Comparison between the maximum likelihood and the bayesian estimation methods for logistic regression model (case study: risk of low birth weight in Indonesia)

P R Sihombing¹,³*, S R Rohimah²,³, A Kurnia³

¹BPS- Statistics Indonesia, Jakarta, Indonesia
²Jakarta State University (UNJ), Jakarta, Indonesia
³Department of Statistics, IPB University, Bogor, Indonesia

*Email: robinsonpardomuan@apps.ipb.ac.id

Abstract. This study aims to compare the efficacy of logistic regression model for identifying the risk factors of low-birth-weight babies in Indonesia using the maximum likelihood estimation (MLE) and the Bayesian estimation methods. The data used in this study is secondary data derived from the 2017 Indonesian Demographic Health Survey with a total sample of 16,344 newborn babies. Selection of the best logistic regression model was based on the smaller Bayesian Schwartz Information Criterion (BIC) value. The logistic regression model with the Bayesian estimation method has a smaller BIC value than the MLE method. Twin births, baby girl, maternal age at risk, birth spacing that is too close, iron deficiency, low education, low economy, inadequate drinking water sources have provided a higher risk of low-birth-weight incidence.

1. Introduction

One target of the Sustainable Development Goals (SDGs) is to reduce the Neonatal Mortality Rate (NMR) to at least 12 per 1,000 live births in 2030. According to the Indonesian Ministry of Health [1], the condition of low birth weight (LBW) is the leading cause of neonatal death. LBW occurs when a baby born weighs less than 2.500 grams. So, it is necessary to know what factors affect LBW so that evidence-based policies can be made to reduce NMR. Several studies related to LBW have been done. Audrey and Candra [2] have investigated the relationship between the anemia status of third-trimester pregnant women with the incidence of LBW babies in the working area of the Halmahera Health Center, Semarang with Fisher's exact test. Ramdaniati and Nastiti [3] also have identified the relationship between the characteristics of children under five, mother's knowledge, and sanitation on the incidence of stunting in children under five in the Labuan sub-district, Pandeglang district with Chi-square statistical test.

In this study, the response variable is dichotomous/binary data, namely babies born with normal weight or LBW. Binary logistic regression is one of the regression methods used to model binary response variables. The estimation of binary logistic regression parameters can be done using the maximum likelihood estimation (MLE) method or the Bayesian estimation method. In estimating parameters, the MLE method only uses the likelihood function, while the Bayesian method uses both the likelihood
function and the prior distribution. In this case, the Bayesian method views the parameters as random variables that describe the initial knowledge of the parameters before the observations are made and expressed in a distribution called the prior distribution.

According to Meliza, Bustan, and Sudarmin [4] have determined the factors that influence the incidence of breast cancer using a logistic regression method. The results have shown that the Bayesian method has a smaller value of variance and errors than the MLE method. Teti, Yanuar and Yozza [5] have also performed logistic regression analysis with the Bayesian estimation method to determine the factors that influence the incidence of LBW. The results showed that maternal age, parity, and multiple births were the significant predictor variables to the LBW. This study, therefore, will compare the efficacy of parameter estimation method between the MLE and the Bayesian in binary logistic regression model for determining the risk of LBW in Indonesia using the Bayesian Schwartz Information Criterion (BIC) values.

2. Method

2.1. Data Sources and Research Variables

This study used secondary data derived from the 2017 Indonesian Demographic Health Survey (IDHS) regarding birth weight in infants born with normal and LBW categories. The number of samples was 16,344 newborns. The response variable in this study was the status of birth weight, namely the baby's weight which was weighed within the first hour delivery. The birth weight of newborns was classified into normal birth weight, i.e., babies born weighing between 2,500 – 3,999 g and low birth weight (LBW), i.e., babies born with birth weight <2,500 g regardless of gestational age or previously referred to as premature [6]. Meanwhile, the predictor variables used partially refer to Setyawan's research [7], as can be seen in Table 1.

| Table 1. Research variables |
|-----------------------------|
| **Response Variable**       | **Category**                           |
| Y  | Birth Weight Status                     | 0: Normal (2,500 – 3,999 g) |
|    |                                           | 1: LBW (< 2,500 g)          |
| **Predictor Variable**      | **Category**                           |
| X1 | Multiple birth (twins)                  | 1: No                        |
|    |                                           | 2: Yes                       |
| X2 | Baby gender                              | 1: Girl                      |
|    |                                           | 2: Boy                       |
| X3 | Birth distance                           | 1: >=2 years                 |
|    |                                           | 2: < 2 years                 |
| X4 | Iron supplements                         | 1: Consumption               |
|    |                                           | 2: No consumption            |
| X5 | Mother's age                             | 1: No risk (what is the range of this age?) |
|    |                                           | 2: Too young (what is the range of this age?) |
|    |                                           | 3: Too old (what is the range of this age?) |
| X6 | Mother's education                       | 1: Middle/high               |
|    |                                           | 2: No school/elementary     |
| X7 | Economic status                          | 1: Middle/high               |
|    |                                           | 2: Low                       |
| X8 | Source of drinking water                 | 1: Adequate                  |
|    |                                           | 2: Inadequate                |
2.2. Logistics Regression
Logistic regression is a regression model that shows the effect of predictor variables, either continuous or categorical, on response variables in the form of categorical data. In binary logistic regression, the response variable consists of two (binary) categories, namely 0 and 1, where the response variable for each observation follows the Bernoulli distribution and mathematically logistics regression model can be written as follows:

\[ g(x) = \ln \left( \frac{\pi(x)}{1-\pi(x)} \right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p \]  \hspace{1cm} (1)

Partial parameter testing to determine whether a predictor variable has a significant effect on the response variable. The hypothesis used in this test:

\[ H_0 : \beta_j = 0 \]
\[ H_1 : \beta_j \neq 0 \] where \( j = 1, 2, \ldots, p \)

with the Wald test statistics

\[ W = \frac{\hat{\beta}_j}{se(\beta_j)} \sim N(0,1) \] \hspace{1cm} (2)

The decision to reject H0 if the test statistic value \(|W| > Z\), which means that the predictor variable significantly affects the response variable.

2.3. Bayes' Theorem
If the occurrence of an event is the intersection of the sample space, then for any event that is \( A_1, A_2, \ldots, A_k \) \( P(A_i) \neq 0 \) for \( i = 1, \cdot \cdot \cdot, k \) and \( P(B) \neq 0 \) apply

\[ P(A_j | B) = \frac{P(A_j) P(B | A_j)}{\sum_{k=1}^{k} P(A_k) P(B | A_k)} \approx P(A_j) P(B | A_j) \] \hspace{1cm} (3)

for \( j = 1, \cdot \cdot \cdot, k \)

\( P(A_j | B) \) is posterior, \( P(A_j) \) is prior and \( P(B | A_j) \) is likelihood function [8]

2.4. Prior and Posterior Distribution
Prior distribution is the initial distribution that provides information about the parameters. To get the posterior distribution, we must first determine the prior distribution of the parameters. Regarding the determination of each parameter, the prior distribution is divided into two parts, namely [9]

a. Informative prior distribution is a prior distribution that refers to the assignment of parameter values from the prior distribution that has been selected, whether it is a conjugate prior distribution or not. Giving parameter values to the prior distribution will significantly affect the shape of the posterior distribution obtained from the existing data.

b. The non-informative prior distribution is the selection of the prior distribution not based on the existing data.

The posterior probability density function of a given sample of observations is \( \beta x_1, x_2, \cdot \cdot \cdot, x_n \) [10]:

\[ f(\beta | x) = \frac{f(x_1, x_2, \cdot \cdot \cdot, x_n | \beta) f(\beta)}{\int f(x_1, x_2, \cdot \cdot \cdot, x_n | \beta) f(\beta) d\beta} \approx f(x_1, x_2, \cdot \cdot \cdot, x_n | \beta) f(\beta) \] \hspace{1cm} (4)

2.5. Markov Chain Monte Carlo (MCMC)
MCMC is a method for determining parameter values of complex analytical integration. There are three kinds of MCMC methods that can be used: the Metropolis method, the Metropolis-Hasting method, and the Gibbs Sampler method. The Metropolis-Hasting method is a technique that is often used by users of the Bayesian method, including in this study. The Metropolis-Hasting (MH) algorithm is used to help generate random samples from the desired posterior distribution. In the MH algorithm, a distribution proposal is needed to generate random sample candidates. The basic steps of this algorithm are \( p(\theta | \theta_{j-1}) \).

Step 1: Take the initial value, i.e for iterations. Raise \( \theta_0 | \theta_{j-1} = \theta \sim p(\theta | \theta_{j-1}) \)
Step 2: Generate a random sample from a uniform distribution $U[0,1]$

Step 3: If $u < \min(1, \frac{p(\theta^*|X,y)p(\theta_{j-1}|\theta^*)}{p(\theta^*|X,y)p(\theta_{j-1}|\theta^*)})$, take $\theta_j = \theta^*$

However, $u > \min(1, \frac{p(\theta^*|X,y)p(\theta_{j-1}|\theta^*)}{p(\theta^*|X,y)p(\theta_{j-1}|\theta^*)})$, take $\theta_j = \theta_{j-1}$

Step 4: Repeat steps 1-3 until the desired amount of sample

The statistics used to measure the degree of dependence between successive retrievals in a Markov chain is autocorrelation. Autocorrelation measures the correlation between two sets of simulated values, and with is the lag or number that separates the two sets. For a certain hyperparameter, the autocorrelation value in the $-th$ lag can be calculated by the formula $r_{lh} = \frac{M}{(M-L)} \sum_{i=1}^{M-L} (\theta_{i+l}-\bar{\theta})(\theta_{i+l}-\bar{\theta})$, where $M$ is the random sample size.

$$r_{lh} = \frac{M}{M-L} \left( \frac{\sum_{i=1}^{M-L} (\theta_{i+l}-\bar{\theta})(\theta_{i+l}-\bar{\theta})}{\sum_{i=1}^{M} (\theta_l-\bar{\theta})^2} \right) \quad (5)$$

The steps of this algorithm can be seen in [11].

2.6. Best Model Selection

The logistic regression models obtained both with MLE and with the Bayesian techniques were compared to obtain the best model that can be used to describe the relationship between the response variable and their predictors. One of the measurements commonly used to select the best model is to use BIC (Bayesian Schwartz Information Criterion). The BIC formula is:

$$BIC = -2 L(\hat{\theta}) + p \ln(n)$$

where $L(\hat{\theta})$ is the likelihood value for the model containing the independent variables, and $p$ is the number of parameters including constants. The best model is the model that has the smallest BIC value.

3. Results and Discussion

Before further analyzing the relationship between the explanatory variable and LBW status, a descriptive analysis was conducted between the predictor variable and the response variable. In Table 2, the percentage of LBW among multiple births (twins) is higher than that of single-births (non-twins). The baby girl has a higher percentage of LBW than the baby boy. The percentage of LBW babies with birth spacing < two years is higher than babies with birth spacing >= two years. The percentage of LBW incidence is higher among mothers whose age are at risk (<20 years old or >35 years old) than mothers whose age are not at risk (20-34 years old). Low economic status and education have a higher percentage of LBW than those with better economic and education status (secondary/high education). The use of inadequate drinking water sources has a higher percentage of LBW newborns than proper drinking water sources.

| Predictor Variable          | BBL Status |
|----------------------------|------------|
|                            | Normal     | Low      |
| Multiple birth (twins)     | No         | 93.1     | 6.9     |
|                            | Yes        | 41.3     | 58.7    |
| Baby gender                | Girl       | 92.2     | 7.8     |
|                            | Boy        | 93.1     | 6.9     |
| Birth distance             | < 2 years  | 92.0     | 8.0     |
|                            | >=2 years  | 93.1     | 6.9     |
| Iron supplements           | No consumption | 92.7  | 7.3     |
|                            | Consumption | 92.9     | 7.1     |
| Mother's age               | Too young (< 20 years old) | 89.3  | 10.7    |
|                            | No risk    | 92.9     | 7.1     |
The next step is to see the association between the predictor variable and the response variable. Because all variables, both predictors and responses, are categorical data, the association test uses a contingency table with Chi-Square testing. The null hypothesis is that there is no relationship/association between the explanatory variable and the response variable. In Table 3, all probability values are below 0.05, meaning there is an association/relationship between the response variable and each predictor variable.

### Table 3. Contingency Table Association Test

| LBW status                        | Chi-Square Tests | P.Value |
|-----------------------------------|------------------|---------|
| Multiple birth (twins)            | 561,348          | 0.000   |
| Baby gender                       | 5.261            | 0.022   |
| Birth distance                    | 6.510            | 0.011   |
| Iron supplements                  | 7.684            | 0.006   |
| Mother's age                      | 8.478            | 0.014   |
| Mother's education                | 36.852           | 0.000   |
| Economic status                   | 40.066           | 0.000   |
| Source of drinking water          | 18.203           | 0.000   |

In Table 4, the results of the logistic regression using the MLE, and Bayesian methods have been given. A Hosmer test was carried out to check the goodness of fit, and a probability value of 0.0586 was obtained. This result indicates that the model used is appropriate. In addition, checking for overdispersion of the model obtained a value of 1.0029 with a probability of 0.896. This result indicates that the logistic binomial model used is correct. In terms of the significance of the test with the MLE method by looking at the probability value and the Bayesian estimation method by looking at the interval of 95 percent, both methods give equally good results where all predictor variables provide significant coefficients at 5 percent negligible. When viewed from the magnitude of the coefficient in general, the result of the Bayesin estimation method is smaller than using the MLE method. Furthermore, from the BIC dan standard error (SE) values, the Bayesian method gives better results, namely a value of 50 percent smaller than the MLE method. In terms of model accuracy, the prediction model with the Bayesian and MLE methods gave the same performance, which is 92.8 percent.

### Table 4. Comparison of MLE and Bayesian Logistics Models

| Variable                          | Category      | Coefisien MLE | Coefisien Bayes | SE MLE | SE Bayes |
|-----------------------------------|---------------|---------------|-----------------|--------|----------|
| Twin births (not twins)           | Twin          | 3.030**       | 3.028**         | 0.175  | 0.015    |
| Baby's gender (boy)               | girl          | 0.142**       | 0.144**         | 0.175  | 0.015    |
| Birth distance (>= 2 years)       | < 2 years     | 0.239**       | 0.229**         | 0.175  | 0.005    |
| Iron supplements (Consumption)    | No Consumption| 0.168**       | 0.160**         | 0.083  | 0.010    |
| Maternal age (no risk)            | too young     | 0.358**       | 0.384**         | 0.167  | 0.023    |
Further, the sample of the posterior distribution of the regression coefficients can be seen by drawing the histogram. Figure 1 shows that despite the unknown prior, the posterior mode is expected to be close to MLE. All posterior distributions are close to normal distributions. Thus, all posterior sample means are expected to approach the posterior mode in absolute values and are expected to be equal to MLE.

Furthermore, the sample of the posterior distribution of the regression coefficients can be seen by drawing the histogram. Figure 1 shows that despite the unknown prior, the posterior mode is expected to be close to MLE. All posterior distributions are close to normal distributions. Thus, all posterior sample means are expected to approach the posterior mode in absolute values and are expected to be equal to MLE.

4. Discussions
Multiple births (twins) have a significant positive relationship with LBW cases, with odds of 20.66. This result means that the chance of LBW babies among multiple births is higher 20.66 times than those single births. This result is in agreement with the research of Siti Masitoh, Syarifuddin, Delmaifanis [12], which states that women with multiple pregnancy have a 22.8 times risk of having LBW babies than women with single pregnancy. The average weight of twins is lower than the weight of a single newborn due to, perhaps, preterm deliveries, which can increase the mortality rate among twins. Rohayati [13] also shown the same results where twin newborns tend to be born with LBW than single newborns.
The sex of the baby has a significant positive relationship with LBW cases, with odds of 1.15. This result means that the risk of having LBW babies will be higher 1.15 times in baby girls compared to baby boys. This result is in line with the research of Pramono and Paramita [14], which states that female babies have a 1.41 times risk for LBW compared to the male babies. During the same pregnancy period, baby girls naturally tend to be smaller than baby boys, so they have a greater risk for LBW.

Birth distance has a significant positive relationship with LBW cases, with odds of 1.26. This result means that the risk of having LBW babies will be higher at 1.26 times at birth intervals of less than two years compared to those with birth interval more than two years. This result is like the the study of Kozuki et al. [15]. The ideal interval between births is more than two years, thus allowing the body to repair its supply and reproductive organs to be ready to conceive again. A disturbed reproductive system will inhibit the growth and development of the fetus.

Iron supplementation has a significant positive relationship with LBW cases, with odds of 1.17. This result means that the risk of having LBW babies is 1.17 times higher among women who do not taking iron supplements than those who take. The same result was obtained in Iriyani's research [16], which states that pregnant women who do not/less consume Fe supplements are more risk of having LBW than pregnant women who consume Fe. Iron (Fe) is a mineral that is needed by all biological systems in the body. Lack of iron intake before or during pregnancy will result in lack of Hb for pregnant women, thus impact on the newborn [2].

Maternal age has a significant positive relationship with LBW cases, with odds of 1.47 for too young (< 20 years old) and 1.09 for too old (> 35 years old). This result means that the risk of having LBW babies will be higher at 1.17 times among too young or too old mothers compared to those who are not at age risk. Similar result was found in Monita et al. [17], where the mother's age during pregnancy affects the condition of the mother's pregnancy because it is related to the maturity of the reproductive organs and psychological conditions. In addition, this result is in line with Misna, Wahyuni, and Santi [18].

Pregnancy at a young age is a risk factor; this is due to the immaturity of the reproductive organs for pregnancy (imperfect endometrium), while at the age of 35 years, the endometrium is less fertile and increases the possibility of suffering from congenital abnormalities so that it can affect the health of the mother and the development and growth of the fetus.

Mother's education has a significant positive relationship with LBW cases, with odds of 1.31. This result means that the risk of having LBW babies will be at 1.31 times higher among mothers with low education compared to those with middle/higher education. The result agreed with the research of Nuryani and Rahmawati [19] that shown a significant relationship between the level of education with the incidence of LBW. This result is also in line with the research of Demelash et al. [20], which shown that low formal education has a risk of giving birth to LBW six times, where education affects a person's perception of behaviour including healthy behaviours such as maternal eating behaviour and utilization of health facilities.

Economic status has a significant positive relationship with LBW cases, with odds of 1.31. This result means that the risk of having LBW babies will be higher at 1.33 times for families with low economic status compared to those with middle/high economic status. Similar result was obtained in the research of Pramono and Paramita [14] and Yongki [21], which stated that families with low economic status have a risk of 1.31 times higher compared to families with upper-middle economic status for having LBW. This result is also related to the fulfilment of nutritional consumption during pregnancy, where families with upper middle income will probably meet nutritional needs. On the other hand, families with low incomes will perhaps more challenging to meet their nutritional needs.

Sources of drinking water have a significant positive relationship with LBW cases, with odds of 1.31. This result means that the risk of having LBW babies will be at 1.31 times higher among families with inadequate/inappropriate drinking water sources compared to those with adequate/proper drinking water sources. This finding was like Ramdaniati and Nastiti [3] research, which stated that families with inadequate water sources have risk of 2.182 times higher of stunting in their toddlers than those with adequate water sources. Water that unfit/polluted, of course, there is a risk of causing various diseases
such as diarrhea, typhus, hepatitis, and others. Improper water consumed during pregnancy will impact the fetus so that it will interfere with the fetus's development and health.

5. Conclusions and Suggestions
The logistic regression model with MLE and Bayesian estimation methods has equally good performance. All explanatory variables gave significant results. In this research, the logistic regression model with the Bayesian estimation method has provided a smaller BIC dan standard error values than the MLE In general, the Bayesian method will give more accurate results if the selection of the prior parameter distribution is correct. Multiple births, baby girl, maternal age at risk, birth spacing that is too close, iron deficiency, low education, low economic status, inadequate drinking water sources were some significant risk factors for the incidence of LBW. Therefore, a comprehensive policy relevant to maternal and child health from an economic and social perspective is needed to improve the health quality of mother and newborn.

References
[1] Kemenkes 2020 Profil Kesehatan Indonesia Tahun 2019 (Kementerian Kesehatan RI: Jakarta)
[2] Audrey H M and Candra A 2016 Hubungan Antara Status Anemia Ibu Hamil Trimester III dengan Kejadian Bayi Berat Lahir Rendah di Wilayah Kerja Puskesmas Halmahera, Semarang J. Kedokteran Diponegoro 5 1-6
[3] Ramdaniati S N and Nastiti D 2019 Hubungan Karakteristik Balita, Pengetahuan Ibu dan Sanitasi terhadap Kejadian Stunting pada Balita di Kecamatan Labuan Kabupaten Pandeglang HEARTY J. Kesehatan Masyarakat 7 47-88
[4] Meliza B D, Bustan M N and Sudarmin 2020 Analisis Regresi Logistik dengan Metode Penduga Bayes untuk Menentukan Faktor-Faktor Yang Berpengaruh Terhadap Kejadian Kanker Payudara (Studi Kasus: Pasien Kanker Payudara di Rumah sakit Dr. Wahidin Sudirohusodo) VARIANSI: Journal of Statistics and Its Application on Teaching and Research 2 52-61
[5] Teti M S, Yanuar F and Yozza H 2015 Analisis Regresi Logistik dengan Metode Penduga Bayes untuk Menentukan Faktor-Faktor yang Mempengaruhi Kejadian Bayi Berat Badan Lahir Rendah J. Matematika UNAND 4 53 – 60
[6] Kosim M D 2009 Buku Ajar Nenatologi (Jakarta: IDA)
[7] Setyawan A 2015 Pemodelan Regresi Logistik pada Kasus Berat Badan Lahir Rendah (BBLR) dan Penagruh Agregrasi Data terhadap Hasil Pendugaan (Bogor: Institut Pertanian Bogor)
[8] Sugito and Ispriyanti D 2010 Distribusi Invers Gamma Pada Inferensi Bayesian Media Statistik 59-68
[9] Bain L and Engelhardt M 1992 Introduction to Probability and Mathematical Statistics Second Edition (California: Duxbury Press)
[10] Ntzoufras I 2009 Bayesian Modeling Using WinBUGS (Ney Jersey: John Wiley & Sons, Inc)
[11] Chen T, Morris J and Martin E 2007 Gaussian Process Regression for Multivariate Spectroscopic Calibration, Chemometrics and Intelligent Laboratory Systems 87 85-97
[12] Masitoh S, Syafrudin and Delmaifanis 2014 Hamil Ganda Penyebab Bermakna Berat Bayi Lahir J. Ilmu dan Teknologi Kesehatan 1 129 - 134
[13] Rohayati D 2004 Hubungan Faktor Bayi dan Faktor Ibu dengan Kejadian BBLR di Jawa Barat (Depok: FKM UI)
[14] Pramono M S and Paramita A 2015 Pola Kejadian dan Determinan Bayi dengan Berat Badan Lahir Rendah (BBLR) di Indonesia Tahun 2013 Buletin Penelitian Sistem Kesehatan 18 1-10
[15] Kozuki N, Lee A C, Silveira M F, Victora C G, Adair L, Humphrey J, Ntozini R, Black R E and Kat J 2013 The associations of birth intervals with small-for-gestational-age, preterm, and neonatal and infant mortality: a meta-analysis *BMC Public Health* **13** 1-9

[16] Iriyani K Hubungan Pemberian Suplemen Zat Besi (Fe) pada Ibu Hamil Dengan Berat Badan Lahir Rendah (BBLR) di Rsud Abdul Wahab Sjahranie Samarinda *J. Imiah Manuntung* **2** 56-59

[17] Monita F, Suhaimi D and Ernalia Y 2016 Hubungan Usia, Jarak Kelahiran dan Kadar Hemoglobin Ibu Hamil dengan Kejadian Berat Bayi Lahir Rendah Di RSUD Arifin Achmad Provinci Riau *Jom FK* **3** 1-17

[18] Misna T, Wahyuni and Santi M 2013 *Determinan Epidemiologi Kejadian BBLR pada Daerah Endemis Kejadian BBLR pada Daerah Endemis Kalimantan Selatan* (Surabaya: Universitas Airlangga)

[19] Nuryani and Rahmawati Kejadian Berat Badan Lahir Rendah di Desa Tinelo *J. Gizi Pangan* **12** 49-54

[20] Demelash H, Motbainor A, Nigatu D, Gashaw K and Melese A 2015 Risk Factors for Low Birth Weight in Bale zone Hospitals, South-East Ethiopia: Case-Control Study *BMC Pregnancy & Children* **15** 264-74

[21] Yongky 2007 Analisis Pertambahan Berat Badan Ibu Hamil Berdasarkan Status Sosial Ekonomi dan Status Gizi Serta Hubungannya dengan Berat Bayi Baru Lahir (Bogor: Institut Pertanian Bogor)