Article

The Daily and Hourly Rainfall Data Modeling using Vector Autoregressive (VAR) with Maximum Likelihood Estimator (MLE) and Bayesian Method (Case Study in Sampean Watershed of Bondowoso, Indonesia)

Suci Astutik[1], Umu Sa'adah[2], Supriatna Adhisuwignjo[3] and Rauzan Sumara[4,*]

[1] Department of Statistics, Brawijaya University, East Java, Indonesia; suci_sp@ub.ac.id

[2] Department of Statistics, Brawijaya University, East Java, Indonesia; u.saadah@ub.ac.id

[3] Department of Electronics Engineering, State Polytechnic of Malang, East Java, Indonesia; supriatna_s@yahoo.com

* Correspondence: rauzan.sumara@yahoo.com; Tel.: +6281909111417

Abstract: The hourly and daily rainfall data which is spatially distributed are required as an input for run-off rain model. Furthermore, the run-off rain model is used to detect early flooding. The daily and hourly rainfall data have characteristics that most of data are zero. Therefore we need a model which can capture the phenomenon. A time series model involving location, which is a model that can be developed to approach the daily and hourly rainfall data, we can call Vector Autoregressive (VAR) model. The VAR model allows us for modeling rainfall data in several areas. However, in certain conditions the VAR model often occurs over-parameterization and reduces degrees of freedom. The aim of this study is to compare the VAR model with Maximum Likelihood Estimator (MLE) and Bayesian to hourly and daily rainfall data in SampeanWatershed of Bondowoso. The results showed that the hourly and daily rainfall data are fitted to VAR process of orde 5 and 1 respectively. Based on the AIC and SBC values indicate that the Bayesian is better than the MLE method. The Bayesian is able to predict parameters by producing a smaller variance covariance matrix than the MLE.

Keywords: VAR; MLE; Bayesian

---

1. Introduction

Disasters in Indonesia increase from 2002 to 2015. According to the National Disaster Management Agency the number of disasters which occurred in Indonesia is 143 disasters in 2002 and 1,681 disasters in 2015. Most of disasters in Indonesia are hydro-meteorological disasters such floods. From 1 January to 8 February 2016, the National Disaster Management Agency recorded 103 floods in Indonesia and 74,369 people were affected by flooding. The areas affected by flooding are East Java about 36 percent, Central Java 21 percent, Aceh 11 percent, West Sumatra 11 percent, Riau 7 percent, Jambi 4 percent, North Sumatra 4 percent, West Java 4 percent and West Nusa Tenggara 4 percent.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 October 2018 doi:10.20944/preprints201810.0302.v1

Therefore, it is necessary to have an early warning system about flooding in the areas. One of solutions is by simulating and predicting rainfall in these locations. Simulations and predictions on time series data such as rainfall data can use statistical models to explain dynamic of data. A statistical model which allows us for modeling rainfall data in several areas at once is the Vector Autoregressive (VAR) model.

Estimating parameter of the VAR model can use the Maximum Likelihood Estimator (MLE). In many cases of using MLE, there are a number of problems such as over-parameterization and collinearity. Therefore, Litterman [4], Sacakli [6] and Tahir [9] use the Bayesian VAR model to avoid these problems. The aim of this study is to model hourly and daily rainfall data with the VAR model using MLE and Bayesian and to compare the two estimation results based on the AIC and SBC values.

2. Materials and Methods

2.1. Materials

We collected secondary data from Sampean Baru Waterhed, Bondowoso. The data are hourly and daily rainfall in 7 rain stations these Sentral, Maesan, Ancar, Kejayan, Pakisan, Maskuning Wetan and Sukokerto during January 2006 and January 2007, which is hourly data e.g Sentral($Y_{1,t}$)and Maesan($Y_{2,t}$), and daily data e.g Sentral($Z_{1,t}$), Maesan($Z_{2,t}$), Ancar($Z_{3,t}$), Kejayan ($Z_{4,t}$), Pakisan ($Z_{5,t}$), Maskuning Wetan ($Z_{6,t}$), dan Sukokerto ($Z_{7,t}$). Stages of the analyses are (1) Stationary test, (3) order VAR(p), (4) estimating parameter of VAR model using MLE and Bayesian, (5) selecting the best model.

2.2. Methods

2.2.1. Order selection

Stationarity of data can be identified by looking at the Matrix Autocorrelation Function (MACF) and Matrix Partial Autocorrelation Function (MPACF). Order selection would be difficult if in the large matrix form, so Tiao and Box in Wei [11] noted the symbols (+), (-), and (.) in the (i, j) position of the sample correlation matrix. The three symbols are explained in Table 1.

Table 1. The meanings of the symbols

Symbol	Summary
+	Denotes a value greater than 2 times the estimated standard errors
-	Denotes a value less than 2 times the estimated standard errors
.	Denotes a value within 2 times the estimated standard errors

2.2.2. Matrix autocorrelation function (MACF)

According to Wei [11], a vector of *time series* is defined as $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_T$ then we can calculate the sample correlation matrix function,

$$\hat{\boldsymbol{\rho}}(k) = [\hat{\rho}_{ij}(k)] \quad (1)$$

Where $\hat{\rho}_{ij}(k)$ are the sample cross-correlations for the i th and j th component series,

$$\hat{\rho}_{ij}(k) = \frac{\sum_{t=1}^{T-k}(Z_{i,t}-\bar{Z}_i)(Z_{j,t+k}-\bar{Z}_j)}{\left[\sum_{t=1}^{T}(Z_{i,t}-\bar{Z}_i)^2 \sum_{t=1}^{T}(Z_{j,t}-\bar{Z}_j)^2\right]^{1/2}} \quad (2)$$

Where $\bar{Z}_i$ and $\bar{Z}_j$ are sample means of the corresponding component series.

2.2. 3. Matrix partial autocorrelation function (MPACF)

Tiao and Box in Wei [11] define the partial aautoregression matrix at lag*s*, denoted by $\boldsymbol{\mathcal{P}}(s)$, to be the last matrix coefficient when then data are fitted to VAR process of orde *s*. Therefore, $\boldsymbol{\mathcal{P}}(s)$ is equal to $\boldsymbol{\Phi}_{s,s}$. The partial autoregression matrix function is defined as,

$$\boldsymbol{\mathcal{P}}(s) = \begin{cases} \boldsymbol{\Gamma}'(1)[\boldsymbol{\Gamma}(0)]^{-1}, & s = 1 \\ \{\boldsymbol{\Gamma}'(s) - \boldsymbol{c}'(s)[\boldsymbol{A}(s)]^{-1}\boldsymbol{b}(s)\}\{\boldsymbol{\Gamma}(0) - \boldsymbol{b}'(s)[\boldsymbol{A}(s)]^{-1}\boldsymbol{b}(s)\}^{-1}, & s > 1 \end{cases} \quad (3)$$

covariance matrices $\boldsymbol{\Gamma}(s)$ can be obtained by the sample covariance matrices $\hat{\boldsymbol{\Gamma}}(s)$,

$$\hat{\boldsymbol{\Gamma}}(s) = \frac{1}{T}\sum_{t=1}^{T-s}(\boldsymbol{Z}_t - \bar{\boldsymbol{Z}})(\boldsymbol{Z}_{t+s} - \bar{\boldsymbol{Z}})', \quad s = 1,2,\ldots \quad (4)$$

where $\bar{\boldsymbol{Z}} = (\bar{Z}_1, \bar{Z}_1, \ldots, \bar{Z}_m)$ is the sample mean vector.

2.2. 4. Vector Autoregressive (VAR) using MLE

Vector Autoregressive (VAR) model is one of multivariate time series models which has dinamic interrelationship among variabels. Wei [11] defines stationary process of VAR(*p*),

$$\boldsymbol{Z}'_t = \boldsymbol{\Theta}' + \sum_{i=1}^{p}\boldsymbol{Z}'_{t-i}\boldsymbol{\Phi}'_i + \boldsymbol{\varepsilon}'_t \quad (5)$$

There are T observations, for $t = p+1, p+2, \ldots, T$ where p is VAR ordo. We have

$$\boldsymbol{Y} = \mathbf{X}\boldsymbol{\Phi} + \boldsymbol{\varepsilon} \quad (6)$$

where,

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Z}'_{p+1} \\ \vdots \\ \boldsymbol{Z}'_T \end{pmatrix}, \boldsymbol{X} = \begin{pmatrix} \mathbf{1} & \boldsymbol{Z}'_p & \cdots & \boldsymbol{Z}'_1 \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{1} & \boldsymbol{Z}'_{T-1} & \ddots & \boldsymbol{Z}'_{T-p} \end{pmatrix}, \boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\Theta}' \\ \boldsymbol{\Phi}'_1 \\ \vdots \\ \boldsymbol{\Phi}'_p \end{pmatrix}, \text{dan } \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}'_{P+1} \\ \vdots \\ \boldsymbol{\varepsilon}'_T \end{pmatrix}$$

where $\boldsymbol{Y}$ and $\boldsymbol{\varepsilon}$ are$(T-p) \times m$ matrices, $\boldsymbol{X}$is$(T-p) \times (1+mp)$ matrix of observations, and $\boldsymbol{\Phi}$ is$(1+mp) \times m$ matrix of unknown parameters. Defined$N = T - p$, likelihood function can be writen,

$$L(\mathbf{Y}|\boldsymbol{\Phi},\boldsymbol{\Sigma}) = (2\pi)^{-\frac{Nm}{2}}|\boldsymbol{\Sigma}|^{-\frac{N}{2}} exp\left\{-\frac{1}{2}tr\left[\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \mathbf{X}\boldsymbol{\Phi})'(\boldsymbol{Y} - \mathbf{X}\boldsymbol{\Phi})\right]\right\} \quad (7)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 October 2018 doi:10.20944/preprints201810.0302.v1

and MLEs of $\mathbf{\Phi}$ and $\mathbf{\Sigma}$ are

$$\hat{\mathbf{\Phi}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{Y}$$

and (8)

$$\hat{\mathbf{\Sigma}} = \frac{\boldsymbol{S}(p)}{N} = \frac{(\boldsymbol{Y}-\mathbf{X\Phi})'(\boldsymbol{Y}-\mathbf{X\Phi})}{T-p}$$

2.2. 5. Vector Autoregressive (VAR) using Bayeisan or Bayesian Vector Autoregressive (BVAR)

Bayesian method is one of the estimation methods used to estimate parameter. There are two important components in estimating using the Bayesian method, prior and posterior distribution. Ntzoufraz [5] noted posterior distribution equal to likelihood function times prior distribution which can be written,

$$f(\boldsymbol{\theta}|\boldsymbol{y}) \propto f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}) \tag{9}$$

A conjugate prior is used in this case. According to Berger [1], prior and posterior distributions in the conjugate prior have similar distributions. It is a multivariate normal distribution for $\mathbf{\Phi}$ parameters and a wishart distribution for $\mathbf{\Sigma}$ parameters. Kadiyala and Karlsson [2] stated that the normal multivariate-wishart prior for BVAR models is better than normal multivariate-diffuse, minnesota, and extended natural conjugate priorbased on RMSE (Root Mean Square Error) and CPU-Time. Tahir [9] noted the normal multivariate-wishartprior is better than minnesota and minnesota-wishart prior based on MSFE (Mean Square Forecast Error). Sims and Zha [7] added that the normal multivariate-wishart prior is suitable in complex models. Koop and Korobilis [3] and Sugita [8] define prior distribution of $\mathbf{\Phi}$ and $\mathbf{\Sigma}$,

$$vec(\mathbf{\Phi}) \sim \mathrm{N}(vec(\mathbf{\Phi}_0), \boldsymbol{V}_0) \tag{10}$$

$$\mathbf{\Sigma}^{-1} \sim \mathrm{Wishart}(\mathbf{S}_0{}^{-1}, n_0)$$

Where $\mathbf{\Phi}_0, \boldsymbol{V}_0, \mathbf{S}_0,$ *and* v_0 are hyperparameter. Initialization of hyperparamter can use non informative prior such as $\mathbf{S}_0, \boldsymbol{V}_0, \mathbf{\Phi}_0,$ and n_0 near to zero. The join posterior is obtained by the join prior

$$f(\boldsymbol{\theta}) \propto f(vec(\mathbf{\Phi}))f(\mathbf{\Sigma}^{-1})$$

$$\propto |\boldsymbol{V}_0|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\left[vec(\mathbf{\Phi}-\mathbf{\Phi}_0)'\boldsymbol{V}_0{}^{-1}vec(\mathbf{\Phi}-\mathbf{\Phi}_0)\right]\right\} \times |\mathbf{S}_0|^{\frac{n_0}{2}}|\mathbf{\Sigma}|^{-\frac{n_0-m-1}{2}} exp\left\{-\frac{1}{2}tr[\mathbf{\Sigma}^{-1}\mathbf{S}_0]\right\}$$

$$\propto |\boldsymbol{V}_0|^{-\frac{1}{2}}|\mathbf{S}_0|^{\frac{n_0}{2}}|\mathbf{\Sigma}|^{-\frac{n_0-m-1}{2}} exp\left\{-\tfrac{1}{2}\left[tr(\mathbf{\Sigma}^{-1}\mathbf{S}_0) + vec(\mathbf{\Phi}-\mathbf{\Phi}_0)'\boldsymbol{V}_0{}^{-1}vec(\mathbf{\Phi}-\mathbf{\Phi}_0)\right]\right\} \tag{11}$$

with the likelihood function (7), so that the join posterior can be written,

$$f(\boldsymbol{\theta}|\boldsymbol{Y}) \propto L(\boldsymbol{Y}|\mathbf{\Phi},\mathbf{\Sigma})f(\boldsymbol{\theta})$$

 doi:10.20944/preprints201810.0302.v1

$$\propto |\boldsymbol{\Sigma}|^{-\frac{N}{2}} exp\left\{-\frac{1}{2} tr\left[\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\mathbf{X\Phi})'(\boldsymbol{Y}-\mathbf{X\Phi})\right]\right\}$$

$$\times |\boldsymbol{V}_0|^{-\frac{1}{2}} |\mathbf{S}_0|^{\frac{n_0}{2}} |\boldsymbol{\Sigma}|^{-\frac{n_0-m-1}{2}} exp\left\{-\frac{1}{2}\left[tr(\boldsymbol{\Sigma}^{-1}\mathbf{S}_0)\right.\right.$$

$$\left.\left. + vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)'\boldsymbol{V}_0^{-1} vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)\right]\right\}$$

$$\propto |\mathbf{S}_0|^{\frac{n_0}{2}} |\boldsymbol{\Sigma}|^{-\frac{N+n_0-m-1}{2}} |\boldsymbol{V}_0|^{-\frac{1}{2}} exp\left\{-\frac{1}{2} tr\left[\boldsymbol{\Sigma}^{-1}\left((\boldsymbol{Y}-\mathbf{X\Phi})'(\boldsymbol{Y}-\mathbf{X\Phi}) +\right.\right.\right.$$

$$\left.\left.\left.\mathbf{S}_0\right)\right]\right\} \times exp\left\{-\frac{1}{2}\left[vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)'\boldsymbol{V}_0^{-1} vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)\right]\right\} \quad (12)$$

From the join posterior (12), we can derive the conditional posterior density for $\boldsymbol{\Sigma}$,

$$f(\boldsymbol{\Sigma}^{-1}|\boldsymbol{Y}, vec(\boldsymbol{\Phi})) \propto \frac{f(\boldsymbol{\theta}|\boldsymbol{Y})}{f(vec(\boldsymbol{\Phi}))}$$

$$\propto |\boldsymbol{\Sigma}|^{-\frac{N+n_0-m-1}{2}} exp\left\{-\frac{1}{2} tr\left[\boldsymbol{\Sigma}^{-1}\left((\boldsymbol{Y}-\mathbf{X\Phi})'(\boldsymbol{Y}-\mathbf{X\Phi}) + \mathbf{S}_0\right)\right]\right\}$$

$$\propto |\boldsymbol{\Sigma}|^{-\frac{n_1-m-1}{2}} exp\left\{-\frac{1}{2} tr[\boldsymbol{\Sigma}^{-1}\mathbf{S}_1]\right\} \quad (13)$$

where $\mathbf{S}_1 = (\boldsymbol{Y}-\mathbf{X\Phi})'(\boldsymbol{Y}-\mathbf{X\Phi}) + \mathbf{S}_0$ and $n_1 = T - p + n_0$ .So that the conditional posterior statistic form is

$$\boldsymbol{\Sigma}^{-1}|\boldsymbol{Y}, vec(\boldsymbol{\Phi}) \sim \text{Wishart}(\mathbf{S}_1^{-1}, n_1) \quad (14)$$

Then we derive the conditional posterior density for $\boldsymbol{\Phi}$,

$$f(vec(\boldsymbol{\Phi})|\boldsymbol{Y}, \boldsymbol{\Sigma}^{-1}) \propto \frac{f(\boldsymbol{\theta}|\boldsymbol{Y})}{f(\boldsymbol{\Sigma}^{-1})}$$

$$\propto exp\left\{-\frac{1}{2}\left[tr\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\mathbf{X\Phi})'(\boldsymbol{Y}-\mathbf{X\Phi})\right) + vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)'\boldsymbol{V}_0^{-1} vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)\right]\right\}$$

$$\propto exp\left\{-\frac{1}{2}\left[vec(\boldsymbol{Y}-\mathbf{X\Phi})'(\boldsymbol{\Sigma}\otimes\mathbf{I})^{-1} vec(\boldsymbol{Y}-\mathbf{X\Phi})\right.\right.$$

$$\left.\left. + vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)'\boldsymbol{V}_0^{-1} vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_0)\right]\right\}$$

$$\propto exp\left\{-\frac{1}{2}\left[vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_*)'\boldsymbol{V}_1^{-1} vec(\boldsymbol{\Phi}-\boldsymbol{\Phi}_*)\right]\right\} \quad (15)$$

where $\boldsymbol{V}_1 = \left[\boldsymbol{V}_0{}^{-1} + (\boldsymbol{\Sigma}^{-1} \otimes (\mathbf{X}'\mathbf{X}))\right]^{-1}$ and $vec(\boldsymbol{\Phi}_*) = \boldsymbol{V}_1\left[\boldsymbol{V}_0{}^{-1}vec(\boldsymbol{\Phi}_0) + (\boldsymbol{\Sigma} \otimes \mathbf{I})^{-1}vec(\mathbf{X}'\mathbf{Y})\right]$, and we get the conditional posterior statistic form is

$$vec(\boldsymbol{\Phi})\,|\boldsymbol{Y}, \boldsymbol{\Sigma}^{-1} \sim N(vec(\boldsymbol{\Phi}_*), \boldsymbol{V}_1) \tag{16}$$

In this study we use Gibbs Sampler Markov Chain Monte Carlo (MCMC) methods to sample the parameter of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Phi}$. The conditional posterior distribution is used to process in Gibbs Sampler algorithm.

2.2. 6. Markov Chain Monte Carlo (MCMC)

Markov Chain Monte Carlo (MCMC) methods are widely used in Bayesian inference. According to Ntzoufraz [5], a Markov chain is a stochastic process $\{\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(3)}, \ldots, \boldsymbol{\theta}^{(R)}\}$ such that

$$f(\boldsymbol{\theta}^{(r+1)}|\boldsymbol{\theta}^{(r)}, \ldots, \boldsymbol{\theta}^{(1)}) = f(\boldsymbol{\theta}^{(r+1)}|\boldsymbol{\theta}^{(r)}) \tag{17}$$

In order to generate a sample from $f(\boldsymbol{\theta}|\boldsymbol{y})$, we must contruct a Markov chain with two desired properties. First, $f(\boldsymbol{\theta}^{(r+1)}|\boldsymbol{\theta}^{(r)})$ should be easy to generate from, and second, equilibrium distribution of the selected Marcov chain must be the posterior distribution of interest $f(\boldsymbol{\theta}|\boldsymbol{y})$. Assuming that we have met with these requirements, we then

1. Select an initial value $\boldsymbol{\theta}^{(0)}$.
2. Generate R values until the equilibrium distribution is reached.
3. Monitor the convergence of MCMC. If convergence diagnostics fail, we then generate more observations.
4. Cut off the first B observations (*Burn-in Priod*).
5. Consider$\{\boldsymbol{\theta}^{(B+1)}, \boldsymbol{\theta}^{(B+1)}, \ldots, \boldsymbol{\theta}^{(R)}\}$as the sample for the posterior analysis.

We refers to analysis of the MCMC, $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(3)}, \ldots, \boldsymbol{\theta}^{(R)}$. From this sample, for any function $G(\boldsymbol{\theta})$ of the parameters of interest $\boldsymbol{\theta}$ we can

1. Obtain a sample of the desired parameter $G(\boldsymbol{\theta})$ by sample considering

$$G(\boldsymbol{\theta}^{(1)}), G(\boldsymbol{\theta}^{(2)}), G(\boldsymbol{\theta}^{(3)}), \ldots G(\boldsymbol{\theta}^{(r)}), \ldots, G(\boldsymbol{\theta}^{(R)})$$

2. Obtain any posterior summary of $G(\boldsymbol{\theta})$ from the sample using traditional sample estimates. For instance, we can estimate the posterior mean by

$$\hat{E}(G(\boldsymbol{\theta})|\boldsymbol{y}) = \frac{1}{R-B}\sum_{r=1}^{R-B} G\left(\boldsymbol{\theta}^{(r)}\right)$$

and the posterior standard deviation by

$$\widehat{SD}(G(\boldsymbol{\theta})|\boldsymbol{y}) = \sqrt{\frac{1}{R-B-1}\sum_{r=1}^{R-B}\left[G(\boldsymbol{\theta}^{(r)}) - \hat{E}(G(\boldsymbol{\theta})|\boldsymbol{y})\right]^2}$$

3. Obtain other measures of interest might be posterior such as 2.5% and 97.5% credible interval.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 October 2018 doi:10.20944/preprints201810.0302.v1

4. Calculate Monte Carlo Error (MC Error).

2.2. 7. Gibbs Sampler

Gibbs Sampler is usually cited as a separate simulation technique because of its popularity and convenience. One advantage of the Gibbs Sampler is that, in each step, random values must be generated from unidimensional distributions for which a wide variety of computational tools exists. Frequently, these conditional distributions have a known form and, thus, random number can be easily simulated using standard function in statistical and computing software. Ntzouftaz [5] noted the algorithm can be summarized by the following steps :

1. Set initial values,

$$\boldsymbol{\theta}^{(0)} = ({\theta_1}^{(0)}, {\theta_2}^{(0)}, \ldots, {\theta_P}^{(0)})$$

2. For $r = 1, \ldots, T$ repeat the following steps

$${\theta_1}^{(r)} \text{ dari } f(\theta_1 | {\theta_2}^{(r-1)}, {\theta_3}^{(r-1)}, \ldots, {\theta_P}^{(r-1)}, \boldsymbol{y})$$

$${\theta_2}^{(r)} \text{ dari } f(\theta_2 | {\theta_1}^{(r)}, {\theta_3}^{(r-1)}, \ldots, {\theta_P}^{(r-1)}, \boldsymbol{y})$$

$${\theta_3}^{(r)} \text{ dari } f(\theta_3 | {\theta_1}^{(r)}, {\theta_2}^{(r)}, {\theta_4}^{(r-1)}, \ldots, {\theta_P}^{(r-1)}, \boldsymbol{y})$$

$$. \quad . \quad .$$

$${\theta_P}^{(r)} \text{ dari } f(\theta_P | {\theta_1}^{(r)}, {\theta_2}^{(r)}, \ldots, {\theta_{r-1}}^{(r)}, \boldsymbol{y})$$

3. Set $\boldsymbol{\theta}^{(r)} = \boldsymbol{\theta}$ and save it as the generated set of $r+1$ iteration of the algorithm.

2.2. 8. Information Criteria

Information criteria have been shown to be effective in selecting a statistical model. In the time series literature, several criteria have been proposed. Two criteria functions are commonly used to determine VAR model, these are AIC (Akaike Information Criterion) and SBC (Schwarzt Bayesian Criterion). The best model is that produce the smallest AIC and SBC values. Tsay[10] provided to calculate the AIC

$$AIC = \ln|\boldsymbol{\Sigma}| + \frac{2}{T} pm^2 \quad (18)$$

and SBC

$$AIC = \ln|\boldsymbol{\Sigma}| + \frac{2}{T} pm^2 \quad (19)$$

Here T is observation, $|\boldsymbol{\Sigma}|$ is determinant of variance-covariance matrix, p is order model and m is variables.

3. Results

For check stationarity in data, we use *Dickey-Fuller* test (DF). The hypotesis of DF test is $H_0 : \delta = 0$ (data are non-stationary) vs $H_1 : \delta < 0$ (data are stationary), as shown on Table 1.

Table 2. Dickey-Fuller test

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 October 2018 doi:10.20944/preprints201810.0302.v1

Data	Variable	df	p-value	Conclusions
low time-scale (hourly data)	$Y_{1,t}$	-768.16	0.0001	Stationary
	$Y_{2,t}$	-1128.20	0.0001	Stationary
high time-scale (daily data)	$Z_{1,t}$	-15.40	0.0046	Stationary
	$Z_{2,t}$	-20.53	0.0008	Stationary
	$Z_{3,t}$	-13.14	0.0094	Stationary
	$Z_{4,t}$	-18.15	0.0019	Stationary
	$Z_{5,t}$	-30.70	0.0001	Stationary
	$Z_{6,t}$	-20.12	0.0010	Stationary
	$Z_{7,t}$	-15.86	0.0040	Stationary

Based on Table 2, *probability values* of statistic tests are less than α (0.05), reject H_0. So the data were stationary. Order selection can be identified by looking at the Matrix Autocorrelation Function (MACF) and Matrix Partial Autocorrelation Function (MPACF). If the time lag of MACF decreases exponential or sinusoid whereas the MPACF is cut off at lag *p*, it is identified as the VAR(p) model. these MACF and MPACF of hourly data based on Table 3 and 4.

Table3. Schematic MACF of hourly data

Schematic Representation of Cross Correlations											
Variable/Lag	0	1	2	3	4	5	6	7	8	9	10
Sentral	+.	++	++	++	++	.+	.+	.+	.+	.+	..
Maesan	.+	.+	+.	.+	..	.+	.+	+.	+.	..	..

Table 4. Schematic MPACF of hourly data

Schematic Representation of Partial Cross Correlations										
Variable/Lag	1	2	3	4	5	6	7	8	9	10
Sentral	+.	+.	++	..	.+	..	..	..	..	..
Maesan	.+	+.	..	..	.+	..	..	..	..	..

And Schematic MACF and MPACF of daily data based onTable 5 and 6,

Table 5. Schematic MACF of daily data

Schematic Representation of Cross Correlations						
Variable/Lag	0	1	2	3	4	5
Sentral	+++++++	+.+....	+.++...	....	.+.....	.+...++
Maesan	+++++++	..+..+.	++++..+	....	...+..+	.+...+.
Ancar	+++++++	++++.+.	++++...	....	.+.....	.+....+
Kejayan	+++++++	....	+......	....	....	.....++

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 October 2018 doi:10.20944/preprints201810.0302.v1

Schematic Representation of Cross Correlations						
Variable/Lag	0	1	2	3	4	5
Pakisan	+++++++	....	....	....	....	....
Maskuning_Wetan	+++++++	+.+....	+.++...	....	....	.+.....
Sukokerto	+++++++	+++..+.	+.++..+	...+...	....	....

Table 6. Schematic MPACF of daily data

Schematic Representation of Partial Cross Correlations					
Variable/Lag	1	2	3	4	5
Sentral	....	....	....	.+.....	....
Maesan	....	....	....	-..+...	....
Ancar	+.+....	....	...-...	....	....
Kejayan	....	....	....	....	.....+.
Pakisan	....	....	....	....	....
Maskuning_Wetan	....	..+....	....	....	....
Sukokerto	....	....	....	....	....

From schematic MACF and MPACF above, we can conclude the hourly data following VAR(5) process and the daily data following VAR(1) process. Estimating parameter is showed on Table 7 and 8.

Table 7. Estimation of VAR(5) parameter using MLE

Parameter	Estimation	Standard Error	p-value
$\hat{\phi}_{11}$	0.32857	0.02602	0.0001
$\hat{\phi}_{12}$	0.03919	0.03044	0.1981
$\hat{\phi}_{13}$	0.06520	0.02739	0.0174
$\hat{\phi}_{14}$	0.10129	0.03060	0.0010
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\hat{\phi}_{27}$	0.01775	0.02309	0.4420
$\hat{\phi}_{28}$	-0.00566	0.02567	0.8256
$\hat{\phi}_{29}$	0.06263	0.02204	0.0046
$\hat{\phi}_{210}$	0.16358	0.02523	0.0001

Table 8. Estimation of VAR(1) parameter using MLE

Parameter	Estimation	Standard Error	p-value
$\hat{\phi}_{11}$	-0.18800	0.16959	0.2591
$\hat{\phi}_{12}$	-0.19337	0.23768	0.3496
$\hat{\phi}_{13}$	0.85982	0.20806	0.6616

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 October 2018 doi:10.20944/preprints201810.0302.v1

$\hat{\phi}_{14}$	-0.67658	0.42041	0.3327
⋮	⋮	⋮	⋮
$\hat{\phi}_{74}$	-0.09275	0.23840	0.6988
$\hat{\phi}_{75}$	0.05637	0.20870	0.7881
$\hat{\phi}_{76}$	-0.13940	0.42170	0.7423
$\hat{\phi}_{77}$	0.23132	0.21750	0.2923

Based onTabel 7 and 8, VAR(5) model of hourly rainfall data is

$$\begin{pmatrix} Y_{1,t} \\ Y_{2,t} \end{pmatrix} = \begin{pmatrix} 0.32857 & 0.03919 \\ 0.02949 & 0.10497 \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0.0652 & 0.10129 \\ 0.00277 & 0.01727 \end{pmatrix} \begin{pmatrix} Y_{1,t-2} \\ Y_{2,t-2} \end{pmatrix} + \begin{pmatrix} 0.08415 & -0.03713 \\ 0,07183 & 0.05306 \end{pmatrix} \begin{pmatrix} Y_{1,t-3} \\ Y_{2,t-3} \end{pmatrix}$$

$$+ \begin{pmatrix} 0.01734 & -0.01373 \\ 0.01775 & -0.00566 \end{pmatrix} \begin{pmatrix} Y_{1,t-4} \\ Y_{2,t-4} \end{pmatrix} + \begin{pmatrix} -0.0406 & 0.0399 \\ 0.06263 & 0.16358 \end{pmatrix} \begin{pmatrix} Y_{1,t-5} \\ Y_{2,t-5} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \quad (20)$$

VAR(1) model of daily rainfall data is

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \\ Z_{5,t} \\ Z_{6,t} \\ Z_{7,t} \end{pmatrix} = \begin{pmatrix} -0.18800 & -0.19337 & 0.85982 & -0.67658 & 0.09860 & -0.03643 & 0.63522 \\ -0.05576 & 0.11628 & 0.22896 & -0.07745 & -0.09388 & -0.11320 & 0.43326 \\ -0.07923 & -0.10828 & 0.88104 & -0.75124 & -0.06305 & 0.17090 & 0.67162 \\ -0.08718 & -0.00664 & 0.41018 & -0.43418 & 0.13383 & -0.21618 & 0.42185 \\ 0.03235 & 0.01375 & 0.19344 & -0.22428 & 0.09159 & -0.41097 & 0.44491 \\ -0.06043 & 0.10110 & 0.13431 & -0.12422 & 0.00324 & -0.15626 & 0.31318 \\ -0.13302 & 0.12277 & 0.26998 & -0.09275 & 0.05637 & -0.13940 & 0.23132 \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \\ Z_{4,t-1} \\ Z_{5,t-1} \\ Z_{6,t-1} \\ Z_{7,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \\ \varepsilon_{5,t} \\ \varepsilon_{6,t} \\ \varepsilon_{7,t} \end{pmatrix}$$

(21)

The simulation process in estimating parameters with Bayesian method uses Gibbs Sampler algorithm. Initial values in the simulation process are approximated by MLE. The first step of simulation process generates the parameter $\mathbf{\Sigma}$ then the second step generates the parameter $\mathbf{\Phi}$. To compile and run the MCMC algorithm for 10000 iterations and 500 burn in. It is divided into 97 batches for calculating MC error. The posterior summary is showed in Table 9 and 10

Table 9. Estimation of BVAR(5) parameter

Node	*Mean*	*Standard Deviation*	*MC Error*	2.50%	Median	97.50%
$\hat{\sigma}_{11}$	3.2674	0.0846	0.0008	3.1040	3.2665	3.4360
$\hat{\sigma}_{12}$	0.0268	0.0505	0.0005	-0.0715	0.0263	0.1272
$\hat{\sigma}_{21}$	0.0268	0.0505	0.0005	-0.0715	0.0263	0.1272
$\hat{\sigma}_{22}$	2.3152	0.0597	0.0006	2.2020	2.3150	2.4370
$\hat{\phi}_{11}$	0.3286	0.0184	0.0002	0.2930	0.3286	0.3646
$\hat{\phi}_{21}$	0.0294	0.0155	0.0002	-0.0011	0.0294	0.0592
$\hat{\phi}_{12}$	0.0392	0.0216	0.0003	-0.0032	0.0392	0.0822
$\hat{\phi}_{22}$	0.1050	0.0180	0.0002	0.0696	0.1049	0.1397
$\hat{\phi}_{13}$	0.0652	0.0192	0.0002	0.0278	0.0651	0.1030
$\hat{\phi}_{23}$	0.0030	0.0164	0.0002	-0.0295	0.0030	0.0342
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\hat{\phi}_{19}$	-0.0409	0.0187	0.0002	-0.0771	-0.0407	-0.0044
$\hat{\phi}_{29}$	0.0624	0.0156	0.0001	0.0314	0.0624	0.0926
$\hat{\phi}_{110}$	0.0400	0.0211	0.0002	-0.0014	0.0402	0.0817

$\hat{\phi}_{210}$	0.1637	0.0179	0.0002	0.1286	0.1638	0.1990

Table 10. Estimation of BVAR(1) parameter

Node	*Mean*	*Standard Deviation*	*MC Error*	2.50%	Median	97.50%
$\hat{\sigma}_{11}$	171.2337	23.4692	0.2617	131.9000	169.4000	223.2525
$\hat{\sigma}_{12}$	100.7099	17.4116	0.1865	70.7800	99.2750	138.9000
$\hat{\sigma}_{13}$	101.3179	17.7741	0.1891	70.6195	99.6850	140.9000
$\hat{\sigma}_{14}$	78.7235	14.3701	0.1585	54.0095	77.3400	110.4525
$\hat{\sigma}_{15}$	88.9727	16.3180	0.1861	60.8948	87.8850	124.9000
$\hat{\sigma}_{16}$	43.7818	8.6913	0.0982	28.4443	43.1200	62.6920
$\hat{\sigma}_{17}$	85.1486	16.3029	0.1820	56.9648	83.6700	120.9000
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\hat{\phi}_{17}$	0.6363	0.1825	0.0019	0.2756	0.6381	0.9976
$\hat{\phi}_{27}$	0.4345	0.1590	0.0015	0.1231	0.4344	0.7508
$\hat{\phi}_{37}$	0.6732	0.1635	0.0015	0.3512	0.6746	0.9953
$\hat{\phi}_{47}$	0.4219	0.1334	0.0012	0.1653	0.4214	0.6883
$\hat{\phi}_{57}$	0.4442	0.1535	0.0016	0.1399	0.4442	0.7458
$\hat{\phi}_{67}$	0.3133	0.0839	0.0008	0.1491	0.3123	0.4809
$\hat{\phi}_{77}$	0.2313	0.1563	0.0014	-0.0718	0.2295	0.5351

Based on table 9 and 10, each of parameters is konvergence, because MC Error values are less than 1% standard deviation. If 2.5% and 97.5% percentiles does not contain a zero, the parameter will be significant. We have BVAR(5) model of hourly data,

$$\begin{pmatrix} Y_{1,t} \\ Y_{2,t} \end{pmatrix} = \begin{pmatrix} 0.3286 & 0.0392 \\ 0.0294 & 0.1050 \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0.0652 & 0.1012 \\ 0.0030 & 0.0172 \end{pmatrix} \begin{pmatrix} Y_{1,t-2} \\ Y_{2,t-2} \end{pmatrix} + \begin{pmatrix} 0.0845 & -0.0374 \\ 0,0716 & 0.0530 \end{pmatrix} \begin{pmatrix} Y_{1,t-3} \\ Y_{2,t-3} \end{pmatrix}$$

$$+ \begin{pmatrix} 0.0174 & -0.0136 \\ 0.0180 & -0.0055 \end{pmatrix} \begin{pmatrix} Y_{1,t-4} \\ Y_{2,t-4} \end{pmatrix} + \begin{pmatrix} -0.0409 & 0.0400 \\ 0.0624 & 0.1637 \end{pmatrix} \begin{pmatrix} Y_{1,t-5} \\ Y_{2,t-5} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \quad (22)$$

and BVAR(1) model of daily data

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \\ Z_{4,t} \\ Z_{5,t} \\ Z_{6,t} \\ Z_{7,t} \end{pmatrix} = \begin{pmatrix} -0.1866 & -0.1923 & 0.8610 & -0.6816 & 0.1009 & -0.0404 & 0.6363 \\ -0.0539 & 0.1167 & 0.2302 & -0.0831 & -0.0927 & -0.1142 & 0.4345 \\ -0.0802 & -0.1078 & 0.8832 & -0.7547 & -0.0598 & 0.1671 & 0.6732 \\ -0.0851 & -0.0074 & 0.4104 & -0.4376 & 0.1351 & -0.2158 & 0.4219 \\ 0.0349 & 0.0137 & 0.1925 & -0.2277 & 0.0933 & -0.4120 & 0.4442 \\ -0.0597 & 0.1012 & 0.1344 & -0.1262 & 0.0041 & -0.1558 & 0.3133 \\ -0.1305 & 0.1229 & 0.2703 & -0.0980 & 0.0584 & -0.1401 & 0.2313 \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ Z_{3,t-1} \\ Z_{4,t-1} \\ Z_{5,t-1} \\ Z_{6,t-1} \\ Z_{7,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \\ \varepsilon_{5,t} \\ \varepsilon_{6,t} \\ \varepsilon_{7,t} \end{pmatrix} \quad (23)$$

The model selection uses AIC and SBC values. The AIC and SBC values are showed in Table 11 bellow,

Table 11. AIC and SBC values

Model	Estimation Method	AIC	SBC
VAR(5)	MLE	3.4257	3.4972
	Bayesian	2.0503	2.1216

VAR(1)	MLE	33.5855	35.2811
	Bayesian	29.7573	31.4385

It shows that Bayesian method is better than MLE, because the AIC and SBC values are smaller than MLE method.

4. Discussion

Based on Table 2 - Table 6 it can be seen that hourly and daily rainfall data have fulfilled the stationary assumption, which is the assumption of the VAR model. The corresponding hourly and daily rainfall models are VAR (5) and VAR (1), respectively, seen from MACF and MPACF. The estimation results of the VAR(1) model parameter for daily rainfall data with the MLE method (Equation 20) and the Bayesian method (Equation 22) produce values that are not much different. So do the VAR(5) model for hourly rainfall data by the MLE method (equation 21) and the bayesian method (Equation 23). This shows that the MLE and bayesian methods provide the results of estimating the parameters of the VAR model that are almost the same. However, based on the AIC and SBC values (Table 11) shows that the Bayesian method produces a VAR model with smaller AIC and SBC values compared to the MLE method for both hourly and daily rainfall data. This means that in hourly and daily rainfall data, the Bayesian method produces a VAR model that is better than the MLE method.

5. Conclusions

We can conclude that the hourly rainfall data and daily rainfall data at Sampean Bondowoso watershed station follow VAR(5) and VAR(2) respectively. The criteria of best models show that parameter estimation using Bayesian method is better than MLE method, because Bayesian method is able to predict parameters with the smaller variance matrix than MLE method.

Acknowledgments: Thank you to Kemenristekdikti for the support of research funding in the Fundamental Research Grant 2018, the Rector of Brawijaya University, the Dean of FMIPA and the Chair of the Mathematics Department of Brawijaya University for the support of the permission and funding assistance for this conference.

References

1. Berger, J.O. *Statistical Decision Theory and Bayesian Analysis*, Second Edition, Springer, New York, 1985.
2. Kadiyala, K.R.; Karlsson, S. *Numerical Methods for Estimation and Inference in Bayesian VAR-Models*, Journal of Applied Econometrics.12 (1997), pp. 99-132.
3. Koop, G.; Korobilis, D. *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics*, University of Strathclyde, Galasgow Scotland, 2010.
4. Litterman, R.B. *Forecasting with Bayesian Vector Autoregressive : Five Years of Experience*, Journal of Business & Economic Statistics. 4 (1986), pp. 25-38.
5. Ntzoufras, I. *Bayesian Modeling Using WinBUGS*, John Wiley & Sons, United States, 2009.
6. Sacakli, I.S. *Do BVAR Models Forecast Turkish GDP Better Than UVAR Models ?*, British Journal of Economics, Management & Trade. 7(2015), pp. 259-268.

 doi:10.20944/preprints201810.0302.v1

7. Sims, C.A.; Zha, T. Bayesian Methods for Dynamic Multivariate Models. *International Economic Review.* 39(1998), pp. 949-968.
8. Sugita, K. Bayesian Analysis of a Vector Autoregressive Model With Multiple Structural Breaks, Economics Bulletin. 3(2008), pp. 1-7.
9. Tahir, M.A. *Analyzing and Forecasting Output Gap and Inflation Using Bayesian Vector Auto Regression (BVAR) Method: A Case of Pakistan*,International Jurnal of Economics and Finance.6 (2014), pp. 233-243.
10. Tsay, R.S. *Multivariate Time Series AnalysisWith R and Financial Applications*, John Wiley & Sons, Canada, 2014.
11. Wei, W.W.S. *Time Series Analysis Univariate & Multivariate*, Second Edition, Addison Wesley, Boston, 2006.