# Spatial generalized linear mixed models in small area estimation

Mahmoud Torabi*

## Abstract

In survey sampling, policy decisions regarding allocation of resources to subgroups in a population, called small areas, are based on reliable predictors of their underlying parameters. However, the information is collected at a different scale than these subgroups. Hence we need to predict characteristics of the subgroups based on the coarser scale data. In view of this, there is a growing demand for reliable small area predictors by borrowing information from other areas. These are commonly based on either the linear mixed models or the generalized linear mixed models (GLMMs). There is a limited literature in the context of small area estimation for GLMMs, assuming small areas are independent from each other, due to some difficulties to develop small area predictors and their corresponding precisions, e.g., mean squared prediction errors (MSPE), from a frequentist perspective. This issue is also added if there is a spatial pattern through the small areas. These models are widely applicable in statistics or health agencies. For example, it is an interest of policy makers (and public) to know the spatial pattern of a rare disease (e.g., chronic disease or cancer) to identify the areas with high risk of disease to implement the prevention. In this paper, we propose small area models in the class of spatial GLMMs to predict small area parameters and also to obtain second-order MSPE estimation of small area predictors using Taylor expansion and parametric bootstrap approaches. Performance of the proposed approach is evaluated through simulation studies and by a real application.

KEY WORDS: generalized linear mixed model, maximum likelihood estimation, parametric bootstrap, small area estimation, spatial model, Taylor expansion

## 1. Introduction

Sample surveys are conducted with the purpose of providing reliable predictors for the finite population characteristics such as totals or means. Methods used in deriving such predictors (direct survey predictors) are based on total sample size. However for the past few decades, there have been increasing demand in using same sample survey data to get predictions for sub-populations, such as counties or gender-age groups. Such sub-populations for which reliable predictions are needed are called small areas in the literature. The traditional area-specific direct predictors tend to have inadequate precision due to small sample sizes corresponding to each small area. Since policy decisions about implementing specific projects to these small areas are made using predictions on underlying characteristics, survey researchers are developing methods to provide more reliable predictions for small areas. To this end, model based estimators (Rao, 2003; Jiang and Lahiri, 2006; Jiang, 2010) have been proposed to borrow strength from other areas where different areas are related to each other by introducing random effects. Depending on the nature of the response variable, either linear mixed model (LMM) (Searle et al. 1992) or generalized linear mixed model (GLMM) (McCulloch and Searle, 2001) is mainly used for small area estimation (Fay and Herriot, 1979; Battese et al. 1988; Kass and Steffey, 1989; MacGibbon and Tomberlin, 1989; Prasad and Rao, 1990; Malec et al. 1997; Ghosh et al. 1998; Singh et al. 1998; Datta and Lahiri, 2000; Ghosh et al. 2009; Torabi et al. 2015). Among other approaches, parameters of the LMM can be estimated using either the maximum likelihood (ML) or restricted ML (REML). Although it is somewhat straightforward to predict the area statistics under the LMM, e.g., using the best linear unbiased predictor (BLUP), obtaining its prediction error and associated prediction interval is difficult. Both parameters estimation and prediction of small area statistics under the GLMM are computationally difficult under the frequentist approach.

In public health, the analysis of disease rates over areas has also received considerable attention due to growing demand for reliable disease rates in small areas. The idea behind developments on spatial and modelling of disease rates is essentially to model variations in true rates and better separate systematic variability from random noise, a component that usually overshadows crude rate maps (Torabi and Rosychuk, 2010; Torabi, 2012). Maps of regional morbidity and mortality rates are useful tools in determining spatial patterns of disease. Disease incidence and mortality rates may differ substantially across geographical areas. A reliable estimate of the underlying disease risk is usually provided by borrowing strength

*University of Manitoba, Winnipeg, MB, R3E 0W3, Canada

from neighbouring geographic sub-areas.

In this paper, we propose a unified approach for Normal and non-Normal responses with spatial patterns in the context of small area estimation. In particular, we provide prediction of small area parameters (Section 2) and obtain corresponding mean squared prediction error (MSPE) in Section 3. We also provide second-order unbiased estimators of MSPE of small area predictors using Taylor expansion and parametric bootstrap approaches (Section 4). In Section 5, we apply our approach to a real dataset esophagus cancer death cases in Minnesota USA from 1991–1998. In Section 6, performance of our proposed approach is also evaluated by two simulation studies (Normal and Poisson responses).

## 2. Statistical Model

The basic model in small area estimation can be described as follows. Let $y_i$ be the variable of interest for the $i$th area ($i = 1, ..., m$). The $y_i$ are assumed to be conditionally given $\eta_i$ independent with exponential family p.d.f.

$$f(y_i|\eta_i, \phi) = \exp[\{y_i\eta_i - a(\eta_i)\}/\phi + b(y_i, \phi)]. \quad (1)$$

The density (1) is parameterized with respect to the canonical parameters $\eta_i$, known scale parameters $\phi$ and functions $a(\cdot)$ and $b(\cdot)$. The natural parameters $\eta_i$ are then modeled as

$$\eta_i = x_i^{'}\beta + z_i^{\prime}u,$$

and $\eta_i = h[E(y_i|u)]$, where $h$ is a strictly increasing function, $x_i^{'}$ is the $i-$th row of known matrix $X(m \times p)$, $\alpha_1 := \beta(p \times 1)$ is a vector of unknown regression coefficient, $z_i^{'}$ is the $i-$th row of the identity matrix $Z(m \times m)$, and $u = (u_1, ..., u_m)'$ are spatial random effects from a multivariate Normal distribution $u|\alpha_2 \sim MVN(0, \Sigma_u(\alpha_2))$. The objective in small area estimation is to make inferences on the small area parameters $\eta_i$ or its variant.

To that end, we first need to obtain the full conditional density of the latent variable $\eta_i$ which can be written as

$$g(\eta_i|y_i, \alpha) \propto \exp\{\frac{-\eta_i^2}{2\sigma_{\eta_i}^2} + \frac{\eta_i(x_i^{\prime}\beta)}{\sigma_{\eta_i}^2} + [y_i\eta_i - a(\eta_i)]/\phi\}, \quad (2)$$

where $\sigma_{\eta_i}^2 = z_i^{\prime}\Sigma_u z_i$ and $\alpha = (\alpha_1, \alpha_2)'$. A Normal approximation, using Laplace approximation (Rue et al. 2009) centred around the point $\eta_i^0 = \arg\max_{\eta_i} f(y_i|\eta_i, \phi)$, to the density (2) is constructed by linearizing the likelihood part of equation (2) at a fixed point $\eta_i^0$. The feasibility of this Normal approximation is evaluated through simulation studies in Section 6. So, one can write the following for each area $i(= 1, ..., m)$ :

$$[y_i\eta_i - a(\eta_i)] \approx [y_i\eta_i^0 - a(\eta_i^0)] + (\eta_i - \eta_i^0)[y_i - a'(\eta_i^0)] - \frac{1}{2}(\eta_i - \eta_i^0)^2 a''(\eta_i^0), \quad (3)$$

where the first and second derivatives can be written in closed form. Inserting (3) into (2), the full conditional density of $\eta_i$ has a Normal approximation with conditional mean $E(\eta_i|y_i, \alpha)$ and conditional variance $var(\eta_i|y_i, \alpha)$ given by

$$E(\eta_i|y_i, \alpha) = x_i^{\prime}\beta + z_i^{\prime}\Sigma_u Z^{'} R^{-1}[l(y, \eta^0) - X\beta],$$

and

$$var(\eta_i|y_i, \alpha) = z_i^{\prime}[\Sigma_u - \Sigma_u Z^{'} R^{-1} Z\Sigma_u]z_i,$$

with $R = Z\Sigma_u Z^{'} + P$, $P$ is a diagonal matrix with entries $P_{i,i} = \phi/a''(\eta_i^0)$, $\eta^0 = (\eta_1^0, ..., \eta_m^0)'$, and $l_i(y_i, \eta_i^0) = [y_i - a'(\eta_i^0) + \eta_i^0 a''(\eta_i^0)]/a''(\eta_i^0)$, $(i = 1, ..., m)$.

When $\alpha$ is known, the best predictor of $\eta_i$ is given by $\tilde{\eta}_i^B(\alpha, y_i) = \tilde{\eta}_i^B = E(\eta_i|y_i, \alpha)$. Moreover, the only sensible prediction variance for $\eta_i$ is given by $E(\tilde{\eta}_i^B - \eta_i)^2 = var(\eta_i|y_i, \alpha) =: g_{1i}(\alpha)$. By estimating the model parameters $\alpha$, called $\hat{\alpha}$, the empirical best (EB) prediction of $\eta_i$ is given by

$$\hat{\eta}_i^{EB} = \tilde{\eta}_i^B(\hat{\alpha}, y_i)\{1 + O_p(m^{-1})\},$$

noting that we estimate the model parameters using maximum likelihood estimation approach via data cloning (see Lele et al. 2010 for more details of the data cloning approach).

## 3. Mean Squared Prediction Error Approximation

We now need to obtain the measure of variability of the $\hat{\eta}_i^{EB}$. To that end, we assume the following regularity conditions (referred to as RC later on) on the estimator $\hat{\alpha}$ and the predictor $\tilde{\eta}_i^B(\alpha, y_i)$ for large $m$ :

1) The dimension of $\alpha$ is bounded and the estimator $\hat{\alpha}$ satisfies that $(\hat{\alpha} - \alpha) = O_p(m^{-1/2})$ and $E(\hat{\alpha} - \alpha) = O(m^{-1/2})$.

2) We have $\eta_i = O_p(1)$ and $\tilde{\eta}_i^B(\alpha, y_i) = O_p(1)$ for $i = 1, ..., m$. In addition, the estimator $\tilde{\eta}_i^B(\alpha, y_i)$ is continuously differentiable w.r.t. $\alpha$, and $\frac{\partial\tilde{\eta}_i^B(\alpha, y_i)}{\partial\alpha} = O_p(1)$.

**Theorem 1** *Under the RC, a second-order approximation to the MSPE of the $\hat{\eta}_i^{EB}$, under the model (1), can be written as*

$$MSPE(\hat{\eta}_i^{EB}) = g_{1i}(\alpha) + g_{2i}(\alpha) + o(m^{-1}), \quad (4)$$

where $g_{2i}(\alpha) = tr\Big\{E[(\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\alpha})(\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\alpha})']E[(\hat{\alpha} - \alpha)(\hat{\alpha} - \alpha)']\Big\}$.

The proof is deferred to the Appendix A.

## 4. Mean Squared Prediction Error Estimation

### 4.1 Linearization Method

Since the approximated $MSPE$ (4) is a function of unknown parameters $\alpha$, it is not computable. We now obtain the estimation of $MSPE(\hat{\eta}_i^{EB})$ which is second-order unbiased in the sense that

$$E[mspe(\hat{\eta}_i^{EB})] = MSPE(\hat{\eta}_i^{EB}) + o(m^{-1}). \quad (5)$$

As shown in Theorem 1, the order of $g_{2i}(\alpha)$ is $O(m^{-1})$, so one can estimate $g_{2i}(\alpha)$ by $g_{2i}(\hat{\alpha})$ unbiasedly up to second-order. To estimate $g_{1i}(\alpha)$, the naive estimator $g_{1i}(\hat{\alpha})$ has a second-order bias due to $g_{1i}(\alpha) = O(1)$. We can then use a Taylor expansion about $\alpha$ for $g_{1i}(\alpha)$ as follows:

$$g_{1i}(\hat{\alpha}) = g_{1i}(\alpha) + (\hat{\alpha} - \alpha)'\frac{\partial g_{1i}(\alpha)}{\partial\alpha} + \frac{1}{2}(\hat{\alpha} - \alpha)'\frac{\partial^2 g_{1i}(\alpha)}{\partial\alpha\partial\alpha'}(\hat{\alpha} - \alpha) + o_p(m^{-1}).$$

We can then write

$$E[g_{1i}(\hat{\alpha})] = g_{1i}(\alpha) + g_{11i}(\alpha) + g_{12i}(\alpha) + o(m^{-1}),$$

where

$$g_{11i}(\alpha, y_i) = (\frac{\partial g_{1i}(\alpha)}{\partial\alpha})'E(\hat{\alpha} - \alpha),$$

and

$$g_{12i}(\alpha) = \frac{1}{2}tr\{(\frac{\partial^2 g_{1i}(\alpha)}{\partial\alpha\partial\alpha'})E[(\hat{\alpha} - \alpha)(\hat{\alpha} - \alpha)']\},$$

noting that the order of $g_{11i}(\alpha)$ and $g_{12i}(\alpha)$ are $O(m^{-1/2})$ and $O(m^{-1})$ under the RC, respectively.

**Theorem 2** *Under the RC, a second-order correct unbiased estimator of the MSPE of $\hat{\eta}_i^{EB}$, under the model (1), can be written as*

$$mspe(\hat{\eta}_i^{EB}) = g_{1i}(\hat{\alpha}) - g_{11i}(\hat{\alpha}) - g_{12i}(\hat{\alpha}) + g_{2i}(\hat{\alpha}).$$

We have derived the EB prediction and corresponding second-order unbiased $MSPE$ estimation of small area parameters for some popular distributions in exponential family (Normal, Poisson, and binomial) based on the above results (see the Appendix B).

### 4.2 Parametric Bootstrap Approach

We now obtain a nearly unbiased estimator of $MSPE(\hat{\eta}_i^{EB})$, in the sense of (5), using the parametric bootstrap approach. We first generate $u^* = (u_1^*, ..., u_m^*)'$ from a multivariate Normal distribution with mean 0 and variance-covariance $\Sigma_u(\hat{\alpha}_2)$. We then have $\eta_i^* = x_i'\hat{\beta} + z_i'u^*, (i = 1, ..., m)$. A bootstrap sample is then generated from $y_i^*|(\eta_i^*, \hat{\alpha}) \sim f(y_i^*|\eta_i^*, \hat{\alpha}); i = 1, ..., m$, noting that we construct the estimator $\hat{\alpha}^*$ from the bootstrap sample $(y_1^*, ..., y_m^*)$ with the same method used to obtain the estimator $\hat{\alpha}$. We then obtain the EB of $\eta_i^*$ using the bootstrap dataset $\{(y_i^*, x_i); i = 1, ..., m\}$ as $\hat{\eta}_i^{EB*} = \tilde{\eta}_i^B(\hat{\alpha}^*, y_i)$ for $i = 1, ..., m$. Hence, the bootstrap MSPE estimator of $\hat{\eta}_i^{EB*}$ is given by

$$mspe_{boot1}(\hat{\eta}_i^{EB}) = E_*\{(\hat{\eta}_i^{EB*} - \eta_i^*)^2\} \equiv \hat{w}_i, \quad (6)$$

where $E_*$ denotes the bootstrap expectation. We also provide a double bootstrap (Hall & Maiti, 2006) by drawing a second-phase bootstrap sample from a given bootstrap sample using the bootstrap model parameters given above. Proceeding as above with the second-phase bootstrap sample to get second-phase bootstrap MSPE as $MSPE_{**}(\hat{\eta}_i^{EB**}) = E_{**}\{(\hat{\eta}_i^{EB**} - \eta_i^{**})^2\}$, where $E_{**}$ denotes the second-phase bootstrap expectation. We have the following bootstrap MSPE estimators proposed by Hall & Maiti (2006):

$$mspe_{boot2}(\hat{\eta}_i^{EB}) \approx \begin{cases} 2\hat{w}_i - \hat{v}_i & \hat{w}_i \geq \hat{v}_i \\ \hat{w}_i exp\{-(\hat{v}_i - \hat{w}_i)/\hat{v}_i\} & \hat{w}_i < \hat{v}_i \end{cases} \quad (7)$$

and

$$mspe_{boot3}(\hat{\eta}_i^{EB}) \approx \hat{w}_i^2/\hat{v}_i, \quad (8)$$

where $\hat{v}_i = E_*[E_{**}\{(\hat{\eta}_i^{EB**} - \eta_i^{**})^2\}]$. In practice, we approximate $\hat{w}_i$ by drawing a large number, $B_1$, of independent bootstrap samples. Similarly, we approximate $\hat{v}_i$ by drawing a large number, $B_2$, of second-phase independent bootstrap samples from each first-phase bootstrap sample.

Table 1: Model parameters estimate and corresponding standard errors using maximum likelihood estimation approach

| Parameter | Estimate | Standard Error |
|---|---|---|
| $\beta$ | -0.041 | 0.054 |
| $\sigma_u^2$ | 0.012 | 0.005 |
| $\lambda_u$ | 0.290 | 0.029 |

## 5. Application

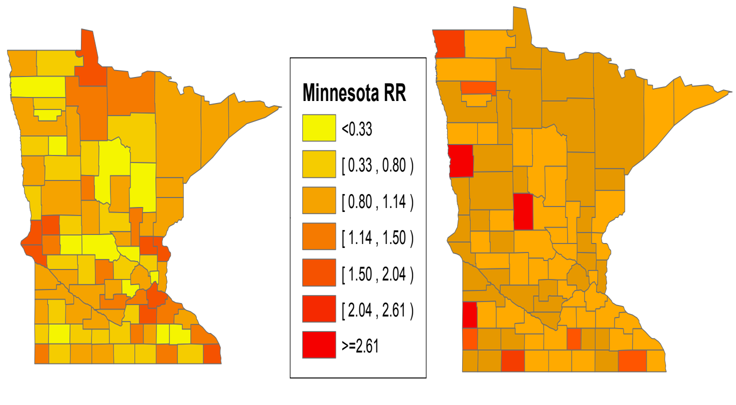
We use a non-Normal response data to evaluate the performance of the proposed approach. The data consists of the number of deaths due to esophagus cancer in the years from 1991 to 1998 at the 87 counties in Minnesota, USA (Jin et al. 2005; Torabi, 2014). A spatial Poisson regression model is used as this disease is assumed to be rare enough relative to the population in each county. The model is then given by

$$y_i \sim Poisson(\lambda_i), \qquad i = 1, ..., 87, \tag{9}$$

$$log(\lambda_i) = \log(E_i) + \beta + z_i' u,$$

where $y_i$ is the observed number of death due to esophagus cancer in county $i$, $E_i$ is the corresponding expected age-adjusted number of deaths, $\beta$ is a fixed effect, $z_i'$ is the $i-$th raw of the identity matrix $Z, u$ are from proper CAR model (see the Appendix B for more details of this model) with parameters $\alpha_2 = (\lambda_u, \sigma_u^2)$. Note that the expected number of deaths $(E_i)$ is calculated by $E_i = \sum_{j=1}^{J} n_{ij} y_j / n_j$ where $n_{ij}$ is the population at risk for the $i-$th county and age group $j, n_j$ is the population at risk for the age group $j$ based on the US Census 2010 dataset, and similarly $y_j$ is the number of deaths for the age group $j$.

We first fit the model (9) to the dataset and provide the model parameters estimate and corresponding standard errors (Table 1). We then provide the prediction of mortality ratio as well as raw ratio $(y_i/E_i)$ of esophagus cancer in each county (Figure 1) with corresponding MSPE estimation of log-ratio of esophagus cancer (Figure 2) using the Taylor expansion and parametric bootstrap approaches; noting that in this paper we consider $B_1 = 1000$ and $B_2 = 100$ for the bootstrap approaches. As shown in Figure 1, our prediction ratios provide smooth estimates of raw ratios.




(a) (b)

Figure 1: Raw (a) and EB prediction (b) of mortality ratio of esophagus cancer in Minnesota, spatial Poisson mixed model.

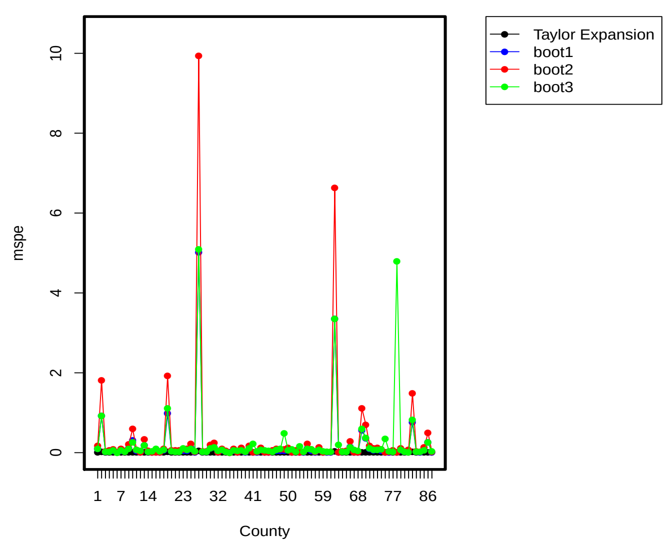



Figure 2: The $mspe$ of log-ratio of esophagus cancer in Minnesota, spatial Poisson mixed model.

# 6. Simulation Study

## 6.1 Normal Mixed Model

We also conduct a simulation study to evaluate the performance of the proposed approach in the Normal mixed model set-up. The simulation set-up is based on the spatial layout of the 87 counties in the state of Minnesota which was used in Section 5. We assume that the data are obtained from a Normal distribution as follows:

$$y_i = \beta + z_i' u + \epsilon_i, \qquad i = 1, ..., 87, \qquad (10)$$

where $\beta$ is a fixed effect, $z_i'$ is the $i-$th raw of the identity matrix $Z, u$ from proper CAR model with parameters $\alpha_2 = (\lambda_u, \sigma_u^2)'$, and $\epsilon_i \overset{i.i.d.}{\sim} N(0, \sigma^2)$ with known $\sigma^2 = 1$. We generate $R = 1000$ independent samples $\{y_i^{(r)}, i = 1, ..., 87; r = 1, ..., R\}$ where $y_i^{(r)} = \beta + z_i' u^{(r)} + \epsilon_i^{(r)}$, $u^{(r)}$ and $\epsilon_i^{(r)}$ are generated from the corresponding Normal distributions of $u$ and $\epsilon_i$ with $\lambda_u = 0.5, \sigma_u^2 = 0.01$, and $\beta = -1$. For each simulated run, we find the MLE of the model parameters to provide the prediction of the EBLUP of small area means $\eta_i^{(r)} = \beta + z_i' u^{(r)}, (r = 1, ..., R)$, using $\hat{\eta}_i^{EB(r)} = \hat{\beta}^{(r)} + z_i' \hat{u}^{(r)}$. We also calculate the empirical MSPE of $\hat{\eta}_i^{EB}$ as

$$EMSPE(\hat{\eta}_i^{EB}) = \frac{1}{R} \sum_{r=1}^{R} [\hat{\eta}_i^{EB(r)} - \eta_i^{(r)}]^2,$$

and the relative bias of an estimator of the MSPE, say $mspe$, as

$$RB[mspe(\hat{\eta}_i^{EB})] = \Big\{ \frac{1}{R} \sum_{r=1}^{R} mspe^{(r)}(\hat{\eta}_i^{EB})$$

$$-EMSPE(\hat{\eta}_i^{EB}) \Big\} / EMSPE(\hat{\eta}_i^{EB}),$$

where $\hat{\eta}_i^{EB(r)}$ and $\eta_i^{(r)}$, and $mspe^{(r)}(\hat{\eta}_i^{EB})$ are the values of $\hat{\eta}_i^{EB}, \eta_i$, and $mspe(\hat{\eta}_i^{EB})$ for the $r-$th simulation batch, respectively. Note that $mspe(\hat{\eta}_i^{EB})$ is calculated for both Taylor expansion and bootstrap approaches.

The result of $EMSPE$ of small area means is reported in Figure 3. As shown in Figure 3, the values of $EMSPE$ are relatively small for the proposed approach. The results of absolute relative bias (ARB) of $mspe$ of small area means for the Taylor expansion and bootstrap approaches are also reported in Figure 4. The proposed approach using Taylor expansion performs very well in terms of ARB ($<$ %3). In addition, the first-phase bootstrap ($\hat{w}_i$) seems to do a better job compared to the both second-phase bootstrap approaches in terms of RB of $mspe$ of small area means.

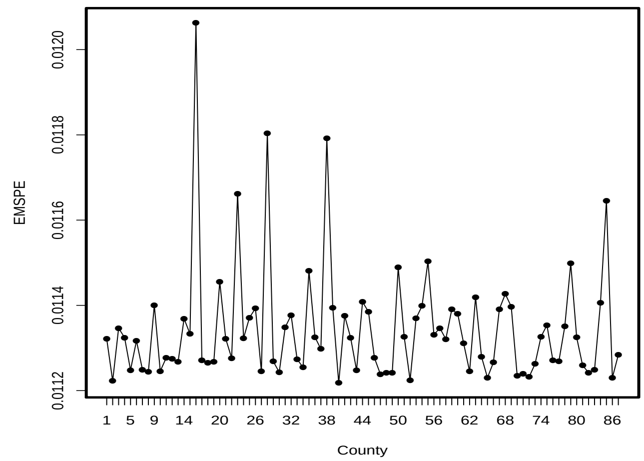



Figure 3: The EMSPE of small area means, spatial Normal mixed model.

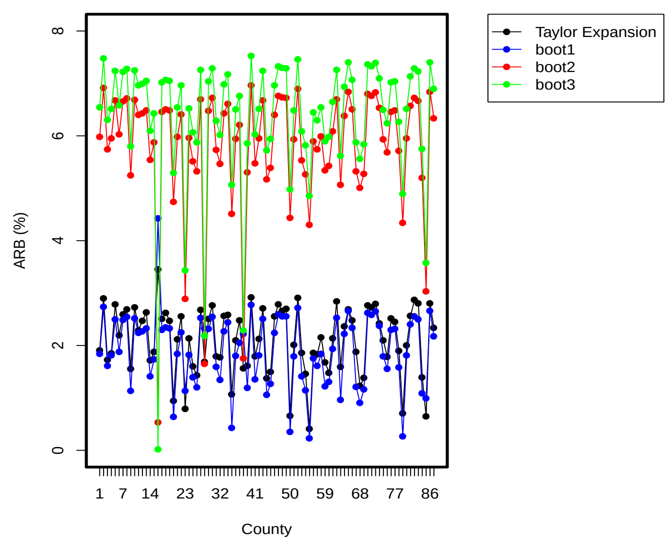



Figure 4: Percent absolute RB of $mspe$ of small area means, spatial Normal mixed model.

## 6.2 Poisson Mixed Model

We also conduct a simulation study to evaluate the performance of the proposed approach in the Poisson mixed model set-up. The spatial structure of the model is also based on the Minnesota county map (Section 5). We assume that the data are obtained from the following model:

$$y_i \sim Poisson(\lambda_i), \qquad i = 1, ..., 87, \qquad (11)$$

$$\log(\lambda_i) = \log(n_i) + \beta + z_i' u,$$

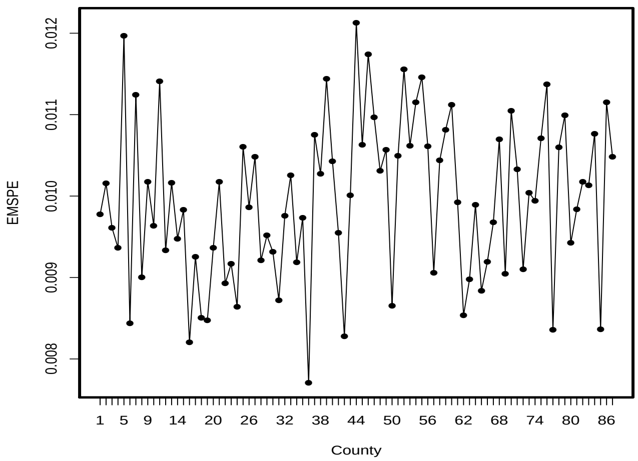



Figure 5: The EMSPE of $\hat{\eta}_i^{EB}$, spatial Poisson mixed model.

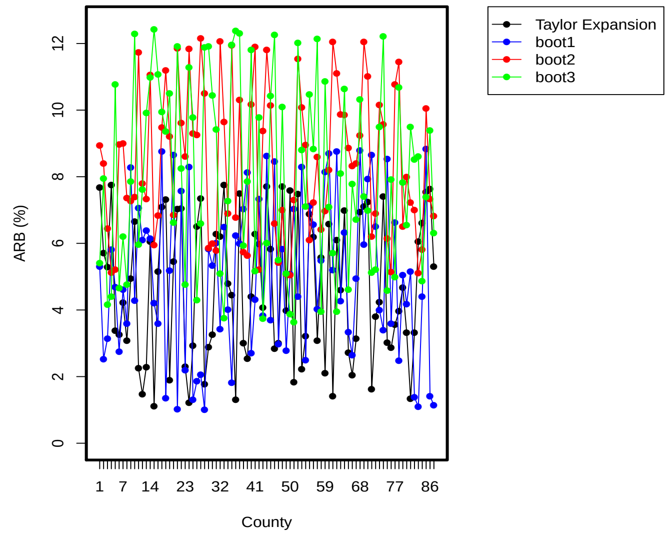



Figure 6: Percent absolute RB of $mspe$ of $\hat{\eta}_i^{EB}$, spatial Poisson mixed model.

where $n_i = 30$ as offset, $\beta$ is a fixed effect, $z_i'$ is the $i-$th raw of the identity matrix $Z, u$ are generated from the proper CAR model with parameters $\alpha_2 = (\lambda_u, \sigma_u^2)'$. We first generate $R = 1000$ independent samples $u^{(r)}, (r = 1, ..., R)$, from the proper CAR model with parameters $\lambda_u = 0.6, \sigma_u^2 = 0.0001$, and then generate $y_i^{(r)} \sim Poisson(\lambda_i^{(r)}), (i = 1, ..., 87; r = 1, ..., R)$, where $log(\lambda_i^{(r)}) = \log(n_i) + \beta + z_i' u^{(r)}$ with $\beta = 0.001$. For each simulated run, we find the MLE of the model parameters to provide the prediction of the small area log-rates $\eta_i^{(r)} = \beta + z_i' u^{(r)}, (r = 1, ..., R)$, using $\hat{\eta}_i^{EB(r)} = \hat{\beta}^{(r)} + z_i' \hat{u}^{(r)}$. We also calculate the $EMSPE(\hat{\eta}_i^{EB})$ and the $RB[mspe(\hat{\eta}_i^{EB})]$ similar to the Normal mixed model in Section 6.1.

The result of $EMSPE$ of $\hat{\eta}_i^{EB}$ is reported in Figure 5. As shown in Figure 5, the values of $EMSPE$ are relatively small for the proposed approach. The results of ARB of $mspe$ of $\hat{\eta}_i^{EB}$ for the Taylor expansion and bootstrap approaches are also reported in Figure 6. The proposed approach using Taylor expansion performs very well in terms of ARB ($< \%8$); noting that the first-phase bootstrap also performs better than the both second-phase bootstrap methods in terms of RB of $mspe$ of $\hat{\eta}_i^{EB}$.

## 7. Conclusions

There is a limited literature in the context of small area estimation for generalized linear mixed models (GLMM), assuming small areas are independent from each other, due to some difficulties to develop small area predictors and their corresponding precisions, e.g., mean squared prediction errors (MSPE), from a frequentist perspective. This issue is also added if there is a spatial pattern through the small areas. These models are widely applicable, for example, in statistics or health agencies. For instance, accurate statistical information concerning the wellbeing of people at regional level is needed to target the policies or programs aimed at reducing poverty in poorer regions; the estimation of poverty at regional or local level is then a really important task for policy making (Marhuenda et al. 2013). As another application, among many others, is when health agencies (e.g., policy making) need to know the spatial pattern of a rare disease (e.g., chronic disease or cancer) to identify the regions with high risk of disease to implement the prevention.

We have proposed a unified approach for Normal and non-Normal responses with spatial patterns in the context of small area estimation. In particular, we have provided prediction of small area parameters and derived second order approximation to the MSPE of small area parameters. We have also obtained second-order MSPE estimation of small area predictors by Taylor expansion as well as parametric bootstrap approaches. We have shown by simulation studies (and a real data application of esophagus cancer dataset in Minnesota) that the proposed approach works very well in terms of small area predictors and their precisions.

## Appendix

### A. Proof of Theorem 1

We can write

$$MSPE(\hat{\eta}_i^{EB}) = E\{(\hat{\eta}_i^{EB} - \eta_i)^2\}$$
$$= E\{(\tilde{\eta}_i^B - \eta_i)^2\} + E\{(\hat{\eta}_i^{EB} - \tilde{\eta}_i^B)^2\}$$
$$= g_{1i}(\alpha) + E\{(\hat{\eta}_i^{EB} - \tilde{\eta}_i^B)^2\} \ ,$$

noting that $E[(\tilde{\eta}_i^B - \eta_i)(\hat{\eta}_i^{EB} - \tilde{\eta}_i^B)] = E[(\hat{\eta}_i^{EB} - \tilde{\eta}_i^B)E(\tilde{\eta}_i^B - \eta_i|y_i)] = 0$. It is also noted that

$$\hat{\eta}_i^{EB} = \tilde{\eta}_i^B(\alpha, y_i) + (\frac{\partial \tilde{\eta}_i^B(\alpha^*, y_i)}{\partial \alpha})'(\hat{\alpha} - \alpha),$$

where $\alpha^*$ is between $\alpha$ and $\hat{\alpha}$. Thus, we obtain

$$E\{(\hat{\eta}_i^{EB} - \tilde{\eta}_i^B)^2\} = E[\{(\hat{\alpha}-\alpha)'\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \alpha}\}^2] + o(m^{-1})$$
$$= trE[(\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \alpha})(\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \alpha})'(\hat{\alpha} - \alpha)(\hat{\alpha} - \alpha)'] + o(m^{-1})$$
$$= tr\Big\{E[(\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \alpha})(\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \alpha})']E[(\hat{\alpha}-\alpha)(\hat{\alpha}-\alpha)']\Big\} + o(m^{-1}),$$

which completes the proof of Theorem 1.

### B. MSPE Estimation of Normal, Poisson, and Binomial Responses

### 1. Normal Response

In the linear mixed model, we have

$$y_i = x_i'\beta + z_i'u + \epsilon_i, \qquad i = 1, ..., m,$$

where $\epsilon_i \sim N(0, \sigma^2)$ and $u \sim N(0, \Sigma_u)$. In the case of proper CAR model for the spatial random effects $u$, we have $\Sigma_u = \sigma_u^2(I_m - \frac{\lambda_u}{a}C)^{-1}$ with $C = I_m - D$ where $D$ is a $m \times m$ matrix with elements $D_{ii} = e_i$, $D_{ij} = -1$ if area $i$ and $j$ are adjacent and $D_{ij} = 0$ otherwise, where $e_i$ is the number of areas which are adjacent to area $i; \sigma_u^2$ is the spatial dispersion parameter; $\lambda_u$ measures the conditional spatial dependence with range $[0, 1)$, $a$ is the maximum eigenvalue of $C$, and $I_m$ is the identity matrix of dimension $m$ (Leroux *et al.* 1999). Our interest is to make an inference about the small area mean $\eta_i = x_i'\beta + z_i'u$ where $y_i|u \sim N(\eta_i, \sigma^2)$. Based on the general result in Section 2, we have $E(u|y, \alpha) = \Sigma_u Z'(Z\Sigma_u Z' + \sigma^2 I_m)^{-1}(y - X\beta)$ and $var(u|y, \alpha) = \Sigma_u - \Sigma_u Z'(Z\Sigma_u Z' + \sigma^2 I_m)^{-1} Z\Sigma_u$, where $y = (y_1, ..., y_m)$. Consequently, we have the best predictor of $\eta_i$ as

$$\tilde{\eta}_i^B(\alpha, y_i) = x_i'\beta + z_i'\{\Sigma_u Z'(Z\Sigma_u Z' + \sigma^2 I_m)^{-1}[y - X\beta]\}, \tag{12}$$

and the EB prediction of $\eta_i$ is given by $\hat{\eta}_i^{EB} = \tilde{\eta}_i^B(\hat{\alpha}, y_i)$.

To get a second-order unbiased estimate of $MSPE(\hat{\eta}_i^{EB})$, we have

$$mspe(\hat{\eta}_i^{EB}) = g_{1i}(\hat{\alpha}) - g_{11i}(\hat{\alpha}) - g_{12i}(\hat{\alpha}) + g_{2i}(\hat{\alpha}), \tag{13}$$

where

$$g_{1i}(\alpha) = z_i'[\Sigma_u - \Sigma_u Z'(Z\Sigma_u Z' + \sigma^2 I_m)^{-1} Z\Sigma_u]z_i, \tag{14}$$

$$\frac{\partial g_{1i}(\alpha)}{\partial \sigma_u^2} = z_i'[\Sigma_u^{\sigma_u^2} - \Sigma_u^{\sigma_u^2} Z'R^{-1}Z\Sigma_u + \Sigma_u Z'R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}Z\Sigma_u^{\sigma_u^2}]z_i, \tag{15}$$

$$\frac{\partial g_{1i}(\alpha)}{\partial \lambda_u} = z_i'[\Sigma_u^{\lambda_u} - \Sigma^{\lambda_u} Z'R^{-1}Z\Sigma_u + \Sigma_u Z'R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}Z\Sigma_u^{\lambda_u}]z_i, \tag{16}$$

$$\frac{\partial^2 g_{1i}(\alpha)}{\partial \sigma_u^2 \partial \sigma_u^2} = z_i'[\Sigma_u^{\sigma_u^2} Z'R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u - \Sigma_u^{\sigma_u^2} Z'R^{-1}Z\Sigma_u^{\sigma_u^2} + \Sigma^{\sigma_u^2} Z'R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}R^{\sigma_u^2}R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}R^{\sigma_u^2}R^{-1}R^{\sigma_u^2}R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u + \Sigma_u Z^{-1}R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma^{\lambda_u} - \Sigma_u^{\sigma_u^2} Z'R^{-1}Z\Sigma_u^{\sigma_u^2} + \Sigma_u Z'R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u^{\sigma_u^2}]z_i, \tag{17}$$

$$\frac{\partial^2 g_{1i}(\alpha)}{\partial \lambda_u \partial \lambda_u} = z_i'[\Sigma_u^{\lambda_u^2} - \Sigma_u^{\lambda_u^2} Z'R^{-1}Z\Sigma_u + \Sigma_u^{\lambda_u} Z'R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u - \Sigma_u^{\lambda_u} Z'R^{-1}Z\Sigma_u^{\lambda_u} + \Sigma_u^{\lambda_u} Z'R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}R^{\lambda_u}R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u + \Sigma_u Z'R^{-1}R^{\lambda_u^2}R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}R^{\lambda_u}R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u + \Sigma_u Z'R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u^{\lambda_u} - \Sigma_u^{\lambda_u} Z'R^{-1}Z\Sigma_u^{\lambda_u} + \Sigma_u Z'R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u^{\lambda_u} - \Sigma_u Z'R^{-1}Z\Sigma_u^{\lambda_u^2}]z_i, \tag{18}$$

$$\frac{\partial^2 g_{1i}(\alpha)}{\partial \sigma_u^2 \partial \lambda_u} = z_i'[\Sigma_u^{\sigma_u^2 \lambda_u} - \Sigma_u^{\sigma_u^2 \lambda_u} Z'R^{-1}Z\Sigma_u + \Sigma_u^{\sigma_u^2} Z'R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u - \Sigma_u^{\sigma_u^2} Z'R^{-1}Z\Sigma_u^{\lambda_u} + \Sigma_u^{\lambda_u} Z'R^{-1}R^{\sigma_u^2}R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}R^{\lambda_u}R^{-1}R^{\sigma_u^2}R^{-1}Z\Sigma_u + \Sigma_u Z'R^{-1}R^{\sigma_u^2 \lambda_u} .R^{-1}Z\Sigma_u - \Sigma_u Z'R^{-1}R^{\sigma_u^2}R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u + \Sigma_u Z'R^{-1}R^{\sigma_u^2}R^{-1}Z .\Sigma_u^{\lambda_u} - \Sigma_u^{\lambda_u} Z'R^{-1}Z\Sigma_u^{\sigma_u^2} + \Sigma_u Z'R^{-1}R^{\lambda_u}R^{-1}Z\Sigma_u^{\sigma_u^2} - \Sigma_u Z'R^{-1}Zz\Sigma_u^{\sigma_u^2 \lambda_u}]z_i, \tag{19}$$

where $R = Z\Sigma_u Z' + \sigma^2 I_m, \Sigma_u^{\sigma_u^2} = (I_m - \frac{\lambda_u}{a}C)^{-1}$, $R^{\sigma_u^2} = Z\Sigma_u^{\sigma_u^2}Z'$, $\Sigma_u^{\lambda_u} = a^{-1}\Sigma_u(I_m - \frac{\lambda_u}{a}C)^{-1}C$, $\Sigma_u^{\lambda_u^2} = [\Sigma_u^{\lambda_u} + (\sigma_u^{-2}/a)\Sigma_u C\Sigma_u](1/a)(I_m - \frac{\lambda_u}{a}C)^{-1}C$, $\Sigma_u^{\sigma_u^2\lambda_u} = (\sigma_u^{-2}/a)(I_m - \frac{\lambda_u}{a}C)^{-1}C\Sigma_u$.

Also, we have $E(\hat{\alpha}_1 - \alpha_1) = 0$ and

$$E(\hat{\alpha}_2 - \alpha_2) =$$

$$\frac{1}{2m}\{I^{-1}(\alpha_2)col_{1\le j\le 2}tr[(XR^{-1}X)^{-1}(XR^{(j)}X)]\},$$

where $R^{(j)} = \frac{\partial R^{-1}}{\partial \alpha_{2(j)}} = -R^{-1}\frac{\partial R}{\partial \alpha_{2(j)}}R^{-1}$ with e.g. $R^{(\lambda_u)} = -R^{-1}R^{\lambda_u}R^{-1}$ and $R^{(\sigma_u^2)} = -R^{-1}R^{\sigma_u^2}R^{-1}$, and $I_{jk}(\alpha_2) = \frac{1}{2}tr[R^{-1}\frac{\partial R}{\partial \alpha_{2(j)}}R^{-1}\frac{\partial R}{\partial \alpha_{2(k)}}]$.

To get $g_{2i}(\hat{\alpha})$, we also need the following terms:

$$E\{[\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \beta}][\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \beta}]'\} = \{x_i - z_i[\Sigma_u Z'(Z\Sigma_u Z'$$

$$+\sigma^2 I_m)^{-1}X]\}\{x_i - z_i[\Sigma_u Z'(Z\Sigma_u Z' + \sigma^2 I_m)^{-1}X]\}',$$

$$E\{[\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \sigma_u^2}]^2\} = z_i'\Sigma_u^{\sigma_u^2}Z'(Z\Sigma_u Z' + \sigma^2 I_m)^{-1}$$

$$.\Big[I_m + \sigma_u^4(Z\Sigma_u^{\sigma_u^2}Z')(Z\Sigma_u Z' + \sigma^2 I_m)^{-2}(Z'\Sigma_u^{\sigma_u^2}Z)$$

$$-2\sigma_u^2(Z\Sigma_u Z' + \sigma^2 I_m)^{-1}(Z\Sigma_u^{\sigma_u^2}Z')\Big]Z\Sigma_u^{\sigma_u^2}z_i,$$

$$E\{[\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \lambda_u}]^2\} = z_i'\Sigma_u^{\lambda_u}Z'(Z\Sigma_u Z' + \sigma^2 I_m)^{-1}$$

$$.[Z\Sigma_u^{\lambda_u} - 2(Z\Sigma_u Z' + \sigma^2 I_m)^{-1}(Z\Sigma_u^{\lambda_u}Z')Z\Sigma_u]z_i$$

$$+z_i'\Sigma_u Z'(Z\Sigma_u^{\lambda_u}Z')(Z\Sigma_u Z' + \sigma^2 I_m)^{-3}(Z'\Sigma_u^{\lambda_u}Z)Z\Sigma_u z_i,$$

$$E\{[\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \lambda_u}][\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \sigma_u^2}]\} = 0,$$

and

$$E[(\hat{\alpha}_1 - \alpha_1)(\hat{\alpha}_1 - \alpha_1)'] = (X'R^{-1}X)^{-1},$$

$$E[(\hat{\alpha}_2 - \alpha_2)(\hat{\alpha}_2 - \alpha_2)'] = I^{-1}(\alpha_2).$$

## 2. Poisson Response

In the case of Poisson, we have $y_i|u_i \sim Poisson(\lambda_i)$ with $\log(\lambda_i) = x_i'\beta + z_i'u \equiv \eta_i$, where $u \sim N(0, \Sigma_u)$ and $u$ are spatial random effects as in the Normal case. Our interest is to make an inference about the small area parameter $\eta_i$. Based on the results of Section 2, we have

$$E(\eta_i|y_i, \alpha) = x_i'\beta + z_i'E(u|y, \alpha),$$

where $E(u|y, \alpha) = \Sigma_u Z'(Z\Sigma_u Z' + P)^{-1}[l(y, u^0) - X\beta]$ with

$$l[y, u^0] = \otimes[y - a'(u^0) + u^0 a''(u^0)]/a''(u^0)$$

$$= \otimes[y - e^{u^0} + u^0 e^{u^0}]/e^{u^0}$$

$$= y \otimes e^{-u^0} + u^0 - 1_m,$$

where $P = diag(e^{-u^0}), u^0 = (u_1^0, ..., u_m^0)'$, and $1_m$ is a vector of ones with dimension $m$. We then have the best prediction of $\eta_i$ as

$$\tilde{\eta}_i^B(\alpha, y_i) =$$

$$x_i'\beta + z_i'\{\Sigma_u Z'(Z\Sigma_u Z' + P)^{-1}[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]\}, \quad (20)$$

and the EB prediction of $\eta_i$ is given by $\hat{\eta}_i^{EB} = \tilde{\eta}_i^B(\hat{\alpha}, y_i)$. We can also have the variance of the best prediction of $\eta_i$ as $g_{1i}(\alpha) = z_i'[\Sigma_u - \Sigma_u Z'R^{-1}Z\Sigma_u]z_i$. Also, to get a second-order unbiased estimate of $MSPE(\hat{\eta}_i^{EB})$, we can use (15)-(19) to get the $mspe(\hat{\eta}_i^{EB})$ given by (13), noting that we should use $P = diag(e^{-u^0})$ and also

$$E\{[\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \beta}][\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \beta}]'\} = \{x_i - z_i[\Sigma_u Z'(Z\Sigma_u Z' + P)^{-1}X]\}$$

$$.\{x_i - z_i[\Sigma_u Z'(Z\Sigma_u Z' + P)^{-1}X]\}',$$

$$E\{[\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \sigma_u^2}]^2\} = z_i'\Sigma_u^{\sigma_u^2}Z'(Z\Sigma_u Z' + P)^{-1}\{E[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]$$

$$.[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]'\}(Z\Sigma_u Z' + P)^{-1}$$

$$.\{Z\Sigma_u^{\sigma_u^2} - 2(Z\Sigma_u Z' + P)^{-1}(Z\Sigma_u^{\sigma_u^2}Z')Z\Sigma_u\}z_i$$

$$+z_i'\Sigma_u Z'(Z\Sigma_u^{\sigma_u^2}Z')(Z\Sigma_u Z' + P)^{-2}\{E[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]$$

$$.[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]'\}(Z\Sigma_u Z' + P)^{-2}(Z\Sigma_u^{\sigma_u^2}Z')Z\Sigma_u z_i,$$

$$E\{[\frac{\partial \tilde{\eta}_i^B(\alpha, y_i)}{\partial \lambda_u}]^2\} = z_i'\Sigma_u^{\lambda_u}Z'(Z\Sigma_u Z' + P)^{-1}\{E[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]$$

$$.[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]'\}(Z\Sigma_u Z' + P)^{-1}$$

$$.\{Z\Sigma_u^{\lambda_u} - 2(Z\Sigma_u Z' + P)^{-1}(Z\Sigma_u^{\lambda_u}Z')Z\Sigma_u\}z_i$$

$$+z_i'\Sigma_u Z'(Z\Sigma_u^{\lambda_u}Z')(Z\Sigma_u Z' + P)^{-2}\{E[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]$$

$$.[y \otimes e^{-u^0} + u^0 - 1_m - X\beta]'\}(Z\Sigma_u Z' + P)^{-2}(Z\Sigma_u^{\lambda_u}Z')Z\Sigma_u z_i,$$

$$E\{[\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\sigma_u^2}][\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\lambda_u}]\} = z_i'[I - \Sigma_u Z'(Z\Sigma_u Z'+P)^{-1}$$

$$.Z][\Sigma_u^{\lambda_u\sigma_u^2} - \Sigma_u^{\lambda_u} Z'(Z\Sigma_u Z'+P)^{-1}Z\Sigma_u^{\sigma_u^2} - \Sigma_u^{\sigma_u^2}Z'$$

$$.(Z\Sigma_u Z'+P)^{-1}Z\Sigma_u^{\lambda_u}][Z'(Z\Sigma_u Z'+P)^{-1}E(l(y,u^0)-X\beta)], \quad (21)$$

where $E(y_i) = \exp(x_i'\beta + z_i'\Sigma_u z_i/2)$, $var(y_i) = \exp(x_i'\beta + z_i'\Sigma_u z_i/2) + \exp(2x_i'\beta)\{\exp(2z_i'\Sigma_u z_i) - \exp(z_i'\Sigma_u z_i)\}$, and $cov(y_i,y_j) = \exp\{(x_i + x_j)'\beta\}\Big[\exp\{(z_i+z_j)'\Sigma_u(z_i+z_j)/2\} - \exp\{z_i'\Sigma_u z_i/2 + z_j'\Sigma_u z_j/2\}\Big]$, for $i \neq j (i = 1, ..., m)$, needed in the above equations. Note that one can also use Monte Carlo to calculate $E(\hat{\alpha}-\alpha)$ and inverse of Fisher information matrix via data cloning to get $E[(\hat{\alpha}-\alpha)(\hat{\alpha}-\alpha)']$.

## 3. Binomial Response

In the case of binomial, we have $y_i|u_i \sim Bin(n_i, p_i)$ with $\log(\frac{p_i}{1-p_i}) = x_i'\beta + z_i'u \equiv \eta_i$, where $u \sim N(0, \Sigma_u)$ and $u$ are proper spatial random effects as in the Normal case. Our interest is to make an inference about the small area parameter $\eta_i$. Based on the results of Section 2, we have

$$E(\eta_i|y_i,\alpha) = x_i'\beta + z_i'E(u|y,\alpha)$$

where $E(u|y,\alpha) = \Sigma_u Z'(Z\Sigma_u Z' + P)^{-1}[l(y,u^0) - X\beta]$ with

$$l[y,u^0] = \otimes[y - a'(u^0) + u^0 a''(u^0)]/a''(u^0)$$

$$= \otimes[y - \frac{e^{u^0}}{1_m + e^{u^0}} + \frac{u^0 e^{u^0}}{(1_m+e^{u^0})^2}]/\frac{e^{u^0}}{(1_m+e^{u^0})^2}$$

$$= \otimes[ye^{-u^0}(1_m + e^{u^0})^2] - (1_m + e^{u^0}) + u^0,$$

and $P = diag\{\otimes(1_m + e^{u^0})^2 e^{-u^0}\}$. We then have the best prediction of $\eta_i$ as

$$\tilde{\eta}_i^B(\alpha,y_i) = x_i'\beta + z_i'\Big[\Sigma_u Z'(Z\Sigma_u Z' + P)^{-1}$$

$$.\{\otimes[ye^{-u^0}(1_m+e^{u^0})^2] - (1_m + e^{u^0}) + u^0 - X\beta\}\Big],$$

and the EB prediction of $\eta_i$ is given by $\hat{\eta}_i^{EB} = \tilde{\eta}_i^B(\hat{\alpha}, y_i)$. We can also have the variance of the best prediction of $\eta_i$ as $g_{1i}(\alpha) = z_i'[\Sigma_u - \Sigma_u Z'R^{-1}Z\Sigma_u]z_i$. Also, to get a second-order unbiased estimate of $MSPE(\hat{\eta}_i^{EB})$, we can use (15)-(19) to get the $mspe(\hat{\eta}_i^{EB})$ given by (13), noting that we should use $P = diag\{\otimes(1_m + e^{u^0})^2 e^{-u^0}\}$, and also

$$E\{[\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\beta}][\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\beta}]'\} = \{x_i - z_i[\Sigma_u Z'(Z\Sigma_u Z'$$

$$+P)^{-1}X]\}\{x_i - z_i[\Sigma_u Z'(Z\Sigma_u Z' + P)^{-1}X]\}',$$

$$E\{[\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\sigma_u^2}]^2\} = z_i'\Sigma_u^{\sigma_u^2}Z'(Z\Sigma_u Z' + P)^{-1}$$

$$.\Big[E\{\otimes[ye^{-u^0}(1+e^{u^0})^2] - (1+e^{u^0}) + u^0 - X\beta\}\{\otimes[ye^{-u^0}(1+e^{u^0})^2]$$

$$-(1+e^{u^0}) + u^0 - X\beta\}'\Big](Z\Sigma_u Z' + P)^{-1}$$

$$.\{Z\Sigma_u^{\sigma_u^2} - 2(Z\Sigma_u Z' + P)^{-1}(Z\Sigma_u^{\sigma_u^2}Z')Z\Sigma_u\}z_i$$

$$+z_i'\Sigma_u Z'(Z\Sigma_u^{\sigma_u^2}Z')(Z\Sigma_u Z' + P)^{-2}$$

$$.\Big[E\{\otimes[ye^{-u^0}(1+e^{u^0})^2] - (1+e^{u^0}) + u^0 - X\beta\}\{\otimes[ye^{-u^0}(1+e^{u^0})^2]$$

$$-(1+e^{u^0}) + u^0 - X\beta\}'\Big](Z\Sigma_u Z' + P)^{-2}(Z\Sigma_u^{\sigma_u^2}Z')Z\Sigma_u z_i,$$

$$E\{[\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\lambda_u}]^2\} = z_i'\Sigma_u^{\lambda_u}Z'(Z\Sigma_u Z' + P)^{-1}$$

$$.\Big[E\{\otimes[ye^{-u^0}(1+e^{u^0})^2] - (1+e^{u^0}) + u^0 - X\beta\}\{\otimes[ye^{-u^0}(1+e^{u^0})^2]$$

$$-(1+e^{u^0}) + u^0 - X\beta\}'\Big](Z\Sigma_u Z' + P)^{-1}\{Z\Sigma_u^{\lambda_u} - 2(Z\Sigma_u Z' + P)^{-1}$$

$$.(Z\Sigma_u^{\lambda_u}Z')Z\Sigma_u\}z_i + z_i'\Sigma_u Z'(Z\Sigma_u^{\lambda_u}Z')(Z\Sigma_u Z' + P)^{-2}$$

$$.\Big[E\{\otimes[ye^{-u^0}(1+e^{u^0})^2] - (1+e^{u^0}) + u^0 - X\beta\}\{\otimes[ye^{-u^0}(1+e^{u^0})^2]$$

$$-(1+e^{u^0}) + u^0 - X\beta\}'\Big](Z\Sigma_u Z' + P)^{-2}(Z\Sigma_u^{\lambda_u}Z')Z\Sigma_u z_i,$$

and $E\{[\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\sigma_u^2}][\frac{\partial\tilde{\eta}_i^B(\alpha,y_i)}{\partial\lambda_u}]\}$ is similar to (21); noting that the expectations involved in the above equations are calculated via Monte Carlo.

## 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," *Journal of the American Statistical Association*, 83, 28–36.

Datta, G.S., and Lahiri, P. (2000), "A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems," *Statistica Sinica*, 10, 613–627.

Fay, R.E., and Herriot, R.A. (1979), "Estimation of income for small places: an application of James-Stein procedures to census data," *Journal of the American Statistical Association*, 74, 269–277.

Ghosh, M., Kim, D., Sinha, K., Maiti, T., Katzoff, M., and Parsons, V.L. (2009), "Hierarchical and empirical Bayes small domain estimation of the proportion of persons without health insurance for minority subpopulations," *Survey Methodology*, 35, 53–66.

Ghosh, M., Natarajan, K., Stroud, T.W.F., and Carlin, B.P. (1998), "Generalized linear models for small-area estimation," *Journal of the American Statistical Association*, 93, 273–282.

Hall, P., and Maiti, T. (2006), "On parametric bootstrap methods for small area prediction," *Journal of Royal Statistical Society, Series B Statistical Methodology*, 68, 221–238.

Jiang, J. (2010), *Large Sample Techniques for Statistics.* New York: Springer.

Jiang, J., and Lahiri, P. (2006), "Mixed model prediction and small area estimation (with discussion)," *Test*, 15, 1–96.

Jin, X., Carlin, B.P., and Banerjee, S. (2005), " Generalized hierarchical multivariate CAR models for areal data," *Biometrics*, 61, 950–961.

Kass, R.E., and Steffey, D. (1989), "Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models)," *Journal of the American Statistical Association*, 84, 717–726.

Leroux, B., Lei, X., and Breslow, N. (1999), "Estimation of disease rates in small areas: A new mixed model for spatial dependence," Springer-Verlag, New York, *Ch. Statistical Models in Epidemiology, the Environment and Clinical Trials*, Halloran, M and Berry, D (eds), pp. 135–178

Lele, S.R., Nadeem, K., and Schmuland, B. (2010), "Estimability and likelihood inference for generalized linear mixed models using data cloning," *Journal of the American Statistical Association*, 105, 1617–1625.

MacGibbon, B., and Tomberlin, T.J. (1989), "Small area estimation of proportions via empirical Bayes techniques," *Survey Methodology*, 15, 237–252.

Malec, D., Sedransk, J., Moriarity, C.L., and LeClere, F.B. (1997), "Small area inference for binary variables in the national health interview survey," *Journal of the American Statistical Association*, 92, 815–826.

Marhuenda, Y., Molina, I., and Morales, D. (2013), "Small area estimation with spatio-temporal Fay-Herriot models," *Computational Statistics and Data Analysis*, 58, 308–325.

McCulloch, C.E., and Searle, S.R. (2001), *Generalized, Linear, and Mixed Models.* New York: John Wiley and Sons.

Prasad, N.G.N., and Rao, J.N.K. (1990), "The estimation of mean squared error of small-area estimators," *Journal of the American Statistical Association*, 85,163–171.

Rao, J.N.K. (2003), *Small Area Estimation.* New York: John Wiley and Sons.

Rue, H., Martino, S., and Chopin, N. (2009), "Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations," *Journal of Royal Statistical Society, Series B Statistical Methodology*, 71,319–392.

Searle, S.R., Casella, G., and McCulloch, C.E. (1992), *Variance Components.* New York: John Wiley and Sons.

Singh, A.C., Stukel, D.M., and Pfeffermann, D. (1998), "Bayesian versus frequentist measures of error in small area estimation," *Journal of Royal Statistical Society, Series B Statistical Methodology*, 60, 377–396.

Torabi, M. (2012), "Hierarchical Bayes estimation of spatial statistics for rates," *Journal of Statistical Planning and Inference*, 142, 358–365.

Torabi, M., Lele, S.R., and Prasad, N.G.N. (2015), "Likelihood inference for small area estimation using data cloning," *Computational Statistics and Data Analysis*, 89, 158–171.

Torabi, M., and Rosychuk, R.J. (2010), "Spatio-temporal modelling of disease mapping of rates," *Canadian Journal of Statistics*, 38, 698–715.