

Comparison of MCEM and Bayesian Correction Methods of Spatially Lagged Covariates Measured with Error : Evidence from Monte Carlo Simulation

Mohammad Masjkur¹, Henk Folmer², Asep Saefuddin³
{masjkur@apps.ipb.ac.id¹, h.folmer@rug.nl², asaefuddin@gmail.com³}

Department of Statistics, Bogor Agricultural University, Bogor, Indonesia^{1,3}, Faculty of Spatial Sciences, University of Groningen, The Netherlands²

Abstract. Measurement errors in (spatially lagged) explanatory variables under the classical-errors-in variables assumption are not routinely accounted for in applied (spatial) research, in spite of their serious consequences. Particularly, the estimator of coefficients of variables measured with error but also of those not measured with error are biased and inconsistent. The purpose of this paper is to analyze and compare by way of Monte Carlo simulation two bias correction methods, i.e. Monte Carlo Expectation-Maximization (MCEM) and Bayesian approach (BA). We consider spatial lag model (SLX) with different spatial correlation of covariate of interest, different measurement error variances and sample sizes. We use relative bias (RelBias) and Root Mean Squared Error (RMSE) as valuation criteria. The main result is that the Bayesian approach and MCEM method outperform the Naive model without measurement error correction. Moreover, the Bayesian approach performs better than MCEM method.

Keywords: Bayesian approach, MCEM, measurement error, Spatial regression.

1 Introduction

Measurement errors in (spatially lagged) explanatory variables are not routinely accounted for in applied (spatial) research, in spite of the fact that their consequences are usually serious under the classical-errors-in variable assumption (assumed throughout this paper). Particularly, the estimator of the coefficients of the variables measured with error but also those of the explanatory variables not measured with error are biased and inconsistent. [1] showed that generally the estimator of model coefficients are attenuated, while the estimator of variance components are inflated, if measurement error in (spatially lagged) covariates variables is ignored. Furthermore, [2] showed that the amount of attenuation depends on the variance of measurement error.

Several approaches to correct for measurement error in spatially lagged exogenous regressors have been proposed. [1] proposed Expectation-Maximization (EM) estimator under a conditional exponential or Gaussian autoregressive (CAR) structure to correct the bias of naive estimator, i.e. estimator that ignores measurement error. [1] considered a structural measurement error modelling approach assuming a parametric model of the error prone covariate. The spatial correlation exists not only in the outcome variable, but also in the error prone covariate with the same spatial correlation structure, but with different parameters. The measurement error variance is assumed known or estimated from data. [1]'s simulations

showed that the EM approach effectively eliminates the biases in naive estimators and has smaller MSE. However, computation time rapidly increases for large data sets. To overcome this problem, [3] proposed a version of the EM algorithm (Monte Carlo EM [MCEM]) for inference of spatial model with covariate measurement error. [3]'s simulation revealed better performance of MCEM approach compared to naive approach measurement error model.

[4] using Bayesian approach analyzed Gaussian spatial model with covariate measurement error under an exponential spatial structure. Unlike [1],[4] consider the functional model, i.e., a fixed true values of covariate. The simulation found that the proposed measurement error model has a better performance than naive model. The advantage of this Bayesian approach is that the proposed model facilitates representing the uncertainties related to model parameter values and able to incorporate prior information.

The purpose of this paper is to analyze and compare by way of Monte Carlo simulation two bias correction methods of spatial regression models with a spatially lagged covariate measured with error, i.e., Method of Monte Carlo Expectation-Maximization (MCEM) and Bayesian approach (BA).

2 Literature review

2.1 Bias Correction Methods

We define the model with spatially lagged exogenous covariates (SLX) as follows:

$$Y = X\beta + WX\gamma + \varepsilon \quad (1)$$

where Y is the $n \times 1$ vector of observations on the dependent variable, X the $n \times k$ matrix of the true unobserved explanatory variables, β the $k \times 1$ vector of regression coefficients (if $x_{i1} = 1$ for all observation i ($i = 1, \dots, n$), the first component of β is the constant term), W then $n \times n$ spatial weight matrix, γ the spatial dependence parameter, and ε is an $n \times 1$ Gaussian error term with zero mean and variance-covariance matrix Σ_ε with $\Sigma_\varepsilon = \sigma_\varepsilon^2 I_n$ with I_n the $n \times n$ identity matrix and σ_ε^2 is the variance of the error term.

Letting $Z = [X, WX]$ and $\tilde{\beta} = [\beta, \gamma]'$, the above model becomes:

$$Y = Z\tilde{\beta} + \varepsilon \quad (2)$$

Let $P = [Q, WQ]$ be the observed error prone explanatory variables related to the true covariate Z according to a classical errors in variables model¹. Then,

$$P = Z + U \quad (3)$$

where U is a $n \times k$ matrix of measurement errors. Its rows are assumed to be *iid* with $U \sim N(0, \Sigma_U)$ and independent of Z and ε . Hence,

$$Y = (P - U)\tilde{\beta} + \varepsilon = P\tilde{\beta} + \omega \quad (4)$$

¹Note that not all X variables need to be measured with errors

where $\boldsymbol{\omega} = \boldsymbol{\varepsilon} - \mathbf{U}\tilde{\boldsymbol{\beta}}$ the error term. This means that the disturbance term $\boldsymbol{\omega}$ is correlated with \mathbf{P} and $E(\boldsymbol{\omega}|\mathbf{P}) \neq \mathbf{0}$. This violates the orthogonality assumption of OLS. The main consequence is that $\tilde{\mathbf{b}}$ and \mathbf{s}_ε^2 are biased and inconsistent estimators of $\tilde{\boldsymbol{\beta}}$ and σ_ε^2 [5].

2.2 Monte Carlo Expectation-Maximization Method

The present method was introduced by [3]. It reads as follows,

The density function of the complete data, including \mathbf{Z}_i , of the SLX model (1), can be decomposed in terms of conditional densities as [6]:

$$f(\mathbf{Y}_i, \mathbf{Z}_i, \mathbf{P}_i) = f(\mathbf{Y}_i|\mathbf{Z}_i, \mathbf{P}_i)f(\mathbf{Z}_i, \mathbf{P}_i) = f(\mathbf{Y}_i|\mathbf{Z}_i)f(\mathbf{P}_i|\mathbf{Z}_i)f(\mathbf{Z}_i) \quad (5)$$

(5) is made up three components: the outcome component, $f(\mathbf{Y}_i|\mathbf{Z}_i)$, the measurement component, $f(\mathbf{P}_i|\mathbf{Z}_i)$, and the covariate component, $f(\mathbf{Z}_i)$. We assume that the covariates \mathbf{Z}_i follow a multivariate normal distribution with k -element mean vector $\boldsymbol{\mu}_Z = (\mu_1, \dots, \mu_k)$ and $k \times k$ covariance matrices $\boldsymbol{\Sigma}_Z$.

The likelihood function for the observed data $(\mathbf{Y}_i, \mathbf{P}_i)$ is:

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^n L(\boldsymbol{\theta}|\mathbf{Y}_i, \mathbf{P}_i) = \prod_{i=1}^n f(\mathbf{Y}_i, \mathbf{P}_i|\boldsymbol{\theta}) = \prod_{i=1}^n \int f(\mathbf{Y}_i|\mathbf{Z}_i)f(\mathbf{P}_i|\mathbf{Z}_i)f(\mathbf{Z}_i)d\mathbf{Z}_i \quad (6)$$

where $\boldsymbol{\theta} = (\tilde{\boldsymbol{\beta}}, \sigma_\varepsilon^2, \boldsymbol{\mu}_Z, \boldsymbol{\Sigma}_Z, \boldsymbol{\Sigma}_U)^T$ denotes the vector of parameters of the joint model. The observed data likelihood therefore depends on the density $f(\mathbf{Y}_i, \mathbf{P}_i)$ which can be found by marginalizing the joint distribution $f(\mathbf{Y}_i, \mathbf{Z}_i, \mathbf{P}_i)$ over \mathbf{Z}_i .

Maximization of the likelihood function of the observed data directly as $L(\boldsymbol{\theta}|\mathbf{Y}_i, \mathbf{P}_i)$ is analytically intractable since it typically involves intractable integrals and is not available in a closed form [7]. The EM algorithm (EM) is a procedure to find ML estimate when direct maximization of the likelihood function is difficult but when maximization of the complete data likelihood is easier. [3] proposed Monte Carlo EM (MCEM) to estimate the model parameter $\boldsymbol{\theta}$ of the joint model. This method is a modification of the EM algorithm where the expectation in the *E-step* is computed numerically through Monte Carlo simulation.

The basic idea of the algorithm is to use the log-likelihood function for the complete data where the unobserved latent variable \mathbf{Z} is treated as missing data. The complete data loglikelihood function is

$$\begin{aligned} \ell(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{P}, \mathbf{Z}) &= \log(f(\mathbf{Y}_i|\mathbf{Z}_i)) + \log(f(\mathbf{P}_i|\mathbf{Z}_i)) + \log(f(\mathbf{Z}_i)) \\ &= -\frac{n}{2}\log(\sigma_\varepsilon^2) - \frac{1}{2\sigma_\varepsilon^2} \|\mathbf{Y} - \mathbf{Z}\tilde{\boldsymbol{\beta}}\|^2 - \frac{n}{2}\log(\sigma_u^2) - \frac{1}{2\sigma_u^2} \|\mathbf{P} - \mathbf{Z}\|^2 - \frac{n}{2}\log(\sigma_Z^2) - \frac{1}{2\sigma_Z^2} \|\mathbf{Z} - \boldsymbol{\mu}_Z\|^2 \end{aligned} \quad (7)$$

where $\|\cdot\|$ denotes the square norm.

The joint distribution of the observed data and the missing data is factored into a part due to the observed data and the conditional distribution of the missing given the observed. It can be shown that each iteration of the EM algorithm monotonically increases the observed likelihood, and MLE $\hat{\boldsymbol{\theta}}$ is reached upon convergence [8].

The MCEM algorithm requires an initial estimate of the model parameters, denote as $\boldsymbol{\theta}^{(0)}$. Let $\boldsymbol{\theta}^{(t)}$ be the current (step t) estimate at the MLE $\hat{\boldsymbol{\theta}}$. In the *E-step* of EM, we find the expected value of the complete data log likelihood function, conditional on the observed data and the current estimate of the parameter vector $\boldsymbol{\theta}^{(t)}$. This is denoted as $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^t)$ and defined as

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^t) = E\{\ell(\boldsymbol{\theta})|\mathbf{Y}, \mathbf{P}, \boldsymbol{\theta}^t\} = \int \ell(\boldsymbol{\theta})f(\mathbf{Z}|\mathbf{Y}, \mathbf{P})d\mathbf{Z} \quad (8)$$

where $\ell(\boldsymbol{\theta})$ denotes the contribution to the log likelihood function from observing $\mathbf{Y}, \mathbf{P}, \mathbf{Z}$. Due to the complexity of conditional expectations in (8), the Monte Carlo simulation algorithm can be used to generate samples from the joint distribution $f(\mathbf{Z}|\mathbf{Y}, \mathbf{P}; \boldsymbol{\theta}^t)$.

At the t th iteration, assume we can randomly draw M values $\mathbf{Z}_1^t, \mathbf{Z}_2^t, \dots, \mathbf{Z}_M^t$, each of size n from $f(\mathbf{Z}|\mathbf{Y}, \mathbf{P}; \boldsymbol{\theta}^t)$. Then the Monte Carlo approximation can be written as [3]

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^t) \approx \frac{1}{M} \sum_{m=1}^M \ell(\boldsymbol{\theta}^t; \mathbf{Y}, \mathbf{P}, \mathbf{Z}^{(m)}) \quad (9)$$

The *M-step* is to maximize Q with respect to $\boldsymbol{\theta}$ to obtain $\boldsymbol{\theta}^{(t+1)} = \arg \max_{\boldsymbol{\theta} \in \Theta} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^t)$ where Θ is the parameter space. The two step process is repeated until convergence is met.

2.3 Bayesian Approach

Considering model (2), assume that the distribution of \mathbf{Z} is $2k$ multivariate normal with $2k$ -element mean vectors $\boldsymbol{\mu}_Z = (\mu_1, \dots, \mu_{2k})$ and $2k \times 2k$ covariance matrices $\boldsymbol{\Sigma}_Z$. The observed value of \mathbf{Z} is the $2k$ -element vector \mathbf{P} and the Gaussian measurement errors on (\mathbf{Y}, \mathbf{P}) have $2k \times 2k$ covariance matrix $\boldsymbol{\Sigma}_U$. Following [9], the statistical model expressed hierarchically is then

$$\mathbf{Z} \sim N_{2k}(\boldsymbol{\mu}_Z, \boldsymbol{\Sigma}_Z) \quad (10)$$

$$\mathbf{Y}|\mathbf{Z} \sim N_k(\tilde{\boldsymbol{\beta}}\mathbf{Z}, \boldsymbol{\Sigma}_\varepsilon) \quad (11)$$

$$\mathbf{Y}, \mathbf{P}|\mathbf{Z} \sim N_k(\mathbf{Z}, \boldsymbol{\Sigma}_\omega) \quad (12)$$

Denoting $\mathbf{A} = (\mathbf{Y}, \mathbf{P})$, the observed data likelihood is

$$p(\mathbf{Y}, \mathbf{P}) = \prod_{i=1}^n (2\pi)^{-(k)/2} |\boldsymbol{\Sigma}_A|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{A} - \boldsymbol{\zeta})' \boldsymbol{\Sigma}_A^{-1} (\mathbf{A} - \boldsymbol{\zeta})\right\} \quad (13)$$

$$\boldsymbol{\zeta} = (\tilde{\boldsymbol{\beta}}\boldsymbol{\mu}_Z, \boldsymbol{\mu}_Z) \quad (14)$$

$$\boldsymbol{\Sigma}_A = \begin{pmatrix} \boldsymbol{\Sigma}_\varepsilon + \tilde{\boldsymbol{\beta}}^2 \boldsymbol{\Sigma}_Z & \tilde{\boldsymbol{\beta}} \boldsymbol{\Sigma}_Z \\ \tilde{\boldsymbol{\beta}} \boldsymbol{\Sigma}_Z & \boldsymbol{\Sigma}_Z + \boldsymbol{\Sigma}_U \end{pmatrix} \quad (15)$$

where $\boldsymbol{\zeta}$ is the k -element mean vector of \mathbf{A} and $\boldsymbol{\Sigma}_A$ is the $k \times k$ covariance matrix of \mathbf{A} .

The observed data likelihood can be decomposed as $p(\mathbf{Y}, \mathbf{P}|\boldsymbol{\theta}) = p(\mathbf{Y}|\mathbf{P}) \times p(\mathbf{P})$ where

$$p(\mathbf{P}) = (2\pi)^{-(k)/2} |\boldsymbol{\Sigma}_P|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu}_P)' \boldsymbol{\Sigma}_P^{-1} (\mathbf{Y} - \boldsymbol{\mu}_P)\right\},$$

$$p(\mathbf{Y}|\mathbf{P}) = (2\pi)^{-1/2} |\text{Var}_{\mathbf{Y}|\mathbf{P}}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}|\mathbf{P}})' \text{Var}_{\mathbf{Y}|\mathbf{P}}^{-1} (\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}|\mathbf{P}})\right\},$$

$\boldsymbol{\theta} = (\tilde{\boldsymbol{\beta}}, \boldsymbol{\mu}_Z, \boldsymbol{\Sigma}_Z, \boldsymbol{\Sigma}_U, \boldsymbol{\Sigma}_\varepsilon)^T$ denotes the vector of parameters of the joint model.

In the Bayesian framework, we need to consider prior distributions for all the unknown parameters expressed in the vectors θ . We consider a multivariate normal prior on the regression parameters $\tilde{\beta}$ and the mean vectors μ_z with mean 0 and large variance, $\pi(\tilde{\beta}) \sim N_{2k}(\mathbf{0}, \sigma_{\tilde{\beta}}^2)$ and $\pi(\mu_z) \sim N_{2k}(\mathbf{0}, \sigma_{\mu_z}^2)$. We assume the inverse Wishart prior of the covariance matrix Σ , i.e., $\Sigma \sim IW_v(\Lambda^{-1})$ where v the degrees of freedom and Λ is a $k \times k$ positive definite scale matrix [10]. It usually assumed that the model parameter are mutually independent. Then, the joint prior distribution of all unknown parameters is

$$p(\theta) = p(\tilde{\beta})p(\mu_z)p(\Sigma_Z)p(\Sigma_U)p(\Sigma_\varepsilon) \quad (16)$$

Combining the likelihood function and the prior distribution, the joint posterior density of all unknowns is

$$p(\theta | \mathbf{P}, \mathbf{Y}) \propto \prod_{i=1}^n (2\pi)^{-(k)/2} |\Sigma_A|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{A} - \zeta)' \Sigma_A^{-1} (\mathbf{A} - \zeta)\right\} p(\theta) \quad (17)$$

The Bayesian inference can then be carried out based on the posterior conditionals by applying appropriate MCMC algorithms [11].

3 Simulation

We consider the common SLX model as follows:

$$Y_i = \alpha + X_i \beta + \gamma \sum_{j \neq i} W_{ij} X_j + V_i' \delta + \varepsilon_i \quad (18)$$

with Y_i the response in location i ($i = 1, \dots, n$); α the intercept, X_i is an unobserved mis-measured true covariates relating to location i with the regression parameter β , $V_i = (V_{i1}, V_{i2}, V_{i3})'$ the 3×1 vector of perfectly-measured covariates with the associated regression parameter $\delta = (\delta_1, \delta_2, \delta_3)'$, W_{ij} is the weights matrix where the weight w_{ij} is 1 if areas i and j are neighbors and 0 otherwise, γ a spatial dependence parameter, and ε_i is the Gaussian error term with zero mean and variance-covariance matrix Σ with $\Sigma = \sigma_\varepsilon^2 \mathbf{I}_n$ with \mathbf{I}_n the $n \times n$ identity matrix and σ_ε^2 is the variance of the error term.

We assume that

$$P = X + U \quad (19)$$

where P is the observed covariates related to the true covariates X according to a classical measurement error model with $U \sim N(0, \sigma_U^2 \mathbf{I}_n)$. We assume $X \sim N(\mu_x, \sigma_x^2)$ with $\mu_x = 0$ and $\sigma_x^2 = 1$. The variables V_1 , V_2 , and V_3 are normal random variables with zero means and unit variance. Path diagram of relationship between observed variables with unobserved variables of error models shown in Figure 1.

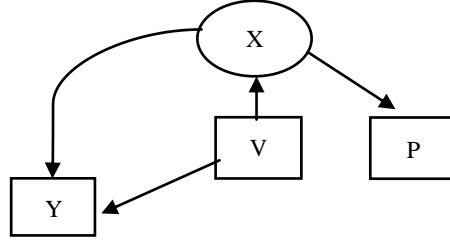


Fig. 1. Path diagram of the error models

We take the data to be on a regular grid with the grid sizes 7 ($T = 7 \times 7$) and 20 ($T = 20 \times 20$) representing small and large sample sizes, respectively. The weights matrix \mathbf{W} is row normalized. We allow different values for spatial dependence parameter γ , namely 0.3, 0.6 and 0.9. The observed error-prone covariate P is generated by adding Gaussian noise with variance $\sigma_U^2 = 0.25$ and 0.75 to X . Outcome data, Y are then generated with regression coefficients parameters set at $(\alpha, \beta, \delta_1, \delta_2, \delta_3)^T = (1, 2, -1, 1, -1)^T$.

We further take $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ with $\sigma_\varepsilon^2 = 1$. For each sample size, γ and σ_U^2 , we generate 100 Monte Carlo simulation datasets. For each generated dataset, we estimate β in the above model under the assumption of

- (1) Naïve model with measurement error ignored (NM)
- (2) Monte Carlo Expectation-Maximization Method (MCEM)
- (3) Bayesian approach (BA)

The following independent priors are considered to perform the Gibbs sampler, $\alpha, \beta, \delta \sim N(0, 100)$, $\sigma_\varepsilon^2 \sim U(0, 100)$, $\sigma_U^2 \sim U(0, 100)$, $\mu_x \sim N(0, 100)$, $\sigma_x^2 \sim U(0, 100)$ [12] and $\gamma \sim U(-1, 1)$ [13]. Three parallel chains are generated with 7000 burn-in iterations and 17000 main iterations with spacing of size 5. To monitor the convergence we use the between and within sequence information to obtain the potential scale reduction factor \hat{R} . The \hat{R} approximately 1 indicates convergence [10], [14].

For each simulation, we compute the relative bias (*RelBias*) and the Root Mean Square Error (*RMSE*) for each parameter estimate over 100 samples. These statistics are defined as

$$RelBias(\beta) = \frac{1}{k} \sum_{j=1}^k \left(\frac{\hat{\beta}_j}{\beta} - 1 \right), RMSE(\beta) = \sqrt{\frac{1}{k} \sum_{j=1}^k (\hat{\beta}_j - \beta)^2}$$

where $\hat{\beta}_j$ is the estimate of β for the j^{th} sample and $k=100$.

4 Results and Discussion

Table 1 and 2 show that for small sample, all measurement error variance and spatial dependence, the regression coefficient β of the naïve estimate (NM) as well as the MCEM and BA estimate is attenuated toward zero or downward biases. The other regression coefficient including intercept α , spatial dependence γ as well as coefficients returned for the perfectly-measured regressors δ are also noticeably biased. The average *RelBias* (in absolute value) of β and the other regression coefficient of naïve estimates are larger than MCEM and BA. We observed the following rankings: $NM > MCEM > BA$. Moreover, based on average *RMSE* we

observed the following rankings : BA < MCEM < NM. Note that for $\sigma_{ij}^2 = 0.25$, the average RMSE of β is as follows : BA = MCEM < NM.

Table 1. *RelBias* and *RMSE* of Regression Coefficients for Spatial Measurement Error Models with Design Methods for sample size N=49 and measurement error variance $\sigma_{ij}^2 = 0.25$.

Method	$\gamma = 0.3$							Average
	α	β	γ	δ_1	δ_2	δ_3		
Naïve	RelBias	-0.0246	-0.3929	-0.1680	0.0053	0.2213	-0.5446	0.2261
	RMSE	0.2278	0.8188	0.3516	0.1916	0.3068	0.5939	0.4151
MCEM	RelBias	-0.0037	-0.0507	-0.1330	0.0061	0.1605	-0.3900	0.1240
	RMSE	0.1712	0.2334	0.3203	0.1636	0.2399	0.4345	0.2605
BA	RelBias	-0.0014	-0.0149	-0.1144	0.0082	-0.0019	-0.0195	0.0267
	RMSE	0.1519	0.2306	0.2723	0.1588	0.1703	0.2336	0.2029
$\gamma = 0.6$								
Naïve	RelBias	0.0041	-0.3918	-0.1725	-0.0007	0.2226	-0.5374	0.2215
	RMSE	0.2323	0.8143	0.3698	0.1874	0.2965	0.5928	0.4155
MCEM	RelBias	-0.0083	-0.0530	-0.0301	0.0107	0.1649	-0.4027	0.1116
	RMSE	0.1756	0.2538	0.3527	0.1564	0.2438	0.4535	0.2726
BA	RelBias	0.0007	-0.0147	0.0117	0.0133	0.0000	-0.0158	0.0094
	RMSE	0.1561	0.2272	0.3186	0.1542	0.1724	0.2302	0.2098
$\gamma = 0.9$								
Naïve	RelBias	0.0046	-0.3918	-0.1854	-0.0014	0.2223	-0.5383	0.2240
	RMSE	0.2459	0.8151	0.3995	0.1876	0.2966	0.5943	0.4232
MCEM	RelBias	-0.0083	-0.0530	-0.0201	0.0107	0.1649	-0.4027	0.1100
	RMSE	0.1756	0.2538	0.3527	0.1564	0.2438	0.4535	0.2726
BA	RelBias	0.0005	-0.0149	0.0078	0.0132	0.0002	-0.0164	0.0088
	RMSE	0.1562	0.2273	0.3184	0.1543	0.1725	0.2301	0.2098

Table 3 and 4 show that for large sample, all measurement error variance and spatial dependence, the regression coefficient β of the naïve estimate (NM) is attenuated toward zero or downward biases. Moreover, the MCEM and BA approach resulted in unbiased estimates. The other regression coefficient as well as coefficients returned for the perfectly-measured regressors δ are also noticeably biased. The average *RelBias* (in absolute value) of all regression coefficient of naïve estimates are larger than MCEM and BA. We observed the following rankings: NM > MCEM > BA. The average *RMSE* of β and the other regression coefficient sequentially is as follows: BA < MCEM < NM. Note also that for $\sigma_{ij}^2 = 0.25$, the average *RMSE* of β is as follows : BA = MCEM < NM. Moreover, for $\sigma_{ij}^2 = 0.75$ and spatial dependence $\gamma = 0.6$ the MCEM method resulted in downward bias and the *RMSE* is the same as NM method.

Figures 1 and 2 show that the average *RelBias* (in absolute value) and the average *RMSE* for all sample sizes, measurement error variance and spatial dependence and the regression coefficient β and the other regression coefficient average of the naïve method (NM) are larger than for BA and MCEM methods. Moreover, the average *RelBias* (in absolute value) and the average *RMSE* of the BA approach is smaller than MCEM method.

Based on the above results, we find that Bayesian approach (BA) and MCEM method are more accurate and efficient than naïve method. Moreover, BA approach more accurate and efficient than MCEM method.

Table 2. *RelBias* and *RMSE* of Regression Coefficients for Spatial Measurement Error Models with Design Methods for sample size $N=49$ and measurement error variance $\sigma_U^2 = 0.75$.

Method	$\gamma = 0.3$							Average
		α	β	γ	δ_1	δ_2	δ_3	
Naïve	RelBias	0.0158	-0.6468	-0.3993	-0.0155	0.3796	-0.9322	0.3982
	RMSE	0.2660	1.3117	0.3339	0.2234	0.4332	0.9641	0.5887
MCEM	RelBias	0.0120	-0.0638	0.0253	-0.0114	0.3370	-0.8616	0.2185
	RMSE	0.2007	0.2941	0.4177	0.1919	0.3859	0.8841	0.3957
BA	RelBias	0.0167	-0.0180	0.0596	-0.0060	0.0009	-0.0403	0.0236
	RMSE	0.1637	0.2084	0.3079	0.1531	0.1491	0.1811	0.1939
$\gamma = 0.6$								
Naïve	RelBias	0.0002	-0.6558	-0.4060	-0.0067	0.3788	-0.9059	0.3922
	RMSE	0.2768	1.3270	0.4047	0.2167	0.4408	0.9441	0.6017
MCEM	RelBias	-0.0153	-0.0532	-0.0865	0.0094	0.3623	-0.8653	0.2320
	RMSE	0.2214	0.3216	0.4386	0.1890	0.4188	0.8918	0.4135
BA	RelBias	0.0004	-0.0143	0.0128	0.0145	-0.0002	-0.0160	0.0097
	RMSE	0.1559	0.2272	0.3184	0.1546	0.1724	0.2307	0.2099
$\gamma = 0.9$								
Naïve	RelBias	0.0044	-0.6464	-0.4601	-0.0211	0.3871	-0.8887	0.4013
	RMSE	0.2472	1.3060	0.5211	0.2373	0.4501	0.9202	0.6137
MCEM	RelBias	0.0198	-0.0457	-0.0474	-0.0372	0.3750	-0.8490	0.2290
	RMSE	0.1922	0.3759	0.3821	0.2038	0.4285	0.8727	0.4092
BA	RelBias	0.0168	-0.0191	-0.0126	-0.0080	0.0226	-0.0184	0.0163
	RMSE	0.1466	0.2347	0.2836	0.1642	0.1851	0.2125	0.2045

Table 3. *RelBias* and *RMSE* of Regression Coefficients for Spatial Measurement Error Models with Design Methods for sample size $N=400$ and measurement error variance $\sigma_U^2 = 0.25$.

Method	$\gamma = 0.3$							Average
		α	β	γ	δ_1	δ_2	δ_3	
Naïve	RelBias	0.0074	-0.3735	-0.2178	0.0013	0.2179	-0.5179	0.2226
	RMSE	0.0708	0.7505	0.1282	0.0670	0.2277	0.5245	0.2948
MCEM	RelBias	0.0019	0.0008	-0.0142	-0.0020	0.1377	-0.3323	0.0815
	RMSE	0.0538	0.0662	0.1113	0.0542	0.1480	0.3385	0.1287
BA	RelBias	-0.0007	0.0004	-0.0037	-0.0018	-0.0044	-0.0012	0.0020
	RMSE	0.0522	0.0687	0.1053	0.0508	0.0554	0.0696	0.0670
$\gamma = 0.6$								
Naïve	RelBias	0.0057	-0.3737	-0.1601	0.0025	0.2234	-0.5218	0.2145
	RMSE	0.0752	0.7519	0.1473	0.0682	0.2331	0.5282	0.3007
MCEM	RelBias	0.0100	0.0030	0.0188	0.0009	0.1413	-0.3332	0.0845
	RMSE	0.0578	0.0855	0.1038	0.0547	0.1515	0.3399	0.1322
BA	RelBias	0.0097	0.0002	0.0149	0.0009	-0.0019	-0.0027	0.0051
	RMSE	0.0543	0.0858	0.0954	0.0491	0.0576	0.0753	0.0696
$\gamma = 0.9$								
Naïve	RelBias	-0.0062	-0.3700	-0.1811	0.0026	0.2249	-0.5163	0.2169
	RMSE	0.0811	0.7443	0.1908	0.0636	0.2350	0.5223	0.3062
MCEM	RelBias	0.0003	0.0008	0.0030	0.0022	0.1415	-0.3299	0.0796
	RMSE	0.0560	0.0771	0.0880	0.0530	0.1507	0.3353	0.1267
BA	RelBias	0.0010	0.0036	0.0033	0.0025	-0.0030	0.0051	0.0031
	RMSE	0.0526	0.0792	0.0812	0.0493	0.0520	0.0726	0.0645

Table 4. *RelBias* and *RMSE* of Regression Coefficients for Spatial Measurement Error Models with Design Methods for sample size $N=400$ and measurement error variance $\sigma_{ij}^2 = 0.75$.

Method	$\gamma = 0.3$							Average
		α	β	γ	δ_1	δ_2	δ_3	
Naïve	RelBias	-0.0076	-0.6463	-0.4142	0.0066	0.3858	-0.9050	0.3943
	RMSE	0.0836	1.2943	0.1663	0.0703	0.3925	0.9093	0.4861
MCEM	RelBias	-0.0131	-0.0140	0.0860	0.0027	0.3427	-0.8047	0.2105
	RMSE	0.0647	0.0963	0.1425	0.0611	0.3481	0.8080	0.2535
BA	RelBias	-0.0124	-0.0066	0.0364	0.0009	0.0036	-0.0078	0.0113
	RMSE	0.0492	0.0834	0.1044	0.0523	0.0546	0.0833	0.0712
$\gamma = 0.6$								
Naïve	RelBias	-0.0068	-0.6396	-0.4108	0.0110	0.3866	-0.8929	0.3913
	RMSE	0.0859	1.2812	0.2655	0.0754	0.3933	0.8971	0.4997
MCEM	RelBias	-0.0006	-0.0095	-0.0040	0.0034	0.3477	-0.8031	0.1947
	RMSE	0.0708	0.0992	0.1189	0.0685	0.3532	0.8058	0.2527
BA	RelBias	0.0044	0.0043	-0.0091	0.0020	-0.0051	0.0035	0.0047
	RMSE	0.0547	0.0862	0.0892	0.0508	0.0550	0.0798	0.0693
$\gamma = 0.9$								
Naïve	RelBias	-0.0027	-0.6415	-0.4161	-0.0048	0.3705	-0.9113	0.3912
	RMSE	0.1005	1.2849	0.3899	0.0769	0.3793	0.9168	0.5247
MCEM	RelBias	0.0018	-0.0141	-0.0046	-0.0017	0.3357	-0.8198	0.1963
	RMSE	0.0725	0.1133	0.1229	0.0676	0.3417	0.8227	0.2568
BA	RelBias	0.0034	0.0006	-0.0069	0.0011	-0.0057	-0.0061	0.0040
	RMSE	0.0523	0.0826	0.0841	0.0574	0.0552	0.0715	0.0672

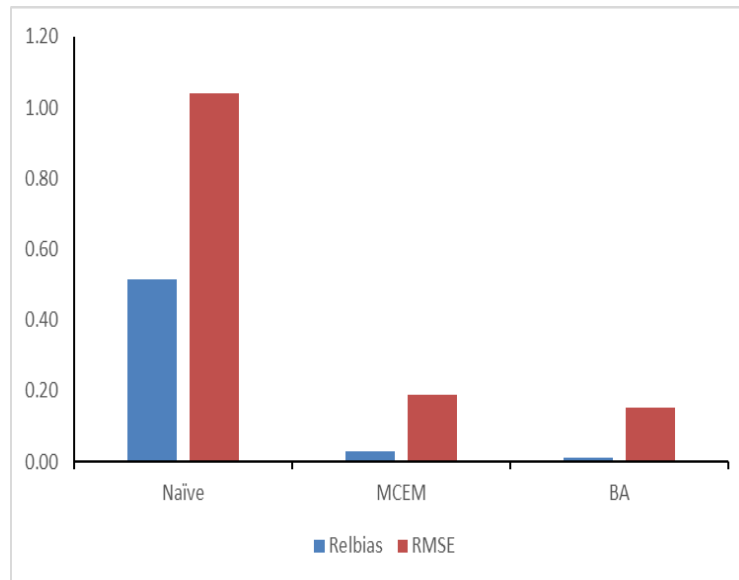


Fig. 2. *RelBias* and *RMSE* of Regression Coefficients β for the Methods

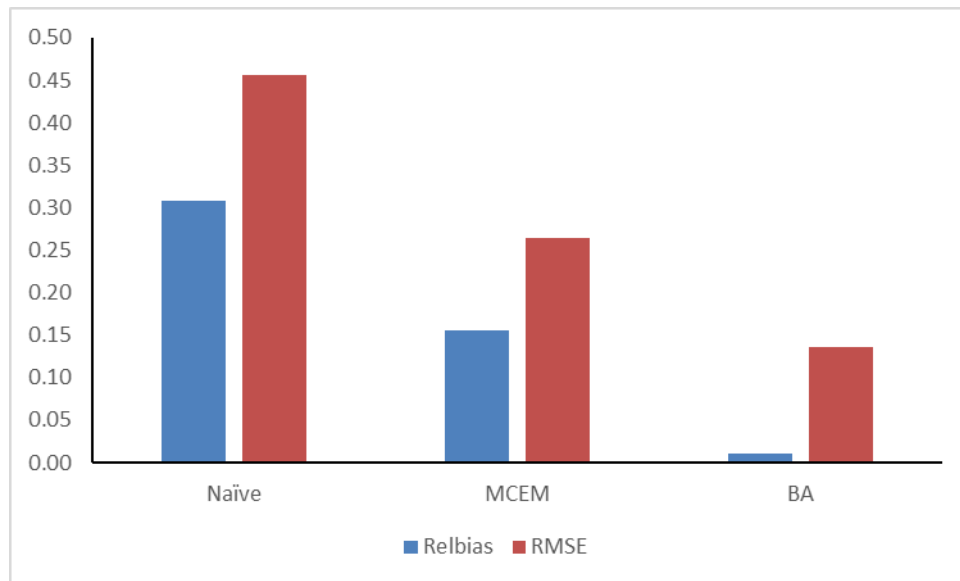


Fig. 3. *RelBias* and *RMSE* of Regression Coefficients Average for the Methods

5 Conclusion

This paper analyze and compare by way of Monte Carlo simulation two bias correction methods of spatial regression models with a spatially lagged covariate measured with error, i.e., Monte Carlo Expectation-Maximization (MCEM) and the Bayesian approach (BA). We consider the spatial lag model (SLX) with different degrees of spatial correlation in the covariate of interest, different measurement error variances and different sample sizes. The simulation result show that Bayesian approach (BA) and MCEM method are more accurate and efficient than naïve method. Moreover, BA approach more accurate and efficient than MCEM method.

References

- [1]Li Y, Tang H, Lin X.: Spatial linear mixed models with covariate measurement errors. *Statistica Sinica* 19(3): 1077-1093(2009)
- [2]Huque MH, Bondell HD, Ryan L.: On the impact of covariate measurement error on spatial regression modelling. *Environmetrics* 25:560-570(2014)
- [3]Tadayon V, Torabi M.: Spatial models for non-Gaussian data with covariate measurement error. *Environmetrics*.:e2545. <https://doi.org/10.1002/env.2545> (2018)
- [4]Tadayon V.: Analysis of Gaussian Spatial Models with Covariate Measurement Error. arXiv:1811.05648v2(2018)

- [5]Wansbeek, T. and E. Meijer.: Measurement Error and Latent Variables. A Companion to Theoretical Econometrics. Edited by Badi H. Baltagi. Blackwell Publishing Ltd. (2007)
- [6]Bartlett, J. W.: Correction for classical covariate measurement error and extensions to life-course studies. PhD thesis. London School of Hygiene and Tropical Medicine. University of London(2010)
- [7]Carroll RJ, Ruppert D, Stefanski LA.: Measurement error in nonlinear models, 2nd ed. Boca Raton, FL, USA: Chapman & Hall/CRC(2006)
- [8]Majumdar A.: Maximum Likelihood Estimation of Measurement Error Models Based on the Monte Carlo EM Algorithm. Dissertation. Faculty of the Graduate School. State University of New York(2007)
- [9]Kelly BC.: Some Aspects of Measurement Error in Linear Regression of Astronomical Data. The Astrophysical Journal, 665:1489 – 1506(2007)
- [10]Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB.: Bayesian Data Analysis, Chapman & Hall/CRC, New York, NY(2014)
- [11]Gray CM.: Use of the Bayesian family of methods to correct for effects of exposure measurement error in polynomial regression models. PhD thesis, London School of Hygiene & Tropical Medicine. DOI: <https://doi.org/10.17037/PUBS.04649757> (2018)
- [12]Gelman, A.: Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 1, 514-534(2006)
- [13]LeSage, J.P., Pace, K.P.: Introduction to Spatial Econometrics. CRC Press(2009)
- [14]Bartlett JW, Keogh RH.: Bayesian correction for covariate measurement error: a frequentist evaluation and comparison with regression calibration. Statistical Methods in Medical Research 0(0) 1–14(2016)