Journal of Statistical Computation & Simulation
Vol. 00, No. 00, December 2012, 1–10

RESEARCH ARTICLE

Hierarchical Bayes estimation in small area estimation using cross-sectional and time-series data

Mahmoud Torabi[a*] and Farhad Shokoohi[b]

[a]*Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada R3E 0W3*; [b]*Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran 1983963113, Iran.*

(*v4.2 released August 2012*)

Bayesian methods have been extensively used in small area estimation. A linear model incorporating autocorrelated random effects and sampling errors was previously proposed in small area estimation using both cross-sectional and time-series data in the Bayesian paradigm. There are, however, many situations that we have time-related counts or proportions in small area estimation; for example monthly dataset on the number of incidence in small areas. This article considers hierarchical Bayes generalized linear models for a unified analysis of both discrete and continuous data with incorporating cross-sectional and time-series data. The performance of the proposed approach is evaluated through several simulation studies and also by a real dataset.

Keywords: Bayesian computation; Hierarchical model; Random effects; Time series.

AMS Subject Classification: 62D99; 62F15; 62J12

1. Introduction

Small area estimation has received a lot of attention in recent years due to growing demand for reliable small area statistics. Rao [15], Jiang and Lahiri [9] and Jiang [8] have given comprehensive accounts of model-based small area estimation. In particular, area level [4] and nested error linear regression models [1, 13] are often used in small area estimation to obtain efficient model-based estimators of small area means.

Most of the research on small area estimation has focused on cross-sectional data at a given point in time, and the research based on time series in the context of small area estimation is limited. Scott and Smith [18], Jones [10] among others used time-series methods to develop efficient estimates of aggregated parameters from repeated surveys. Tiller [23] used the idea of Kalman filter to combine a current-period state-wide estimate from the U.S. Current Population Survey with past estimate for the same state. However, non of them studied small area estimation by combining cross-sectional and time-series data.

Pferrermann and Burck [12] and Singh et al. [20] among others studied cross-sectional and time-series models for small area estimation using Kalman filter by assuming specific models for the sampling errors over time. Rao and Yu [16, 17]

*Corresponding author. Email: torabi@cc.umanitoba.ca; fax: +1-204-789-3905

ISSN: 0094-9655 print/ISSN 1563-5163 online

DOI: 10.1080/0094965YYxxxxxxxx
http://www.informaworld.com

proposed a combined cross-sectional and time-series linear model involving autocorrelated random effects and sampling errors using Bayesian and frequentist approaches, respectively. Using Bayesian approach, Datta et al. [3] applied same model as Rao-Yu model but replacing autoregressive (AR) random effects with random walk model. Datta et al. [2] considered a similar model but added extra terms to reflect seasonal variation in their application. Torabi [24] extended Datta et al. [2] model to account for spatial variation over regions.

The main purpose of this paper is to extend the Rao-Yu model for non-Normal data in the Bayesian framework. There are many applications in small area estimation where the responses are time-related counts or proportions. For instance, we may be interested to analyze monthly or yearly dataset of number of incidence in small areas. Indeed, these types of models fall in the class of Generalized Linear Mixed Models (GLMMs).

In this paper, we use Bayesian approach to propose a combined cross-sectional and time-series model with AR(1) for non-Normal data. In the next section, we describe the combined cross-sectional and time-series models. We then describe how Bayesian paradigm can be used to make inference for the small area parameters. The performance of proposed approach is reported in several simulation studies with a corresponding evaluation of sensitivity of such type of analysis to prior assumptions, and also by a real dataset. Finally, some concluding remarks are given.

2. Cross-sectional and time-series models

The basic model for the combined cross-sectional and time-series data can be described as follows. Let y_{it} be the variable of interest for the ith area in given time $t (t = 1, ..., T; i = 1, ..., m)$. The y_{it} are assumed to be conditionally independent with exponential family p.d.f.

$$f(y_{it}|\theta_{it}, \phi_{it}) = \exp[\{y_{it}\theta_{it} - a(\theta_{it})\}/\phi_{it} + b(y_{it}, \phi_{it})], \tag{1}$$

$(t = 1, ..., T; i = 1, ..., m)$. The density (1) is parameterized with respect to the canonical parameters θ_{it}, known scale parameters ϕ_{it} and functions $a(\cdot)$ and $b(\cdot)$. The exponential family (1) covers well-known distributions including Normal, binomial and Poisson distributions. The natural parameters θ_{it} are then modeled as

$$h(\theta_{it}) = \boldsymbol{x}'_{it}\boldsymbol{\beta} + v_i + u_{it},$$

where h is a strictly increasing function, $\boldsymbol{x}_{it}(p\times 1)$ are known design vectors, $\boldsymbol{\beta}(p\times 1)$ is a vector of unknown regression coefficient, $v_i \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$, and u_{it}'s are assumed to follow a common AR(1) process for each i, that is,

$$u_{it} = \rho u_{i,t-1} + \epsilon_{it}, \quad |\rho| < 1,$$

with $\epsilon_{it} \overset{i.i.d.}{\sim} N(0, \sigma_\epsilon^2)$.

As a special case, under Normal distribution $h(\theta_{it}) = \theta_{it}$, the Rao-Yu model is given by

$$\hat{\theta}_{it} = \theta_{it} + e_{it},$$

where e_{it}'s are sampling errors normally distributed, given the θ_{it}'s, with zero means and a known block diagonal covariance matrix Ψ with blocks Ψ_i. The errors $(v_i, \epsilon_{it}, e_{it})$ are also assumed to be independent of each other. Our goal is to make inference for small area parameters θ_{ij} or function of θ_{ij}.

3. Bayesian inference

The Bayesian approach is employed to estimate the small area parameters. The Gibbs sampler (e.g., [5, 7]) may be used to obtain the posterior mean and posterior variance of small area parameters. To generate samples from the posterior distribution using Markov chain Monte Carlo (MCMC) method via the Gibbs sampler, we need to sample from the full conditional distributions. Note that in our application, all of these full conditional distributions are standard distributions that can be easily sampled. To implement our application in the hierarchical Bayes (HB) setup, we use the WinBUGS software [22].

4. Simulation study

4.1. *Linear mixed model*

We conduct a simulation study to evaluate the performance of Bayesian approach in the linear mixed model set up. Note that Rao and Yu [16] studied a linear model with incorporating cross-sectional and time-series data in the Bayesian framework, however, they did not evaluate its performance. We consider the following model:

$$\begin{aligned} y_{it} &= v_i + u_{it} + e_{it}(t = 1, ..., T; i = 1, ..., m), \\ u_{it} &= \rho u_{i,t-1} + \epsilon_{it}, \qquad |\rho| < 1 \end{aligned}$$

with $\rho = 0.2, 0.4, e_{it} \overset{i.i.d.}{\sim} N(0, 1), v_i \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$ and $\epsilon_{it} \overset{i.i.d.}{\sim} N(0, \sigma_\epsilon^2)$. Similar to Rao and Yu [17], we set $m = 40$ small areas and $T = 5$, and then generate $R = 5000$ independent samples $\{y_{it}^{(r)}; t = 1, ..., T; i = 1, ..., m; r = 1, ..., R\}$ for each selected pair $(\sigma_v^2, \sigma_\epsilon^2)$ and ρ, and keep Ψ_i as an identity matrix. For each simulated sample, we apply MCMC method to get Bayesian inference of the small area parameters $\theta_{it} = v_i + u_{it}$.

In this paper, the proper priors are used for variance components. In particular, the gamma distribution was used for the inverse of variance components with shape and scale parameter 0.001. We also considered uniform distribution $U(-1, 1)$ for ρ. To monitor the convergence of the model parameters, we used several diagnostic methods implemented in the Bayesian output analysis (BOA) program [21], a freely available package created for R [14]. For this simulation set up, we used two chains and the average number of iterations for convergence of the model parameters was about 20,000.

Similar to Rao and Yu [17], we report the estimator of mean squared prediction error (MSPE) for only $\hat{\theta}_{1T}$. The true MSPE (TMSPE) of $\hat{\theta}_{1T}$, and relative bias (RB) of an estimator of the MSPE, say mspe, are given by

$$\text{TMSPE}(\hat{\theta}_{1T}) = \frac{1}{R} \sum_{r=1}^{R} \{\hat{\theta}_{1T}^{(r)} - \theta_{1T}^{(r)}\}^2,$$

Table 1. Percent relative bias of estimators of MSPE in linear mixed model.

		σ_ϵ^2			
	σ_v^2	0.25	0.5	1.0	2.0
$\rho = 0.2$	0.25	-12.4	-7.1	-2.2	-1.1
	0.5	-11.0	-6.6	-2.1	-1.0
	1.0	-11.0	-6.7	-2.3	-1.0
	2.0	-11.3	-7.2	-2.6	-1.2
$\rho = 0.4$	0.25	-12.1	-5.4	-1.1	0.2
	0.5	-11.0	-5.8	-1.2	0.0
	1.0	-10.4	-6.0	-1.4	-0.1
	2.0	-11.2	-6.4	-1.6	-0.2

and

$$\text{RB}\{\text{mspe}(\hat{\theta}_{1T})\} = \Big\{\frac{1}{R}\sum_{r=1}^{R}\text{mspe}^{(r)}(\hat{\theta}_{1T}) - \text{TMSPE}(\hat{\theta}_{1T})\Big\}/\text{TMSPE}(\hat{\theta}_{1T}),$$

where $\hat{\theta}_{1T}^{(r)}, \theta_{1T}^{(r)}$, and $\text{mspe}^{(r)}(\hat{\theta}_{1T})$ are the values of $\hat{\theta}_{1T}, \theta_{1T}$, and $\text{mspe}(\hat{\theta}_{1T})$ for the rth simulation study, respectively. Note that $\text{mspe}(\hat{\theta}_{1T})$ is the posterior variance of $\hat{\theta}_{1T}$.

The results of RB of $\text{mspe}(\hat{\theta}_{1T})$ are reported in Table 1 for different ρ and pair of $(\sigma_v^2, \sigma_\epsilon^2)$. As shown, the estimator of MSPE performs well for higher between-time variation for both $\rho = 0.2$ and $\rho = 0.4$; noting that the RB is slightly smaller for $\rho = 0.4$ compared to $\rho = 0.2$ in most scenarios. However, the RB is stable with increasing area-specific variation.

We also study the performance of the prediction interval of $\hat{\theta}_{1T}$. To this end, for each simulation run r, we calculate $\theta_{1T}^{(r)} = v_1^{(r)} + u_{1T}^{(r)}$ and compute appropriate quantiles α and $(1-\alpha)$ of the posterior means $\hat{\theta}_{1T}^{(r)}$. The coverage probabilities of the $\hat{\theta}_{1T}$ is the proportion of the times (over $R = 5000$) that $\theta_{1T}^{(r)}$ falls within $(\hat{\theta}_{1T}^{(r)(\alpha)}, \hat{\theta}_{1T}^{(r)(1-\alpha)})$. Table 2 shows the coverage probabilities of the estimates of θ_{1T}. The Bayesian method performs very well in terms of coverage probabilities of the $\hat{\theta}_{1T}$ for different confidence coefficients for both $\rho = 0.2$ and $\rho = 0.4$. In particular, for different σ_v^2, the coverage probabilities reach the nominal values with increasing between-time variation.

4.2. *Logistic mixed model*

We also conduct a simulation study to evaluate the performance of Bayesian approach in the logistic mixed model set up. To that end, we first generate $R = 5000$ independent samples:

$$y_{it,s}^{(r)} \sim Binomial(n, p_{it}^{(r)}) \tag{2}$$

$$\log\Big(\frac{p_{it}^{(r)}}{1-p_{it}^{(r)}}\Big) = v_i^{(r)} + u_{it}^{(r)} (t = 1, ..., T; i = 1, ..., m),$$

where y_{it} is the number of "success" in the ith area at time t with corresponding success rate p_{it} and sample size $n, v_i^{(r)} \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$, and $u_{it}^{(r)}$ is generated from

Table 2. Coverage probabilities of the $\hat{\theta}_{1T}$ in linear mixed model with different confidence coefficients.

	σ_v^2	σ_ϵ^2	Confidence coefficient			
			0.90	0.95	0.98	0.99
$\rho = 0.2$	0.25	0.25	0.859	0.925	0.966	0.982
		0.5	0.879	0.934	0.973	0.985
		1.0	0.892	0.944	0.979	0.988
		2.0	0.898	0.948	0.978	0.988
	0.5	0.25	0.866	0.927	0.967	0.983
		0.50	0.883	0.934	0.973	0.986
		1.0	0.896	0.944	0.978	0.988
		2.0	0.896	0.946	0.978	0.987
	1.0	0.25	0.867	0.929	0.969	0.984
		0.50	0.883	0.934	0.972	0.986
		1.0	0.894	0.943	0.978	0.987
		2.0	0.898	0.947	0.978	0.987
	2.0	0.25	0.871	0.929	0.968	0.984
		0.50	0.883	0.935	0.972	0.985
		1.0	0.891	0.944	0.977	0.987
		2.0	0.898	0.947	0.977	0.988
$\rho = 0.4$	0.25	0.25	0.863	0.923	0.967	0.982
		0.5	0.883	0.939	0.975	0.987
		1.0	0.894	0.949	0.979	0.990
		2.0	0.901	0.951	0.980	0.990
	0.5	0.25	0.866	0.930	0.970	0.982
		0.50	0.881	0.939	0.975	0.987
		1.0	0.897	0.946	0.980	0.990
		2.0	0.900	0.952	0.980	0.990
	1.0	0.25	0.868	0.931	0.968	0.983
		0.50	0.884	0.939	0.973	0.987
		1.0	0.898	0.946	0.979	0.990
		2.0	0.900	0.951	0.979	0.990
	2.0	0.25	0.870	0.929	0.969	0.981
		0.50	0.883	0.939	0.975	0.986
		1.0	0.899	0.945	0.980	0.990
		2.0	0.881	0.939	0.978	0.989

AR(1) with appropriate ρ. We also generate $R = 5000$ independent non-samples:

$$y_{it,ns}^{(r)} \sim Binomial(N - n, p_{it}^{(r)}). \tag{3}$$

We set $N = 100, n = 5, \rho = 0.4$, and consider $T = 5$ for each selected pair $(\sigma_v^2, \sigma_\epsilon^2)$. To evaluate the role of number of areas (m) in the performance of Bayesian approach particulary in terms of RB, we consider three different number of areas $m = 20, 40$ and 80. For each simulation run r, the true small area proportion is $P_{it}^{(r)} = N^{-1}(y_{it,s}^{(r)} + y_{it,ns}^{(r)})$. We compute the small area proportions $\hat{p}_{it}$ from (2), for each simulation run r, called $\hat{p}_{it}^{(r)}$. For this simulation set up, with two chains, the average number of iterations for convergence of the model parameters was about 20,000. The TMSPE of $\hat{p}_{it}$ and RB of mspe$(\hat{p}_{it})$ are then given by

$$\begin{aligned} \text{TMSPE}(\hat{p}_{it}) &= R^{-1} \textstyle\sum_{r=1}^{R} (\hat{p}_{it}^{(r)} - P_{it}^{(r)})^2, \\ \text{RB}[\text{mspe}(\hat{p}_{it})] &= \left\{ \tfrac{1}{R} \textstyle\sum_{r=1}^{R} \text{mspe}(\hat{p}_{it}^{(r)}) - \text{TMSPE}(\hat{p}_{it}) \right\} / \text{TMSPE}(\hat{p}_{it}), \end{aligned}$$

Table 3. True MSPE of $\hat{p}_{1T}$ in logistic mixed model for different number of small areas (m) and variance components ($\sigma_v^2, \sigma_\epsilon^2$).

σ_v^2	σ_ϵ^2	m		
		20	40	80
1	1	0.020	0.020	0.020
	2	0.022	0.021	0.021
2	1	0.019	0.019	0.018
	2	0.020	0.020	0.020

Table 4. Percent relative bias of estimators of MSPE of $\hat{p}_{1T}$ in logistic mixed model for different number of small areas (m) and variance components ($\sigma_v^2, \sigma_\epsilon^2$).

σ_v^2	σ_ϵ^2	m		
		20	40	80
1	1	-5.3	-3.6	-2.5
	2	-1.8	1.2	1.1
2	1	-8.9	-5.1	-1.5
	2	0.6	-0.3	-0.8

Table 5. Coverage probability (and average length) for $\hat{p}_{1T}$ in logistic mixed model for different number of small areas (m) and variance components ($\sigma_v^2, \sigma_\epsilon^2$).

σ_v^2	σ_ϵ^2	m	Confidence coefficient			
			0.90	0.95	0.98	0.99
1	1	20	0.876(0.442)	0.931(0.515)	0.964(0.593)	0.977(0.641)
		40	0.883(0.447)	0.934(0.520)	0.967(0.598)	0.980(0.647)
		80	0.886(0.451)	0.941(0.524)	0.974(0.603)	0.984(0.651)
	2	20	0.888(0.461)	0.938(0.537)	0.965(0.617)	0.976(0.667)
		40	0.884(0.461)	0.935(0.537)	0.961(0.618)	0.972(0.667)
		80	0.891(0.463)	0.938(0.539)	0.969(0.620)	0.978(0.670)
2	1	20	0.864(0.414)	0.916(0.484)	0.951(0.561)	0.963(0.610)
		40	0.872(0.417)	0.929(0.488)	0.961(0.566)	0.970(0.614)
		80	0.879(0.422)	0.926(0.493)	0.960(0.571)	0.972(0.619)
	2	20	0.872(0.434)	0.919(0.508)	0.947(0.588)	0.957(0.639)
		40	0.874(0.436)	0.923(0.511)	0.956(0.591)	0.965(0.641)
		80	0.877(0.436)	0.920(0.510)	0.948(0.591)	0.958(0.641)

where mspe($\hat{p}_{it}$) is the posterior variance of $\hat{p}_{it}$. We also study the coverage probabilities of $\hat{p}_{it}$.

We report the TMSPE for only $\hat{p}_{1T}$ which is stable for different number of small areas (m) and variance components ($\sigma_v^2, \sigma_\epsilon^2$), (Table 3). The RB of mspe($\hat{p}_{1T}$) is reported in Table 4 which performs very well and also, in general, the RB is decreased with increasing number of small areas, as expected. The results of the coverage probabilities and average length (in parenthesis) of the $\hat{p}_{1T}$ for different number of small areas and confidence coefficients are reported in Table 5, which also provide good coverage probabilities of the $\hat{p}_{1T}$ for different confidence coefficients.

5. Sensitivity analysis

We now investigate the choice of priors through a sensitivity study for our simulation study, for example, for the logistic mixed model set up. Full details of the prior sensitivity and choice of models appear in [11]. The hyperprior distributions of the variance components are generally set to be vague to get the most information from the data. In general, the prior for the precision of the random effects (σ^{-2})

Table 6. True MSPE and RB(%) of $\hat{p}_{1T}$ for sensitivity analysis of prior distributions in the case $m = 40$.

Prior	A	B	C	D	E	F	G	H
TMSPE	0.0218	0.0202	0.0201	0.0200	0.0222	0.0220	0.0297	0.0198
RB(%)	-17.75	-3.58	-2.86	-2.50	-21.67	-17.85	-67.85	-0.70

is often specified as a gamma distribution with scale and shape parameters both equal to 0.001. One may also use a uniform prior for the standard errors [6].

To investigate the influence of hyperprior specifications in the logistic context, we conduct a sensitivity analysis with respect to the prior distributions for the precision of random effects parameters σ_v^{-2} and σ_ϵ^{-2}, assuming a variety of different gamma priors $G(a, b)$, whose mean and variance are a/b and a/b^2, respectively. Following [19], [24] and [25], we use the following combinations in our experimental design: $(a, b) = (0.5, 0.0005)$, $(0.001, 0.001)$, $(0.01, 0.01)$, $(0.1, 0.1)$, $(2, 0.001)$, $(0.2, 0.0004)$, and $(10, 0.25)$, which are denoted by A, B, C, D, E, F, and G, respectively. We also consider the uniform distribution $U(0, 100)$ for standard errors $(\sigma_v, \sigma_\epsilon)$ denoted by H. We consider the same set up as in our simulation study for logistic mixed model for $\rho = 0.4$ and $\sigma_v^2 = \sigma_\epsilon^2 = 1$; noting that the ρ is generated from uniform distribution $U(-1, 1)$.

Table 6 provides the TMSPE and RB(%) of $\hat{p}_{1T}$ for different sceneries. As shown, the TMSPE is stable for different scenarios of gamma and uniform distributions for variance components. It seems that the RB is similar for scenarios B, C, and D in the case of gamma distribution, however, with increasing the mean and variance of gamma priors (A,E,F, and G), the RB values are also increasing. The RB for our uniform distribution (H) is also better than all other scenarios.

6. Application

We also consider the Bayesian analysis by using a real dataset of logistic mixed model. We use a yearly dataset of childhood (age $\leq$ 20 years) asthma visits to hospital in the Canadian province of Manitoba during 2000-2010 fiscal years. The population of Manitoba was stable during the study period from 1.15 million in 2000 to 1.20 million in 2010, with an average population of children of around 335,000. The province consisted of eleven Regional Health Authorities that were responsible for the delivery of health care services. These eleven regions were further sub-divided into 56 Regional Health Authorities Districts (RHAD) and these RHAD are used as small areas in our model. The number of children asthma visits totaled 14,690 over the study period with mean and median number of yearly cases per region of 26 and 17 (range 3 to 422), respectively. The region children population sizes varied from 290 to 175,300, with mean and median numbers of 5,998 and 2,488, respectively. We ignore the variation of geographical regions in this data analysis, and our focus is to apply our cross-sectional and time-series binomial mixed model to this dataset. We consider the following model

$$\log(\frac{p_{it}}{1 - p_{it}}) = \alpha + v_i + u_{it}(t = 1, ..., 10; i = 1, ..., 56)$$

where α is overall mean over area and time, $v_i \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$, and $u_{it} = \rho u_{i,t-1} + \epsilon_{it}$, with $|\rho| < 1$ and $\epsilon_{it} \overset{i.i.d.}{\sim} N(0, \sigma_\epsilon^2)$; noting that y_{it}, children asthma visits to hospital in the ith area at time t, has binomial distribution with parameters p_{it} and n_{it} where n_{it} is the corresponding population size. We first consider the estimates of model parameters by applying Bayesian approach. The estimates of the model

Table 7. Parameter estimates and standard errors (SE) of yearly children asthma visits to hospital 2000-2010 using Bayesian approach, logistic mixed model.

Parameter	α	σ_v^2	ρ	σ_ϵ^2
Estimate	-5.089	0.196	0.881	0.067
SE	0.093	0.167	0.070	0.010

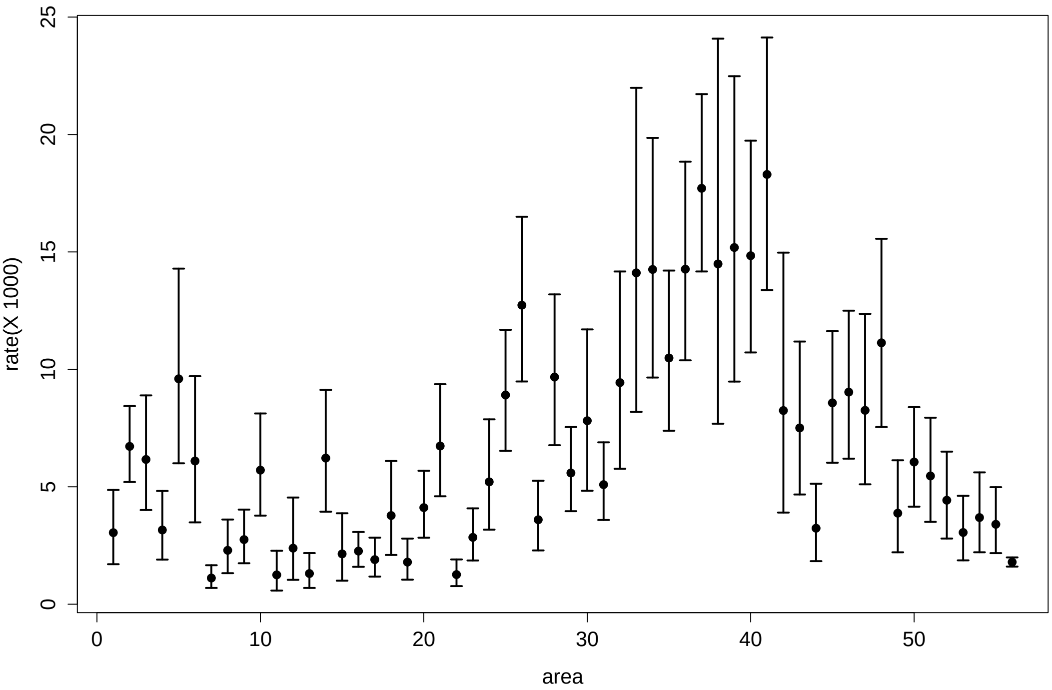

Figure 1. The 95% credible interval of the rate of children asthma visits to hospital in 2010 using Bayesian approach, logistic mixed model.

parameters and associated standard errors are reported in Table 7. We also provide 95% credible interval of the rates of children asthma visits to hospital for different areas in 2010 (Figure 1).

7. Concluding remarks

In small area estimation, there are many situations where observations are time-related counts or proportions. Using Bayesian approach, we have proposed a generalized model involving autocorrelated random effects and sampling errors for small area estimation with utilizing both cross-sectional and time-series data. Under the GLMM, our simulation results have shown that Bayesian approach performs very well in terms of relative bias of estimators of MSPE of small area parameters. The Bayesian based prediction approach also provided very good coverage probabilities of the small area parameters. In a separate manuscript (Torabi and Shokoohi [26]), we have also proposed a frequentist approach in small area estimation for generalized model with utilizing both cross-sectional and time-series data.

We studied the convergence of the samples obtained through diagnostic methods, and concluded that convergence was achieved. Our sensitivity analysis using different priors for the variance components pointed out that this hierarchical Bayesian analysis for cross-sectional and time-series data yields good results in terms of RB and coverage probabilities with using uniform distribution or gamma distribution with relatively small variances for precision of random effects. However, in general, we got large RB with using gamma distribution with large variances.

Acknowledgments

This work was done while second author visited the first author as a Ph.D. student. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. The comments of a referee are gratefully acknowledged.
Disclaimer: The interpretations, conclusions and opinions expressed in this paper are those of the authors and do not necessarily reflect the position of Manitoba Health. This study is based in part on data provided by Manitoba Health through Manitoba Centre for Health Policy. The interpretation and conclusions contained herein are those of the researchers and do not necessarily represent the views of the government of Manitoba.

References

[1] Battese, G., Harter, R., and Fuller, W. An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data. *Journal of the American Statistical Association 83*, 401 (1988), 28–36.
[2] Datta, G., Lahiri, P., Maiti, T., and Lu, K. Hierarchical Bayes Estimation of Unemployment Rates for the States of the U.S. *Journal of the American Statistical Association 94*, 448 (1999), 1074–1082.
[3] Datta, G. S., Lahiri, P., and Maiti, T. Empirical Bayes Estimation of Median Income of Four-Person Families by State Using Time Series and Cross-Sectional Data. *Journal of Statistical Planning and Inference 102*, 1 (2002), 83 – 97.
[4] Fay, I. R. E., and Herriot, R. A. Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association 74*, 366 (1979), 269–277.
[5] Gelfand, A. E., and Smith, A. F. M. Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association 85*, 410 (1990), 398–409.
[6] Gelman, A. Prior Distributions for Variance Parameters in Hierarchical Models. *Bayesian Analysis 1*, 3 (2006), 515–533.
[7] Gelman, A., and Rubin, D. B. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science 7*, 4 (1992), 457–472.
[8] Jiang, J. *Large Sample Techniques for Statistics.* Springer, 2010.
[9] Jiang, J., and Lahiri, P. Mixed model prediction and small area estimation (with discussion). *TEST 15*, 1 (2006), 1–96.
[10] Jones, R. H. Maximum Likelihood Fitting of ARMA Models to Time Series with Missing Observations. *Technometrics 22*, 3 (1980), 389–395.
[11] Pascutto, C., Wakefield, J., Best, N., Richardson, S., Bernardinelli, L., Staines, A., and Elliott, P. Statistical issues in the analysis of disease mapping data. *Statistics in Medicine 19*, 17-18 (2000), 2493–2519.
[12] Pfeffermann, D., and Burck, L. Robust Small Area Estimation Combining Time Series and Cross-Sectional Data. *Survey Methodology 16*, 1 (1990), 217–237.
[13] Prasad, N. G. N., and Rao, J. N. K. The Estimation of the Mean Squared Error of Small-Area Estimators. *Journal of the American Statistical Association 85*, 409 (1990), 163–171.
[14] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.
[15] Rao, J. N. K. *Small Area Estimation.* Wiley-Interscience, 2003.
[16] Rao, J. N. K., and Yu, M. Small area estimation combining time series and cross-sectional data. In *Proc. Survey Research Methods Section* (1992), American Statistical Association, pp. 1–9.
[17] Rao, J. N. K., and Yu, M. Small-Area Estimation by Combining Time-Series and Cross-Sectional Data. *The Canadian Journal of Statistics 22*, 4 (1994), 511–528.
[18] Scott, A. J., and Smith, T. M. F. Analysis of Repeated Surveys Using Time Series Methods. *Journal of the American Statistical Association 69*, 347 (1974), 674–678.
[19] Silva, G. L., Dean, C. B., Niyonsenga, T., and Vanasse, A. Hierarchical bayesian spatiotemporal analysis of revascularization odds using smoothing splines. *Statistics in Medicine 27*, 13 (2008), 2381–2401.
[20] Singh, A., Mantel, H., and Thomas, B. Time Series Generalizations of Fay-Herriot Estimator for Small Areas. In *Proc. Survey Research Methods Section* (1991), American Statistical Association, pp. 455–460.

[21] Smith, B. J. BOA: An R Package for MCMC Output Convergence Assessment and Posterior Inference. *Journal of Statistical Software 21*, 11 (2007), 1–37.

[22] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Linde, A. v. d. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 64*, 4 (2002), 583–639.

[23] Tiller, R. B. Time Series Modeling of Sample Survey Data from the U.S. Current Population Survey. *Journal of Official Statistics 8*, 2 (1992), 149–166.

[24] Torabi, M. Hierarchical Bayes Estimation of Spatial Statistics for Rates. *Journal of Statistical Planning and Inference 142*, 1 (2012), 358 – 365.

[25] Torabi, M., and Rosychuk, R. J. Hierarchical Bayesian spatiotemporal analysis of childhood cancer trends. *Geographical Analysis 44*, 2 (2012), 109–120.

[26] Torabi, M., and Shokoohi, F. Likelihood inference in small area estimation by combining time-series and cross-sectional data. *Journal of Multivariate Analysis 111* (2012), 213–221.