Variance Component Estimation Under a General Area-level Model

Yuting Chen[1] **and Hanqing Li**[2]

[1]Eastern Kentucky University, USA, yuting.chen@eku.edu
[2]University of Calgary, Canada, hanqing.li@ucalgary.ca

Abstract

Small area estimation (SAE) provides reliable inference for domains with limited survey sample data by borrowing strength across areas through modeling. The typical area-level model assumes normally distributed random effects, an assumption that may not hold in practice. This paper empirically examines the performance of residual maximum likelihood (REML) and adjusted REML estimators under general area-level models with non-normal random effects. Using simulations with heavy-tailed and asymmetric distributions, we evaluate point estimation and prediction interval performance. REML remains reasonably robust in estimating the variance component and supporting reliable predictions, but zero boundary estimates can degrade interval performance when the number of domains is small. Adjusted REML reduces boundary issues and yields more reliable interval coverage while maintaining competitive estimation accuracy. These results highlight adjusted likelihood methods as a practical and robust option even when the normality assumption is uncertain.

Keywords: data linkage, survey statistics, uncertainty quantification.

1 Introduction

In survey sampling, researchers often aim to estimate population parameters such as totals, means, or proportions based on data from a representative sample. In many practical settings, however, it is also of interest to estimate similar characteristics for specific subpopulations or domains (e.g., regions, demographic groups, or institutions). Large-scale surveys are typically designed to yield reliable estimates for large domains, but for smaller domains, the sample sizes may be too small or even zero, to produce direct estimates with acceptable precision. This situation gives rise to the small area problem.

To address this challenge without increasing sample sizes, small area estimation (SAE) techniques have been developed to “borrow strength” across related areas. Model-based SAE methods achieve this by linking data from different areas through statistical models that include area-specific random effects and auxiliary information. These approaches enable more precise and reliable estimation of small area parameters.

Suppose that the population of interest, U, is partitioned into m areas (or subpopulations), denoted by $U_1, \cdots, U_m$ and that we are interested in estimating the corresponding area means $\{\theta_i,\ i = 1, \cdots, m\}$. Let s_i denote the sample drawn from area U_i. When the sample size n_i is small, we may encounter the small area issue. A widely used framework in SAE is the two-level area-level model, which for area $i = 1, \cdots, m$, can be expressed as:

$$\begin{aligned} &\text{Level 1: (Sampling model): } \hat{y}_i|\theta_i \overset{\text{ind}}{\sim} \mathcal{N}(\theta_i, D_i); \\ &\text{Level 2: (Linking model): } \theta_i \overset{\text{ind}}{\sim} \mathcal{G}(\mathbf{x}_i'\beta, A, \phi). \end{aligned} \tag{1}$$

The Level 1 model represents the sampling distribution of the direct estimator $\hat{y}_i$, which may be a weighted or unweighted estimate for area i. For example, $\hat{y}_i$ could be the sample mean based on n_i observations from area i with sampling variance $D_i = \sigma^2/n_i$, where σ^2 is known or reliably estimated from all areas (Fay and Herriot, 1979; Otto and Bell, 1995; Hawala and Lahiri, 2018). The Level 2 model links the true small area means θ_i to a vector of known auxiliary variables $\mathbf{x}_i = (x_{i1}, \cdots, x_{ip})'$, often obtained from administrative records, census data, or other external sources. We assume that the Level 2 distribution $\mathcal{G}$ is a fully parametric distribution, not necessarily normal, with mean $\mathrm{E}(\theta_i) = \mathbf{x}_i'\beta$, variance $\mathrm{Var}(\theta_i) = A \geq 0$, and any additional parameters ϕ. The coefficient vector $\beta \in \mathcal{R}^p$ and the variance component A are unknown and must be estimated from the data.

The classical area-level model proposed by Fay and Herriot (1979) assumes normality at both levels. The normality assumption at Level 1 may not be considered as restrictive as the normality of θ_i, due to the central limit theorem's effect on direct estimator $\hat{y}_i$ (Rao and Molina, 2015; Jiang and Torabi, 2022). To relax this assumption, recent studies have explored non-normal alternatives for the Level 2 distribution $\mathcal{G}$ (Chen, Hirose, and Lahiri, 2024). For instance, Bell and Huang (2006) used a t-distribution to mitigate the influence of outliers; Fabrizi and Trivisano (2010) proposed exponential power and skewed exponential power distributions to handle heavy-tailed or asymmetric effects; and Jiang and Torabi (2022) employed a skewed normal distribution.

The above two-level model can equivalently be expressed as the linear mixed model:

$$\hat{y}_i = \theta_i + e_i = \mathbf{x}_i'\beta + u_i + e_i, \quad i = 1, \cdots, m, \tag{2}$$

where random effects u_i's and sampling errors e_i's are independent with $u_i \overset{\text{i.i.d}}{\sim} \mathcal{G}(0, A, \phi)$ and $e_i \overset{\text{ind}}{\sim} \mathcal{N}(0, D_i)$. The small area parameter of interest is $\theta_i = \mathbf{x}_i'\beta + u_i,\ i = 1, \cdots, m$. When A is known, the best linear unbiased predictor (BLUP) of θ_i that minimize the mean squared prediction error (MSPE) among all linear unbiased predictors, is given by

$$\tilde{\theta}_i = (1 - B_i)\hat{y}_i + B_i\mathbf{x}_i'\tilde{\beta}, \tag{3}$$

where $B_i = D_i/(A + D_i)$ is the shrinkage factor, $\tilde{\beta} = \tilde{\beta}(A)$ is the standard weighted least squares estimator of β. The BLUP effectively shrinks the direct estimator y_i toward the regression synthetic estimator $\mathbf{x}_i'\tilde{\beta}$, with the degree of shrinkage determined by B_i. In this paper, we assume $A > 0$. In practice, since A is unknown, it must be estimated from the data, leading to the empirical BLUP (EBLUP):

$$\hat{\theta}_i = (1 - \hat{B}_i)y_i + \hat{B}_i\mathbf{x}_i'\hat{\beta}, \tag{4}$$

where $\hat{B}_i = D_i/(\hat{A} + D_i)$ and $\hat{\beta} = \tilde{\beta}(\hat{A})$.

When $\mathcal{G}$ is normal, several methods have been proposed to estimate A, including the Fay-Herriot method-of-moments (FH) estimator (Fay and Herriot, 1979), the Prasad-Rao simple method-of-moments (PR) estimator (Prasad and Rao, 1990), the maximum likelihood (ML) estimators and the residual maximum likelihood (REML) estimators (Datta and Lahiri, 2000). When the number of areas m is small, standard variance estimation methods, particularly the PR estimator, often produce boundary estimate $A = 0$, leading to $\hat{B}_i = 1$ for all i, even when some of the true B_i are not close to 1 (Li and Lahiri, 2010; Chen, Hirose, and Lahiri, 2024). This causes an overshrinkage problem in EBLUP, since now the EBLUP of θ_i reduces to the regression synthetic estimator. Moreover, with $\hat{A} = 0$, it also causes the problem of degenerate distribution and prevents the use of parametric bootstrap methods for uncertainty quantification, such as estimating MSPE or constructing prediction intervals.

To address this issue under normal random effects, several adjusted likelihood methods have been developed to guarantee positive estimates of A (Li, 2007; Yoshimori and Lahiri, 2014; Hirose and Lahiri, 2018). These methods solve the two problems above simultaneously in SAE applications. In addition, they show that the biases of the adjusted ML and REML estimators are of order $O(m^{-1})$ (Li and Lahiri, 2010), and those of the parametric bootstrap MSPE being $o(m^{-1})$ (Hirose and Lahiri, 2018). However, the performance of these adjusted estimators when the random effects are non-normal remains largely unexplored.

In this study, we investigate methods for estimating variance components under a general area-level model that allows for possibly non-normal random effects. Laird and Ware, 1982 and Cressie, 1990, among others, have favored the REML method over the ML method for variance component estimation in complex small area models. This preference was later supported by Datta and Lahiri (2000), in which they showed that the REML estimator has a lower order of bias than the ML estimator. Therefore, in this paper, we focus on the REML approach. Following Jiang (1996), we define the REML estimator of variance components as the solution to the REML equations, which we introduce in the next section. Although Jiang (1996) theoretically showed that REML estimates are consistent under certain identifiability and information conditions, their empirical performance under non-normal random effects has not been well studied in SAE. We therefore (i) empirically evaluate the performance of REML estimators under various non-normal settings, and (ii) extend the adjusted REML methods of Li and Lahiri (2010) to the general area-level model, assessing their performance through Monte Carlo simulations.

The remainder of this paper is organized as follows. Section 2 provides the list of notations and regularity conditions. Section 3 reviews the estimation methods for variance components, including REML and adjusted REML estimators. Section 4 presents Monte Carlo simulation results comparing different estimators under various model settings. Section 5 concludes with a summary and discussion.

2 A list of notations and regularity conditions

We introduce the following notations that will be used throughout the paper:

$\mathbf{y} = (\hat{y}_1, \cdots, \hat{y}_m)'$, a $m \times 1$ column vector of direct estimates;

$X' = (\mathbf{x}_1, \cdots, \mathbf{x}_m)$, a $p \times m$ known matrix of rank p;

$\Sigma = \mathrm{diag}(A + D_1, \cdots, A + D_m)$, a $m \times m$ diagonal matrix;

$\tilde{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\mathbf{y}$, weighted least square estimator of β with known A;

$\mathbf{P} = \Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$.

We assume the following regularity conditions throughout the paper:

r.1 rank$(X) = p$ is fixed;

r.2 $\sup_{i\geq 1} h_{ii} = O(m^{-1})$, where $h_{ii} = \mathbf{x}_i'(X'X)^{-1}\mathbf{x}_i$;

r.3 $0 < \inf_{i\geq 1} D_i \leq \sup_{i\geq 1} D_i < \infty$.

3 REML and adjusted REML estimators

The REML approach introduced by Patterson and Thompson (1971), eliminates dependence on nuisance parameters by basing inference on linear transformations of the data that remove the fixed effects. Under normality at both levels, the restricted likelihood function is given by:

$$L_{\mathrm{RE}}(A) = c|X'\Sigma^{-1}X|^{-\frac{1}{2}}|\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right) \tag{5}$$

where c is a constant independent of A. Let $l_{\mathrm{RE}}(A)$ denote the corresponding restricted log-likelihood. The REML estimator $\hat{A}_{\mathrm{RE}}$ satisfies:

$$\begin{aligned}\frac{\partial l_{\mathrm{RE}}(A)}{\partial A} &= \frac{1}{2}\left[\mathbf{y}'\mathbf{P}^2\mathbf{y} - \mathrm{tr}(\mathbf{P})\right] \\ &= 0\end{aligned} \tag{6}$$

In general (without assuming normality), the REML estimate $\hat{A}_{\mathrm{RE}}$ is defined as solution of (6).

Following Li and Lahiri (2010), we also consider the same adjusted restricted likelihood under the general area-level model:

$$L_{\mathrm{adj}}(A) = A \times L_{\mathrm{RE}}(A). \tag{7}$$

The adjusted maximum likelihood estimator $\hat{A}_{\mathrm{adj}}$ is obtained by maximizing $L_{\mathrm{adj}}(A)$ or its logarithmic form, $l_{\mathrm{adj}}(A)$.

Since $L_{\mathrm{RE}}(A)$ is a continuous positive function of A and $\lim_{A\to\infty} A \times L_{\mathrm{RE}}(A) = 0$ for $m > p + 2$, it follows from Lemma 2.1 of Li (2007) that the maximizer $\hat{A}_{\mathrm{adj}}$ is strictly positive.

Specifically, because $L_{\mathrm{RE}}(A) > 0$ for all A, we have $A \times L_{\mathrm{RE}}(A) \leq 0$ for $A \leq 0$ and $A \times L_{\mathrm{RE}}(A) > 0$ for $A > 0$. Moreover, since $A \times L_{\mathrm{RE}}(A) \to 0$ as $A \to \infty$, there exists some $A_0 > 0$ such that

$$A_0 \times L_{\mathrm{RE}}(A_0) = \max_A\{A \times L_{\mathrm{RE}}\},$$

which ensures that the maximizer A_0 is positive.

3.1 Parametric bootstrap prediction intervals

A traditional prediction interval for θ_i is of the form $\hat{\theta}_i \pm z_{\alpha/2}\sqrt{\mathrm{mspe}}$, where $z_{\alpha/2}$ is the $100(1-\alpha/2)$th standard normal percentile and mspe is an estimate of the mean squared prediction error of $\hat{\theta}_i$. However, such intervals have coverage errors of order $O(m^{-1})$, which may be inadequate for small area applications. Chatterjee, Lahiri, and Li, 2008 proposed a parametric bootstrap method that constructs intervals from the bootstrap distribution approximation of $\hat{\sigma}_1^{-1}(\theta_i - \hat{\theta}_i)$ under a normal linear mixed model, where $\hat{\sigma}_1^2 = D_i(1 - \hat{B}_i)$. This method achieves improved coverage error of order $O(m^{-3/2})$.

Chen, Hirose, and Lahiri (2024) extended this method to the general area-level model (1) with non-normal level-2 distributions, and interestingly found that the bootstrap intervals can exhibit overcoverage under certain conditions. Their simulations also showed that there was high percentage of zero estimates in $\hat{A}_{\mathrm{PR}}$ estimator which affects the performance of associated bootstrap intervals. The result is consistent with the findings in Li and Lahiri, 2010.

In this paper, we assess the performance of the similar parametric bootstrap procedures under non-normal models using the REML and adjusted REML estimators of A. Specifically, let

$$\hat{y}_i^* = \mathbf{x}_i'\hat{\beta} + u_i^* + e_i^*$$

where $u_i^* \overset{\text{iid}}{\sim} \mathcal{G}(0, \hat{A}, \hat{\phi})$ and $e_i^* \overset{\text{ind}}{\sim} N(0, D_i)$ for $i = 1, \cdots, m$. Denote by $\hat{\beta}^*$, $\hat{A}^*$, $\hat{\theta}_i^*$, and $\hat{\sigma}_1^*$ the quantities computed from bootstrap samples $\mathbf{y}^* = \{\hat{y}_i^*,\ i = 1, \cdots, m\}$, and let $\theta_i^* = \mathbf{x}_i'\hat{\beta} + u_i^*$. The bootstrap distribution of $\hat{\sigma}_1^{*-1}(\theta_i^* - \hat{\theta}_i^*)$ is then used to approximate the distribution of $\hat{\sigma}_1^{-1}(\theta_i - \hat{\theta}_i)$. For a given significance level α, let q_l and q_u denote the $\alpha/2$ and $1 - \alpha/2$ quantiles of the bootstrap distribution, respectively. The parametric bootstrap prediction interval for θ_i is then given by $(\hat{\theta}_i + q_l\hat{\sigma}_1, \hat{\theta}_i + q_u\hat{\sigma}_1)$.

4 Monte Carlo Simulations

To empirically evaluate the performance of various variance estimators and their associated prediction intervals in small m settings, we consider $m = 10$ and $m = 15$. Following Li and Lahiri, 2010, we use an unbalanced pattern for the sampling variances (D_i), consisting of five groups of small areas with common D_i values within each group. Specifically, we set $D_i \in \{4.0, 0.6, 0.5, 0.4, 0.2\}$ and fix $A = 1$. Without loss of generality, we take $\mathbf{x}_i'\beta = 0$. To reflect practical conditions, we still estimate the mean even when it is theoretically zero. Since areas within each group are exchangeable, we summarize results by group means in the tables.

We consider two non-normal Level 2 distributions in the area-level model (1): (i) a t-distribution with 5 degrees of freedom (symmetric case), and (ii) a shifted exponential (SE) distribution (asymmetric case). For each distributional scenario, we generate $N = 1,000$ independent datasets $\{y_i,\ i = 1, \ldots, m\}$ and use 1,000 bootstrap samples to construct the parametric bootstrap prediction intervals.

We examine three estimators of A: the PR estimator $\hat{A}_{\text{PR}}$ which does not rely on distributional assumptions, the REML estimator $\hat{A}_{\text{RE}}$ and the adjusted REML estimator $\hat{A}_{\text{AR}}$. We use both bias and mean squared error to compare different estimators. Let $\hat{A}^{(j)}$ be the estimate for the jth simulation run. We compute the following Monte Carlo measures:

$$\text{Bias}(\hat{A}) = \frac{1}{N}\sum_{j=1}^{N}(\hat{A}^{(j)} - A), \quad \text{RMSE}(\hat{A}) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\hat{A}^{(j)} - A)^2}.$$

Table 1 shows the percentages of zero estimates in $\hat{A}$ and $\hat{A}^*$. For $m = 10$, the PR estimator yields the highest rate of zero estimates in both $\hat{A}$ and $\hat{A}^*$. Under the shifted exponential distribution, REML also result in a zero estimate in $\hat{A}$ although the percentage of 0 is relatively low (about 0.1%). All methods can produce zero estimates in $\hat{A}^*$, and the adjusted REML estimator exhibits the lowest percentage in all cases. As m increases to 15, the chance of zero estimate decreases across all methods.

Table 2 summarizes the small-sample performance of the three variance estimators in terms of bias and RMSE. Both PR and REML generally show smaller bias than adjusted REML. Overall, REML achieves the best performance in terms of both bias and RMSE under both distributions. The performance of adjusted REML estimator improves as m increases in terms of both bias and RMSE.

In SAE applications, prediction is often the primary objective. To investigate prediction accuracy of EBLUP with different plug-in variance estimates, we approximate the true MSPE through Monte Carlo

Table 1: Percentages of zero estimates in $\hat{A}$ and $\hat{A}^*$ for different estimation methods.

m	$\hat{A}_{\mathrm{PR}}$	$\hat{A}_{\mathrm{RE}}$	$\hat{A}_{\mathrm{AR}}$	$\hat{A}^*_{\mathrm{PR}}$	$\hat{A}^*_{\mathrm{RE}}$	$\hat{A}^*_{\mathrm{AR}}$
			$t\ \{u_i\}$			
10	21.900	0	0	33.411	0.015	<0.001
15	11.700	0	0	25.120	0.001	0
			Shifted exponential $\{u_i\}$			
10	26.200	0.100	0	36.673	0.022	0.001
15	16.100	0	0	27.714	0.001	<0.001

Table 2: Comparison of different estimators of A for $m = 10$ and $m = 15$ with true value of $A = 1$.

	Monte Carlo Bias			**Monte Carlo RMSE**		
m	PR	RE	AR	PR	RE	AR
			$t\ \{u_i\}$			
10	0.035	-0.015	0.720	1.225	0.901	1.409
15	0.074	0.011	0.437	1.031	0.762	1.002
			Shifted exponential $\{u_i\}$			
10	0.053	-0.043	0.690	1.418	1.119	1.653
15	0.086	0.014	0.443	1.148	0.917	1.173

simulations. Let $\theta_i^{(j)}$ and $\hat{\theta}_i^{(j)}$ be simulated true value and the EBLUP for area i in the jth simulation respectively, $i = 1, \cdots, m;\ j = 1, \cdots, N$. We also compute the Monte Carlo mean squared prediction error of $\hat{\theta}_i$:

$$\text{MSPE}(\hat{\theta}_i) = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_i^{(j)} - \theta_i^{(j)})^2.$$

Figure (1) shows the simulated MSPE results. When $m = 10$, $\hat{\theta}_i(\hat{A}_{\mathrm{RE}})$ tends to have the smallest MSPE when the sampling variance is large ($D_i = 4$), and $\hat{\theta}_i(\hat{A}_{\mathrm{RE}})$ and $\hat{\theta}_i(\hat{A}_{\mathrm{AR}})$ outperform $\hat{\theta}_i(\hat{A}_{\mathrm{PR}})$ in the remaining groups. When m = 15, $\hat{\theta}_i(\hat{A}_{\mathrm{RE}})$ and $\hat{\theta}_i(\hat{A}_{\mathrm{AR}})$ perform similarly across all groups and better than $\hat{\theta}_i(\hat{A}_{\mathrm{PR}})$.

For interval estimation, we compare two traditional intervals of the form $\hat{\theta}_i \pm z_{\alpha/2}\sqrt{\mathrm{mspe}}$ based on $\hat{A}_{\mathrm{PR}}$ and $\hat{A}_{\mathrm{RE}}$, and three parametric bootstrap intervals based on $\hat{A}_{\mathrm{PR}}$, $\hat{A}_{\mathrm{RE}}$, and $\hat{A}_{\mathrm{AR}}$. Derivations of $\mathrm{mspe}(\hat{\theta}_i)$ using $\hat{A}_{\mathrm{PR}}$ and $\hat{A}_{\mathrm{RE}}$ appear in Prasad and Rao (1990) and Datta and Lahiri (2000), respectively.

Tables 3 and 4 present the empirical coverage probabilities and average lengths for nominal 95% intervals. When $m = 10$, the parametric bootstrap method using $\hat{A}_{AR}$ (PB-AR) performs the best in terms of the coverage probabilities and the average lengths. The PR-based traditional interval (PR) and PB-PR show severe undercoverage across all groups. The traditional REML interval also undercovers, especially for group 1. PB-RE achieves good coverage but yields substantially longer intervals than PB-AR. This may be because the REML method sometimes produces zero estimates. Since the estimate $\hat{A}^*$ appears in the denominator of the term $\hat{\sigma}_1^{*-1}(\theta_i^* - \hat{\theta}_i^*)$ used in our parametric bootstrap method, this quantity becomes undefined whenever $\hat{A}^*_{\mathrm{RE}} = 0$. To address this issue, we replaced those zero estimates with 0.01. In such cases, the resulting values can be extremely large, which may in turn lead to overly wide prediction intervals. As m increases, all methods improve, although PR, RE, and PB-PR still exhibit undercoverage. Overall, PB-AR provides competitive cov-

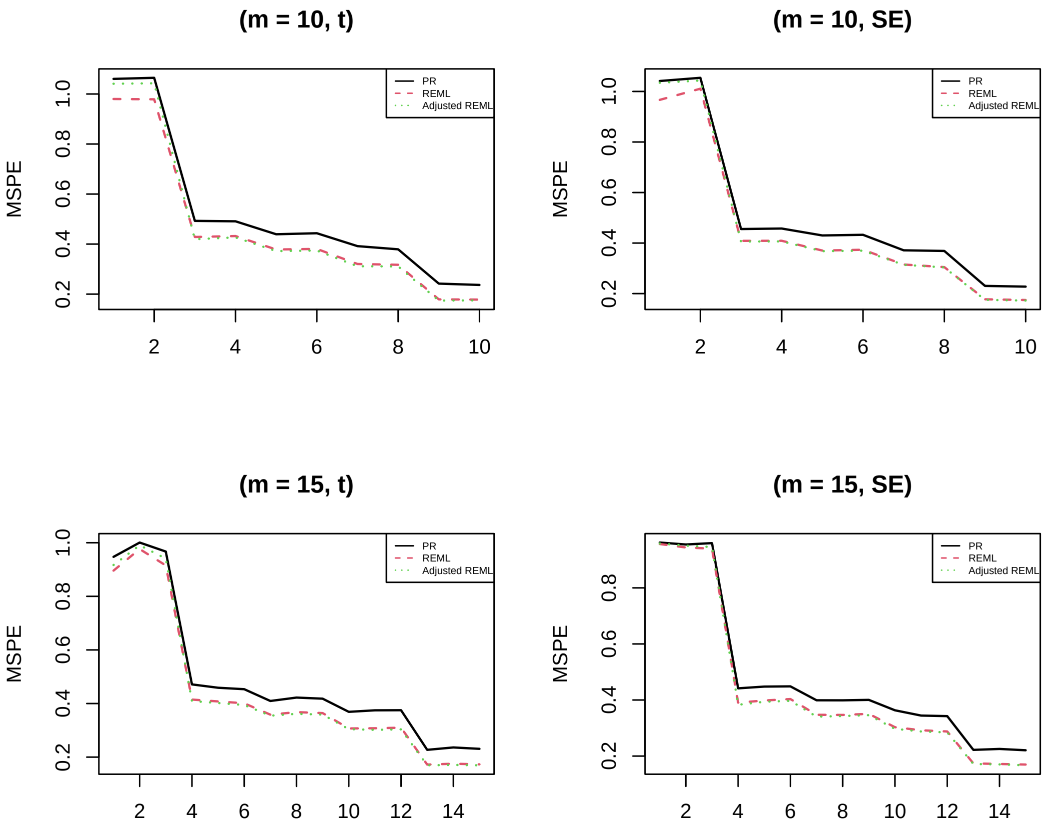

Figure 1: Simulated mean squared prediction error of $\hat{\theta}_i(\hat{A})$.

erage probabilities and interval lengths, showing only slight undercoverage for Group 1 under the shifted exponential distribution.

Table 3: Average Monte Carlo coverage and average length of different intervals for $m = 10$ with nominal coverage = 95% under t-distribution and shifted exponential distribution.

	PR	RE	PB-PR	PB-RE	PB-AR
			t $\{u_i\}$		
G1	84.10 (12.78)	88.70 (3.52)	83.90 (11.18)	96.25 (9.12)	95.10 (4.26)
G2	85.65 (11.22)	93.90 (2.53)	85.65 (6.74)	96.75 (5.39)	95.15 (2.56)
G3	85.65 (11.22)	93.90 (2.53)	85.65 (6.74)	96.75 (5.39)	95.15 (2.56)
G4	86.00 (10.76)	94.60 (2.24)	86.10 (5.69)	96.70 (4.42)	94.65 (2.19)
G5	86.25 (9.65)	96.30 (1.77)	86.60 (3.97)	97.25 (2.94)	94.70 (1.63)
			Shifted exponential $\{u_i\}$		
G1	84.20 (12.06)	87.85 (3.31)	84.10 (10.41)	94.55 (9.52)	94.85 (4.18)
G2	83.85 (10.52)	93.70 (2.44)	83.90 (6.19)	95.50 (5.41)	94.80 (2.48)
G3	83.85 (10.52)	93.70 (2.44)	83.90 (6.19)	95.50 (5.41)	94.80 (2.48)
G4	84.30 (10.14)	95.45 (2.20)	84.75 (5.26)	95.95 (4.42)	95.00 (2.14)
G5	87.05 (9.08)	96.80 (1.80)	87.35 (3.69)	96.00 (3.02)	94.70 (1.61)

Table 4: Average Monte Carlo coverage and average length of different intervals for $m = 15$ with nominal coverage = 95% under t-distribution and shifted exponential distribution.

	PR	RE	PB-PR	PB-RE	PB-AR
			t $\{u_i\}$		
G1	90.40 (10.76)	90.60 (3.52)	90.13 (9.52)	97.97 (6.46)	95.00 (4.07)
G2	90.67 (9.82)	93.50 (2.46)	90.70 (6.02)	97.60 (3.98)	95.17 (2.50)
G3	89.93 (9.74)	93.87 (2.33)	90.00 (5.65)	97.30 (3.69)	94.80 (2.34)
G4	90.37 (9.61)	93.73 (2.17)	90.20 (5.18)	97.13 (3.34)	94.43 (2.15)
G5	91.10 (8.99)	95.27 (1.68)	91.17 (3.75)	97.57 (2.35)	94.67 (1.62)
			Shifted exponential $\{u_i\}$		
G1	87.73 (10.63)	89.23 (3.40)	87.73 (9.36)	95.83 (7.14)	93.73 (4.09)
G2	89.13 (9.68)	93.60 (2.40)	89.20 (5.85)	96.50 (4.24)	94.97 (2.45)
G3	88.07 (9.56)	93.63 (2.28)	88.10 (5.47)	96.27 (3.94)	94.40 (2.29)
G4	89.03 (9.44)	94.03 (2.13)	88.67 (5.02)	95.97 (3.56)	94.60 (2.11)
G5	88.83 (8.85)	95.60 (1.69)	88.97 (3.64)	96.10 (2.51)	94.37 (1.60)

5 Discussion

This study provides empirical evidence on variance component estimation in general area-level models that allow non-normal random effects. The results indicate that the REML estimator can remain reasonably robust to deviations from normality, even when the number of areas is relatively small (for example, $m = 10$). Under both heavy-tailed and asymmetric random effect distributions, according to our simulation results, the bias of the REML estimator is similar to the PR estimator and its RMSE is smaller than both PR and adjusted REML estimators. Moreover, associated EBLUP based on REML estimate tends to perform well in prediction accuracy.

The simulation results also show that the effectiveness of parametric bootstrap prediction intervals depends heavily on the variance component estimator. When zero estimates are frequent, particularly when using the PR variance estimator, bootstrap intervals become unreliable due to the induced degeneracy. In contrast, the adjusted REML estimator reduces boundary estimates and supports stable bootstrap inference, leading to improved coverage across all simulation settings considered. This indicates that parametric bootstrap intervals based on adjusted REML estimates could be an effective alternative, when m is small.

There are promising directions for future work. For example, a deeper theoretical investigation of adjusted REML under non-normal random effects, including refined bias corrections and accurate MSPE estimation of EBLUP with adjust REML variance estimate, would strengthen its methodological foundations. Overall, the findings highlight that positive and stable estimation of variance components is essential for reliable small area prediction and inference. Adjusted likelihood methods offer a practical and robust alternative in applications where the normality assumption for random effects may not hold.

References

Bell, W. R. and E. T. Huang (2006). "Using the t-distribution to deal with outliers in small area estimation". In: *Proceedings of statistics Canada symposium*.

Chatterjee, S., P. Lahiri, and H. Li (2008). "Parametric bootstrap approximation to the distribution of EBLUP and related prediction intervals in linear mixed models". In: *The Annals of Statistics* 36.3, pp. 1221 –1245. DOI: 10.1214/07-AOS512. URL: https://doi.org/10.1214/07-AOS512.

Chen, Y., M. Y. Hirose, and P. Lahiri (2024). "Impact of existence and nonexistence of pivot on the coverage of empirical best linear prediction intervals for small areas". In: *arXiv preprint arXiv:2410.11238*.

Cressie, N. (1990). *REML estimation in empirical Bayes smoothing of census undercount*. Iowa State University. Department of Statistics. Statistical Laboratory.

Datta, G. S. and P. Lahiri (2000). "A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems". In: *Statistica Sinica*, pp. 613–627.

Fabrizi, E. and C. Trivisano (2010). "Robust linear mixed models for small area estimation". In: *Journal of Statistical Planning and Inference* 140.2, pp. 433–443.

Fay, R. E. and R. A. Herriot (1979). "Estimates of income for small places: an application of James-Stein procedures to census data". In: *Journal of the American Statistical Association* 74.366a, pp. 269–277.

Hawala, S. and P. Lahiri (2018). "Variance modeling for domains". In: *Statistics and Applications* 16.1, pp. 399–409.

Hirose, M. Y. and P. Lahiri (2018). "Estimating variance of random effects to solve multiple problems simultaneously". In: *The Annals of Statistics* 46.4, pp. 1721–1741.

Jiang, J. (1996). "REML estimation: asymptotic behavior and related topics". In: *The Annals of Statistics* 24.1, pp. 255–286.

Jiang, J. and M. Torabi (2022). "Goodness-of-fit test with a robustness feature". In: *TEST* 31.1, pp. 76–100.

Laird, N. M. and J. H. Ware (1982). "Random-effects models for longitudinal data". In: *Biometrics*, pp. 963–974.

Li, H. (2007). *Small area estimation: An empirical best linear unbiased prediction approach*. University of Maryland, College Park.

Li, H. and P. Lahiri (2010). "An adjusted maximum likelihood method for solving small area estimation problems". In: *Journal of multivariate analysis* 101.4, pp. 882–892.

Otto, M. C. and W. R. Bell (1995). "Sampling error modelling of poverty and income statistics for states". In: *American Statistical Association, Proceedings of the Section on Government Statistics*, pp. 160–165.

Patterson, H. D. and R. Thompson (1971). "Recovery of inter-block information when block sizes are unequal". In: *Biometrika* 58.3, pp. 545–554.
Prasad, N. N. and J. N. Rao (1990). "The estimation of the mean squared error of small-area estimators". In: *Journal of the American statistical association* 85.409, pp. 163–171.
Rao, J. N. and I. Molina (2015). *Small area estimation*. John Wiley & Sons.
Yoshimori, M. and P. Lahiri (2014). "A second-order efficient empirical Bayes confidence interval". In: *The Annals of Statistics* 42.4, pp. 1233 –1261. DOI: 10.1214/14-AOS1219. URL: https://doi.org/10.1214/14-AOS1219.