VOL. 9, NO. 9, SEPTEMBER 2014 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

BAYESIAN APPROACH ON PARAMETER ESTIMATION IN HIDDEN MARKOV MODEL

Dwi Agustin N.S[1], Septiadi Padmadisastra[2] and Sudartianto[2]
[1]Departement of Industrial Engineering, Widya Mandala Catholic University, Surabaya, Indonesia
[2]Departement of Statistics, Padjadjaran University, Bandung, Indonesia
E-Mail: dwi.agustinns@gmail.com

ABSTRACT

This paper presents study about the parameter estimation in hidden markov model. The approach is taken from a Bayesian method, there will be two sources of information,there are information from the likelihood function and the prior function. This approach will be applied to daily rainfall data in Darajat, Garut. The numbers of hidden states are used in this paper based on Schmidth and Fergusson's climate classification which are suitable to the local conditions. This classification was obtained three types of division in the period of one year where the condition called wet months when monthly rainfall > 100 mm per month, moist months when monthly rainfall between 100 - 60 mm and the dry months when monthly rainfall <60 mm per month. The process estimation of hidden markov parameters is using Gibbs Sampler algorithm.

Keywords: hidden Markov models, Bayesian, Schmidth and Fergusson's Climate classification, Gibbs sampler.

INTRODUCTION

Modeling of the precipitation has been developed by many researchers; one of them is Coe and Stern (1982) who tried to modeling the daily rainfall data in Zinder region, Nigeria and the Kharja, Jordan. The method used at their research was Generalized Linear Models (GLM) and markov chain. In the 1970's the mathematician Baum and Petrie introduced markov chain development, namely Hidden Markov Model (HMM). HMM was increasingly popular applied in various fields? Rabiner (1989) applied the methods of HMM in speech recognition. Zucchini and Guttorp (1991) applied HMM in the precipitation phenomena. In their research, Zucchini and Guttorp introduced unobserved climate states that influence the occurrence of rain. Thyer and Kuezera (2000) developed a method of HMM to simulate the long-term hydroclimatic data for water resource planning in Australia.

In previous studies, the number of hidden states was observed only two kinds, while in Indonesia itself more widely used climate classification based on Schmidth and Fergusson. In Sudrajat's research (2009) mentioned that according Schmidth and Fergusson (1951), classification of climate is based on a comparison between the dry and wet months, from this relationship was obtained three types of division in the period of one year where the condition called wet months when monthly rainfall > 100 mm per month, moist months when monthly rainfall between 100 - 60 mm and the dry months when monthly rainfall <60 mm per month.

Garut's topographic itself is the mountains area where there are protected forests and plantations. The climate clasification which are suitable for this area is climate classification based on Schmidth and Fergusson (Sudrajat, 2009).

The process of parameter estimation in hidden markov models in several previous studies were based on a frequentist approach using Baum-Welch algorithm or EM Algorithm. At this approach, the parameters only considered as a fixed value. However, there are other approaches that assume these parameters will form a random variable. It can be happened, because the parameters of hidden markov models, particular in the case of hydrology (rainfall), the influence of time would make the values of these parameters to form a random variable having a probability distribution. So the Bayesian approach can be used in the process of parameter estimation in hidden markov model. So in this paper want to study about parameter estimation in hidden markov models with three different types of hidden climate states in Garut through a Bayesian approach.

HIDDEN MARKOV MODEL (HMM) WITH THREE HIDDEN CLIMATE STATES

Basic Model of HMM

HMM are models in which the distribution that generates an observations depends on the state of an underlying and unobserved markov process (Zucchini and MacDonald, 2009).

The basic model of HMM was illustrated by Ingmar Visser (2011) as shown below:

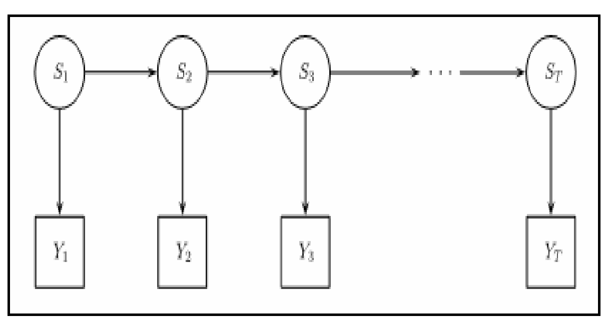

Figure-1. Hidden Markov Model Illustrated.

Zucchini and MacDonald (2009) are formulating the basic model above becomes:

$$\Pr\left(S_t \middle| S^{(t-1)}\right) = \Pr\left(S_t \middle| S_{t-1}\right), \qquad t = 2,3,... \qquad (1)$$

$$\Pr\left(Y_t \middle| Y^{(t-1)}, S^{(t)}\right) = \Pr\left(Y_t \middle| S_t\right), \qquad t = 2,3,... \qquad (2)$$

with $\mathrm{Y}^{(t)} = \left(Y_1, Y_1, ..., Y_T\right)$ and $S^{(t)} = \left(S_1, S_2, ..., S_T\right)$.

The Equilibrium Probability from HMM with Three Hidden Climate States

Climate states will be following a markov chain with an illustration of this process can be seen in Figure-2.

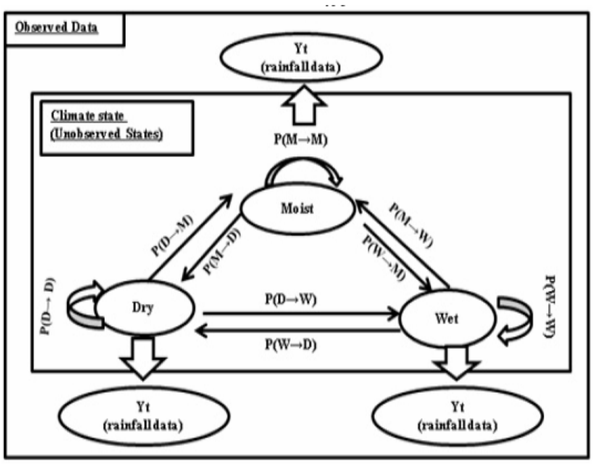

Figure-2. Hidden Markov Model Framework.

Because of climate states are following the markov chain, so the climate states at the time t only depend on the climate state at the time *t-1* or mathematically can be written as:

$$\Pr\left(S_t \middle| S^{(t-1)}\right) = \Pr\left(S_t \middle| S_{t-1}\right), \qquad t = 2,3,... \qquad (3)$$

with $S^{(t)} = \left(S_1, S_2, ..., S_T\right)$

In Zucchini and Guttorp (1991), transition probabilities from the climate states are:

$$P_{ij} = P\left\{S_t = j \middle| S_{t-1} = i\right\} \qquad i, j = \text{Dry,Moist,Wet} \qquad (4)$$

so that, the transition probability matrix of the climate process, we have :

$$\mathbf{P} = \begin{bmatrix} P_{DD} & P_{DM} & P_{DW} \\ P_{MD} & P_{MM} & P_{MW} \\ P_{WD} & P_{WM} & P_{WW} \end{bmatrix} \qquad (5)$$

And the equilibrium probability from each of climate states are:

$$\boldsymbol{\delta} = \left[\delta_D \quad \delta_M \quad \delta_W\right] \qquad (6)$$

with $\delta_D + \delta_M + \delta_W = 1$

The relationship between the transition probability matrix with the equilibrium probability from each of climate states are:

$$\boldsymbol{\delta}\ \mathbf{P} = \boldsymbol{\delta} \qquad (7)$$

Parameters in HMM with Three Hidden Climate States

According to Tyer dan Kuezera (2000) explained that there is a unknown parameters (θ) in the HMM. Parameters (θ) will be consist of: the parameters of each climate states distribution (μ dan σ), the transition probability of the climate states, and climate states sequence $S_N = \left\{s_1, s_2, ..., s_n\right\}$.

So that, its parameters will be consist of:

$$\boldsymbol{\theta} = \left(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\right) \qquad (8)$$

where: $\boldsymbol{\mu} = \left(\mu_D, \mu_M, \mu_W\right)$, $\boldsymbol{\sigma}^2 = \left(\sigma_D^2, \sigma_M^2, \sigma_W^2\right)$, $\mathbf{q}_i$ is the *i*-th row of P.

Parameter Estimation in HMM with Three Hidden Climate States

Lambert at al (2003) were estimating parameters through a bayesian equation, by calculating μ and σ from posterior distribution of **θ**. Based on bayes theorem can be written as:

$$p\left(\theta \middle| Y_n\right) = \frac{P\left(Y_n \middle| \theta\right) \times p\left(\theta\right)}{P\left(Y_n\right)} \qquad (9)$$

where:

$p\left(\theta \mid Y_n\right)$ = conditional on the entire set of observation Y_n

$P\left(Y_n \mid \theta\right)$ = likelihood function of Y_n

$p\left(\theta\right)$ = prior distribution of θ

$P\left(Y_n\right)$ = marginal probability of Y_n

In Chib (1995) likelihood function of Y_n can be written as:

$$\begin{aligned} P\left(y_t \middle| Y_{t-1}, \boldsymbol{\theta}\right) = {} & P\left(z_t = K \middle| Y_{t-1}, \boldsymbol{\theta}\right) \phi\left(y_t \middle| \mu_K, \sigma_K^2\right) + P\left(z_t = L \middle| Y_{t-1}, \boldsymbol{\theta}\right) \\ & \phi\left(y_t \middle| \mu_L, \sigma_L^2\right) + \left(1 - P\left(z_t = K \middle| Y_{t-1}, \boldsymbol{\theta}\right) - P\left(z_t = L \middle| Y_{t-1}, \boldsymbol{\theta}\right)\right) \\ & \phi\left(y_t \middle| \mu_B, \sigma_B^2\right) \end{aligned} \qquad (10)$$

where Y_{t-1} represents all the observation up to time *(t-1)*.

The joint density of all the data is then:

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

$$P(\mathbf{y}|\theta) = \prod_{t=1}^{n} P(y_t | Y_{t-1}, \theta)$$

$$P(\mathbf{y}|\theta) = P(y_1|\theta) \prod_{t=2}^{n} P(y_t | Y_{t-1}, \theta) \quad (11)$$

Prior Distribution of $\boldsymbol{\theta}$ or $p(\theta)$

Chib (1995) define the prior information through the distribution:

$$\mu_j \sim N(\mu_0, A^{-1})$$

$$\sigma_j^2 \sim IG\left(\frac{\upsilon_0}{2}, \frac{\delta_0}{2}\right)$$

$$\mathbf{q} \sim Dirichlet(\alpha_1, \alpha_2, \alpha_3)$$

Posterior Distribution of $\boldsymbol{\theta}$ or $p(\boldsymbol{\theta} | Y_t)$

Chib (1995) define the posterior distribution of $\boldsymbol{\theta}$ through the distribution:

$$\pi(\boldsymbol{\mu} | y, \mathbf{z}, \boldsymbol{\sigma}^2) = \prod_{j=1}^{d} \phi(\mu_j | \hat{\mu}_j, B_j)$$

with $\hat{\mu}_j = (A + \sigma_j^{-2} n_j)^{-1} (A\mu_0 + \sigma_j^{-2} \mathbf{i}_j^T \mathbf{y}_j)$ and $B_j = (A + \sigma_j^{-2} n_j)^{-1}$.

$$\pi(\boldsymbol{\sigma}^2 | y, \mathbf{z}, \boldsymbol{\mu}) = \prod_{j=1}^{d} \phi\left(\sigma_j^2 \left| \frac{\{\upsilon_0 + n_j\}}{2}, \frac{\{\delta_0 + \delta_j\}}{2}\right.\right)$$

with $\delta_j = (\mathbf{y}_j - \mathbf{i}_j \mu_j)^T (\mathbf{y}_j - \mathbf{i}_j \mu_j)$.

$$\pi(\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3} | y, \mathbf{z}) = \prod_{i=1}^{3} \phi(\mathbf{q}_i | \alpha_{i1} + n_{i1}, \alpha_{i2} + n_{i2}, \alpha_{i3} + n_{i3})$$

So that, the posterior density can be computed from the decomposition:

$$\pi(\boldsymbol{\theta}^* | \mathbf{y}) = \pi(\boldsymbol{\mu}^* | \mathbf{y}) \times \pi(\boldsymbol{\sigma}^{2*} | \boldsymbol{\mu}^*, \mathbf{y}) \times \pi(\mathbf{q}^* | \boldsymbol{\sigma}^{2*}, \boldsymbol{\mu}^*, \mathbf{y}) \quad (12)$$

It is difficult when solving bayes equation above, so that required simulation method to obtain the value of parameter $\boldsymbol{\theta}$. Tyer and Kuezera (2000) using Markov Chain Monte Carlo (MCMC) in solving this problem.

MCMC Gibbs Sampler

MCMC Gibbs Sampler aims to find the estimated value of θ_i using a conditional posterior distribution which is assumed that θ_{i-1} has known. The Gibbs sampler algoritm can be summarized by the following steps:

- ❖ set up data
- ❖ set number of iterations (*T*)
- ❖ set prior parameters : $\mu_0, A, \nu_0, \delta_0, \alpha_1, \alpha_2, \alpha_3$
- ❖ initialize vectors of sampled values μ, σ, α
- ❖ set initial current $\mu^{(0)}, \sigma^{(0)}, \alpha^{(0)}$
- ❖ for $t = 1, 2, 3, ..., T$ and $j = 1, 2, 3$ (climate states) repeat:
  - ▪ calculate $\hat{\mu}_j = (A + \sigma_j^{-2} n_j)^{-1} (A\mu_0 + \sigma_j^{-2} \mathbf{i}_j^T \mathbf{y}_j)$
  - ▪ calculate $B_j = (A + \sigma_j^{-2} n_j)^{-1}$
  - ▪ generate $\mu \sim N(\hat{\mu}_j, B_j)$
  - ▪ calculate $a = \frac{\{\upsilon_0 + n_j\}}{2}$
  - ▪ calculate $\delta_j = (\mathbf{y}_j - \mathbf{i}_j \mu_j)^T (\mathbf{y}_j - \mathbf{i}_j \mu_j)$
  - ▪ calculate $b = \frac{\{\delta_0 + \delta_j\}}{2}$
  - ▪ generate $\tau \sim G(a, b)$
  - ▪ calculate $\sigma = \sqrt{1/\tau}$
  - ▪ calculate $e = \alpha_j + n_j$
  - ▪ generate $q \sim Dirichlet(e)$
  - ▪ Set $\mu^{(t)} = \mu, \sigma^{(t)} = \sigma, \alpha^{(t)} = q$
  - ▪ end of for loop
- ❖ make autocorrelation plot, trace plots, and ergodic mean plots to see the convergence of algoritm and determine its burn-in period.
- ❖ calculate average parameters of θ after burn-in period up to the last iteration:

$$\hat{\theta} = \frac{\sum_{t=1}^{m} \theta^{(t)}}{m}$$

- ❖ calculate standard deviation parameters of θ after burn-in period up to the last iteration:

$$s(\theta) = \frac{\sum_{t=1}^{m} (\theta^{(t)} - \hat{\theta})^2}{m-1}$$

- ❖ finish

APPLICATION OF HMM WITH THREE HIDDEN CLIMATE STATES

The data used in this study were daily rainfall data in Darajat Garut from March 1, 2011 until March 27, 2012 and this daily rainfall data expressed in millimeters (mm).

Simulations Results for Parameter μ and σ^2 in Dry Climate State

Initialization value of each prior parameter distribution for dry climate state is:

$$\mu_{dry} \sim N(0.4, 0.25)$$

$$\sigma_{dry}^2 \sim IG(10, 10)$$

and iterations were performed as 5000 times. The results of the convergence simulations can be seen in Figure-3 and Figure-4.

VOL. 9, NO. 9, SEPTEMBER 2014 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

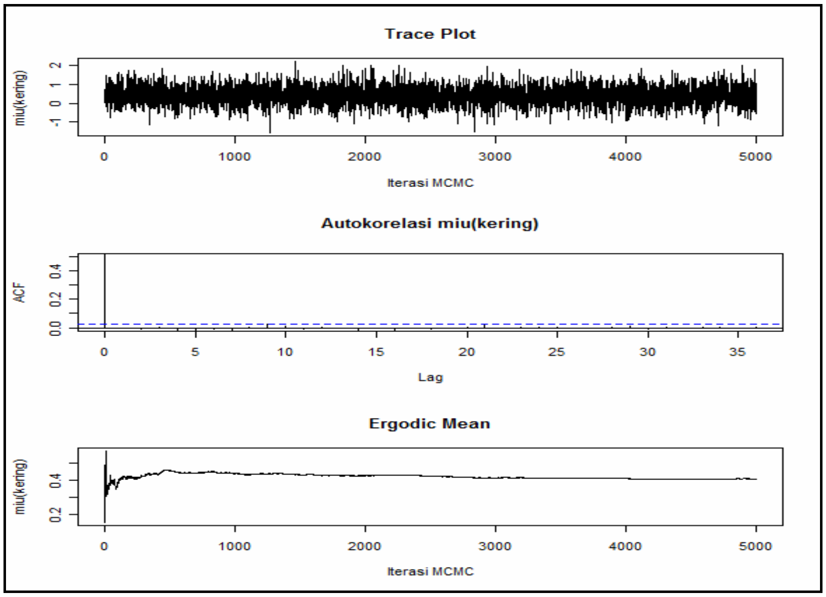

Figure-3. Convergence Simulations for Parameter μ_{dry}

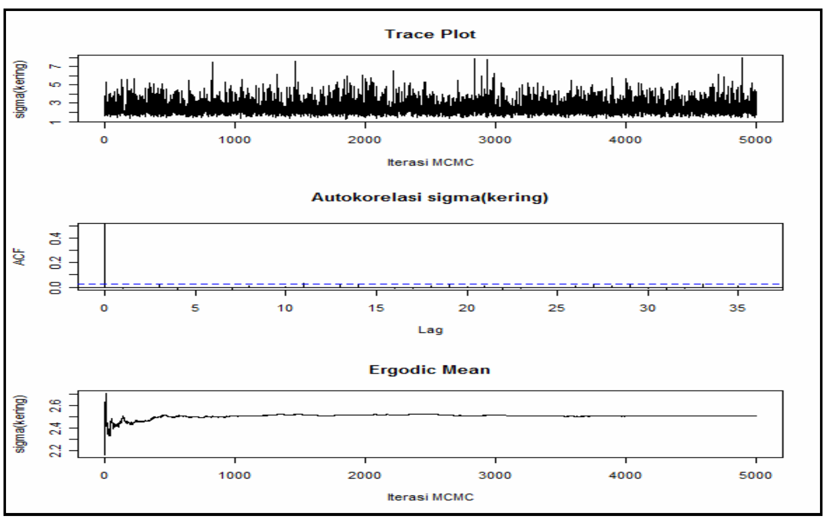

Figure-4. Convergence Simulations for Parameter σ^2_{dry}.

Based on Figure-3 and Figure-4 above, MCMC sampling scheme can be said convergence in terms of stationarity. It was seen from the trace plots of each parameter that indicates the sample parameter values have the same periodesitas and also the ergodic means plot has shown a steady value after the first 1000 times iterations or at this case burn in periode is 1000 times iterations. So that, the results of estimation parameter μ and σ^2 for dry climate state is:

Table-1. Summary of Results for Parameter μ and σ^2 in dry climate state.

	prior		Posterior	
parameter	Mean	stdev	mean	Stdev
μ_{dry}	0.4	0.5	0.399	0.493
σ^2_{dry}	1.111	0.154	2.511	0.774

Same steps performed for each parameter in HMM, so that the result value of the parameter as:

VOL. 9, NO. 9, SEPTEMBER 2014 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-2. Summary of Results for All Parameters in HMM with Three Hidden Climate State.

	prior		**posterior**	
parameter	Mean	stdev	mean	stdev
μ_{dry}	0.4	0.5	0.399	0.493
σ^2_{dry}	1.111	0.154	2.511	0.774
μ_{moist}	2.5	0.3	2.505	0.285
σ^2_{moist}	1.111	0.154	1.078	0.178
μ_{wet}	20	17.7	20.445	16.73477
σ^2_{wet}	1.111	0.154	64.977	21.669
δ_{dry}			0.565	0.0248
δ_{moist}			0.050	0.0112
δ_{wet}			0.3840	0.0244

REFERENCES

[1] Chib S. 1995. Marginal Likelihood from the Gibbs Output. *American Statistical Association.* 90, 1313-1321.

[2] Coe R and Stern R.D. 1982. Fitting Models to Rainfall Data. *Journal of Applied Meteorology*. 21, 1024-1031.

[3] Lambert. M. F., Whiting. J. P. and Metcalfe. A. V. 2003. A Non-Parametric Hidden Markov Model for Climate State Identification. *Hydrology and Earth System Sciences.* 7(5): 652-667.

[4] Rabiner L. R. 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proc. IEEE*. 77, 257-286.

[5] Sudrajat A. 2009. Pemetaan Klasifikasi Iklim Oldeman dan Schmidth-Fergusson Sebagai Upaya Pemanfaatan Sumber Daya Iklim Dalam Pengelolaan Sumberdaya Alam Di Sumatera Utara. Universitas Sumatera Utara.

[6] Taylor H.M dan Karlin S. 1994. *An Introduction to Stochastic Modeling*. 3rd Edition. Academic Press,Inc., New York.

[7] Thyer M. A. and Kuezera G. A. 2000. Modelling Long-Term Persistence in Hydrological Time Series Using a Hidden State Markov Model. *Water Resources Research*. 36, 3301-3310.

[8] Visser I. 2011. Seven Things to Remember about Hidden Markov Models: A Tutorial on Markovian Models for Time Series. *Journal of Mathematical Prychology*, 55, 403-415.

[9] Zucchini W. and Guttorp P. 1991. A Hidden Markov Model for Space Time Precipitation. *Water Resources Research*. 27, 1917-1923.

[10] Zucchini W. and MacDonald I.L. 2009. *Hidden Markov Models for Time Series An Introduction Using R*. CRC Press. New York.