Atmospheric Research 92 (2009) 236–243

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Mixed log series geometric distribution for sequences of dry days

Sayang Mohd Deni [a,*], Abdul Aziz Jemain [b,1]

[a] Center for Statistical Studies, Faculty of Information Technology & Quantitative Sciences, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia
[b] School of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 9 May 2008
Received in revised form 4 August 2008
Accepted 30 October 2008

Keywords:
Dry spells
Mixture probability model
Log series distribution
Geometric distribution
Daily rainfall occurrence

ABSTRACT

The development of rainfall occurrence models is continuously being explored by a number of researchers in the field since the 20th century. In order to further enhance the development of rainfall occurrence modeling, this present study is aimed to propose a new probability model which is able to describe the daily series of rainfall occurrence, particularly on the duration of consecutive dry or wet days in Peninsular Malaysia. In selecting the most successful model to describe the distribution of the rainfall event, the model with the less number of parameters is preferred. In addition, there have been cases where the model with more parameters did not show a significant fit. In this situation, it is necessary to develop a more appropriate model which can be used for data generation purposes and other applications. Based on several probability models developed previously, the mixture of log series with geometric distribution (MLGD) is proposed as the alternative probability model to describe the distribution of dry (wet) spells in daily rainfall events. This study aims to fit nine types of probability models including the MLGD to dry spells data for 16 selected rainfall stations in Peninsular Malaysia. The sequence of dry days will be analyzed separately at each station using daily rainfall observations for the period of 1975 to 2004. The adequacy of the MLGD and the existing probability models in fitting the observed distribution of dry spells at each station are evaluated using a chi square goodness-of-fit test. The results demonstrated that all the data sets were found to successfully fit the new proposed model, the MLGD, in representing the sequence of dry days over the peninsula. Moreover, this model was also found to best fit the three data sets which were not able to fit the existing models.

1. Introduction

Due to the rapid growth in the population as well as the development in industrialization, water supply which is mainly from rainfall is highly in demand not only in Malaysia, but also throughout the world. Therefore, analyzing rainfall behavior particularly the characteristics of dry spells is very important to the water related sectors in a way to manage water supply more efficiently especially during the dry periods. In order to provide informative resources to the relevant sectors, the analysis on the distribution of dry spells should be conducted comprehensively. Various types of rainfall occurrence models had been developed by the previous researchers in the field. In determining the most successful model to describe the distribution of the rainfall event, the model with the less number of parameters is preferred. In addition, there have been cases where the model with more parameters did not show a significant fit. In this situation, it has been necessary to develop a more appropriate model not only for data generation purposes, but also for providing some useful information in various applications including water resource management in the hydrological and agricultural sectors.

Various types of probability models which have been applied for the distribution of dry (wet) spells were introduced in the literature by previous researchers such as the log series distribution (LSD) by Williams (1952),

* Corresponding author. Tel.: +60 3 5543 5460; fax: +60 3 5543 5501.
E-mail addresses: sayan929@salam.uitm.edu.my (S.M. Deni), azizj@ukm.my (A.A. Jemain).
[1] Tel.: +60 3 8921 5724; fax: +60 3 8925 6086.

0169-8095/$ – see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.atmosres.2008.10.032

geometric distribution (GD) by Gabriel and Neumann (1957), modified log series distribution (MLD) by Green (1970), compound geometric distribution (CGD) by Yap (1973), truncated negative binomial distribution (TNBD) by Buishand (1978), and mixed geometric series with LSD (GMLS) by Srinivasan (1958). By considering the mixed probability models, Racsko et al. (1991) proposed the mixture of the two geometric distributions (MGD) and it was observed that the distribution of dry and wet spells at Hungary fitted well with this model. Moreover, their findings indicated that the duration of wet spells could be approached by a single GD, while for long dry spells the mixed distributions were more appropriate. Meanwhile, the mixture of geometric and Poisson distribution (MGPD) which was first introduced by Dobi-Wantuch et al. (2000) successfully fitted the distribution of dry (wet) spells for the two stations in Hungary.

The study on identifying the best models for dry and wet spells is continuously being explored by many researchers in the field. More recently, Deni et al. (2008) applied seven types of probability models not including MGD and MGPD for each of the 10 selected rainfall stations over Peninsular Malaysia. They found that TNBD and CGD fitted well to the distribution of dry and wet spells, respectively. Di Giuseppe et al. (2005) reported that PLD and TNBD were found more efficient in fitting the observed dry (wet) spells at four Italian sites. Moreover, the negative binomial distribution had been successfully fitted to the distribution of wet and dry spells in Greece (Tolika and Maheras, 2005; Anagnostopoulou et al., 2003). Modeling extreme dry spells was another important aspect in rainfall occurrence event that received attention and was discussed in the literature (Lana et al., 2006a,b; Vicente-Serrano and Begueria-Portugues, 2003; Wilks, 1999).

This present study is aimed to fit various types of probability models to represent the observed distribution of dry spells over Peninsular Malaysia. The performance of the new proposed model, MLGD and the existing eight probability models will be evaluated based on the successfulness and the best fitting models in describing the underlying observed frequency of dry spells at each rainfall station. In the following section, the description of the data sets together with the homogeneity tests will be discussed briefly. This is followed by the explanation on the probability models applied and the methods used for

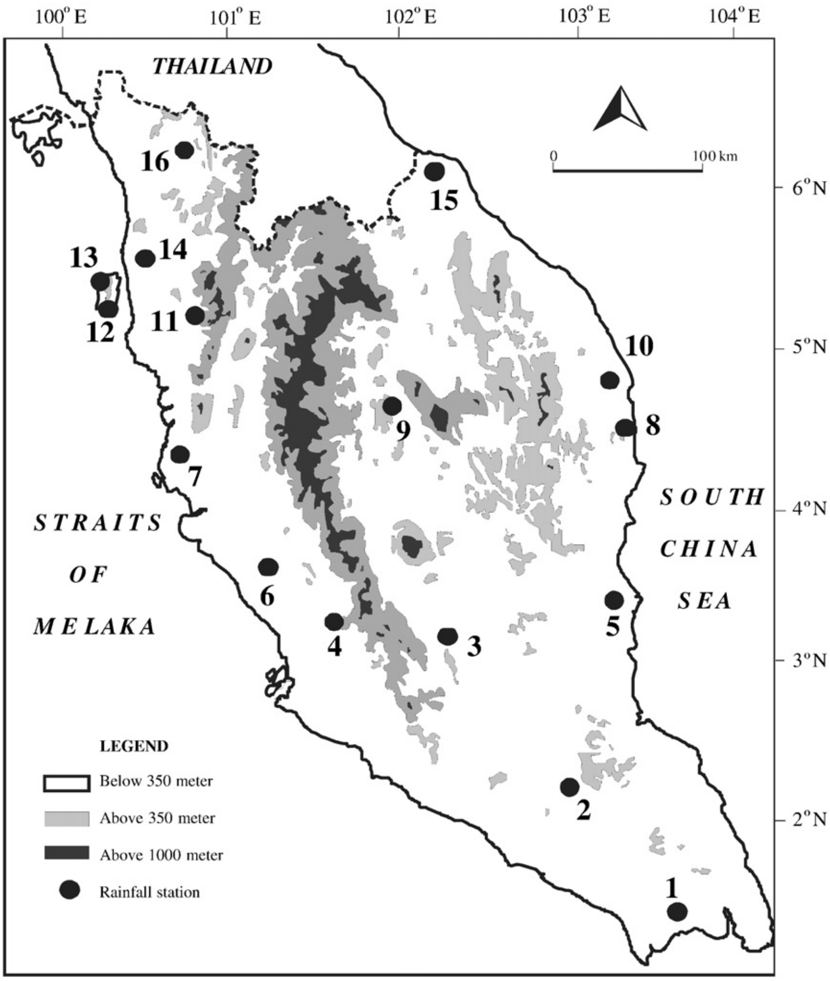

Fig. 1. The physical map showing the locations of the 16 selected rainfall stations in Peninsular Malaysia.

Table 1
Geographical coordinates, the percentage of the missing data and the results of homogeneity tests for the mean dry spells at each of the 16 selected rainfall stations in Peninsular Malaysia

Station code	Station name	Latitude (°N)	Longitude (°E)	% missing	Homogeneity Tests		
					SNHT	BRT	VonNR
1	Senai	1.63	103.67	0.0	1.59	0.94	2.17
2	Kluang	2.02	103.32	0.0	6.22	1.33	1.88
3	Sg. Lui Halt	3.08	102.37	2.0	6.06	1.27	1.45
4	Gombak	3.27	101.72	0.5	2.99	1.16	2.29
5	Pekan	3.55	103.35	4.3	2.66	1.08	1.90
6	Bagan Terap	3.72	101.07	9.3	2.39	1.27	2.11
7	Sitiawan	4.22	100.70	0.0	3.97	1.08	2.10
8	Dungun	4.75	103.42	3.5	3.91	1.12	1.46
9	Gua Musang	4.87	101.97	7.2	3.22	0.98	1.63
10	Kg. Menerong	4.93	103.05	4.3	4.43	1.22	1.82
11	Selama	5.13	100.68	0.8	4.26	0.94	1.75
12	Bayan Lepas	5.30	100.27	0.0	3.70	0.78	2.20
13	Air Itam	5.38	100.25	4.4	8.16	1.26	1.61
14	Bumbong Lima	5.55	100.43	3.4	6.64	1.42	1.66
15	Kota Bharu	6.17	102.28	0.0	3.71	1.27	1.44
16	Ampang Pedu	6.23	100.77	3.8	5.37	1.49	2.00
			Critical value (5%)		**7.65**	**1.50**	**1.42**

estimating the parameter(s) for each model. In Sections 3 and 4, the results of the analysis and the conclusion of this study are provided respectively.

2. Data and methodology

2.1. The study area and data

Peninsular Malaysia lies entirely in the equatorial zone which is situated in the northern latitude between 1° and 6° N and the eastern longitude from 100° to 103° E. There are two types of monsoons that influence the climate of the country, namely, the Southwest monsoon (May to August) and the Northeast monsoon (November to February). The data used in this study were collected from the database of the Malaysian Meteorological Department (MMD) and the Drainage and Irrigation Department (DID), for the period of records that ranged from 1975 to 2004. Fig. 1 shows the location of the 16 selected rainfall stations over Peninsular Malaysia. Moreover, the homogeneity of the data series were checked using three out of the four types of homogeneity tests as recommended by Wijngaard et al. (2003), the standard normal homogeneity test (SHNT), the Buishand range test (BRT) and the Von Neumann (VonNR) ratio test. Table 1 displays the geographical coordinates, the percentage of missing values together with the results of the homogeneity test for each station. It is observed that the mean dry spells of the data series were homogeneous at 5% level of significance. The data used in this present study can be considered good quality data with less than 10% missing values throughout the 30-year period. The missing values in the data series for the period of 1975 to 2004 were estimated using various types of weighting methods such as the inverse distance, the normal ratio and the correlation between the target and the neighboring stations (Suhaila et al., 2008; Teegavarapu and Chandramouli, 2005; Sullivan and Unwin, 2003; Eischeid et al., 2000).

Table 2b
The list of probability models with three parameters and the probability functions used for fitting the distribution of dry spells, $x = 1,2,\ldots$

Probability Models	Probability Function
MGD Racsko et al. (1991)	$\mathrm{P}(x) = Wp_1\,(1-p_1)^{x-1} + (1-W)p_2(1-p_2)^{x-1}$
MGPD Dobi-Wantuch et al. (2000)	$P(x) = Wp(1-p)^{x-1} + (1-W)\frac{\lambda^{x-1}}{(x-1)!}e^{-\lambda}$
MLGD (New Model Proposed)	$P(x) = Wp_1^x[x\log(1/(1-p_1))]^{-1}$ $+ (1-W)p_2(1-p_2)^{x-1}$

2.2. Probability models for dry spells

A dry day is defined as a day with a rainfall amount of less than 0.1 mm. A dry spells is a period of consecutive days of exactly, say x, dry days immediately preceded and followed

Table 2a
The list of methods, probability function, estimators and the method of estimation (e.g., ML—Maximum Likelihood, FM—Factorial moment, M—moment) used for fitting the distribution of dry spells

Probability models	Probability function	Estimators	Method of estimation
LSD (One-parameter model) Williams (1952)	$P(x) = -p^x/x\log(1-p) \quad x = 1, 2, \ldots$	$\bar{x} = \left[\left(1-\hat{p}^{-1}\right)log(1-\hat{p})\right]^{-1}$	ML
GD (One-parameter model) Gabriel and Neumann (1957)	$P(x) = p(1-p)^{x-1} \quad x = 1, 2, \ldots$	$\hat{p} = 1/\bar{x}$	ML
MLD (Two-parameter model) Green (1970)	$P(x) = \frac{cp^x}{x+h} \quad 0 \le h < \infty, \quad x = 1, 2, \ldots$ $c = (\sum p^x/(x+h))^{-1}$ is normalizing constant.	$\bar{x} = c\sum_{x=1}\frac{x\hat{p}^x}{x+\hat{h}}$ and $\frac{1}{N}\sum_{x=1}\frac{N_x}{x+\hat{h}} = c\sum_{x=1}\frac{\hat{p}^x}{(x+\hat{h})^2}$	ML
CGD (Two-parameter model) Yap (1973)	$P(0) = b(a-1)^{-1}$ $P(x) = \frac{a+x-2}{a+b+x-1}P(x-1) \quad x = 1, 2, \ldots$	$\hat{a} = (\bar{x}-1)\left(\hat{b}-1\right)$ and $\hat{b} = \frac{2s^2}{\bar{x}+s^2-\bar{x}^2}$	FM
TNBD (Two-parameter model) Buishand (1978)	$P(x) = {}^{r+x-1}C_x(1-p)^x p^r/(1-p^r) \quad x = 1, 2, \ldots$	$\hat{p} = \frac{\bar{x}}{s^2}\left(1-\frac{N_1}{N}\right)$ and $\hat{r} = \frac{1}{1-\hat{p}}\left(\bar{x}\hat{p}-\frac{N_1}{N}\right), r \ge -1$	M
GMLS (Three-parameter model) Srinivasan (1958)	$P(x) = \left(A_1 + \frac{A_2}{x}\right)p^x \quad x = 1, 2, \ldots$	$\hat{p} = \frac{T_3 - T_2 \pm \sqrt{4T_2^2 - 2T_1(T_2+T_3)}}{T_3+T_2}$, $T_j = \sum_j x^j N_x, \hat{A}_1 = \left(\frac{1-\hat{p}}{\hat{p}}\right)^2(T_2(1-\hat{p})-T_1)$, $\hat{A}_2 = \frac{T_1(1-\hat{p})^2 - \hat{A}_1\hat{p}}{\hat{p}(1-\hat{p})}$.	M

Table 3
Main characteristics of distribution of dry spells in Peninsular Malaysia

	Dry Spells			
Station	Mean	S.D.	Max	Nd
1	2.34	2.41	25	4913
2	2.49	2.74	39	5355
3	4.69	5.24	55	8238
4	2.69	2.91	35	5602
5	3.45	4.14	45	6514
6	3.21	3.49	35	5931
7	2.79	2.85	34	5798
8	3.61	4.24	37	6318
9	2.97	4.38	111	5485
10	2.73	3.44	38	4810
11	2.94	3.16	31	6125
12	2.87	3.39	41	5619
13	3.24	4.17	42	5854
14	3.91	5.17	63	7021
15	3.30	3.85	41	6122
16	3.73	6.42	71	6331

Abbreviations:
Mean—Mean length of the dry spells.
S.D—Standard deviation of the dry spells.
Max—Maximum length of the dry spells.
Nd—Total number of dry days.

by a wet day. The three-parameter model, GMLS, was one of the alternative probability models introduced by Srinivasan (1958). However, GMLS had a disadvantage in the estimation of parameter, p, which was obtained by solving the quadratic equation. In some cases, the real value of p could not be produced due to the imaginary term existing in the quadratic function. Based on the assumption made on LSD by Williams (1952), that the dry (wet) spells had certain characteristics whereby, the longer the spell lasted, the more likely it was to last another day and the successfulness of GD in representing the sequences of dry (wet) days, the new probability models, MLGD is proposed in the present study. The new three-parameter model is more appropriate compared to GMLS due to the estimation of the parameters. This model is developed to provide a statistical description of the observation which is meant for various purposes not only for data generation but also for modeling climatic events in order to describe the physical explanation of the rainfall occurrence.

In fitting a particular probability model to the observed distribution of dry spells, the parameters are estimated by using either the methods of maximum likelihood or factorial moment. The parameters of the LSD, MLD and GD are estimated using the maximum likelihood method. However, for mathematical convenience due to the complexity of the three other distributions, CGD, TNBD and GMLS, the moment and factorial moment methods will be applied when estimating the parameters. Throughout the analysis the following notations will be used, is the observed frequencies of dry spells of length x days, $\sum_{x=1} N_x$ or N is the total observed frequencies of dry spells, while and

Table 4
The chi square goodness of fit test of various types of probability models for the dry spells in each of the 16 selected rainfall stations over Peninsular Malaysia

		Number of Parameter(s)								
		One		Two			Three			
Station		GD	LSD	TNBD	CGD	MLD	GMLS	MGD	MGPD	MLGD
1	Chi sq.	255.87	22.52	21.48	*16.51*	21.97	na	***7.92***	255.87	*15.33*
	p-values	0.00	0.02	0.03	*0.09*	0.02	na	***0.64***	0.00	*0.12*
2	Chi sq.	237.56	*18.14*	*20.45*	*15.65*	*19.29*	na	*13.69*	75.17	***10.99***
	p-values	0.00	*0.11*	*0.06*	*0.15*	*0.06*	na	*0.25*	0.00	***0.36***
3	Chi sq.	114.14	89.53	***20.05***	*27.50*	*17.99*	86.58	114.14	59.67	*19.34*
	p-values	0.00	0.00	***0.46***	*0.07*	*0.52*	0.00	0.00	0.00	*0.37*
4	Chi sq.	255.87	*9.36*	***9.01***	27.93	*13.48*	na	*10.81*	77.31	*9.16*
	p-values	0.00	*0.75*	***0.77***	0.00	*0.34*	na	*0.55*	0.00	*0.61*
5	Chi sq.	213.41	29.81	*22.98*	17.39	*16.56*	na	213.41	65.49	***8.48***
	p-values	0.00	0.03	*0.11*	0.24	*0.35*	na	0.00	0.00	***0.86***
6	Chi sq.	199.13	*18.68*	*13.34*	40.91	*15.94*	*16.41*	***12.96***	54.83	*18.41*
	p-values	0.00	*0.29*	*0.50*	0.00	*0.32*	*0.23*	***0.53***	0.00	*0.19*
7	Chi sq.	201.34	*21.36*	*15.01*	26.14	*14.99*	*19.86*	201.34	42.21	***12.28***
	p-values	0.00	*0.09*	*0.24*	0.01	*0.24*	*0.07*	0.00	0.00	***0.34***
8	Chi sq.	282.86	***21.79***	*23.64*	57.12	*21.82*	25.80	26.97	81.88	*22.11*
	p-values	0.00	***0.19***	*0.10*	0.00	*0.15*	0.04	0.04	0.00	*0.11*
9	Chi sq.	374.88	32.47	39.31	26.85	23.43	na	***13.41***	104.47	*12.64*
	p-values	0.00	0.00	0.00	0.01	0.04	na	***0.49***	0.00	*0.40*
10	Chi sq.	276.97	43.80	28.55	42.26	26.66	na	276.97	102.68	***17.97***
	p-values	0.00	0.00	0.01	0.00	0.01	na	0.00	0.00	***0.08***
11	Chi sq.	236.35	*11.71*	***9.81***	39.73	*9.64*	*11.08*	236.35	80.70	*9.22*
	p-values	0.00	*0.70*	***0.78***	0.00	*0.72*	*0.60*	0.00	0.00	*0.76*
12	Chi sq.	278.76	26.64	34.16	28.70	31.27	na	***17.58***	67.06	*19.57*
	p-values	0.00	0.02	0.00	0.00	0.00	na	***0.17***	0.00	*0.08*
13	Chi sq.	427.72	33.75	27.35	68.58	32.22	na	427.72	140.78	***22.86***
	p-values	0.00	0.01	0.04	0.00	0.01	na	0.00	0.00	***0.06***
14	Chi sq.	275.96	*26.36*	*27.66*	37.17	***17.22***	na	*21.54*	na	*22.82*
	p-values	0.00	*0.12*	*0.09*	0.00	***0.44***	na	*0.20*	na	*0.12*
15	Chi sq.	214.41	*17.29*	*14.04*	36.99	*17.85*	na	*12.73*	84.96	***11.44***
	p-values	0.00	*0.37*	*0.52*	0.00	*0.21*	na	*0.55*	0.00	***0.65***
16	Chi sq.	457.85	53.85	51.52	50.32	43.72	na	27.40	166.82	***9.56***
	p-values	0.00	0.00	0.00	0.00	0.00	na	0.04	0.00	***0.73***

The italics and bold face indicate the successfully fitted and best fitted models at 5% level of significance, respectively.
na Represents no solution exists for the GMLS and MGPD.

denote the sample values of and, respectively. Table 2a briefly describes the probability models applied including the probability functions, estimators and also the method of estimating the parameters for the six models tested.

Moreover, for the other three probability models, MGD, MGPD, and MLGD, the estimators of the parameters are not shown due to the complexity in the mathematical functions. The parameters of these mixture probability models were estimated using the quasi-Newton method with the maximum likelihood estimation. Table 2b displays the probability models applied and the probability functions of the three models tested. The parameter of p, p_1 and p_2 for each probability model applied ranges from 0 to 1, while W is the weight factors, where the sum of W and $(1-W)$ is unity.

The performance of the new alternative model proposed, MLGD, together with the existing eight probability models will be evaluated based on the most successfulness and the best fitted probability model in representing the underlying frequency of observed distribution of dry spells. In selecting the most successful and the best fitted model for each data set, the chi-square goodness-of-fit test is considered. The test will be computed based on the differences between the observed and expected frequencies of various lengths of dry spells (Snedecor and Cochran, 1989). The classes should be grouped for adjacent spells, whenever frequencies are less than 5. The Pearson's chi-square goodness-of-fit test is shown below,

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

where O_i and E_i are the observed and expected frequencies of dry spells, and n is the number of classes. The chi-square test of 5% level of significance with the degree of freedom, $v=n-t-1$, where t is the number of parameters estimated from each probability model, will be considered. Note that a higher probability associated to the lower chi-square statistics value produces a better fit.

3. Results and discussion

3.1. Main characteristics of the length of dry days

Table 3 shows the usual statistical parameters of the distribution of dry days for each station. These include the mean, standard deviation, maximum duration of dry spell length and the total number of dry days. The average duration of dry

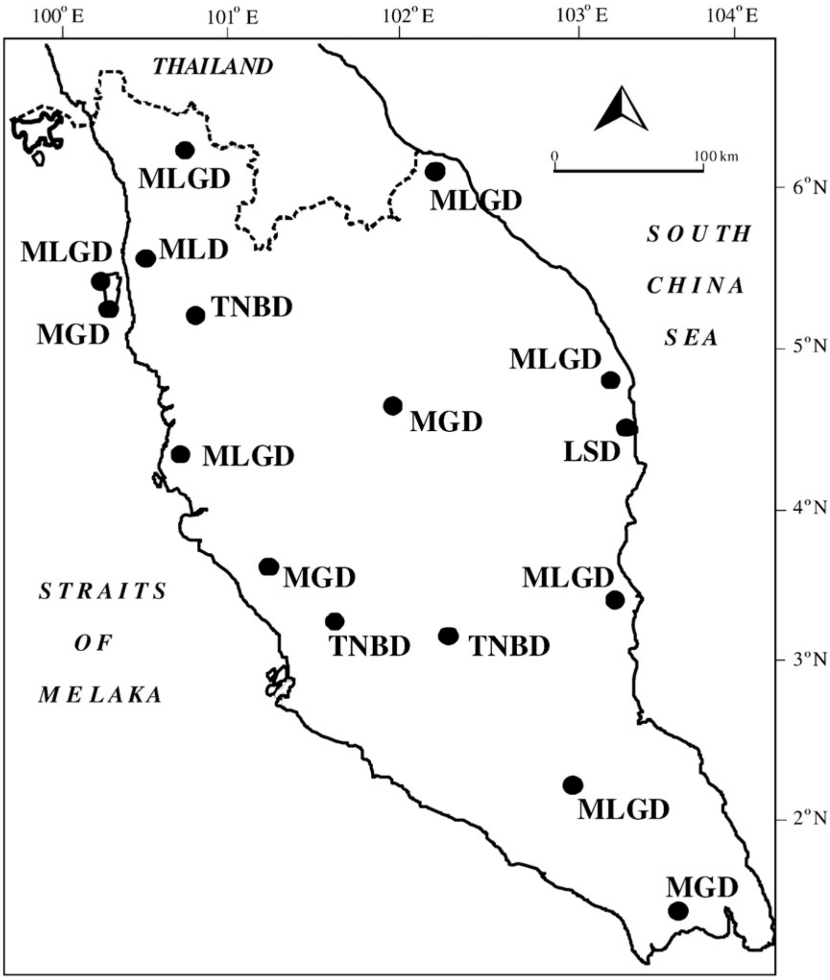

Fig. 2. The best fitted probability models for dry spells at each of the selected rainfall stations in Peninsular Malaysia.

periods is varied from 3 to 5 days, whereas the variability of the length of dry days is varied from 3 to 7 days over the peninsula. Over the study period, Station 9 experienced the longest duration of dry spells, with a maximum of 111 days. There were also two stations located in the northwestern area, namely, Station 14 and Station 16, that had experienced no rainy days for the maximum of more than two months consecutively. Although Station 3 had no experienced of having more than two months of dry spells consecutively, however, the highest total number of dry days, 8238 days, was observed at this station.

3.2. The most successful probability models for dry spells

The most successful models as shown in Table 4 (in italics) were based on the highest number of stations that successfully

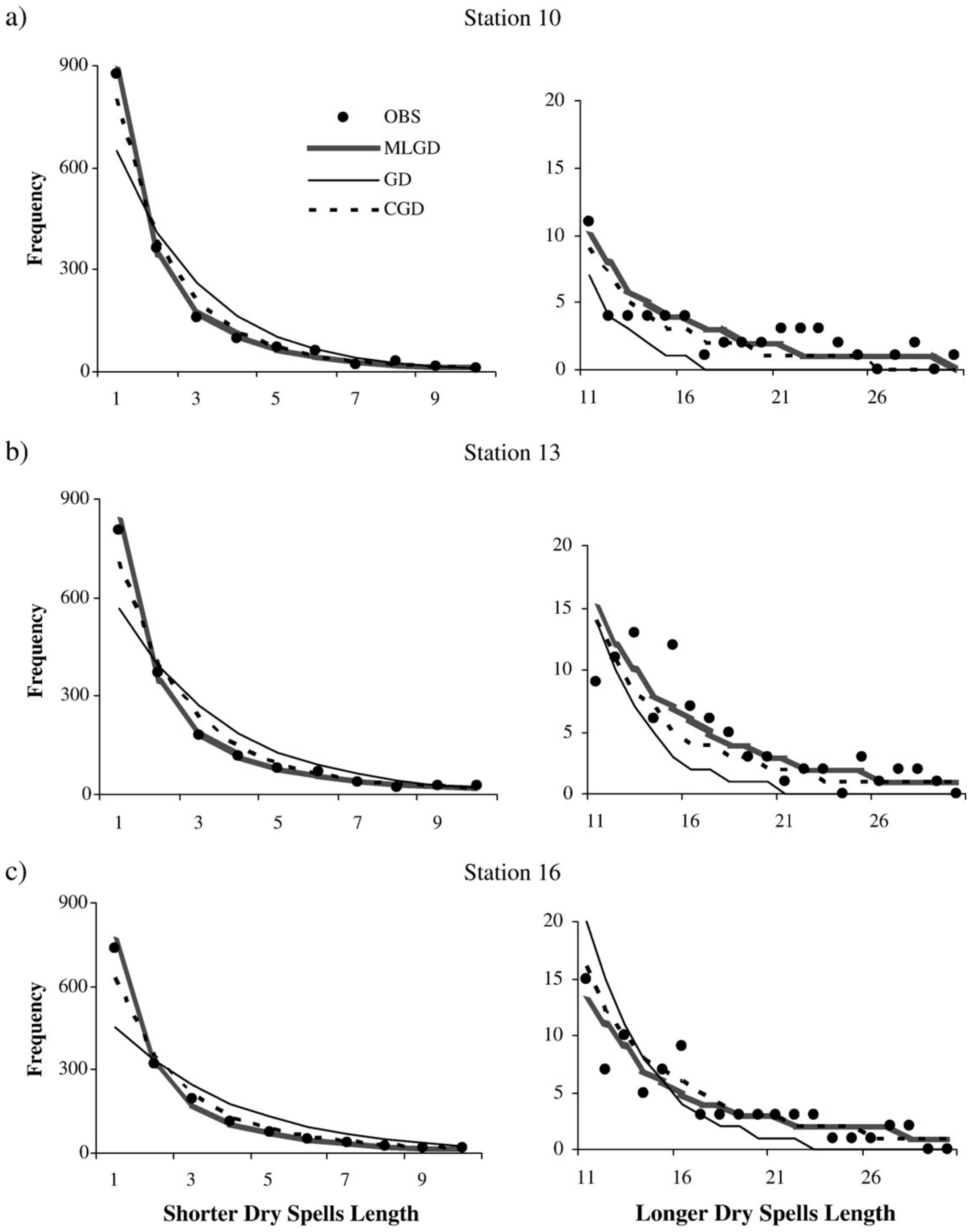

Fig. 3. The observed and estimated frequencies of GD, CGD and MLGD for the distribution of shorter (1 to 10 days) and longer (11 to 30 days) length of dry spells at Stations 10, 13 and 16.

fitted the distribution of dry spells to a particular probability model at 5% level of significance. In determining the most successful model to describe the distribution of dry (wet) spells, the model with the less number of parameters is always a main interest in this field of study. However, not all of the one or two parameter models can be fitted to the distribution of interest. For example, the results in Table 4 indicated that there were 3 out of the 16 data sets that failed to fit both the one and two parameter models. In this situation, the model with the higher number of parameters is recommended. Surprisingly, none of the existing probability models with three parameters successfully fitted these three rainfall stations, Station 10, Station 13, and Station 16. Thus, the main concern of this present study is to propose a more appropriate model which is able to fit the distribution of dry spells of the tested data sets. Based on the results shown in Table 4, the abovementioned three rainfall stations were found to successfully fit the MLGD. It is observed that the MLGD was found to be the most frequent model selected where this new model was observed to successfully fit the distribution of dry spells at each of the rainfall stations used in this present study. The results of this present study seem to be in agreement with the findings of Deni and Jemain (2008) who found that GD and MGPD failed to fit the distribution of dry spells over Peninsular Malaysia. Moreover, the failure of GMLS was also proven here where only five out of the 16 data sets were able to produce the parameter, p. Among the two-parameter models, TNBD and MLD were the second most frequently selected models which fitted the observed dry spells at 10 stations successfully.

3.3. The best fitting probability models for dry spells

The best fitting models were based on the largest probability associated to the lower chi square values (Table 4 in bold) for the distribution of dry spells at each station. The results in Table 4 and Fig. 2 revealed that the new model performed well in representing the distribution of dry spells where 7 out of the 16 rainfall stations demonstrated the MLGD as the best fitted model. The findings of this present study indicated that none of the geometric distribution family, GD, CGD, GMLS or MGPD was found to be the best fitted model to describe the distribution of dry spells, except for the MGD and the MLGD. The study also revealed that the distribution of dry spells at the three rainfall stations which did not successfully fit the existing models, were however able to fit the MLGD. In addition, Fig. 3 displayed the observed and the estimated distributions for shorter (1 to 10 days) and longer (11 to 30 days) length of dry spells using the one-parameter model, GD, the two-parameter model, CGD, and the three-parameter model, MLGD, at Stations 10, 13 and 16, which were unable to fit the existing probability models. It can be seen that the estimated frequency of dry spells which were obtained from the MLGD actually described the observed distributions for both lengths of spells. On the other hand, it was observed that GD and CGD slightly overestimated and underestimated the shorter and the longer distribution of dry spells, respectively.

4. Conclusions

The development of the rainfall occurrence model is necessary in providing a more comprehensive analysis of the behavior of daily rainfall event which can be used for various applications. In determining the best fitting distribution of dry (wet) spells, the models with the less number of parameters estimated are preferred. But, not all of the cases of the one parameter, two parameter or even three parameter models were able to fit the observed data. Thus, this present study is aimed to propose the new probability model, MLGD, which is able to describe the distribution of dry and wet spells, to be used as an input to the climate monitoring system to obtain a better prediction for future climatic events.

The improvement of the existing model, GMLS, was needed since this model had a disadvantage in estimating the parameter, p, due to the existence of the imaginary term in solving the quadratic equation. The results of this present study have proven that the performance of the MLGD is more convincing than the GMLS which is also a three-parameter model. It can be seen in Table 4 that 11 out of the total 16 rainfall stations were found unable to produce the parameter p for GMLS due to the imaginary term in the quadratic equation.

The investigation on the optimal probability models for dry and wet spells for the Malaysian data is extended in this present study. The findings of this present study indicated that the new probability model, MLGD, was superior to the existing probability models in describing the distribution of dry spells over the peninsula for the period of 1975 to 2004. Fig. 2 clearly indicated that 7 out of the total 16 rainfall stations were best fitted with the MLGD to represent the sequences of dry days. The results indicated that all the data sets used in this present study were found to successfully fit the model, including the data sets from the three stations, Station 10, Station 13 and Station 16, which were unable to fit the existing eight probability models. It is shown in Fig. 3 that the MLGD adequately fitted the shorter and longer duration of dry spells at those three stations.

The new probability model proposed in this present study is developed to further enhance the development of the rainfall occurrence models. It is meant to provide a better or an alternative probability model in describing the sequences of dry (wet) days in daily rainfall events. These models can be applied at wider regions or for different monsoon periods with the existence of very long duration of dry (wet) spells.

Acknowledgements

The authors are indebted to the staff of the Drainage and Irrigation Department and the Malaysian Meteorological Department for providing the daily rainfall data for this study. This research would not have been possible without the sponsorship of the Universiti Teknologi MARA (UiTM), Malaysia. They also acknowledge their sincere appreciation to both reviewers for their valuable suggestions and remarks which improved the manuscript. This research was funded by the Universiti Kebangsaan Malaysia (UKM-GUP-PI-08-34-318).

References

Anagnostopoulou, Chr., Maheras, P., Karacostas, T., Vafiadis, M., 2003. Spatial and temporal analysis of dry spells in Greece. Theor. Appl. Climatol. 74 (1–2), 77–91.

Buishand, T.A., 1978. Some remarks on the use of daily rainfall models. J. Hydrol. 36, 295–308.

Deni, S.M., Jemain, A.A., Ibrahim, K., 2008. The spatial distribution of wet and dry spells over Peninsular Malaysia. Theor. Appl. Climatol 94 (3–4), 163–173.

Deni, S.M., Jemain, A.A., 2008. Probability models for dry spells in Peninsular Malaysia. Asia-Pacific J. Atmos. Sci. 44 (1), 37–47.

Di Giuseppe, E., Vento, D., Epifani, C., Esposito, S., 2005. Analysis of dry and wet spells from 1870 to 2000 in four Italian sites. Geophys. Res. Abst. 7, 07712.

Dobi-Wantuch, I., Mika, J., Szeidl, L., 2000. Modeling wet and dry spells with mixture distributions. Meteor. Atmos. Phy. 73, 245–256.

Eischeid, J.K., Pasteris, P.A., Diaz, H.F., Plantico, M.S., Lott, N.J., 2000. Creating a serially complete, national daily time series of temperature and precipitation for the Western United States. J. Appl. Meteor. 39, 1580–1591.

Gabriel, K.R., Neumann, J., 1957. On the distribution of weather cycles by length. Quar. J. Roy. Meteor. Soc. 83, 375–379.

Green, J.R., 1970. A generalized probability model for sequences of wet and dry days. Mon. Wea. Rev. 98, 238–241.

Lana, X., Martinez, M.D., Burgueno, A., Serra, C., Martin-Vide, J., Gomez, L., 2006a. Distributions of long dry spells in the Iberian Peninsula, years 1951–1990. Int. J. Climatol. 26, 1999–2021.

Lana, X., Martinez, M.D., Burgueno, A., Serra, C., 2006b. Statistical distributions and sampling strategies for the analysis of extreme dry spells in Catalonia (NE Spain). J. Hydrol. 324, 94–114.

Racsko, P., Szeidl, L., Semenov, M., 1991. A serial approach to local stochastic weather models. Ecol. Mod. 57, 27–41.

Snedecor, G.W., Cochran, W.G., 1989. Statistical Methods, Eighth Edition. Iowa State: University Press. 503 pp.

Srinivasan, T.R., 1958. A statistical model suitable for describing weather persistence. Ind. J Meteor. Geophys. 10, 321–324.

Suhaila, J., Deni, S.M., Jemain, A.A., 2008. Revised spatial weighting methods for estimation of missing rainfall data. Asia-Pacific J. Atmos. Sci. 44 (2), 93–104.

Sullivan, D.O., Unwin, D.J., 2003. Geographic Information Analysis. Wiley, Hoboken, NJ.

Teegavarapu, R.S.V., Chandramouli, V., 2005. Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. J. Hydrol. 312 (1–4), 191–206.

Tolika, K., Maheras, P., 2005. Spatial and temporal characteristics of wet spells in Greece. Theor. Appl. Climatol. 81 (1–2), 71–85.

Vicente-Serrano, S.M., Begueria-Portugues, S., 2003. Estimating extreme dry-spell risk in the middle Ebro Valley (Northeastern Spain): a comparative analysis on spatial duration series with a general Pareto distribution and annual maxima with a Gumbel distribution. Int. J. Climatol. 23, 1103–1118.

Wijngaard, J.B., Klien Tank, A.M.G., Konnen, G.P., 2003. Homogeneity of 20th century European daily temperature and precipitation series. Int. J. Climatol. 23, 679–692.

Wilks, D.S., 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. Agric. Forest Meteorol. 93 (3), 153–169.

Williams, C.B., 1952. Sequences of wet and of dry days considered in relation to the Logarithmic Series. Quar. J. Roy. Meteor. Soc. 78, 511–516.

Yap, W.C., 1973. The persistence of wet and dry spells in Sungai Buloh, Selangor. Meteor. Mag. 102, 240–245.