Journal of Applied Sciences 7 (14): 1880-1886, 2007
ISSN 1812-5654


# Fitting Daily Rainfall Amount in Malaysia Using the Normal Transform Distribution

[1]Jamaludin Suhaila and [2]Abdul Aziz Jemain
[1]Department of Mathematics, Faculty of Science, Universiti Teknologi Malaysia, 81310, Skudai, Johor, Malaysia
[2]Sciences Mathematics Studies Center, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600, Bangi, Selangor, Malaysia

**Abstract:** This research presents a statistical study of rainfall which compares several types of normal transform distributions that describe rainfall distribution in Malaysia over a multi-year period. The normal transform distribution meaning that all tested distributions were established based on some modification or transformation of the normal distribution through several methods. The lognormal, skew normal and mixed lognormal distributions are among the normal transform distributions that are proposed and tested to identify the optimal model for daily rainfall amount in several rain gauge stations in Malaysia. The selected model will be chosen based on the minimum error produced by seven criteria of goodness of-fit (GOF) tests, namely the median of absolute difference (MAD) between the empirical and hypothesized distribution function, the traditional Empirical Distribution Function (EDF) Statistics which include Kolmogorov-Smirnov statistic D, Anderson Darling statistic $A^2$ and Cramer-von-Mises statistic $W^2$ and the new method of EDF Statistics based on the likelihood ratio statistics. Of the three models tested, the mixed lognormal is found to be the most appropriate distribution for describing the daily rainfall amount in Malaysia.

**Key words:** Goodness-of-fit-tests, mixed distribution, lognormal, mixed lognormal, skew normal

## INTRODUCTION

Modeling of daily rainfall using various mathematical models has been done throughout the world to develop a better understanding about the rainfall pattern and its characteristics. Various positively skewed distributions have been used in the analysis on fitting daily rainfall amount on wet days. The gamma, weibull and mixed exponential are widely used in this study. However, less attention is given to the distributions in the form of the normal transform distributions.

Among the studies that focused on the normal transform distributions is the one by Lloyd and Schreuder (1981) in which the lognormal and $S_B$ distribution were employed together with the gamma, weibull and beta distribution. In their study, the $S_B$ distribution was found to be the best fitting distribution in analyzing the rainfall amount based on the Kolmogorov-Smirnov and the log likelihood criterion test.

In addition, Chapman (1997, 1998) used another form of normal distribution known as the skewed normal. This distribution was obtained through normal fitting of data that has been normalized by the Box-Cox transformation. In his study, the data for 22 islands in the Western Pacific with minimum 18 years complete record have been chosen and analyzed. This distribution was also tested together with the gamma, weibull, mixed exponential and kappa. The test results showed the complete dominance of the mixed exponential and the skewed normal models based on the Akaike Information Criterion (AIC).

Even though many studies in analyzing rainfall data involve single distribution but several researchers also have began to examine the effect of the mixing distribution in fitting rainfall data especially the mixed exponential (Woolhiser and Roldan, 1982; Wilks, 1999). Kedem *et al.* (1990) has employed the univariate mixed lognormal distribution in his study to determine the linear relationships between the area average and the fractional area which he found that the distribution provided a very close fit for area-average rain rate. Meanwhile, Shimizu (1993) has proposed a bivariate mixed lognormal distribution as a probability model for describing rainfalls which containing zeroes measured at two monitoring sites.

Mathematical models of rainfall amount are very useful in a variety of water resource applications especially in the studies on agriculture and crop planning. Sharda and Das (2005) have tested the lognormal with

**Corresponding Author:** Jamaludin Suhaila, Department of Mathematics, Faculty of Science, Universiti Teknologi Malaysia, 81310, Skudai, Johor, Malaysia Tel: (607)-5534317 Fax: (607)-5566162

other distributions to describe the weekly rainfall data in India where the results from the model have been used to study the effect of rainfall variability during the cropping season. Besides that, the lognormal distribution also has been widely used in the study of flood frequency (Kroll and Vogel, 2002). Therefore, the importance of these mathematical models in rainfall studies should not be neglected.

The selection of the best fitting distribution is always a main interest in the study of rainfall amount. In this study, we would like to find the best fitting distribution for daily rainfall amount based on several goodness-of-fit criteria. The scope of this study is to focus on the distributions that are in the normal transform which means that all tested distributions are established based on some modification or transformation to the normal distributions through several methods. To clarify this meaning, let's take the skew normal distribution as an example. This distribution is obtained by transforming the rainfall data using the Box-Cox transformation and then fit the transform data by using normal distribution. This is what we called normal transform. In addition to the single distribution, the mixed distribution will also be employed in this study but only limited to two components. If more than two components are used, it will involve many parameters and things will get complicated especially in the estimation process. Furthermore, it does not bring much difference to the fitted model.

The new method of GOF tests based on the likelihood ratio statistics developed by Zhang (2002) and Zhang and Wu (2005) will be engaged together with the traditional GOF tests in finding the best fitting distribution. Referring to Zhang's approach, these new GOF tests are more powerful than the traditional GOF. Other additional criterion included in the analysis is the median of absolute difference between the hypothesized and the empirical distribution function. The best distribution will be chosen based on the minimum error specified by all those criteria.

## TOPOGRAPHY AND CLIMATE

Malaysia is situated in the tropics between 1° and 7° degrees north of the equator. It occupies a total area approximately 330,400 square km and is separated by the South China Sea into West Malaysia, which is known as Peninsular Malaysia and East Malaysia (Sabah and Sarawak). Malaysia is generally formed by highland, floodplain and coastal zones. In Peninsular Malaysia the Main Range known as Banjaran Titiwangsa runs from the Malaysia-Thai border in the north to south which spans a distance of 483 km with an elevation of 914-243 m above sea level separating the eastern part of Peninsula from the western. About four-fifth of Peninsular Malaysia is covered by forest and swamps. The same features are also common in East Malaysia.

Malaysia in general, experiences a wet and humid tropical climate throughout the year that is characterized by high annual rainfall, humidity and temperature. Malaysia has uniform temperatures throughout the year of 25.5° to 32° Celsius. Normally, the annual rainfall amount is between 2000 and 4000 mm while the annual number of wet days ranges from 150 to 200 days. There are major differences of climate observed within the country especially between the west and east coasts of Malaysia and slightly between north and south. These differences arise from the discrepancy of altitude and the exposure of the coastal lowlands to the southwest and northeast monsoon winds. The southwest monsoon is usually occurred in mid of May and ends in September. The wind is generally light below 15 knots. Meanwhile, the northeast monsoon usually begins in early November and ends in March with speed between 10 to 20 knots. During this season, the more severely affected areas are the east coast states of Peninsular Malaysia where the wind may reach more than 30 knots. The coast that exposed to the northeast monsoon in Malaysia tends to be wetter than those exposed to the southwest monsoon. The period of the south-west monsoon is a drier period for the whole country, particularly for the other states of the west cost of the Peninsula. The period of change between the two monsoon is the inter monsoon which occur in April and October. These two inter monsoons usually associate with heavy rainfall. Thus, in general the rainfall distribution in Malaysia is governed by those monsoons.

## RAINFALL DATA

Daily rainfall series data for this study have been obtained from the Malaysian Meteorology Department for the periods ranging from 21 to 35 years. For this study, eighteen rain gauge stations were chosen based on the completeness of the data. The stations are selected to represent rainfall pattern for the whole Peninsular Malaysia. The locations for these stations are on map in Fig. 1.

## MODELING RAINFALL AMOUNT

In this study, three models which are in the normal transform distribution were tested. There are lognormal, skew normal and mixed lognormal. These three distributions are described as follow.

- The lognormal with two parameters, $\mu$ and $\sigma$ denote the location and scale parameters, respectively. Meanwhile, X represents the daily rainfall amount. The probability density function is given as follows.

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right), \quad (1)$$

$$x > 0, \quad \sigma > 0, \quad -\infty < \mu < \infty$$

The maximum likelihood estimation for those two parameters is given as below.

$$\hat{\mu} = \frac{\sum_{i=1}^{n} \ln X_i}{n} \text{ and } \hat{\sigma} = \left[\frac{\sum_{i=1}^{n}\left(\ln X_i - \hat{\mu}\right)^2}{n}\right]^{1/2} \quad (2)$$

- The skew normal distribution involves three parameters which are the location ($\mu$), scale ($\sigma$) and shape parameter ($\lambda$). The skew normal distribution is obtained by transforming the rainfall data using the Box-Cox transformation described below.

$$y = \begin{cases} \frac{(x+c)^{\lambda} - 1}{\lambda}, & \lambda \neq 0 \\ \log(x+c), & \lambda = 0 \end{cases} \quad (3)$$

We select $\lambda$ that maximizes the log likelihood function in equation (4) by varying $\lambda$ between -1 and 1;

$$\log L = -\frac{n}{2}\log\left[\sum_{i=1}^{n}\frac{\left(y_i - \bar{y}\right)^2}{n}\right] + (\lambda - 1)\sum_{i=1}^{n}\log(y_i) \quad (4)$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

is the arithmetic mean of the transform data.
and this maximized lambda is substituted in Eq. (3) to get the transform data. Then the transform data is fitted using the normal distribution. The maximum likelihood estimation for normal distribution is given below.

$$\hat{\mu} = \bar{x} \text{ and } \hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \quad (5)$$

- The mixed log-normal with five parameters is the mixture of two-parameter lognormal where p denote the mixing probability while $\mu_1$, $\mu_2$, $\sigma_1$ and $\sigma_2$ represent the location and scale parameters respectively. Meanwhile, X represents the daily rainfall amount.

$$f(x) = \frac{p}{x\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu_1}{\sigma_1}\right)^2\right) + \frac{1-p}{x\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu_2}{\sigma_2}\right)^2\right) \quad (6)$$

$$0 \le p \le 1, \quad \sigma_1 > 0, \; \sigma_2 > 0, \quad -\infty < \mu_1 < \infty, \; -\infty < \mu_2 < \infty \quad x > 0$$

The maximum likelihood equation for this distribution is in implicit form and complicated and we will not discuss details in this paper. However, Newton-Raphson's method has been employed in the iteration procedure to solve the equation. Different initial values have been tested for this procedure. If the initial values converge to the same values and have the largest likelihood, it is considered to be the chosen estimated parameter.

## GOODNES-OF-FIT TESTS (GOF)

Seven different GOF tests have been used to identify the best fit models. The tests are based on the degree of discrepancy between the empirical distribution function $F_n(x_{(i)})$ and the hypothesized distribution function $F(x_{(i)}, \hat{\theta})$. The chosen distribution that best fits the daily rainfall amount is based on the minimum error indicate by all these seven tests.

- First criteria involve the median absolute difference (MAD) between $F(x_{(i)}, \hat{\theta})$ and $F_n(x_{(i)})$. The formula is given as below.

$$\text{MAD} = \text{Med}|F_n(x_{(i)}) - F(x_{(i)}, \hat{\theta})| \quad (7)$$

In this equation, $x_{(i)}$ represents the ordered data whereas the vector of estimated parameters is represented by $\hat{\theta}$.

- Second criteria involve the classical EDF statistics also called as classical GOF tests. They are usually divided into two classes; the supremum and the quadratic. This research concentrates on the well known Kolmogorov-Smirnov statistic D which belongs to the supremum class of EDF statistics.

This statistic is based on the largest vertical differences between $F_n(x_{(i)})$ and $F(x_{(i)}, \hat{\theta})$. For the quadratic class, this paper focuses on the Cramer-von-Mises statistic $W^2$ and the Anderson Darling statistic $A^2$ test. This class of statistics is based on the squared difference $[F_n(x_{(i)}) - F(x_{(i)}, \hat{\theta})]^2$.

Let $Z_{(i)}$ with $F(x_{(i)}, \hat{\theta})$ represents the ordered data whereas the vector of estimated parameters is represented by $\hat{\theta}$. Then the above EDF statistics have the following computing formula:

$$\begin{aligned} D^+ &= \max_i \{i/n - Z_{(i)}\}, \\ D^- &= \max_i \{Z_{(i)} - (i-1)/n\} \\ D &= \max(D^+, D^-) \end{aligned} \quad (8)$$

$$W^2 = \sum_{i=1}^{n}\left\{Z_{(i)} - (2i-1)/2n\right\}^2 + \frac{1}{12n} \quad (9)$$

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}\begin{bmatrix}(2i-1)\log Z_{(i)} + \\ (2n+1-2i)\log\left(1-Z_{(i)}\right)\end{bmatrix} \quad (10)$$

- Third criteria involve the new powerful GOF tests which are based on the likelihood ratio statistics between the hypothesized distribution $Z_{(i)} = F(x_{(i)}, \hat{\theta})$ and the empirical distribution function $F_n(x_{(i)})$. The computed formulas as in Zhang (2002) and Zhang and Wu (2005) are given as follow.

a) New Kolmogorov-Smirnov test

$$Z_K = \max_{1\le i\le n}\begin{bmatrix}(i-0.5)\log\dfrac{i-0.5}{nZ_{(i)}} + \\ (n-i+0.5)\log\dfrac{n-i+0.5}{n(1-Z_{(i)})}\end{bmatrix} \quad (11)$$

b) New Cramer-von-Mises test

$$Z_C = \sum_{i=1}^{n}\left[\log\frac{Z_{(i)}^{-1}-1}{(n-0.5)/(i-0.75)-1}\right]^2 \quad (12)$$

c) New Anderson-Darling test

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\log Z_{(i)}}{n-i+0.5} + \frac{\log(1-Z_{(i)})}{i-0.5}\right] \quad (13)$$

## RESULTS AND DISCUSSION

First, we would like to discuss the result of descriptive statistics for each of the eighteen rain gauge stations. Next we will proceed with the discussion on fitting the distributions.

**Descriptive statistics:** A summary of the basic statistics of daily rainfall amount for eighteen rain gauge stations is displayed in Table 1. The differences among the stations can be compared through the mean, standard deviation, coefficient of variations, skewness, kurtosis, number of wet days and maximum amount of daily rainfall.

From Table 1, it can be shown that Kota Bharu, Kuantan, Kuala Trengganu and Mersing which are located along the east coast of Peninsular Malaysia have the highest mean and standard deviation of daily rainfall amount compared to other stations. These four stations also have the highest coefficient of variations which ranged between 1.8 and 2.0. Meanwhile, the coefficient of variations for the other stations range between 1.3 and 1.6. This shows that the rainfall variability in the studied stations is not quite homogenious.

Kota Bharu indicates the highest value of skewness and kurtosis followed by other stations in the east coast of Peninsular Malaysia. Kluang which is in the south coast of Peninsular Malaysia also has high value of skewness and kurtosis. This may be due to the effect of extreme values in the rainfall series or the maximum amount of rainfall that was received by each station. Kota Bharu, Kuantan, Mersing, Kluang and Senai are among the stations that received the highest amount of maximum daily rainfall. That explains why the shape of the rainfall distribution for those stations was skewed so much.

The question that usually arises in the study of rainfall amount is whether the amount of rainfall is correlated to the number of wet days of the stations. Table 1 indicates that Kota Bharu has a smaller number of wet days compared to other stations, but has the highest mean daily rainfall amount. Meanwhile, Petaling Jaya has the highest number of wet days but a smaller mean daily rainfall amount compared to the other stations. This indicates that the higher mean amount of rainfall is not due to the large number of wet days, but it is possibly contributed by extreme rainfall.

Table 1: Statistics of daily rainfall amount on wet days for eighteen rain gauge stations

| Stations | Mean | St.dev | CV | Skewness | Kurtosis | Number of Wet days | Maximum amount of rainfall (mm) |
|---|---|---|---|---|---|---|---|
| Alor Star | 12.06 | 16.75 | 1.39 | 2.85 | 12.52 | 5878 | 178.8 |
| Batu Embun | 11.41 | 16.82 | 1.47 | 2.73 | 9.84 | 4375 | 160.8 |
| Bayan Lepas | 13.42 | 19.79 | 1.48 | 3.06 | 15.61 | 6192 | 288.2 |
| Cameron Highlands | 11.82 | 14.18 | 1.20 | 2.08 | 5.58 | 5329 | 107.6 |
| Chuping | 10.72 | 15.81 | 1.48 | 3.65 | 26.68 | 4461 | 267.2 |
| Ipoh | 12.59 | 16.88 | 1.34 | 2.24 | 6.04 | 6829 | 135.4 |
| Kluang | 11.22 | 17.79 | 1.59 | 4.88 | 63.89 | 5935 | 433.4 |
| Kota Bharu | 15.87 | 31.02 | 1.95 | 6.13 | 63.90 | 5660 | 591.5 |
| Kuala Krai | 13.15 | 22.19 | 1.69 | 4.59 | 35.02 | 3928 | 356 |
| Kuala Trengganu | 15.66 | 30.31 | 1.94 | 5.44 | 44.87 | 3494 | 432.9 |
| Kuantan | 15.84 | 29.01 | 1.83 | 5.42 | 47.89 | 6493 | 527.5 |
| Malacca | 11.50 | 16.96 | 1.48 | 3.11 | 18.17 | 5982 | 275.2 |
| Mersing | 14.38 | 26.28 | 1.83 | 5.50 | 48.32 | 6468 | 430 |
| Petaling Jaya | 13.83 | 18.40 | 1.33 | 2.39 | 8.01 | 7001 | 177.2 |
| Senai | 11.94 | 17.53 | 1.47 | 4.06 | 40.25 | 6415 | 364.4 |
| Sitiawan | 10.40 | 15.76 | 1.52 | 2.93 | 12.45 | 5964 | 178.7 |
| Subang | 12.43 | 17.07 | 1.37 | 2.58 | 9.43 | 6926 | 171.5 |
| Temerloh | 11.39 | 17.15 | 1.51 | 3.07 | 14.40 | 4661 | 200.1 |

**Fitting distribution:** The values of the seven goodness-of-fit criteria have been calculated. The best distribution was chosen based on the minimum error of GOF tests. The distributions were then ranked in ascending order based on those values. Unfortunately, when many criteria are used to identify the best distribution, it is more difficult to make the selection. The selected statistical distribution for the same data may be different for different analysis. In this study, we chose the best fitting distribution based on the majority of the tests since we did not investigate which is the most powerful test. However, based on Zhang's approach the new GOF tests which are based on the likelihood ratio statistics are found to be more powerful than the traditional GOF tests.

Overall results of the analysis are plotted onto the map of Peninsular Malaysia as shown in Fig. 1. Based on the results, it showed the complete dominance of the mixed lognormal based on the values of GOF tests. Of the three distributions tested, the mixed lognormal was found to be the best fitting distribution for all studied stations. Most of the GOF criteria indicated the minimum error for the mixed lognormal. These were followed by the skew normal which ranked second and finally the lognormal. In general, the outcomes shown that the mixture of two distributions is better than the single distributions in describing daily rainfall in Malaysia.

For better understanding about this mixture distribution, their estimated parameters are displayed in Table 2 for each eighteen rain gauge station. The estimated mean for the first and second component have been calculated and we noticed that the estimated mean for the second component is larger than the first component. We could say that the rainfall pattern in Malaysia can be classified into two types of components, probably light and heavy rainfall. Thus, we concluded that the first component represents the light rains while the second component describes the heavy rains.

Based on the values of mixing probability presented in Table 2, approximately 50 to 60% of daily rainfall amount series represent light rains for most of the stations in the west coast of Peninsular Malaysia. Meanwhile, more than 60% of daily rainfall amount series in Kota Bharu, Kuantan, Kuala Krai and Mersing which are in the east coast of Peninsular Malaysia represent heavy rainfall. This result shows a major difference between the rains at rain gauge stations located in the east coast and west coast of Peninsular Malaysia. As stated before, the monsoons controlled the rainfall distribution in Malaysia. One of the reasons that this might happened because of the east coast of Peninsular Malaysia is mostly affected by the northeast monsoon which bring heavy rainfalls compared to the stations that are located in the west coast.

For further explanation, two rain gauge stations; Kota Bharu in the east coast of Peninsular Malaysia which has the highest mean rainfall amount and Sitiawan in the west coast of Peninsular Malaysia that has the lowest mean rainfall amount are taken as examples. About 29% of rainfall amount series in Kota Bharu described light rains, while approximately 71% described heavy rains. Comparing the estimated mean for each component, we can say that 98% of the total estimated mean rainfall amount in Kota Bharu is contributed by heavy rains while only 2% is influenced by light rains. Meanwhile the rainfall amount series in Sitiawan comprised of 59% of light rains and 41% of heavy rains. However,

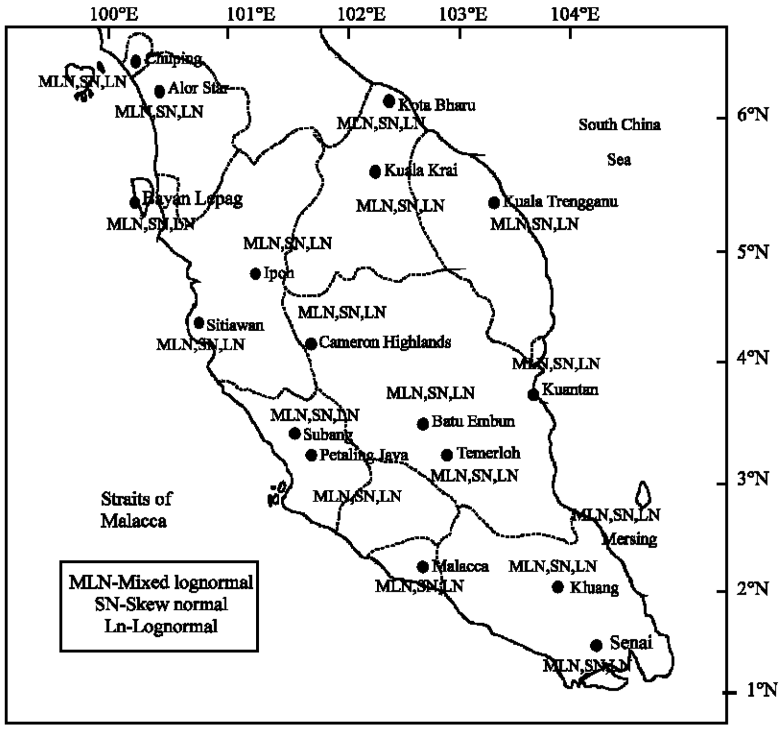


Fig. 1: Map of Peninsular Malaysia showing the eighteen rain gauge stations and the ranking distributions for each station

Table 2: Estimated parameters for mixed lognormal as the best fitting distribution

| Stations | Mixing probability ($p_1$) | Location 1 ($\mu_1$) | Scale 1 ($\sigma_1$) | Estimated mean (1) | Location 2 ($\mu_2$) | Scale2 ($\sigma_2$) | Estimated Mean (2) |
|---|---|---|---|---|---|---|---|
| Alor Star | 0.503 | 0.309 | 1.366 | 1.739 | 2.674 | 0.874 | 10.566 |
| Batu Embun | 0.349 | -0.623 | 1.095 | 0.341 | 2.307 | 1.074 | 11.646 |
| Bayan Lepas | 0.495 | 0.234 | 1.368 | 1.593 | 2.74 | 0.94 | 12.177 |
| Cameron Highlands | 0.505 | 0.546 | 1.39 | 2.289 | 2.687 | 0.777 | 9.835 |
| Chuping | 0.506 | 0.109 | 1.361 | 1.425 | 2.552 | 0.9 | 9.499 |
| Ipoh | 0.546 | 0.409 | 1.419 | 2.247 | 2.808 | 0.836 | 10.684 |
| Kluang | 0.383 | -0.398 | 1.147 | 0.497 | 2.345 | 1.043 | 11.084 |
| Kota Bharu | 0.285 | -0.412 | 1.044 | 0.326 | 2.353 | 1.226 | 15.935 |
| Kuala Krai | 0.317 | -0.416 | 1.114 | 0.389 | 2.324 | 1.128 | 13.174 |
| Kuala Trengganu | 0.464 | 0.081 | 1.258 | 1.11 | 2.709 | 1.106 | 14.831 |
| Kuantan | 0.355 | -0.227 | 1.159 | 0.553 | 2.557 | 1.12 | 15.589 |
| Malacca | 0.58 | 0.262 | 1.385 | 1.967 | 2.793 | 0.836 | 9.722 |
| Mersing | 0.31 | -0.48 | 1.098 | 0.35 | 2.377 | 1.145 | 14.326 |
| Petaling Jaya | 0.577 | 0.605 | 1.442 | 2.991 | 2.974 | 0.776 | 11.176 |
| Senai | 0.559 | 0.406 | 1.409 | 2.264 | 2.741 | 0.859 | 9.898 |
| Sitiawan | 0.585 | 0.145 | 1.322 | 1.621 | 2.703 | 0.863 | 8.977 |
| Subang | 0.538 | 0.377 | 1.38 | 2.035 | 2.789 | 0.835 | 10.639 |
| Temerloh | 0.558 | 0.207 | 1.392 | 1.806 | 2.699 | 0.899 | 9.849 |

approximately 85% of the total estimated mean rainfall amount in Sitiawan is still contributed by heavy rains while the remainder is influenced by light rains. In general, we can summarize that even though there are large differences in the proportion of light and heavy rains between stations, the heavy rainfalls still becomes a major contribution to the mean rainfall amount on wet days in Malaysia.

## CONCLUSIONS

The search for the best distribution in fitting daily rainfall amount has become a main interest in several studies. Various forms of distributions have been tested in order to find the best fitting distribution. Different tests of goodness-of-fit have been attempted in the studies.

In this study, a comparison of the lognormal, skew normal and mixed lognormal show that for eighteen Malaysia rain gauge stations, the mixed lognormal best describes the distribution of daily rainfall amount based on seven different goodness-of-fit tests. The skew normal and log-normal ranked second and third respectively. In conclusion, we showed that mixed distribution is better than single distribution for describing the daily rainfall amount. Therefore, we recommend using a mixed distribution for fitting daily rainfall amount. Even though there are major differences in the proportion of heavy and light rains between rain gauge stations but the mean amount of rainfall in Malaysia is contributed mainly by heavy rains. Besides, the large number of wet days does not contribute to the large amount of rainfall.

## ACKNOWLEDGMENTS

Authors are thankful to the staff of Malaysian Meteorology Department for providing daily rainfall data for this paper. This study would not have been possible without the sponsorship from Universiti Teknologi Malaysia.

## REFERENCES

Chapman, T.G., 1997. Stochastic models for daily rainfall in the Western Pacific. Mathematics and Computers in Simulation, 43: 351-358.

Chapman, T.G., 1998. Stochastic modelling of daily rainfall: The impact of adjoining wet days on the distribution of rainfall amounts. Environ. Modelling and Software, 13: 317-324.

Kedem, B., L.S. Chiu *et al*., 1990. An analysis of the threshold method for measuring area-average rainfall. J. Applied Meteorol., 29: 3-20.

Kroll, C.N. and R.M. Vogel, 2002. Probability distribution of low streamflow series in the United States. J. Hydrologic Eng., 7: 137-146.

Sharda, V.N. and P.K. Das, 2005. Modeling weekly rainfall data for crop planning in a sub-humid climate of India. Agricultural Water Management, 76: 120-138.

Shimizu, K., 1993. A bivariate mixed log-normal distribution with an analysis of rainfall Data. J. Applied Meteorol., 32: 161-171.

Lloyd, S.W., Jr. and H.T. Schreuder, 1981. Fitting daily precipitation amounts using the $S_B$ Distribution. Monthly Weather Review, 109: 2535-2540.

Wilks, D.S., 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. Agric. For. Meteorol., 93: 153-169.

Woolhiser, D.A. and J. Roldan, 1982. Stochastic daily precipitation models 2. A comparison of distribution of amounts. Water Resour. Res., 18: 1461-1468.

Zhang, J., 2002. Powerful goodness-of-fit tests based on the likelihood ratio. J. Royal Statistical Society Series B., 64: 281-294.

Zhang, J. and Y. Wu, 2005. Likelihood ratio tests for normality. Comput. Statistics and Data Anal., 49: 709-721.