimation. *Journal of Statistical Sciences* 28(1): 40-68
19. Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error component model for prediction of county crop area using survey and satellite data. *Journal of American Statistical Association* 83: 28-36.
20. Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Euclid Online Journal of Mathematical and Statistics*.
21. Rao, J.N.K. (2003(a)). 'Some new developments in small area estimation', *Journal of Iranian Statistical Society* 2(2): 145-169.