Models for Analyzing Economical Crop Yields in Thailand

Pitsanu Tongkhow[1], Pornpit Sirima[1], Krisada Lekdee[2], Rachadasak Supengcum[1*]

[1]Department of Industrial Engineering, Faculty of Engineering, Rajamangala University of Technology Phra Nakhon, Thailand.
[2]Department of Mathematics and Statistics, Faculty of Science and Technology, Rajamangala University of Technology Phra Nakhon, Thailand.
*Corresponding author. E-mail: rachadasak.s@rmutp.ac.th

ABSTRACT

This research is to propose an appropriate and efficient model for analyzing economical crop yields, rice, sugarcane and corn, in Thailand. The data were collected from provinces in the central region of Thailand, 22 provinces for rice, 9 provinces for corn and 14 provinces for sugarcanes, from 2007 to 2019. A linear mixed model (LMM) including spatial relationship explained by a conditional autoregressive (CAR) model was adopted. The Bayesian method was used for parameter estimation. The factors related to the production of rice, sugarcane and corn were also investigated. The results showed that the amount of rain, temperature, region and spatial effects affected the yields. When comparing the LMM with spatial relationship and the one without spatial relationship, it was found that the LMM model spatial relationship was more suitable than the one without spatial relationship. For each yield, the mean square error (MSE) of the LMM was greater than the MSE of the proposed model (rice yields: 6.07 vs 1.59, corn yields: 7226.11 vs 6287.44 and sugarcane yields: 76.76 vs 46.69).

Keywords: Crop yields, Linear mixed model, Conditional autoregressive model, Spatial relationship, Bayesian estimation

1. INTRODUCTION

Field crops that are considered economic crops bring much income to Thailand such as rice as food, sugarcane as raw materials for sugar production, corn as raw materials for animal feed production, cassava as raw materials for starch and animal feed production. Rice is one of the most important economic crops in the world. Thailand is the top rice producer and has exported the most rice in the world continuously for a long time. Rice is also an agricultural product that is traded to advance in the agricultural futures exchange of Thailand [1]. Advance price data analysis is useful for those who are involved in crop yields, whether directly or indirectly, to prevent price risk of the investment as well as to plan production (Office of the Commission). The world rice production was increased by 2.0 percent from the previous year while Thai rice production was decreased by 3.1 percent from the previous year due to reducing rice cultivated area as farmers switched to other crop yields. Sugarcane is one type of grass that is very important when considering productivity. Sugarcane can be used for growth factors such as energy, water and nutrients effectively. In addition, sugarcane is also a plant that is easy to grow and once grown, it can be harvested many times. Sugarcane prefers hot and humid weather, so the countries that grow sugarcane well

Received 04-03-2020
Revised 01-05-2020
Accepted 05-05-2020

include Thailand as well. Corn is an economic crop that is important to the consumption and feed industry. Most Corns are grown for animal feed. Corn is an economic crop that is important to the animal feed industry. In Thailand, there is about 94% of the corn production being for use in the animal food industry. Thailand has an increasing demand for corn every year. Some years, therefore, Thailand has to import corn for use in the country. Each year, most farmers will grow corn for 2 seasons. Namely, rainy corn will be planted from March to October, and harvested around June to January of next year, while the dry season corn will start in November to February of the next year and harvest during March to April of the next year. Corns for consumption will also be planted. For the above reasons, the researchers are interested in creating an effective model for analyzing the various influences that affect the continuous production of rice, corn and sugarcane. There are various in-depth mathematical and statistical models. For example, the method most commonly used for repetitive measurement data is a generalized estimating equation (GEE) model, which was presented by Liang and Zeger [2]. It is a model that shows the relationship between dependent variables and the independent variables, without the assumption that the dependent variables must be independent. The dependence occurs when the data are collected repeatedly in the same sample unit. The dependent variables can be both continuous and discontinuous. For the independent structure, we assume that the relationship between the data from repeated measurements is zero. For the exchangeable structure, we assume that the relationship is constant. For the autoregressive structure, it requires that the relationship decreases over time passed by and for the unstructured pattern, we assume that the relationship of each pair of data is formatted. The equation showing the relationship between the independent variables and the dependent variables is called GEE. The model also has the regression coefficients describing the magnitude of the influence of factors in the population as a whole, the population-averaged model. Limmun and Ingsrisawang [3] used GEE to find factors that affected road traffic accidents. Lekdee, K. and Ingsrisawang [4] used the GEE to analyze malaria risk data. Apart from the GEE model, there is another model that can take spatial relationships into consideration, with the most widely used spatial data analysis. The model is based on a linear mixed model, or LMM model. It is the model that shows the relationship of dependent variables and the independent variables. The model is highly flexible, so it can easily add some more variables to the model, such as variables showing spatial relationship. Therefore, it can be applied to analyze the data with spatial relationships caused by the fact that things that are close to one another are more related than far apart. Saengseedam and Kantanantha [5] proposed a linear mixed model (LMM) with spatial effects, trend, seasonality and outliers for spatio-temporal time series data. Sammatat and Lekdee [6] used LMM with spatial effects for estimation and detection of rice yield in Thailand. LMMs are available in commercial programs such as SPSS (mixed), SAS, S-PLUS, MLwin, or ASReml. A LMMgui is a free, graphic user interface that uses lme4 (Bates et al. [7]). Another one is a package in the free, open-source program which is called R Core Team [8]. Magezi [9] explained linear mixed-effects models for within-participant psychology experiments. There are several methods for estimating parameters in a GLMM, but the method which is widely used when the model is complex is a Bayesian method. The model is complex when some more variables are added to show both spatial and time relations, for example. The Bayesian method for estimating parameters will determine the distribution pattern of the data. The distributions of parameters that are random variables are called priors. The product of the distribution of data with prior is called posterior. The parameter estimation in the posterior widely used is Markov Chain Monte Carlo (MCMC) using Gibbs sampling. An LMM model and Bayesian method for parameter estimation appeared in Pedroza [10]. As mentioned above, the researchers are interested in using the principles of modeling LMM to analyze the data of rice, corn and sugarcane yields. The data have been collected continuously from provinces in the central region of Thailand, 22 provinces for rice, 9 provinces for corn and 14 provinces for sugarcanes, from 2007 to 2019. Therefore, the researchers are interested to study and develop a model for analyzing various influences that accurately affects the production of rice, corn and sugarcane. It is useful for further planning in Thailand with the most effective mathematical model, with examples of preliminary factors to be considered, including rainfall and average temperature. After that comparing the LMM model without spatial relationship is performed. As mentioned above, the researchers are interested in using the principles of modeling LMM to analyze the data of rice, corn and sugarcane yields. The data have been collected continuously. Therefore, the researchers are interested to study and develop a model for analyzing various influences that accurately affects the production of rice, corn and sugarcane. It is useful for further planning in Thailand with the most effective mathematical model, with examples of preliminary factors to be considered, including rainfall, average temperature, central region, northern region, northeastern region, southern region, eastern region and western region. After that comparing the LMM model without spatial relationship is performed.

2. MATERIALS AND METHODS

A linear mixed model (LMM) has the following details.

$y_{ij} \mid \mathbf{b} \sim N(\mu_{ij}, \sigma^2)$,where $i = 1,\ldots,\mathrm{m}$, $j = 1,\ldots,\mathrm{n}$, when μ_{ij} i shown in Eq. (1),

$$\mu_{ij} = \beta_0 + \beta_1(\text{Rain}) + \beta_2(\text{Temp}) + \beta_3(\text{Region}) + \mathrm{b_i} + \phi_i, \quad (1)$$

$\beta_0, \beta_1, \beta_2, \beta_3 \sim N(0, 1.0E06)$, $\mathrm{b_i} \sim N(0, \tau_b^{\,2})$

$, \phi_i \mid \mathbf{\Phi}_{(-i)} \sim N\left(0, \frac{\tau_\phi^{\,2}}{w_{i+}}\right),$

$\mathrm{var}(\mathbf{\Phi}) = (\mathbf{D}_w - \rho\mathbf{W})^{-1}$, $\rho \sim \text{Uniform}(0,1)$,

$\tau_b^{\,2}, \sigma^2, \tau_\phi^{\,2} \sim \text{IG}(0.5, 0.0005)$.

West et al. [11] described the LMM model in details. The important characteristics of LMM are as follows. The LMMs model is a model that starts with the conditional distribution of the vector of variables $\mathbf{y}$, given the vector values of random variables $\mathbf{b}$ that show random influences on the vector $\mathbf{y}$.

It is denoted by $\mathbf{y} \mid \mathbf{b}$ where $i = 1,\ldots,m$, $j = 1,\ldots,n_i$ and $y_{ij} \mid \mathbf{b} \overset{\text{iid}}{\sim} N(\mu_{ij}, \sigma^2)$.

The expected value is $E(y_{ij} \mid \mathbf{b}) = \mu_{ij}$ and $\mu_{ij} = \mathbf{x}_{ij}^{\,T}\boldsymbol{\beta} + \mathbf{z}_{ij}^{\,T}\mathbf{b}$, where $\mathbf{x}_{ij}^{\,T}$ is the member in row i of the fixed effects matrix, $\boldsymbol{\beta}$ is the vector of the constant parameter and $\mathbf{z}_{ij}^{\,T}$ is the member in row i of the random effects matrix, $\mathbf{b}$, the vector of parameters that are random variables.

μ_{ij} is the expected value or the average of the conditional distribution of $y_{ij} \mid \mathbf{b}$. Since $\mathbf{b}$ is the random variable, it must be specified the form of the distribution. In general, the distribution of $\mathbf{b}$ is $\mathbf{b} \sim \mathrm{N}(\mathbf{0}, \mathbf{B})$. The variance of $y_{ij} \mid \mathbf{b}$ is $\mathrm{var}(y_{ij} \mid \mathbf{b}) = \sigma^2$, which has a constant value.

For the LMM model, when spatial relationship is accounted, the variable that shows spatial influence can be added, as shown in Eq.(2).

$$\mu_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{b} + \phi_i \qquad (2)$$

The model the most widely used for spatial influence is a conditional autoregressive (CAR) model which has the following details.

$$\phi_i \mid \mathbf{\Phi}_{(-i)} \sim N\left(\sum_{j=1}^{m} a_{ij}\phi_j, \tau_i^2 \right),$$

where $\phi_{(-i)} = \{\phi_j : j \neq i\}$. Banerjee et al. [12] described the details of the CAR model as follows.

Let $\mathbf{\Phi} = (\phi_1, ..., \phi_m)^T$ represent the vector of the random variable that changes in area, which has the influence on the variable Y_i, which measures the value in each area $i,\ i = 1, ..., m$. It is a conditional distribution in which τ^2 is the conditional variance. a_{ij} is the constant that represents the relationship of the area i and j, where $a_{ii} = 0$. $\mathbf{A} = (a_{ij})$ and $\mathbf{M} = Diag(\tau_1^2, ..., \tau_m^2)$. By Brook's Lemma, the joint distribution of $(\phi_1, ..., \phi_m)$ is

$$p(\mathbf{\Phi}) \propto \exp\left\{ -\frac{1}{2} \mathbf{\Phi}^T \mathbf{M}^{-1} (\mathbf{I} - \mathbf{A}) \mathbf{\Phi} \right\},$$

$$\mathrm{var}(\mathbf{\Phi}) = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{M}. \qquad (3)$$

$\mathbf{M}^{-1}(\mathbf{I} - \mathbf{A})$ must be a symmetrical matrix and $p(\mathbf{\Phi})$ must be a probability distribution. That is the condition $\frac{a_{ij}}{\tau_i^2} = \frac{a_{ji}}{\tau_j^2}$ must be true. This condition can be true when $a_{ij} = \frac{w_{ij}}{w_{i+}}$ and $\tau_i^2 = \frac{\tau^2}{w_{i+}}$.

Therefore, the enumeration of $\phi_i \mid \mathbf{\Phi}_{(-i)}$ which is a probability distribution. Therefore it has the form; $\phi_i \mid \mathbf{\Phi}_{(-i)} \sim N\left(\sum_{j=1}^{m} \frac{w_{ij}\phi_j}{w_{i+}}, \frac{\tau^2}{w_{i+}} \right)$ which is called the CAR model. The joint distribution of $(\phi_1, ..., \phi_m)$ is

$$p(\mathbf{\Phi}) \propto \exp\left\{ -\frac{1}{2\tau^2} \mathbf{\Phi}^T (\mathbf{D}_w - \mathbf{W}) \mathbf{\Phi} \right\}.$$

$\mathbf{W} = (w_{ij})$ is a weighted matrix showing the relationship between areas. $w_{ij} = 1$ if the areas are adjacent and $w_{ij} = 0$ if the areas are not adjacent to each other. $\mathbf{D}_w = Diag(w_{i+})$ is the main diagonal matrix whose members are in diagonal and $w_{i+} = \sum_j w_{ij}$. Since $(\mathbf{D}_w - \mathbf{W})$ is a singular matrix, $p(\mathbf{\Phi})$ is improper, i.e., the total area under the graph is not equal to 1. Therefore, the CAR model is improper. To The solution for $(\mathbf{D}_w - \mathbf{W})$ to have inverse is to add a parameter ρ to $(\mathbf{D}_w - \mathbf{W})$, then $\mathrm{var}(\mathbf{\Phi}) = (\mathbf{D}_w - \rho\mathbf{W})^{-1}$. The distribution of $\phi_i \mid \mathbf{\Phi}_{(-i)}$ has a new form,

$\phi_i \mid \boldsymbol{\Phi}_{(-i)} \sim N\left(\rho \sum_{j=1}^{m} \frac{w_{ij}\phi_j}{w_{i+}}, \frac{\tau^2}{w_{i+}}\right)$ and is called spatial parameter. This distribution is called proper CAR model. Clayton and Kaldor [13] analyzed the incidence of the disease and created a disease map using the GLMM model that has spatial relationships of data in the form of CAR. The model of Clayton and Kaldor [13] was widely used and the estimation of parameters in the model was performed based on the Bayesian method using the Markov Chain Monte Carlo method with Gibbs sampling. The Bayesian method was obtained by writing programs in OpenBUGS and processing simulation in R program with the package of R2OpenBUGS [14-15].

The data in this study were at a provincial level and collected yearly from the central region of Thailand, 22 provinces for rice, 9 provinces for corns and 14 provinces for sugarcanes, from 2007 to 2019. The dependent variables were rice, corn and sugarcane yields obtained from the department of agriculture, office of agricultural Statistics. The independent variables were the amount of rainfall and temperature obtained from the meteorological department [16-17].

3. RESULTS AND DISCUSSION

Data analysis was performed by programming in OpenBugs and R. The general characteristics of the data and the factors affecting the production of rice, corn and sugarcane were analyzed. The mean and standard deviation were used to present the general characteristics of the data. The LMM with spatial influence was used for the estimation of the rice, corn, sugarcane production and to investigate the factors related to the products. Then the LMM with spatial influence was compared to LMM models without spatial influence. The Bayesian method was used for parameter estimation. The details are shown in Eq. (4-6).

$$y_{ij} \mid \mathbf{b} \sim N(\mu_{ij}, \sigma^2),\ i = 1,\dots,22,\ j = 1,\dots,13$$

$$\mu_{ij} = \beta_0 + \beta_1 * rain_{ij} + \beta_2 * temp_{ij} + \beta_3 * k + b_{1i} + b_{2ij} + v_i \qquad (4)$$

$$y_{ij} \mid \mathbf{b} \sim N(\mu_{ij}, \sigma^2),\ i = 1,\dots,9,\ j = 1,\dots,13$$

$$\mu_{ij} = \beta_0 + \beta_1 * rain_{ij} + \beta_2 * temp_{ij} + \beta_3 * k + b_{1i} + b_{2ij} + v_i \qquad (5)$$

$$y_{ij} \mid \mathbf{b} \sim N(\mu_{ij}, \sigma^2),\ i = 1,\dots,14,\ j = 1,\dots,12$$

$$\mu_{ij} = \beta_0 + \beta_1 * rain_{ij} + \beta_2 * temp_{ij} + \beta_3 * k + b_{1i} + b_{2ij} + v_i \qquad (6)$$

v is the spatial influence following the distribution of CAR which has the following form.

$$v_i \mid \mathbf{v}_{(-i)} \sim \mathrm{N}\left(\sum_{k=1}^{m} \frac{w_{ik}v_k}{w_{i+}}, \frac{\tau_v^{\,2}}{w_{i+}}\right) \text{ and}$$

$$\mathbf{v} \sim \mathrm{N}\left(\mathbf{0}, \tau_v^{\,2}(\mathbf{D}_w - \mathbf{W})^{-1}\right) \text{ or}$$

$$\mathrm{p}(\mathbf{v}) \propto \exp\left\{-\frac{1}{2\tau_v^{\,2}} \mathbf{v}^T(\mathbf{D}_w - \mathbf{W})\mathbf{v}\right\}.$$

Under the Bayesian method, the prior distributions are assumed to be non-informative priors that do not affect the posterior.

$$\beta_0, \beta_1, \beta_2, \beta_3 \sim N(0, 100000000)$$

$$b_{1i} \sim N(0, \tau_{b1}^{\,2})$$

$$b_{2it} \sim N(0, \tau_{b2}^{\,2})$$

$$\tau_{b1}^{\,2}, \tau_{b2}^{\,2}, \tau_v^{\,2} \sim InvGamma(0.0001, 0.0001)$$

The parameters were estimated by the Bayesian method using MCMC Gibbs sampling, programming in OpenBUGS and simulating the situation in R2OpenBUGS.

Some examples of typical data on the production of rice, sugarcane, and corn are as follows. The top 5 rice yields ranked from the highest to the lowest value are in Nakhon Sawan, Phichit, Suphan Buri, Phitsanulok and Kamphaeng Phet, respectively. The top 5 provinces with the estimated average annual corn yield ranked from the highest to the least value are Pichit, Chai Nat, Suphan Buri, Kamphaeng Phet and Nakhon Sawan, respectively. The top 5 provinces with estimated sugar cane yields per year, ranked from the highest to lowest value, are Nakhon Sawan, Suphan Buri, Kamphaeng Phet, Phetchabun and Lop buri, respectively. The estimated magnitude of factors affecting the average annual rice yield, corn yield and sugarcane yield are shown in Table 1-3.

From Table 1-3, the factors influencing rice yields, corn yields and sugarcane yields in the central region are rainfall, average temperature trends. If the amount of rain increases by 1 mm, the average annual rice yield will decrease by 26.69 tons; if the average temperature increases by 1 Celsius, the average annual rice yield will decrease 213.55 tons and when the time increases by 1 year, the rice yield likely to increase by 14.85 tons. For corn yields, if the rainfall increases by 1 mm, the average corn yield per year will increase 71.61 tons; if the average temperature rises 1 Celsius, the average annual corn yield will be reduced by 42.28 tons; if the time passes by 1 year, the corn yield tends to decrease by 3.58 tons. Finally, for sugar cane production, if the amount of rain increases by 1 mm, the average sugar cane yield per year will decrease by 53.19 tons. If the average temperature rises 1 Celsius, the average sugar cane yield per year will increase by 119.46 tons and with an increase of 1 year, the rice production tends to increase by 28.09 tons.

Table 1. Estimated parameters of rice yields

P.	Mean	S.D	95% C.I	
β_0	26.06	43.88	-71.12	68.86
β_1	-26.69	1.04	-26.74	-24.56
β_2	-213.55	2.59	-217.55	-210.15
β_3	14.85	6.47	-3.56	20.05

[a] P., S.D, C.I is the abbreviation for the word of parameter, standard deviation and credible interval respectively.

Table 2. Estimated parameters of corn yields

P.	Mean	S.D	95% C.I	
β_0	14.43	61.83	-141.34	141.26
β_1	71.61	1.18	67.49	80.32
β_2	-42.28	42.9	-94.98	96.26
β_3	-3.58	96.56	-166.04	103.56

Table 3. Estimated parameters of sugarcane yields

P.	Mean	S.D	95% C.I	
β_0	48.25	52.2	-93.88	125.46
β_1	-53.19	-0.4	-54.43	-51.41
β_2	119.46	82.33	16.53	193.36
β_3	28.09	99.26	-100.47	120.46

The spatial influence of each province that has rice yields, corn yields and sugarcane yields in the central region in Thailand is shown in Table 4-6. From Table 4, it is found that the spatial influence values of the province on the average annual rice production in Thailand ranked in order from highest to lowest value are Bangkok, Samut Sakhon, Nonthaburi, Nakhon Pathom, Pathum Thani, Saraburi, Ang Thong, Samut Prakan, Phra Nakhon Si Ayutthaya and Nakhon Nayok, respectively. From Table 5, it is found that the spatial influence of the province on average corn yield per year in Thailand ranked from the highest to the least value are, namely, Chai Nat, Suphan Buri, Nakhon Sawan, Phichit, Kamphaeng Phet, Lop Buri, Saraburi, Sukhothai and Phitsanulok, respectively. From Table 6, it is found that the spatial influence of the province on average sugarcane yield per year ranked in order from the highest to the least value are Sukhotha, Phichit, Phitsanulok, Kamphaeng Phet, Phetchabun, Nakhon Sawan, Uthai Thani, Lop Buri, Ang Thong, Sing Buri, Chai Nat, Suphan Buri, Nakhon Pathom, and Saraburi, respectively. The results showed that the rainfall, the average temperature and the tendency influence the production of rice, sugarcane and corn. For rice, if the amount of rainfall increases, the average monthly rice yield will decrease. One of the reasons is that the central region is a river basin and there are rivers which are the source of rice barns in Thailand. The central region is the best place to produce rice because there is enough water throughout the year and the area is irrigated. If there is too much rain, it will cause yield reduction. For the comparison of the proposed model, LMM with spatial effects, and the one without spatial effects, it was is found that the proposed model gives the lowest error value (MSE) for all kinds of crop yields, (rice: 1.59 vs 6.07, corns : 6287.44 vs 7226.11 and sugarcanes: 46.69 vs76.76). This refers to the proposed model is more effective than the LMM without spatial effects. The comparison results are shown in Tables 7-9.

Table 4. Spatial influence of each province on rice yield

P.	spatial influence			
	Mean	**S.D**	**95% C.I**	
1	0.42	5.03	-4.55	9.94
16	0.40	4.60	-5.00	9.15
7	0.30	3.94	-5.07	8.51
5	0.27	4.07	-5.33	8.53
8	0.20	4.50	-7.33	8.20
17	0.14	4.10	-4.92	6.70
21	0.12	3.90	-6.54	6.32
14	0.10	8.03	-9.47	15.89
9	0.09	2.87	-4.51	4.63
4	0.08	6.40	-6.87	10.51
13	-0.01	3.28	-5.32	5.14
22	-0.04	5.03	-7.46	5.13
19	-0.05	6.78	-9.43	10.16
20	-0.06	3.29	-5.98	4.38
2	-0.11	5.27	-9.22	9.12
15	-0.12	9.18	-15.55	15.69
18	-0.16	3.64	-7.24	4.29
10	-0.23	5.18	-8.35	6.85
11	-0.25	5.23	-8.26	7.21
6	-0.27	3.46	-6.75	3.46
3	-0.36	3.86	-9.27	3.87
12	-0.46	4.13	-9.32	4.14

[a] P., S.D, C.I is the abbreviation for the word of province, standard deviation and credible interval respectively.

Table 5. Spatial influence of each province on corn yield

P.	spatial influence			
	Mean	**S.D**	**95% C.I**	
3	109,900	16,880	99,400	163,200
20	79,660	11,260	44,820	91,270
6	48,090	8,723	21,170	55,530

Table 5. (Continue) Spatial influence of each province on corn yield

P.	spatial influence			
	Mean	S.D	95% C.I	
10	-2,065	8,299	-17,620	19,750
2	-19,060	10,740	-26,120	18,400
13	-31,260	13,420	-37,970	18,160
17	-37,190	14,480	-81,420	-23,030
19	-71,040	11,310	-109,900	-65,960
11	-77,070	8,657	-104,100	-65,860

Table 6. Spatial influence of each province on Sugarcane yield

P.	spatial influence			
	Mean	S.D	95% C.I	
19	32.1	313.2	-36.1	95.2
10	24.0	242.8	-23.2	58.8
11	23.5	230.7	-18.2	70.8
2	21.5	206.0	-20.6	71.4
12	3.0	154.1	-32.7	21.7
6	1.5	129.1	-18.2	15.1
22	-3.2	171.9	-28.9	32.0
13	-7.1	138.6	-42.1	14.1
21	-7.8	142.9	-40.5	25.6
18	-9.9	132.7	-39.3	20.1
3	-11.0	180.5	-30.6	26.1
20	-13.0	182.3	-39.1	27.6
5	-22.3	322.2	-57.7	60.6
17	-31.4	314.0	-77.4	26.3

Table 7. Model comparison for rice yield

Rice yield	MSE
LMM	6.07
LMM with spatial	1.59

Table 8. Model comparison for corn yield

Corn yield	MSE
LMM	7226.11
LMM with spatial	6287.44

Table 9. Model comparison for sugarcane yield

Sugarcane yield	MSE
LMM	76.76
LMM with spatial	46.69

4. CONCLUSIONS

The linear mixed model with CAR spatial effects in this research was used for yield data. This model is very suitable for the data of rice, corn and sugarcane since it considers spatial relationship of the data as well. For all 3 types of crops, the influence of the adjacent or non-adjacent provinces is influenced by spatial relationships. For corn yields, if the amount of rainfall increases, the average monthly corn yield will increase. One of the reasons is that most corn plantations grow well during the rainy season. The period of corn growth requires only a little water and gradually more water with age and requires maximum water during flowering. The period of seed production requires less water again; therefore, lacking water during the flowering period will greatly reduce the yield, so the planting date must be estimated to avoid the corn from affecting the drought during flowering. Lastly, sugarcanes are an important economic crop for food and other industries. Moreover, they are raw materials for ethanol production as a renewable energy. If the amount of rain increases, the average monthly sugarcane yield will decrease. One of the reasons is that the central region has enough water to grow sugarcane. If there is additional rainfall, it may reduce the yield. When adding spatial influences to the model, the model becomes more complex in which the maximum likelihood (ML) cannot be used for parameter estimation.

The Bayesian method and MCMC process are gaining widespread popularity since it is able to solve complex model problems. It has many advantages. One is that the estimation of that parameter can be performed even though the sample is small where the classical ML desires large sample size. The MCMC process is a computer-based sampling. Even though we do not know the probability distribution of the data, we can estimate the parameters such as the mean and standard deviation by using the process of some random samplings, such as Gibb sampling.

5. ACKNOWLEDGMENTS

This research was well accomplished with the support and the help from many people. The research team would like to thank the president of Rajamangala University of Technology Phra Nakhon, the dean of the Faculty of Engineering and director and all the staff of the Research and Development Institute for supporting this research. Finally, the research team would like to thank Rajamangala University of Technology Phra Nakhon for providing financial support until this research was accomplished well.

6. REFERENCES

[1] The agricultural futures exchange of Thailand, the Thailand Futures Exchange (TFEX)[Internet]. [updated 2019 Aug 19; cited 2019 Aug 9]. Available from: https://www.tfex.co.th/th/about/glance.html.

[2] Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika. 1986; 73: 13-22.

[3] Limmun W, Ingsrisawang, L. Study of factors affecting road traffic accidents using generalized estimating equations and generalized linear models. Journal of KMUTNB. 2010; 20(2): 311-21.

[4] Lekdee K, Ingsrisawang L. Risk factors for malaria in Thailand using generalized estimating equations (GEE) and genalized linear mixed model (GLMM). J Health Sci. 2010; 19: pp. 364-73.

[5] Saengseedam P, Kantanantha N. Spatio-temporal model for crop yield forecasting. J Appl Stat. 2017; 44(3):427-40.

[6] Sammatat, S. and Lekdee K. Estimation and Detection of Rice Yield in Thailand Using Spatial and Longitudinal Data Analysis. Applied Mathematical Sciences, 2019; 13(18): 877 - 84.

[7] Bates DM, Mächler M, Bolker BM, et al. Fitting Linear Mixed-Effects Models Using lme4.2014[Internet]. [updated 2002 May 19; cited 2002 Jul 30]. Available from:http://arxiv.org /abs/1406.5823.

[8] R Core Team. R: A Language and Environment for Statistical Computing. 2014[Internet]. [updated 2002 May 19; cited 2002 Jul 30] Available from: http://www.R-project.org

[9] Magezi DA. Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). Switzerland: University of Fribourg: Department of Medicine, Faculty of Science; 2015.

[10] Pedroza CA. Bayesian forecasting model: predicting U.S. male mortality. Biostatistics. 2006; 7(4):530-50.

[11] West BT, Welch KB, Galecki AT. Linear mixed models: A practical guide to using statistical software. NY : Chapman & Hall/CRC; 2007.

[12] Banerjee S, Carlin BP, Gelfand AE. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC Press. FL. Janeway CA, Travers P, Walport M, Shlomchik M.

Immunobiology. 5th ed. New York: Garland Publishing; 2004.
[13] Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics. 1987;43:671-81.
[14] Cowles MK. Applied Bayesian statistics: with R and OpenBUGS, New York: springer; 2013.
[15] Baio G, Berardi A, Heath A. Bayesian cost-effectiveness analysis with the R package BCEA. New York: springer; 2017.
[16] Temperature data. [Internet]. the thai meteorological department [updated 2019 Aug 19; cited 2019 Aug 9]. Available from: https://www.tmd.go.th/en/.
[17] Yield data. [Internet]. the Office of Agricultural Economics [updated 2019 Aug 19; cited 2019 Aug 9]. Available from: http://www.oae.go.th/view/1/%E0%B8%AB%E0%B8%99%E0%B9%89%E0%B8%B2%E0%B9%81%E0%B8%A3%E0%B8%81/TH-TH.