Sains Malaysiana 41(11)(2012): 1345–1353

# Application of Loglinear Models in Estimating Wet Category in Monthly Rainfall

(Penggunaan Model Loglinear dalam Penganggaran Kategori Basah Hujan Bulanan)

WAHIDAH SANUSI* & KAMARULZAMAN IBRAHIM

ABSTRACT

*Climate changes have become serious issues that have been widely discussed by researchers. One of the issues concerns with the study in changes of rainfall patterns. Changes in rainfall patterns affect the dryness and wetness conditions of a region. In this study, the three-dimensional loglinear model was used to fit the observed frequencies and to model the expected frequencies of wet class transition on eight rainfall stations in Peninsular Malaysia. The expected frequency values could be employed to determine the odds value of wet classes of each station. Further, the odds values were used to estimate the wet class of the following month if the wet class of the previous month and current month were identified. The wet classification based on SPI index (Standardized Precipitation Index). For station that was analyzed, there was no difference found were between estimated and observed wet classes. It was concluded that the loglinear models can be used to estimate the wetness classes through the estimates of odds values.*

*Keywords: Loglinear models; odds; Standardized Precipitation Index (SPI); wet classification*

ABSTRAK

*Perubahan iklim merupakan isu yang banyak diperbincangkan oleh penyelidik. Salah satunya ialah tentang kajian perubahan corak hujan. Perubahan corak hujan membawa kesan terhadap keadaan kering ataupun basah sesebuah rantau. Dalam kajian ini digunakan model loglinear tiga dimensi untuk menyuaikan kekerapan dicerap dan untuk memodelkan kekerapan dijangka peralihan kelas basah pada lapan stesen hujan di Semenanjung Malaysia. Nilai kekerapan dijangka dapat digunakan untuk menentukan nilai kemungkinan kelas basah setiap stesen. Selanjutnya, anggaran nilai kemungkinan yang telah diperoleh dapat digunakan untuk menganggar kelas basah satu bulan ke hadapan, jika diketahui kelas basah bulan sebelum dan bulan semasa. Pengelasan basah yang digunakan adalah berdasarkan indeks SPI (indeks hujan dipiawai). Bagi stesen hujan yang dianalisis, hasil bandingan antara anggaran kelas basah dengan cerapan didapati tidak ada perbezaan. Hasil kajian ini memperlihatkan bahawa model loglinear dapat digunakan untuk menganggar kelas kebasahan melalui anggaran nilai kemungkinan.*

*Kata kunci: Indeks Hujan Dipiawai (SPI); kemungkinan; model loglinear; pengelasan basah*

## INTRODUCTION

Weather and climate conditions tend to have varying impacts on human life. One of the important indicators of climate change is rainfall pattern. Changes in rainfall may lead to flood or drought. This can cause problems for the survival of humans and other living creatures. The impacts can lead to the declining in agricultural production, health risks, water availability and forest fires. These can cause enormous losses both in material and life loss that needs serious attention from various parties. One thing can be done to anticipate the impact of climate change is through weather estimating techniques. Paulo and Pereira (2007) predicted the drought class transition through a stochastic approach. Similarly, Cancelliere et al. (2007) predicted drought using the standardized precipitation index. Meanwhile, Cai (2010) estimated the wet and low water of precipitation with the weighted Markov chain methods. The estimated results are very useful as to provide early warning for the community regarding the presence of a disaster so that the losses due to climate change can be minimized (Heriah 2007).

Contingency tables and loglinear models are statistical methods that can be applied in qualitative data (Agresti 2002). Through contingency tables, the relationship between qualitative variables can be determined, while the loglinear analysis can be used to determine the risks or effects of each variable category to the other variables. Loglinear models can be used to describe the pattern of relationships among several categorical variables (Agresti 2002). In this study, the three-dimensional loglinear models were used. The goals were to fit the observed frequencies and to model the expected frequencies transition of wet categories. Wet category is obtained by means of SPI index, i.e. the standardized precipitation index (Cancelliere et al. 2007). Moreira et al. (2006, 2008) and Paulo et al. (2005) have shown that the loglinear model is an adequate tool for predicting the SPI drought class transitions. SPI index was first developed by McKee et al. (1993) for detecting and monitoring drought. The advantages of this index is only based on precipitation data (Mishra & Desai 2005). SPI has been widely used by researchers to study about the drought in their region (Cancelliere et al. 2007; Mishra & Desai 2005; Turkes & Tath 2009). SPI has also been used to detect rainfall patterns in northeast Spain (Lana et al. 2001).

The objective of this study was to estimate the wet category for one-month ahead, if the wet category for the current month and one-month before it are given. This estimation is obtained by the odds values of the best loglinear models.

The paper is structured as follows. First, a brief description of precipitation data for all stations are given in Data and Methods section. The explanation of how to get the SPI values of the precipitation data and determining the SPI wet categories according to the SPI index are given. This section also presents the best loglinear model selection, the odds estimation and confidence intervals for odds. In the Results and Discussion section, the best loglinear model for each station is obtained and the results of analysis for the selected station is provided. In the final section, conclusion and suggestion for future research are stated.

## DATA AND METHODS

### DATA

For this study, the monthly precipitation (mm) were used in which data were gathered from eight rainfall stations in Peninsular Malaysia (Figure 1). Selection of the station was based on the completeness of the data for the period from February 1974 to January 2005. The data were obtained from the Drainage and Irrigation Department and the Malaysian Meteorological Department. They were collected using the automatic and manual rain gauges. Table 1 denotes the geographical coordinates of the stations and the statistical properties of the annual precipitation data. This table indicates that Chin-Chin station has the lowest annual precipitation mean (1667.74 mm) and Gombak station has the highest annual precipitation mean (23994.42 mm). Figure 2 shows that the stations in Ampang, Alor Setar, Chin-Chin and Sg. Pinang have the similar monthly mean rainfall. Likewise, Gombak, Johor Bahru and Kg. Sg. Tua stations also have the same monthly mean rainfall. Meanwhile, the monthly mean rainfall values of Johol station lies between the monthly mean rainfall values of the two groups of stations.

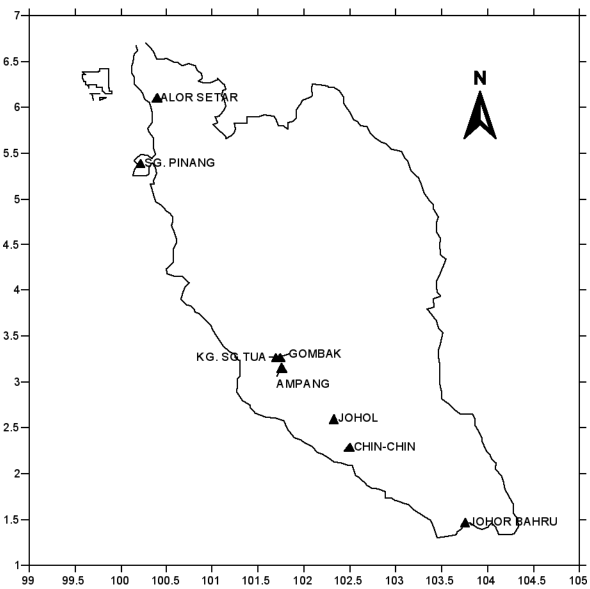


FIGURE 1. The location of the stations

TABLE 1. Rain gauge stations, geographic coordinates and statistical properties of annual rainfall

| No. | Rain gauges stations | Latitude (N) | Longitude (E) | Min (mm) | Max (mm) | Mean (mm) |
|---|---|---|---|---|---|---|
| 1 | Ampang | 03 09 20 | 101 45 00 | 1476.10 | 3138.80 | 2403.06 |
| 2 | Alor Setar | 06 06 20 | 100 23 30 | 1452.60 | 2597.60 | 1903.76 |
| 3 | Chin-Chin | 02 17 20 | 102 29 30 | 1116.90 | 2078.80 | 1667.74 |
| 4 | Gombak | 03 16 05 | 101 43 45 | 17613.00 | 28312.00 | 23994.42 |
| 5 | Johol | 02 36 10 | 102 19 10 | 12344.00 | 22261.00 | 16854.87 |
| 6 | Johor Bahru | 01 28 15 | 103 45 10 | 14915.00 | 30621.00 | 23252.32 |
| 7 | Kg. Sg. Tua | 03 16 20 | 101 41 10 | 14725.00 | 29341.00 | 23450.07 |
| 8 | Sg. Pinang | 05 23 30 | 100 12 45 | 1641.90 | 3547.00 | 2599.68 |

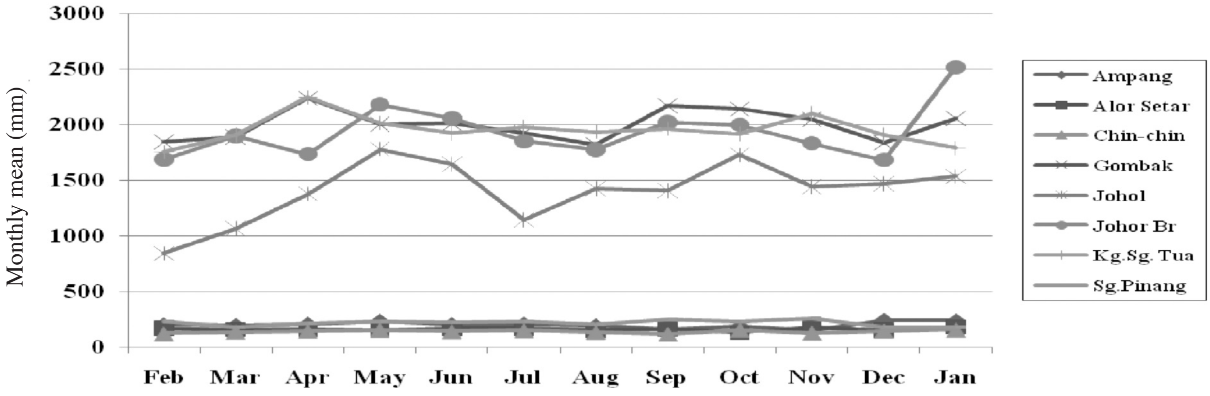


FIGURE 2. Monthly mean rainfall for all stations (1974-2005)

## METHODS

### STANDARDIZED PRECIPITATION INDEX (SPI)

Before applying the loglinear modeling, the calculation of the SPI is performed. SPI is calculated based on the precipitation that has been summed over the time scale of 3 months for each station. The SPI 3-month time scale is a short time scale which reflects the seasonality of the data and is more appropriate to identify the wet/drought impact on agriculture (Labedzki 2007; Moreira et al. 2008; Sene 2010). The cumulative distribution function of precipitation totals are formed from the gamma distribution. Then each probability density function is transformed into a standardized normal distribution by using inverse standard normal distribution (Durdu 2010; Turkes & Tath 2009). The gamma distribution has the probability density function as follows:

$$\mathrm{f}(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{x/\beta}, \; x > 0, \tag{1}$$

where, $\alpha > 0$ is a shape parameter, $\beta > 0$ is a scale parameter, and $x$ is the amount of precipitation. The gamma function $\Gamma(\alpha)$ is defined by:

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy, \; y > 0. \tag{2}$$

Edwards and McKee (1997) suggest to estimate the $\alpha$ dan $\beta$ parameters through the maximum likelihood estimation using the approximation of Thom (1958) as follows:

$$\hat{\alpha} = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right). \tag{3}$$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}}, \tag{4}$$

where for $n$ observations:

$$A = \ln(\bar{x}) - \frac{\sum \ln(x)}{n}, \tag{5}$$

$n$ is number of precipitation observations, and $\bar{x}$ is mean precipitation over the time scale of 3 months.

The resulting parameters were then used to determine the cumulative probability of an observed precipitation:

$$F(x) = \int_0^x \mathrm{f}(u)\, du = \frac{1}{\hat{\beta}^{\hat{\alpha}}\Gamma(\hat{\alpha})} \int_0^x u^{\hat{\alpha}-1} e^{-u/\hat{\beta}} du. \tag{6}$$

Since the gamma distribution is not defined for $x = 0$ and a precipitation distribution may contain zero, then the cumulative probability distribution function becomes:

$$H(x) = q + (1 - q)F(x), \tag{7}$$

where $q = \mathrm{Prob}(x = 0) > 0$, $\mathrm{Prob}(x = 0)$ is the probability of zero precipitation. The cumulative probability distribution $H(x)$ is transformed into the standard normal random variable Z with mean zero and variance one. It is the value of SPI. For the calculation of SPI, the approximation provided by Abramowitz and Stegun (1970) can be used as an alternative:

$$SPI = -\left(w - \frac{c_0 + c_1 w + c_2 w^2}{1 + d_1 w + d_2 w^2 + d_3 w^3}\right), \text{ for } 0 < H(x) \le 0.5, \tag{8}$$

$$SPI = \left(w - \frac{c_0 + c_1 w + c_2 w^2}{1 + d_1 w + d_2 w^2 + d_3 w^3}\right), \text{ for } 0.5 < H(x) < 1, \tag{9}$$

where

$$w = \sqrt{\ln\left[\frac{1}{\left(H(x)\right)^2}\right]}, \text{ for } 0 < H(x) \le 0.5, \tag{10}$$

$$w = \sqrt{\ln\left[\frac{1}{\left(1 - H(x)\right)^2}\right]}, \text{ for } 0.5 < H(x) < 1, \tag{11}$$

and

$c_0 = 2.515517$, $c_1 = 0.802853$, $c_2 = 0.010328$

$d_1 = 1.432788$, $d_2 = 0.189269$, $d_3 = 0.001308$.

After that the SPI values were categorized based on the classification of wet periods.

TABLE 2. Wet classification based on the SPI index (modified from Cancelliere et al. 2007)

| Code | SPI values | Wet classes |
|---|---|---|
| 1 | SPI ≤ 0 | Non wet (NW) |
| 2 | 0 < SPI <1 | Near normal (NN) |
| 3 | 1 ≤ SPI < 1.5 | Moderate (M) |
| 4 | SPI ≤ 1.5 | Very wet (VW) |

Cancelliere et al. (2007) distinguished the classification between the extreme category (SPI ≥ 2) and very wet category (1.5 ≤ SPI <2), while in this study the extreme and very wet categories are grouped.

## LOGLINEAR MODELS

The data that has been formed in the contingency tables can be analyzed using loglinear models (Agung 2001). Loglinear models used in this study is three-dimensional loglinear models (Agresti 2002; Moreira et al. 2006, 2008; Paulo et al. 2005) which is defined by:

$$\begin{aligned} \text{Log}\, E_{ijk} = \;& \lambda + \lambda_i^A + \lambda_J^B + \lambda_K^C + \beta_0 u_i v_i + \beta_1 u_i w_k \\ & + \beta_2 v_j w_k + \beta_3 u_i v_j w_{l'} + \eta_{1i} I(i=j) \\ & + \eta_{2i} I(i=k) + \eta_{3i} I(j=k) + \eta_{4i} I(i=j=k), \end{aligned} \tag{12}$$

where $E_{ijk}$ is expected frequencies for each cell (i, j, k), A is wet classes at month (t-1) with levels i, i = 1, 2, 3, 4. B is wet classes at month t with levels j, j = 1, 2, 3, 4. $C$ is wet classes at month (t+1) with levels k, k = 1, 2, 3, 4. The first category (1) refers to non wet, the second category (2) refer to near normal, the third category (3) refer to moderate and the fourth category (4) refer to very wet. $\lambda$ is total mean, $\lambda_i^A$ is the parameter associated with $i$th level of the category $A$, $\lambda_J^B$ is the parameter associated with $j$th level of the category B, $\lambda_K$ is the parameter associated with $k$th level of the category $C$, $u_i$ is the $i$th level score of category A, $v_j$ is the $j$th level score of category $B$ and $w_k$ is the $k$th level score of category $C$. $\beta_0, \beta_1$, $\beta_2$ and $\beta_3$ are the linear association parameters. $\eta_{1i}$, $\eta_{2i}$ and $\eta_{4i}$ are parameters associated to the $i$th diagonal element of category $A$ and $\eta_{3j}$ is associated with the $j$th diagonal element of category $B$. The indicator function I(.) (Agresti 2002; Paulo et al. 2005) is defined by:

$$\text{I}(.) = \begin{cases} 1, \text{ if true} \\ 0, \text{ if false} \end{cases}. \tag{13}$$

The parameters estimation of loglinear models are obtained by the maximum likelihood methods, while the goodness of fit test based on the residual deviance value (Agresti 2002; Moreira et al. 2006, 2008) which is defined by:

$$D = 2\, \Sigma_i \Sigma_j \Sigma_k\, O_{ijk} \log(O_{ijk}/E_{ijk})\,, \tag{14}$$

where $O_{ijk}$ denote observed frequencies of transitions between wet class i at month (t-1), wet class j at month t and wet class k at month (t+1). $D$ has the chi-square distribution with degree of freedom equal to the number of cells in contingency table minus the number of linearly independent estimated model parameters (Moreira et al. 2006, 2008). The best model is selected by backward elimination methods (Agresti 2002; Kuruppumullage & Sooriyarachchi 2007; Moreira et al. 2006, 2008; Paulo et al. 2005). Backward elimination begins with a complex model and sequentially removes terms. The reason is that it is safer to delete terms from an overly complex model than to add term to an overly simple one (Agresti 2002).

To estimate the wet category we used the odds estimate. An odds is a ratio of expected frequencies. The odds for the three-dimension models (Moreira et al. 2006, 2008) were defined as:

$$\theta_{kl\backslash ij} = \frac{E_{ijk}}{E_{ijl}},\; k \neq l,\, (k, l = 1, 2, 3, 4). \tag{15}$$

The odds $\theta_{kl\backslash ij}$ means that, one month ahead, it is $\theta_{kl\backslash ij}$ times more, less, or equally probable that a specific station is in class k instead of class l, given that at present it is in class j and one month before it was in class i. For large sample, $\theta_{kl\backslash ij}$ has the asymptotic normal distribution and log ($\theta_{kl\backslash ij}$) also converges to a normal distribution. The confidence intervals for log ($\theta_{kl\backslash ij}$) with a probability $(1-\alpha)$ is given by:

$$\begin{aligned} &\Big[ Log\left(\theta_{kl\backslash ij}\right) - z_{1-\alpha/2}\sqrt{\text{Var}\left(Log\left(\theta_{kl\backslash ij}\right)\right)}, \\ &Log\left(\theta_{kl\backslash ij}\right) + z_{1-\alpha/2}\sqrt{Var\left(Log\left(\theta_{kl\backslash ij}\right)\right)}\Big], \end{aligned} \tag{16}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of a standard normal variable (Moreira et al. 2006, 2008).

## RESULTS AND DISCUSSION

Table 3 and Figure 3 indicate that there are almost no difference of the wet classes frequencies between one station to another. However, the Johor Bahru station has the highest very wet class frequency (28), while the Johol station has the lowest very wet class frequency (10). Figure 4 shows that the wet classes have equal behavior of frequencies for all stations. The frequencies of wet classes decreased with the degree of severity. These results seem similar to the findings of Paulo et al. (2005) and Paulo and Pereira (2007), but for the drought cases.

TABLE 3. The frequencies of wet classes for all stations

| Rain gauge stations | Wet classes | | | | Rain gauge stations | Wet classes | | | |
|---|---|---|---|---|---|---|---|---|---|
| | NW | NN | M | VW | | NW | NN | M | VW |
| Ampang | 171 | 143 | 39 | 19 | Johol | 168 | 159 | 35 | 10 |
| Alor Setar | 175 | 139 | 37 | 21 | Johor. Br | 187 | 117 | 40 | 28 |
| Chin-Chin | 174 | 138 | 37 | 23 | Kg. Sg. Tua | 184 | 126 | 40 | 22 |
| Gombak | 173 | 140 | 38 | 21 | Sg. Pinang | 171 | 153 | 36 | 12 |

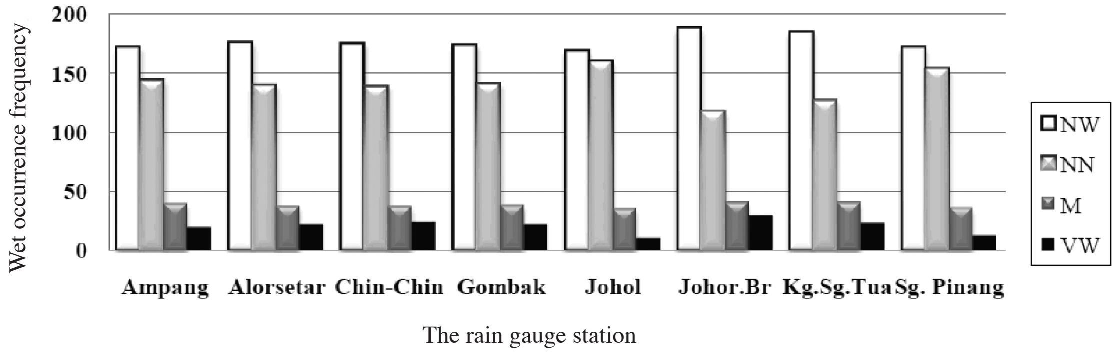


FIGURE 3. Frequencies of wet classes according to the SPI values on a 3-month scale time for each station

For each station, the selection of the loglinear models was conducted through the backward elimination methods and the *p*-value exceeding $\alpha = 0.05$ (Table 4). This table shows that Sg. Pinang station has the complete loglinear models with the lowest residual deviance value (35.94) compared with the residual deviance of the stations at Alor Setar (59.94), Chin-Chin (47.70) and Kg. Sg. Tua (53.19). Therefore, this station was chosen as an example in this paper. The estimation of the model parameters for Sg. Pinang station (Table 5) were obtained using maximum likelihood methods. The results showed that the interaction and main factor were significant with significance level ($\alpha$) is 0.05, except the total mean factor ($\lambda$) was not significant.

Based on the estimation of the model parameters for Sg. Pinang station, we can obtain the expected frequencies as presented in Table 6. Figures 5(a), 5(b) and Table 6 exhibit that there are similarities between observed and expected frequencies. The highest expected frequencies values occurred for transitions NW→NW→NW and

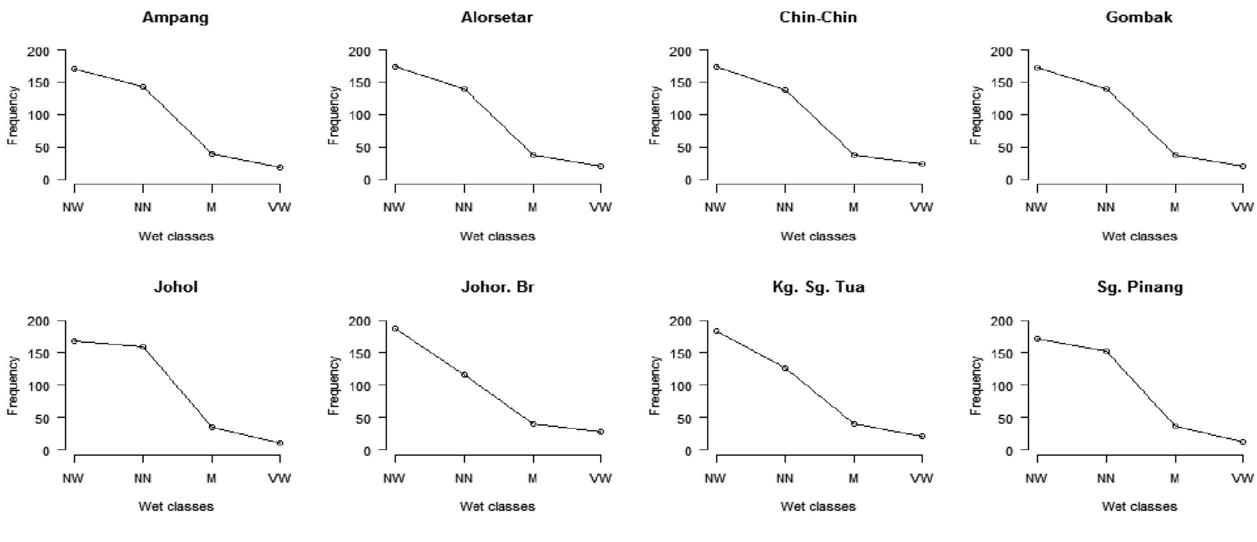


FIGURE 4. Frequencies of the wet classes for all stations

TABLE 4. Selected loglinear models, residual deviance, degrees of freedom and $p$-values

| Stations | Selected models | Res. Dev | df | $p$-values |
|---|---|---|---|---|
| Johol | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_2 v_j w_k$ | 53.10 | 52 | 0.43 |
| Ampang | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_2 v_j w_k + \beta_3 u_i v_j w_k$ | 64.79 | 51 | 0.09 |
| Alor Setar | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_1 u_i w_k + \beta_2 v_j w_k + \beta_3 u_i v_j w_k$ | 59.94 | 50 | 0.16 |
| Chin-Chin | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_1 u_i w_k + \beta_2 v_j w_k + \beta_3 u_i v_j w_k$ | 47.70 | 50 | 0.57 |
| Gombak | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_2 v_j w_k + \beta_3 u_i v_j w_k$ | 40.93 | 51 | 0.84 |
| Johor Br | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_2 v_j w_k + \beta_3 u_i v_j w_k$ | 32.88 | 51 | 0.98 |
| Kg. Sg. Tua | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_1 u_i w_k + \beta_2 v_j w_k + \beta_3 u_i v_j w_k$ | 53.19 | 50 | 0.35 |
| Sg. Pinang | $Log\ E_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \beta_0 u_i v_j + \beta_1 u_i w_k + \beta_2 v_j w_k + \beta_3 u_i v_j w_k$ | 35.94 | 50 | 0.93 |

NN→NN→NN. The two consecutive transitions showed that the non-wet and near normal classes have the self-perpetuating characteristic. Meanwhile, the lowest expected frequencies values (near zero) occurred for transitions NW→VW→NW and VW→NW→VW, as shown in Table 6. These transitions show that they have a very low probability to occur, because the wet occurrence did not initiate or dissipate suddenly.

On the other hand, Figures 5(c) and 5(d) indicated the difference between observed and expected frequencies for moderate class and very wet class, respectively. The expected frequencies for each cell are not zero. The estimate of odds for Sg Pinang station can be calculated (Table 7).

Interpretation of the odds estimation and confidence intervals expressed the following, for instance, the estimate of odds $\theta_{\text{NN:M|M:VM}} = 40.8$ with the confidence interval [7.29; 228.18], because the value 1 was not included in the interval. For example in December 1998 wet class was moderate (M) and January 1999 was very wet (VW), then on February 1999 it is more likely to change to the near normal (NN) category 40.8 times than to the moderate category (M). For the estimate of odd $\theta_{\text{NN:M|M:NN}} = 1.29$ with confidence interval [0.43, 3.85]. The confidence interval includes the value 1, this means that, if given wet class on April 1999 was moderate (M) and May 1999 was near normal (NN), then June 1999, it is likely wet class will remain at the near normal category (NN) or changes to the moderate category (M).

The estimates of odds obtained can be used to predict wet classes one month ahead (t+1), if wet class at last month (t-1) and this month t are given. Table 8 displays the comparison between observed (actual) and estimated wet classes 1 month ahead as validation of loglinear model for Sg. Pinang station. The results show that between the estimated classes and the observed classes are almost

TABLE 5. The parameter estimation for models, standard error, z-values and Prob(>|z|) for Sg. Pinang Station

| Parameters | Estimate | Std. error | z-values | Prob(>\|z\|) |
|---|---|---|---|---|
| $\lambda$ | -0.72 | 0.55 | -1.31 | 0.19 |
| $\lambda_i^A$ ($i$ = NN) | -4.33 | 0.50 | -8.59 | <2.00E-16 |
| $\lambda_i^A$ ($i$ = M) | -11.09 | 1.11 | -9.99 | <2.00E-16 |
| $\lambda_i^A$ ($i$ = VW) | -18.37 | 1.78 | -10.35 | <2.00E-16 |
| $\lambda_j^B$($j$ = NN) | -5.95 | 0.56 | -10.68 | <2.00E-16 |
| $\lambda_j^B$($j$ = M) | -15.13 | 1.30 | -11.66 | <2.00E-16 |
| $\lambda_j^B$($j$ = VW) | -25.44 | 2.13 | -11.93 | <2.00E-16 |
| $\lambda_k^C$($k$ = NN) | -4.36 | 0.51 | -8.64 | <2.00E-16 |
| $\lambda_k^C$($k$ = M) | -11.15 | 1.11 | -10.02 | <2.00E-16 |
| $\lambda_k^C$($k$ = VW) | -18.47 | 1.78 | -10.38 | <2.00E-16 |
| $\beta_0$ | 2.24 | 0.26 | 8.66 | <2.00E-16 |
| $\beta_1$ | 1.24 | 0.27 | 4.65 | 3.36E-06 |
| $\beta_2$ | 2.26 | 0.26 | 8.70 | <2.00E-16 |
| $\beta_3$ | -0.45 | 0.10 | -4.56 | 5.09E-06 |

TABLE 6. Observed versus expected frequencies of wet class transitions from month t-1 to month t to month t+1 for Sg. Pinang Station

| Wet class at t-1 | Wet class at t | Observed frequencies | | | | Expected frequencies | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Wet class at t+1 | | | | Wet class at t+1 | | | |
| | | NW | NN | M | VW | NW | NN | M | VW |
| NW | NW | 95 | 29 | 0 | 0 | 96.70 | 25.93 | 0.62 | 0.01 |
| | NN | 16 | 21 | 6 | 0 | 14.41 | 23.42 | 3.38 | 0.29 |
| | M | 0 | 0 | 0 | 0 | 0.09 | 0.84 | 0.73 | 0.38 |
| | VW | 0 | 0 | 0 | 0 | 0.00 | 0.01 | 0.05 | 0.16 |
| NN | NW | 30 | 14 | 0 | 0 | 26.55 | 15.72 | 0.83 | 0.03 |
| | NN | 22 | 53 | 9 | 3 | 23.73 | 54.18 | 10.97 | 1.30 |
| | M | 0 | 10 | 6 | 2 | 0.84 | 7.39 | 5.77 | 2.64 |
| | VW | 0 | 1 | 2 | 1 | 0.01 | 0.33 | 0.98 | 1.73 |
| M | NW | 0 | 0 | 0 | 0 | 0.64 | 0.84 | 0.10 | 0.01 |
| | NN | 4 | 10 | 2 | 1 | 3.43 | 10.98 | 3.12 | 0.52 |
| | M | 0 | 5 | 6 | 3 | 0.73 | 5.72 | 3.99 | 1.63 |
| | VW | 0 | 2 | 2 | 1 | 0.05 | 0.96 | 1.64 | 1.65 |
| VW | NW | 0 | 0 | 0 | 0 | 0.01 | 0.03 | 0.01 | 0.00 |
| | NN | 1 | 3 | 1 | 0 | 0.30 | 1.33 | 0.53 | 0.12 |
| | M | 0 | 2 | 2 | 0 | 0.38 | 2.64 | 1.64 | 0.60 |
| | VW | 0 | 2 | 0 | 1 | 0.15 | 1.69 | 1.64 | 0.94 |

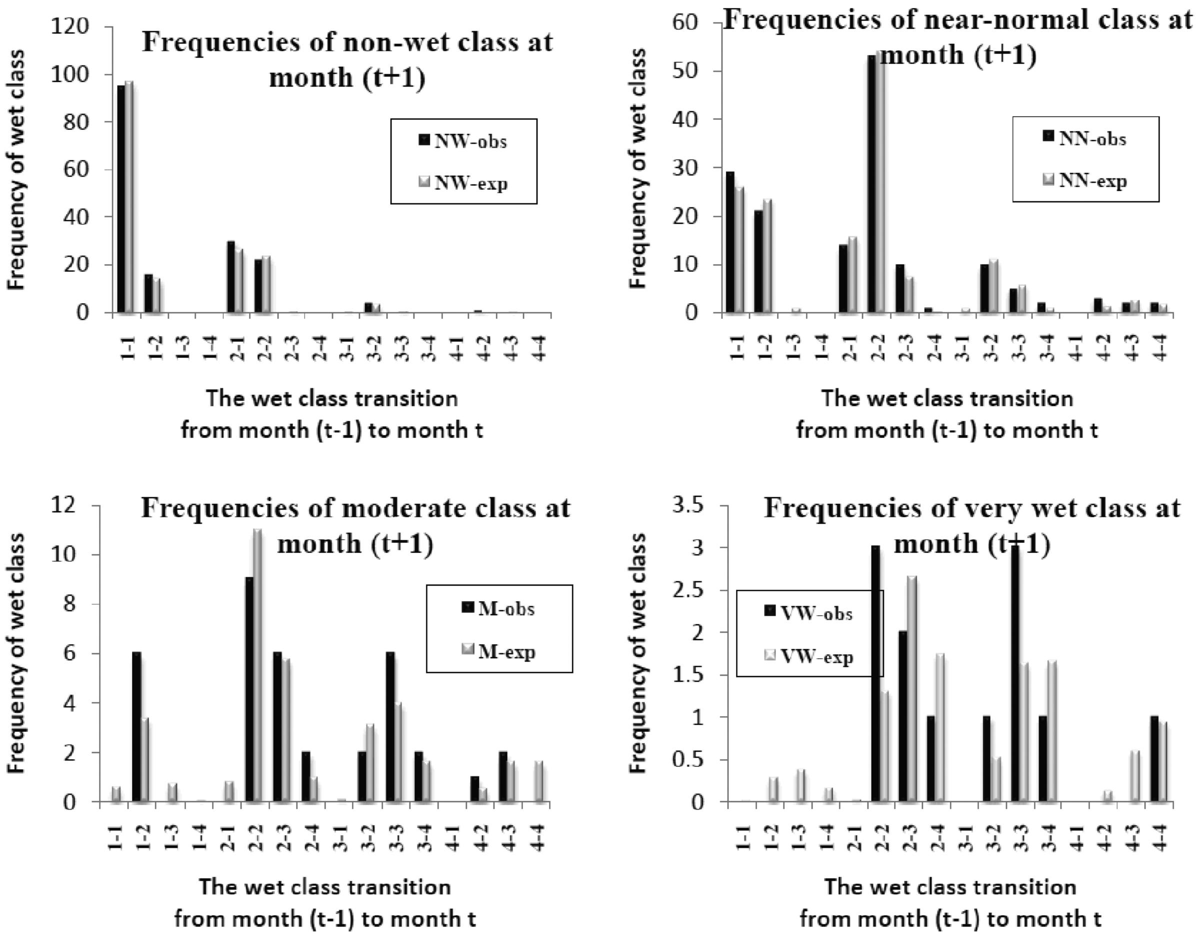


FIGURE 5. The observed and expected frequencies for two consecutive transitions between wet classes for Sg. Pinang station, 1: non-wet; 2: near-normal; 3: moderate and 4: very wet

TABLE 7. The estimates of odds $\theta_{NN:Mlij}$ and correspondent confidence intervals for Sg. Pinang station referring to estimates for month t+1, when the wet classes are known for month t-1 and t

| Wet class at t-1 | Odds estimates | | | | Lower bound | | | | Upper bound | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Wet class at t | | | | Wet class at t | | | | Wet class at t | | | |
| | NW | NN | M | VW | NW | NN | M | VW | NW | NN | M | VW |
| NW | 0.36 | 0.11 | 0.09 | 2.13 | 0.00 | 0.01 | 0.00 | **1.04** | 1.05E+10 | 1.09 | 15.47 | **4.37** |
| NN | 0.23 | 997.23 | 2.3 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 8 | 1.91E+08 | 1808.11 | 2.04E+06 |
| M | 3.57 | 1.29 | 1.05 | 40.8 | 0.00 | 0.43 | 0.12 | **7.29** | 1.16E+12 | 3.85 | 8.86 | **228.18** |
| VW | 4.41 | 32.12 | 4.2 | 0.57 | 0.27 | 0.00 | 0.01 | 0.05 | 72.7 | 7.37E+06 | 2056.05 | 6.95 |

(In bold the cases when the confidence intervals are not include the value 1)

TABLE 8. The comparison between observed and estimated wet classes 1 month ahead (t+1)

| Year | Month t | Wet class at t-1 | Wet class at t | Fitted Wet class at t+1 | Obs. Wet class at t+1 |
|---|---|---|---|---|---|
| 1999 | Jan | M | VW | NN | M |
| | Feb | VW | M | NN or M | M |
| | Mar | M | M | NN or M | M |
| | Apr | M | M | NN or M | NN |
| | May | M | NN | NN or M | NN |
| | Jun | NN | NN | NN or M | NN |
| | Jul | NN | NN | NN or M | NN |
| | Aug | NN | NN | NN or M | M |
| | Sep | NN | M | NN or M | NN |
| | Oct | M | NN | NN or M | NN |
| | Nov | NN | NN | NN or M | NN |
| | Dec | NN | NN | NN or M | M |

equal. NN or M in Table 8 indicates that the probabilities for transitions into the classes NN or M are similar.

## CONCLUSION

The three-dimensional loglinear models have been used to determine the expected frequencies of wet classes of the two consecutive transitions. Classification of wet categories was based on the SPI precipitation data. The results of the expected frequencies showed that the non-wet category and the near-normal category have self-perpetuating characteristic, for station that was analyzed. However, the transition from the non-wet category to very wet and then moved to the non-wet category, and the transition from the category of very wet to non-wet, then move into the very wet category has a very low probability to occur. Meanwhile, the observed frequencies of the wet classes have the same behavior, that they decreased with the degree of severity.

It was found that the loglinear models can be used to estimate wet classes through the odds estimation. Similarly, the study by Moreira et al. (2006, 2008) showed that the loglinear models were powerful tools to predict drought class transitions.

Further analysis could be conducted on other rainfall stations in Peninsular Malaysia to obtain information about other areas wet classes. In the future, the use of loglinear models in estimating drought category in Peninsular Malaysia can be analyzed.

## ACKNOWLEDGEMENT

We are grateful to two anonymous reviewers for their helpful comments that substantially improved the manuscript. We thank Prof. Dr. Abdul Aziz Jemain for his useful suggestions. We would also like to acknowledge the Research Centre for Tropical Climate Change System (IKLIM) Faculty of Science and Technology, Universiti Kebangsaan Malaysia and the goverment of South Sulawesi province for the financial support.

## REFERENCES

Abramowitz, M. & Stegun, I.A. 1970. Handbook of mathematical functions with formulas, graphs and mathematical tables. 9th ed. In *Development of the SPI Drought Index for Greece using Geo-Statistical Methods*. Republic of Macedonia: Balwois, edited by Chortaria, C., Karavitis, C.A. & Alexandris, S.

Agresti, A. 2002. *Categorical Data Analysis*. New York: John Wiley & Sons.

Agung, I.G.N. 2001. *Statistika: Analisis Hubungan Kausal Berdasarkan Data Kategorik*. Jakarta: PT. RajaGrafindo Persada.

Cai, M. 2010. Study on variation in wet and low water of precipitation prediction based on Markov with weights theory. 2010 Sixth International Conference on Natural Computation (ICNC). *IIEEE Journal* 8: 4296-4300.

Cancelliere, A., Mauro, G.D., Bonaccorso, B. & Rossi, G. 2007. Drought forecasting using the Standardized Precipitation Index. *Journal of Water Resources Management* 21: 801-819.

Durdu, F.O. 2010. Application of linear stochastic models for drought forecasting in the Buyuk Menderes river basin, western Turkey. *Journal of Stoch. Environ. Res. Risk. Assess.* 24: 1145-1162.

Edwards, D.C. & McKee, T.B. 1997. Characteristics of 20th century drought in the United States at multiple time scales. *Journal of Water Resour. Manage*. 23: 881-897.

Heriah, K. 2007. *Perubahan Iklim Gobal: Dampak dan Bahayanya*. Jurusan Tanah, Fakultas Pertanian, Universitas Brawijaya, Malang.

Kuruppumullage, P. & Sooriyarachchi, R. 2007. Log-linear models for ordinal multidimensional categorical data. *Journal of the National Science Foundation of Sri Lanka* 35(1): 29-40.

Labedzki. 2007. Estimation of lokal drought frequency in central Poland using the Standardized Precipitation Index SPI. *Irrigation and Drainage* 56: 67-77.

Lana, X., Serra, C. & Burgueno A. 2001. Patterns of monthly rainfall shortage and excess in terms of the standardized precipitation index. *International Journal of Climatology* 21: 1669-1691.

McKee, T.B., Doesken, N.J. & Kleist, J. 1993. The relationship of drought frequency and duration to time scale. In *Proceeding of the Ninth Conference on Applied Climatology,* Boston: American Meteorological Society.

Mishra, A.K. & Desai, V.R. 2005. Drought forecasting using stochastic models. *Stoch. Environ. Res. Risk. Assess.* 19: 326-339.

Moreira, E.E., Paulo, A.A., Pereira, L.S. & Mexia, J.T. 2006. Analysis of SPI drought class transitions using loglinear models. *Journal of Hydrology* 331: 349-359.

Moreira, E.E., Coelho, C.A., Paulo, A.A., Pereira, L.S. & Mexia, J.T. 2008. SPI-based drought category prediction using loglinear models. *Journal of Hydrology* 354: 116-130.

Paulo, A.A., Ferreira, E., Coelho, C. & Pereira, L.S. 2005. Drought class transition analysis through Markov and Loglinear models, an approach to early warning. *Journal of Agricultural Water Management* 77: 59-81.

Paulo, A.A. & Pereira, L.S. 2007. Prediction of SPI drought class transition using Markov chains. *Journal of Water Resour. Manage*. 21: 1813-1827.

Sene, K. 2010. *Drought*. *Hydrometeorology*. DOI 10.1007/978–90–481–3403-8_8. Springer Science+Bisnis Media B.V.

Thom, H.C.S. 1958. A note on the gamma distribution. *Monthly Weather Review* 86: 117-122.

Turkes, M. & Tath, H. 2009. Use of the Standardized Precipitation Index (SPI) and a modified SPI for shaping the drought probabilities over Turkey. *International Journal of Climatology* 29: 2270-2282.

Wahidah Sanusi*
Department of Mathematics
Faculty of Mathematics and Natural Science
Universitas Negeri Makassar
90224, Parangtambung Makassa
Sulawesi Selatan
Indonesia

Kamarulzaman Ibrahim
School of Mathematical Sciences
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43600, Bangi, Selangor
Malaysia

*Corresponding author; email: w_sanusi@yahoo.com

Received: 28 September 2011
Accepted: 31 May 2012