

Variations of characteristics of consecutive rainfall days over northern Thailand

P. Klongvessa¹ · M. Lu¹ · S. Chotpantararat^{2,3}

Received: 30 May 2016 / Accepted: 19 June 2017 / Published online: 14 July 2017
© Springer-Verlag GmbH Austria 2017

Abstract The Chao Phraya basin, Thailand, is frequently inundated by flooding during the southwest monsoon period. Most floods coincide with consecutive rainfall days. This study investigated consecutive rainfall days during the southwest monsoon period at 11 stations over northern Thailand, the upstream area of this basin. The Markov chain probability model was used to study the consecutiveness of days with at least 0.1, 10.1, and 35.1 mm of rainfall. The consecutive length of rainfall days from the model showed good agreement with the observed value. A chi-square test of independence was applied to assess the significance of the consecutiveness, and it was found that days with at least 10.1 mm of rainfall tend to be consecutive over the entire area. Moreover, days with at least 35.1 mm of rainfall were found to be consecutive over the joint area where the mountainous region meets the plain area. However, the consecutiveness of days with less than 10.1 mm of rainfall was not obvious. The rainfall amount on days with at least 10.1 mm of rainfall was also calculated and it showed lower

values over the mountainous region than over the plain. Hence, this study established the characteristics of consecutive rainfall days over the plain, mountainous region, and joint area.

1 Introduction

Flooding in the Chao Phraya basin has been a regular occurrence for several decades. This basin is the largest and most populated basin in Thailand. It has an important role in the economic development of the country. Therefore, floods in this area are of great interest to researchers and to the Thai government. There have been at least five major flood events since 1980, i.e., 1983, 1995, 1996, 2006, and 2011, which typically have caused losses and damage on the scale of tens of millions of US dollars (Prajamwong and Supparataratn 2007); however, the losses and damage associated with the 2011 flood alone were more than 40 billion US dollars (The World Bank 2012). These floods usually occur between May and October because of the high rainfall caused by tropical cyclones and the moist air brought to the area by the southwest monsoon.

According to the records of the Hydro and Agro Informatics Institute of Thailand, many flood events coincide with consecutive rainfall days (Hydro and Agro Informatics Institute 2008a, b, 2009). Previous studies have investigated the rainfall in this basin and most have mentioned the role of topography on rainfall variation. Okumura et al. (2003) found that the phase delay of the rainfall peak corresponds to the distance from the leeward side (eastern side) of Dawna Range, which is the western boundary of the Chao Phraya basin. Yokoi and Satomura (2008) found that the intraseasonal variation of precipitation on the eastern side of this mountain range is different from that on the western

✉ P. Klongvessa
nineboon@hotmail.com
S. Chotpantararat
csrilert@gmail.com

¹ Department of Civil and Environmental Engineering, Nagaoka University of Technology, Nagaoka, Japan

² Department of Geology, Chulalongkorn University, Bangkok, Thailand

³ Research Program in Control of Hazardous Contaminants in Raw Water Resources for Water Scarcity Resilience, Center of Excellence on Hazardous Substance Management (HSM), Chulalongkorn University, 254 Phyathai Road, Pathumwan, Bangkok 10330, Thailand

side. Kuraji et al. (2009) found that the counts of observed rainfall hours and total rainfall amount have positive correlation with altitude. Takahashi et al. (2010) suggested that rainfall has an afternoon peak over the mountain crest and an evening peak over the plain, which is due to the diurnal wind. However, the seasonal variation of surface wind seems not to affect the diurnal variation of rainfall, but to affect the rainfall amount (Takahashi 2010). Mahavik et al. (2014) revealed that rainfall tends to concentrate over the windward side of the mountain.

With regard to consecutive rainfall days, Dahale et al. (1994) suggested that the likelihood of whether a particular day would be a dry day or a rainfall day is related to whether the previous day was a dry day or a rainfall day. Furthermore, Szyniszewska and Waylen (2012) found that the probability of consecutive rainfall days increases with the total amount of monthly rainfall. However, few studies have investigated the subject of consecutive rainfall days, even though the amount of rainfall on three consecutive days has been proposed as one of the criteria for flood warning (Hydro and Agro Informatics Institute 2007). Given this context, the objective of this study was to identify the spatial characteristics of consecutive rainfall days in the upstream area of the Chao Phraya basin, known as the Ping, Wang, Yom, and Nan basins.

2 Study area

The study area is in the northern part of Thailand, encompassing the Ping, Wang, Yom, and Nan basins with a total area of approximately 100,000 km², as shown in Fig. 1. These basins represent the upstream area of the Chao Phraya basin. The northern part of the study area is a mountainous region with some plain areas along the valleys, and the southern part of the study area comprises mostly plain area with a small proportion of mountains along the eastern and western boundaries. The elevation of the area ranges from <100 MASL over the plain area to >2500 MASL at Doi Inthanon in the mountainous region. The rainfall in this area, as well as in most parts of Thailand, is high between May and October because of the moist air brought from the Gulf of Thailand and Andaman Sea by the southwest monsoon. The average annual rainfall in northern Thailand during 1971–2010 was 1240.1 mm, of which 952.1 mm could be attributed to the southwest monsoon (Thai Meteorological Department 2012).

3 Method

The daily rainfall data between May and October recorded by Thai Meteorological Department (TMD) rain gauges

have been used to investigate consecutive rainfall days. Data suitable for this study started to become available in 1981; therefore, we decided to use data obtained during the latest climatological period, i.e., 1981–2010. Data from 11 rain gauges were available for the period 1981–2010; the locations of these rain gauges are shown in Fig. 1, and their specific details are shown in Table 1.

Based on the daily rainfall amount, the TMD classifies days into one of four types: a dry day (rainfall 0.0 mm), light rainfall day (rainfall 0.1–10.0 mm), moderate rainfall day (rainfall 10.1–35.0 mm), and heavy rainfall day (rainfall >35.1 mm). In this study, a Markov chain probability model was applied to study the transition of days in three cases. The first case studied the transition between days with 0.0 mm and at least 0.1 mm of rainfall. The second case studied the transition between days with less than 10.1 mm and at least 10.1 mm of rainfall. The third case studied the transition between days with less than 35.1 mm and at least 35.1 mm of rainfall. A chi-square test of independence was applied to test for significance. The results show the consecutiveness of days with at least 10.1 mm of rainfall (see Section 4); therefore, the spatial distribution of the rainfall amount on those days was analyzed further using the Student's *t* test.

3.1 Markov chain probability model

A Markov chain probability model is a model used to describe the transition probability among states. In studying rainfall, this model has been applied to investigate the transition probability between dry days and rainfall days (Dahale et al. 1994; Moon et al. 1994; Dastidar et al. 2010; Hossain and Anam 2012; Sonnadara and Jayewardene 2015). This transition probability can be used for calculating the probability of consecutive dry or rainfall days. However, as this study focused on consecutive rainfall days, only those calculations used to calculate consecutive rainfall days are discussed. The transition probability from a dry day to a rainfall day is defined as the probability that a dry day subsequently changes state into a rainfall day on the following day. Similarly, the transition probability from a rainfall day to a rainfall day is defined as the probability that the rainfall day retains the state of a rainfall day on the following day. These two transition probabilities can be calculated using Eqs. 1 and 2:

$$P_{01} = \frac{n_{01}}{n_0} \quad (1)$$

$$P_{11} = \frac{n_{11}}{n_1} \quad (2)$$

where P_{01} is the transition probability from a dry day to a rainfall day, P_{11} is the transition probability from a rainfall day to a rainfall day, n_{01} is the number of dry days that are

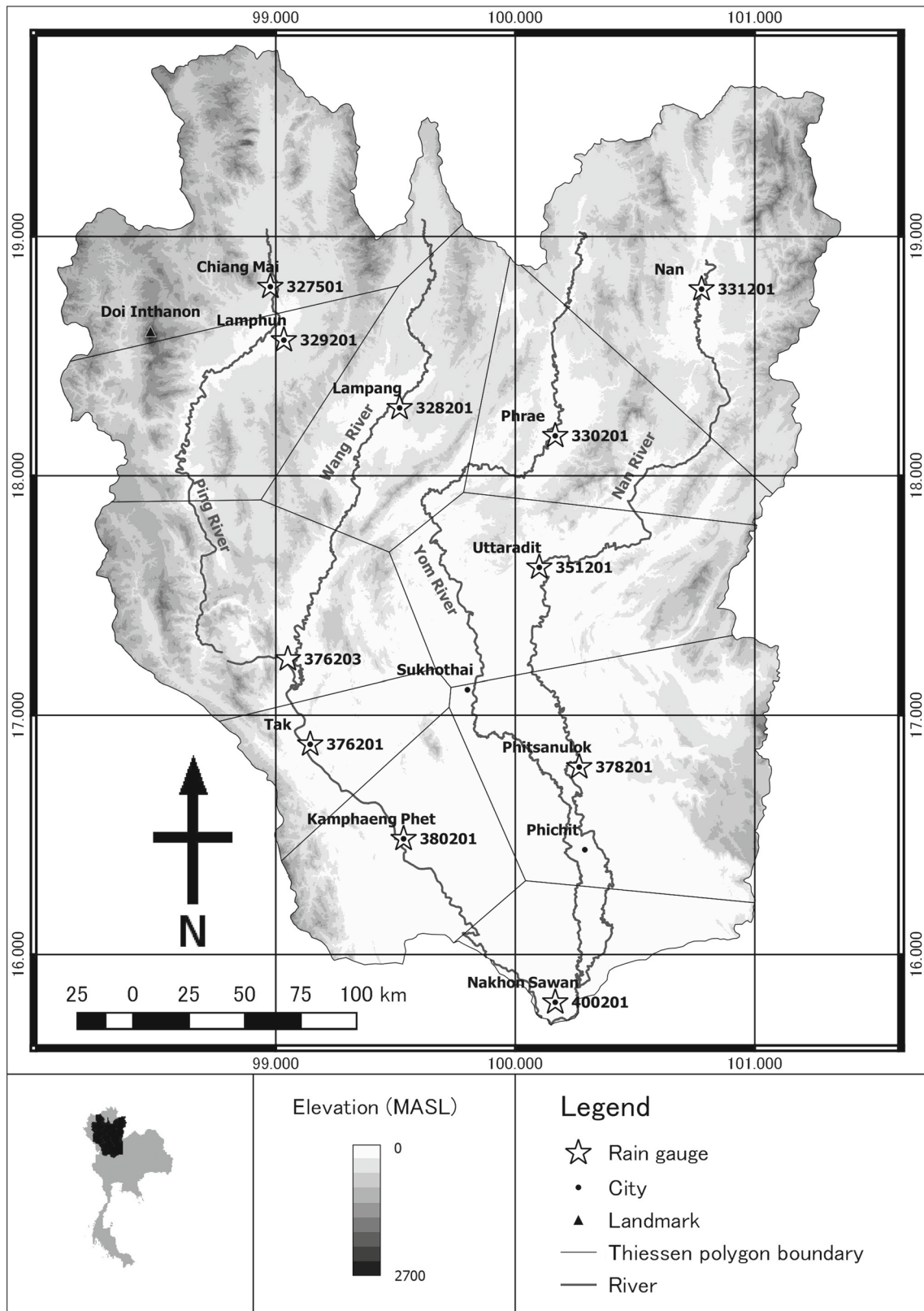


Fig. 1 Study area. Elevation data is from GMTED10 derived by United States Geological Surveys

Table 1 Location, elevation (MASL), and coverage area (km²) of the rain gauges

Station	Latitude	Longitude	Elevation ^a	Area ^b
327501	18.7900	98.9769	316	13031.9
331201	18.7797	100.7778	204	14912.1
329201	18.5667	99.0333	298	8153.8
328201	18.2833	99.5167	257	8549.7
330201	18.1667	100.1667	163	8451.6
351201	17.6167	100.1000	67	10262.8
376203	17.2333	99.0500	162	11011.8
376201	16.8783	99.1433	124	3965.8
378201	16.7833	100.2667	47	13555.2
380201	16.4833	99.5333	78	7381.4
400201	15.8000	100.1667	25	4388.2

^aElevation as of GMTED10

^bCoverage area determined by Thiessen polygon within the study area

followed by rainfall days in the time series, n_{11} is the number of rainfall days that are followed by rainfall days in the time series, n_0 is the total number of dry days in the time series, and n_1 is the total number of rainfall days in the time series.

The period of consecutive rainfall days with length L is defined as the period of $L + 1$ days, the first L days of which are rainfall days and the last day of which is a dry day. Given this definition, the mean length of consecutive rainfall days can be calculated using Eq. 3 (Sonnadara and Jayewardene 2015):

$$\bar{L} = \frac{1}{1 - P_{11}} \quad (3)$$

where \bar{L} is the mean length of the consecutive rainfall days.

The return period of consecutive rainfall days with length of more than L can be calculated using Eq. 4 (Weiss 1964):

$$T_L = \frac{1 - P_{11} + P_{01}}{s P_{01}(1 - P_{11})P_{11}^L} \quad (4)$$

where T_L is the return period of consecutive rainfall days with length of more than L , and s is the number of studied days in each year. For this study, s was 184, which represents the number of days between May and October, and the unit of T_L is year.

The classification of days for the Markov chain probability model is not limited to only dry days where the rainfall amount is 0.0 mm and rainfall days where the rainfall amount is at least 0.1 mm. Some studies have applied different criteria for the classification of days, e.g., Sonnadara and Jayewardene (2015) used threshold values of 1.0 and 3.0 mm, as well as 0.1 mm, to determine rainfall days. In this study, the threshold values of 0.1, 10.1, and

35.1 mm were used because the TMD uses these criteria to classify rainfall days, as mentioned above.

3.2 Chi-square test of independence

A chi-square test of independence is a test used to prove that two variables are independent of each other (Sheskin 1996). For the application of a Markov chain probability model on rainfall, some studies have applied this test to check whether the state of each day (dry or rainfall) affects the state of the following day (Moon et al. 1994; Hossain and Anam 2012). This application is known as a Markov chain analysis (Hiscott 1981). The null hypothesis of this test is that the state of each day does not affect the state of the following day. Most studies have found that the null hypothesis is not true because a dry day usually leads to a subsequent dry day and a rainfall day usually leads to a subsequent rainfall day (Dahale et al. 1994; Moon et al. 1994; Dastidar et al. 2010). In other words, dry days tend to be consecutive as are rainfall days. In this study, the states of days were classified using the threshold values of 0.1, 10.1, and 35.1 mm, and a chi-square test of independence was applied to test whether each day with 0.1–10.0, 10.1–35.0, and at least 35.1 mm of rainfall showed more consecutiveness than the others.

In the daily rainfall time series, for each two consecutive days, we let $i \in \{0, 1, 2, 3\}$ denote the state of the previous day and $j \in \{0, 1, 2, 3\}$ denote the state of the following day where 0, 1, 2, and 3 represent the states of dry, light rainfall, moderate rainfall, and heavy rainfall, respectively. Let n_{ij} denote the number of days with state i that are followed by days with state j in the time series. With the null hypothesis that the state of each day does not affect the state of the following day, the expected value of n_{ij} can be calculated using Eq. 5 (Sheskin 1996):

$$E(n_{ij}) = \frac{n_i n_j}{n} \quad (5)$$

where $E(n_{ij})$ is the expected value of n_{ij} under the null hypothesis, n_i and n_j are the numbers of days with states i and j , respectively, and n is the total number of days in the time series.

The test value can be calculated using Eqs. 6 and 7:

$$\chi^2 = \sum_i \sum_j R_{ij}^2 \quad (6)$$

$$R_{ij} = \frac{n_{ij} - E(n_{ij})}{\sqrt{E(n_{ij})}} \quad (7)$$

where χ^2 is the test value. R_{ij} is the residual of the transition from state i to state j , which indicates the level of dominance of the transition. Large positive values of R_{ij} suggest

that a day with state i usually leads to state j on the following day, while large negative values suggest that a day with state i rarely leads to state j on the following day. Small values of R_{ij} suggest that a day with state i rarely affects whether the following day will have state j .

In order to test the null hypothesis, the criterion value can be determined from the chi-square distribution with $(r - 1)(c - 1)$ degrees of freedom, where r represents a number of states on the previous day and c represents a number of states on the following day. If the test value (χ^2) is larger than the criterion value, the null hypothesis is rejected, implying that the state of each day does affect the state of the following day. Otherwise, the null hypothesis is accepted, which implies that the state of each day does not affect the state of the following day.

In the case that the null hypothesis is rejected, the residual (R_{ij}) can indicate how the state of each day affects the state of the following day. As mentioned previously, the value of R_{ij} indicates the level of dominance of the transition from state i to state j state through its magnitude and its sign. To test qualitatively whether the magnitude of R_{ij} is sufficiently high to conclude that the transition from state i to state j state is dominant, a z test is applied (Sheskin 1996). The R_{ij} value is compared with the criterion value determined from a two-tailed standard normal distribution. If R_{ij} is more than the positive criterion value from the standard normal distribution, the day with state i tends to lead to state j on the following day. If R_{ij} is less than the negative criterion value from the standard normal distribution, the day with state i tends not to lead to state j on the following day. If R_{ij} is between the positive and negative criterion values from the standard normal distribution, the transition from state i to state j is not the component that supports the rejection of the null hypothesis.

3.3 Student's t test

As discussed in detail in Section 4.1.2, the result of the chi-square test of independence reveals the pattern of consecutive days with at least 10.1 mm of rainfall over the entire area. Therefore, the rainfall amounts on these days were also investigated at each station to determine whether any part of the area has higher or lower amounts of rainfall on consecutive rainfall days. This investigation could be done using the two-tailed Student's t test (Sheskin 1996). This test has been applied in many studies on rainfall, e.g., Raju et al. (2002) and Lee et al. (2014). In this test, the rainfall amount on a day with at least 10.1 mm of rainfall at each station was compared with the areal average. The null hypothesis of this test is that the rainfall amount at the station is equal to the areal average, and rejection of the null hypothesis implies that the rainfall amount at the station is higher or lower than the areal average.

For each year, the average value of daily rainfall amount on days with at least 10.1 mm of rainfall (X) at each station was calculated using Eq. 8:

$$X = \frac{\sum p_{10.1}}{n_{10.1}} \tag{8}$$

where $\sum p_{10.1}$ was the total rainfall amount from days with at least 10.1 mm of rainfall in the year, and $n_{10.1}$ was the number of days with at least 10.1 mm of rainfall in the year.

The areal rainfall amount on days with at least 10.1 mm of rainfall for each year was determined as the average value among all the stations weighted by a Thiessen polygon, as shown in Fig. 1. At each station, the test value for the Student's t test can be calculated from Eq. 9:

$$t = \frac{\bar{X} - \mu}{\tilde{s}/\sqrt{n_s}} \tag{9}$$

where t is the test value, \bar{X} and \tilde{s} are the average and standard deviation of X during 1981–2010, μ is the areal rainfall amount on days with at least 10.1 mm of rainfall averaged from 1981 to 2010, and n_s is the number of years, which for this study was 30.

The criterion value is determined from t -distribution with $n_s - 1$ degrees of freedom. The station where t is higher than the positive criterion value is the station where the average rainfall amount on days with at least 10.1 mm of rainfall is higher than the areal average. The station where t is lower than the negative criterion value is the station where the average rainfall amount on days with at least 10.1 mm of rainfall is lower than the areal average. The station where t is between the negative and positive criterion values is the station where the average rainfall amount on days with at least 10.1 mm of rainfall is similar to the areal average.

4 Results and discussion

4.1 Consecutive rainfall days

4.1.1 Model accuracy

The Markov chain probability model developed in this study appears to be accurate because the observed frequency of the consecutive rainfall days shows good agreement with the modeled value for all days with at least 0.1, 10.1, and 35.1 mm of rainfall. The largest overestimations of mean consecutive lengths (\bar{L}) are at station 376201, where the overestimations are 0.006, 0.002, and 0.001 days for days with at least 0.1, 10.1, and 35.1 mm of rainfall, respectively. The largest underestimations of mean consecutive length (\bar{L}) are at station 351201 for days with at least 0.1 and 35.1 mm of rainfall, and at station 331201 for days with at least 10.1 mm of rainfall. The values of the largest

Table 2 Transition probabilities from dry day to rainfall day (P_{01}) and from rainfall day to rainfall day (P_{11}), modeled and observed mean consecutive length of rainfall days (\bar{L}), and error between modeledand observed \bar{L} 's when a 0.1-mm threshold value is used to define the rainfall day

Parameter	Station										
	327501	331201	329201	328201	330201	351201	376203	376201	378201	380201	400201
P_{01}	0.392	0.380	0.371	0.392	0.382	0.358	0.331	0.324	0.389	0.423	0.407
P_{11}	0.669	0.667	0.635	0.631	0.659	0.684	0.655	0.641	0.663	0.656	0.612
Modeled \bar{L}	3.025	3.001	2.739	2.713	2.931	3.164	2.901	2.784	2.966	2.910	2.573
Observed \bar{L}	3.040	3.011	2.741	2.717	2.943	3.185	2.903	2.778	2.966	2.921	2.576
Error ^a	-0.014	-0.010	-0.002	-0.003	-0.012	-0.021	-0.002	0.006	0.000	-0.011	-0.001

^aNegative error represents an underestimation and positive error represents an overestimation

underestimations are 0.021, 0.002, and 0.002 days for days with at least 0.1, 10.1, and 35.1 mm of rainfall, respectively. Tables 2, 3, and 4 show the values of transition probabilities (P_{01} and P_{11}), and modeled and observed \bar{L} 's at each station. Figure 2 shows the observed and modeled frequencies of consecutive rainfall days at the stations where there are largest overestimations or underestimations (stations 331201, 351201, and 376201). For other stations, the differences between the observed and modeled frequencies of consecutive rainfall days are less.

4.1.2 Consecutive rainfall days probability

As the Markov chain probability model has been proven accurate, this model can be used to describe the probability of consecutive rainfall days over northern Thailand. Here, the probability is described in terms of the return period (T_L). Figure 3 shows the return period for days with at least 0.1, 10.1, and 35.1 mm of rainfall that are consecutive for more than 2 days, and Fig. 4 shows the return period for days with at least 0.1, 10.1, and 35.1 mm of rainfall that are consecutive for more than 3 days. It can be seen that when the threshold value of 35.1 mm of rainfall is applied, the consecutive rainfall days probability is obviously high at stations in the joint area between the plain and mountainous region

(stations 351201, 376201, and 376203). At these three stations, the return period for days with at least 35.1 mm of rainfall that are consecutive for more than 3 days is on the scale of 10 years, while the return periods at other stations are on the scale of at least 100 years. However, with the threshold value of 10.1 mm of rainfall, the differences in the consecutive rainfall days probabilities are less obvious, and no differences can be seen when the threshold value of 0.1 mm of rainfall is applied. This finding suggests that the joint area between the plain and mountainous region has higher probability of consecutive heavy rainfall days than other areas.

In order to explain the results qualitatively counts of transitions among no rainfall days, 0.1–10.0 mm rainfall days, 10.1–35.0 mm rainfall days, and at least 35.1 mm rainfall days were applied to the chi-square test of independence, as mentioned in Section 3.2. The results at all 11 stations are similar. Table 5 shows the results at selected stations over the mountainous region (331201), joint area (376203), and plain area (400201). The counts of transitions from no rainfall days to no rainfall days, and from rainfall days to rainfall days (i.e., all of the 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days) are higher than the expected value under the null hypothesis of the chi-square test of independence, while the counts of transitions from no rainfall days

Table 3 Same as Table 2 but for when a 10.1-mm threshold value is used to define the rainfall day

Parameter	Station										
	327501	331201	329201	328201	330201	351201	376203	376201	378201	380201	400201
P_{01}	0.144	0.144	0.134	0.130	0.136	0.156	0.103	0.115	0.171	0.168	0.144
P_{11}	0.269	0.330	0.260	0.274	0.280	0.354	0.361	0.341	0.288	0.274	0.256
Modeled \bar{L}	1.368	1.492	1.352	1.378	1.390	1.547	1.565	1.518	1.405	1.378	1.345
Observed \bar{L}	1.370	1.494	1.351	1.378	1.391	1.548	1.564	1.516	1.404	1.380	1.346
Error ^a	-0.002	-0.002	0.001	0.000	-0.001	-0.001	0.001	0.002	0.001	-0.002	-0.002

^aNegative error represents an underestimation and positive error represents an overestimation

Table 4 Same as Table 2 but for when a 35.1-mm threshold value is used to define the rainfall day

Parameter	Station										
	327501	331201	329201	328201	330201	351201	376203	376201	378201	380201	400201
P_{01}	0.027	0.036	0.029	0.026	0.028	0.046	0.030	0.030	0.042	0.042	0.039
P_{11}	0.077	0.086	0.044	0.090	0.084	0.147	0.172	0.137	0.091	0.106	0.075
Modeled \bar{L}	1.083	1.094	1.046	1.099	1.092	1.172	1.208	1.158	1.100	1.118	1.081
Observed \bar{L}	1.083	1.095	1.046	1.100	1.092	1.174	1.208	1.157	1.100	1.118	1.082
Error ^a	0.001	-0.001	0.000	-0.001	-0.001	-0.002	0.000	0.001	0.000	0.000	-0.001

^aNegative error represents an underestimation and positive error represents an overestimation

to rainfall days and vice versa are less than the expected value. In other words, rainfall days tend to be consecutive rather than isolated. This finding regarding the consecutiveness of rainfall days is similar to the results of many other studies in different areas (Moon et al. 1994; Hossain and Anam 2012; Szyniszewska and Waylen 2012; Sonnadara and Jayewardene 2015).

From Table 5, it can be seen that the counts of transitions among rainfall days are high irrespective of whether they are for 0.1–10.0, 10.1–35.0, or at least 35.1 mm rainfall days, and the counts of transitions from no rainfall days to rainfall days are comparatively low. Hence, the differences in the transitions from 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days can rarely be seen when the transition from no rainfall days is included in the comparison. To investigate the differences among the transitions from 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days, the chi-square test of independence was performed again with the exclusion of the transition from no rainfall days (i.e., the first row of Table 5).

The results at all 11 stations are still similar. Table 6 shows the results at the same sample stations (331201, 376203, and 400201). It can be seen that the count of transitions from 0.1–10.0 mm rainfall days to no rainfall days is higher than the expected value, while the counts of transitions from 0.1–10.0 mm rainfall days to at least 10.1 mm rainfall days (i.e., both 10.1–35.0 and at least 35.1 mm rainfall days) are lower than the expected value. Conversely, the counts of transitions from at least 10.1 mm rainfall days to no rainfall days are lower than the expected value, and the counts of transitions within at least 10.1 mm rainfall days are higher than the expected value. Hence, at this step, it can be concluded that days with at least 10.1 mm of rainfall tend to be more consecutive than days with 0.1–10.0 mm of rainfall over the entire area.

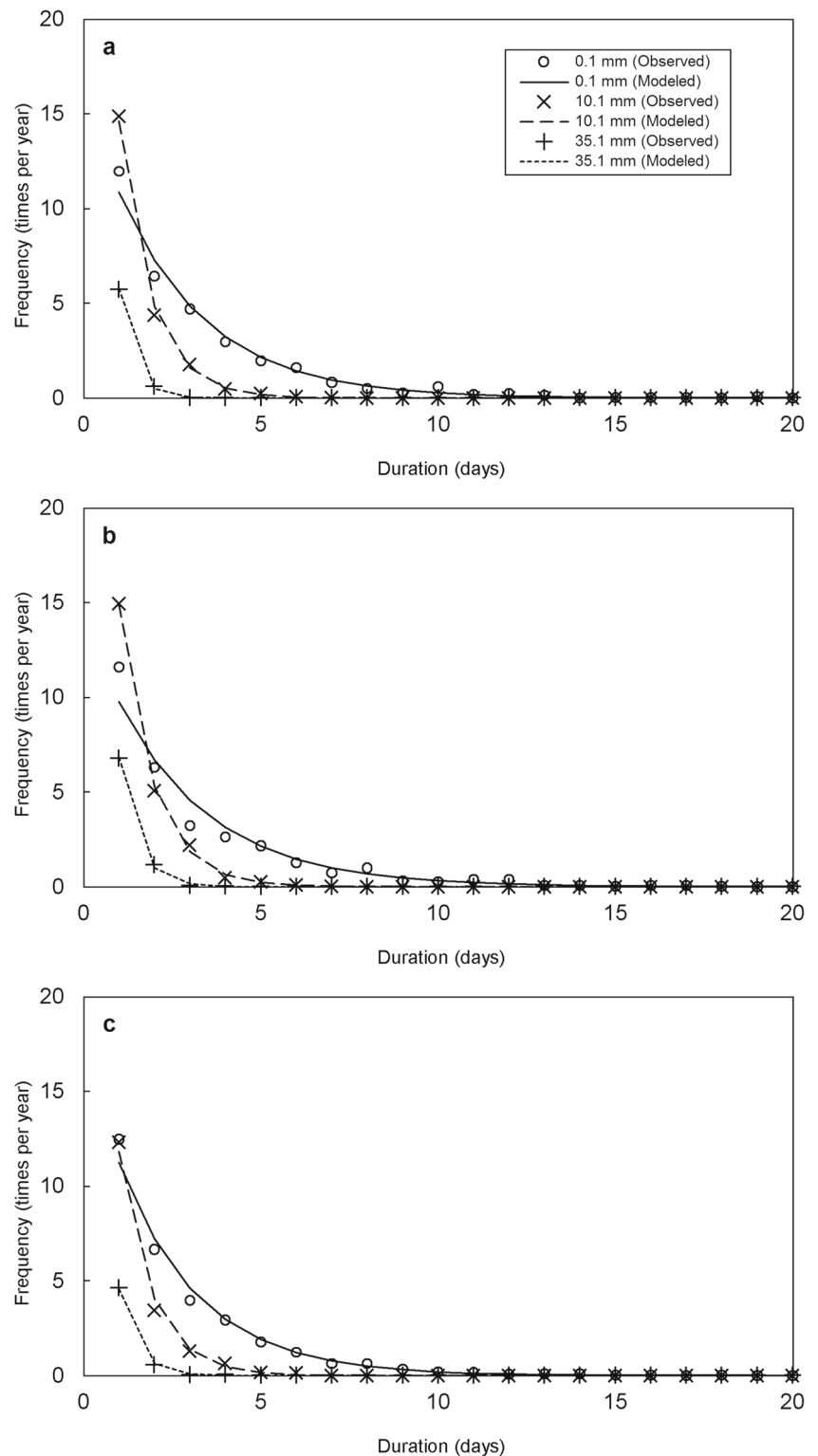
The differences between the transitions from 10.1–35.0 mm rainfall days and at least 35.1 mm rainfall days can rarely be seen since the counts of the transitions within at least 10.1 mm rainfall days, irrespective of whether 10.1–35.0 mm or at least 35.1 mm rainfall days, are all high and

the counts of the transitions from 0.1–10.0 mm rainfall days to at least 10.1 mm rainfall days are comparatively low. Hence, the chi-square test of independence was performed again with the exclusion of the transition from 0.1–10.0 mm rainfall days (i.e., the first row of Table 6) to investigate the differences between the transitions from 10.1–35.0 mm rainfall days and at least 35.1 mm rainfall days.

At this step, the differences between the transitions from 10.1–35.0 and at least 35.1 mm rainfall days can be seen clearly over the joint area. Table 7 shows the results at the stations over the joint area (351201, 376203, and 376201). At these three stations, the test values from the chi-square test of independence suggest the rejection of the null hypothesis under the 0.05 significance level. The residual analysis suggests that the count of transitions within at least 35.1 mm rainfall days is higher than the expected value, while the count of transitions from at least 35.1 mm rainfall days to no rainfall days is lower than the expected value at stations 351201 and 376203. For station 376201, probably because of its location, which is closer to the plain, the residual analysis does not show significant difference between any of the observed and expected counts of transitions. However, in comparison with stations outside the joint area (discussed later), station 376201 still has a higher count of transitions within at least 35.1 mm rainfall days and a lower count of transitions from at least 35.1 mm rainfall days to no rainfall days. Hence, we may conclude that consecutive days with at least 35.1 mm of rainfall are evident over the joint area.

Over the mountainous region, the differences between the transitions from 10.1–35.0 and at least 35.1 mm rainfall days are not obvious. The test values from the chi-square test of independence suggest the acceptance of the null hypothesis at all stations, except station 327501. For station 327501, even though the test value suggests the rejection of the null hypothesis, the test value is not high in comparison with the stations over the joint area, and the residual analysis does not suggest any difference between the observed and expected counts of transitions. Hence, we may conclude that consecutiveness of days with at least 35.1 mm

Fig. 2 Comparison of observed and modeled frequencies of consecutive days with at least 0.1, 10.1, and 35.1 mm of rainfall at stations **a** 331201, **b** 351201, and **c** 376201



of rainfall over the mountainous region is not as obvious as that over the joint area. Table 8 shows the results at sample stations over the mountainous region (327501, 331201, and 328201). The results at other stations (329201 and 330201) are similar stations 331201 and 328201.

Over the plain area, as in the mountainous region, the differences between the transitions from 10.1–35.0 and at least 35.1 mm rainfall days is not obvious. The test values from the chi-square test of independence suggest the acceptance of the null hypothesis at all stations, except station 380201.

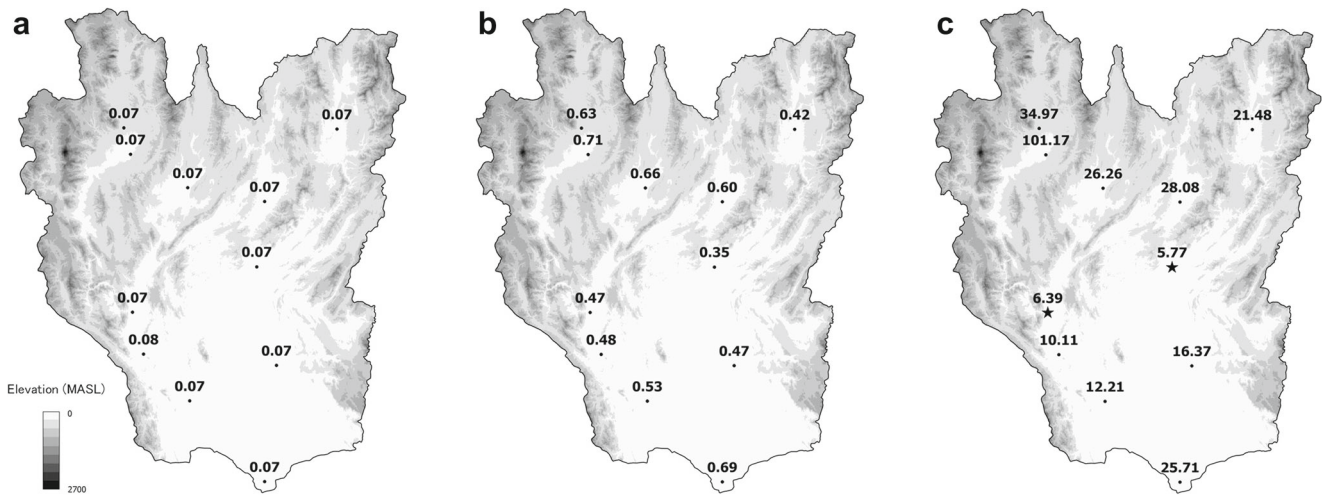


Fig. 3 Return period (years) for days with **a** at least 0.1 mm of rainfall, **b** at least 10.1 mm of rainfall, and **c** at least 35.1 mm of rainfall that are consecutive for more than 2 days. For **c**, stars indicate stations where the comparison between the transitions from days with

10.1–35.0 mm of rainfall and from days with at least 35.1 mm of rainfall show that days with at least 35.1 mm of rainfall are obviously consecutive according to the residual analysis at the significance level of 0.05

For station 380201, probably because its location is close to station 376201 in the joint area, the test value and the result of the residual analysis are similar to station 376201 (see Table 7). Hence, this station might have some characteristics of consecutive days with at least 35.1 mm of rainfall similar to the joint area. From Figs. 3 and 4, it can also be noted that the return period of days with at least 35.1 mm of rainfall at station 380201 is shorter in comparison with the other stations over the plain area (378201 and 400201), but that the return period is not as short as for the stations over

the joint area (351201, 376203, and 376201). Hence, we can conclude that the consecutiveness of days with at least 35.1 mm of rainfall over the plain area is not as obvious as over the joint area. Table 9 shows the results for the stations over the plain area (378201, 380201, and 400201).

4.2 Spatial distribution of rainfall amount

From Section 4.1, the Markov chain probability model and the chi-square test of independence have shown that days

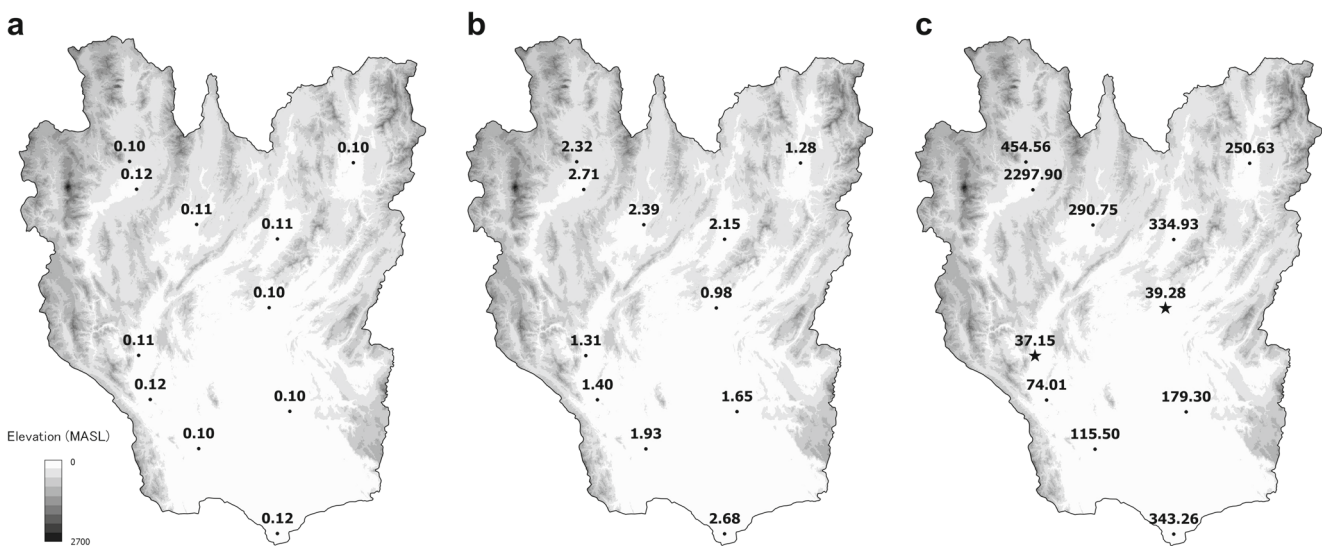


Fig. 4 Return period (years) for days with **a** at least 0.1 mm of rainfall, **b** at least 10.1 mm of rainfall, and **c** at least 35.1 mm of rainfall that are consecutive for more than 3 days. For **c**, stars indicate stations where the comparison between the transitions from days with

10.1–35.0 mm of rainfall and from days with at least 35.1 mm of rainfall show that days with at least 35.1 mm of rainfall are obviously consecutive according to the residual analysis at the significance level of 0.05

Table 5 Counts of transitions from 0.0, 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days to 0.0, 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days, and test values from the chi-square test of independence at sample stations (331201, 376203, and 400201)

Previous day rainfall (mm)	Latter day rainfall (mm)											
	Station 331201				Station 376203				Station 400201			
	0	0–10	10–35	>35	0	0–10	10–35	>35	0	0–10	10–35	>35
0	1596 ⁺ (1202)	707 ⁻ (916)	215 ⁻ (358)	55 ⁻ (97)	1884 ⁺ (1436)	731 ⁻ (990)	158 ⁻ (291)	42 ⁻ (98)	1596 ⁺ (1314)	780 ⁻ (943)	238 ⁻ (326)	79 ⁻ (110)
0–10	764 ⁻ (919)	819 ⁺ (700)	305 (274)	79 (74)	750 ⁻ (991)	904 ⁺ (683)	228 (201)	60 (68)	799 ⁻ (942)	784 ⁺ (676)	264 ⁺ (234)	83 (79)
10–35	182 ⁻ (360)	346 ⁺ (274)	186 ⁺ (107)	56 ⁺ (29)	153 ⁻ (291)	229 ⁺ (200)	131 ⁺ (59)	57 ⁺ (20)	229 ⁻ (327)	275 ⁺ (235)	120 ⁺ (81)	46 ⁺ (27)
>35	36 ⁻ (98)	93 ⁺ (75)	63 ⁺ (29)	18 ⁺ (8)	29 ⁻ (98)	76 (67)	54 ⁺ (20)	33 ⁺ (7)	70 ⁻ (111)	93 (79)	47 ⁺ (28)	17 ⁺ (9)
Test value	587.62 *				873.87 *				269.50 *			

States of 0.0, 0.1–10.0, 10.1–35.0, and at least 35.1 mm of rainfall are written as 0, 0–10, 10–35, and >35, respectively. The observed counts are shown outside the parentheses and the expected counts under the null hypothesis of the chi-square test of independence are shown inside the parentheses

⁺The count is significantly higher than the expected value according to the residual analysis under the 0.05 significance level

⁻The count is significantly lower than the expected value according to the residual analysis under the 0.05 significance level

*The test value suggests the rejection of the null hypothesis of the chi-square test under the 0.05 significance level

with at least 10.1 mm of rainfall tend to be consecutive over the entire area and days with at least 35.1 mm of rainfall are consecutive over the joint area; however, they cannot show differences in the characteristics of consecutive rainfall days between the plain area and the mountainous region. Hence, in this section, we investigate the rainfall amount on days with at least 10.1 mm of rainfall at each station

to identify whether there are some different characteristics of consecutive rainfall days between the plain area and the mountainous region. Figure 5 shows the average rainfall amount on days with at least 10.1 mm of rainfall at each station, as well as the significance of the difference from the areal average value determined by the Student's *t* test under the 0.05 significance level. Among the five stations

Table 6 Counts of transitions from 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days to 0.0, 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days, and test values from the chi-square test of independence at sample stations (331201, 376203, and 400201)

Previous day rainfall (mm)	Latter day rainfall (mm)											
	Station 331201				Station 376203				Station 400201			
	0	0–10	10–35	>35	0	0–10	10–35	>35	0	0–10	10–35	>35
0–10	764 ⁺ (655)	819 (840)	305 ⁻ (370)	79 ⁻ (102)	750 ⁺ (669)	904 (868)	228 ⁻ (297)	60 ⁻ (108)	799 (750)	784 (786)	264 (294)	83 (100)
10–35	182 ⁻ (257)	346 (329)	186 ⁺ (145)	56 ⁺ (40)	153 ⁻ (196)	229 (255)	131 ⁺ (87)	57 ⁺ (32)	229 (260)	275 (273)	120 (102)	46 (35)
>35	36 ⁻ (70)	93 (90)	63 ⁺ (39)	18 ⁺ (11)	29 ⁻ (66)	76 (86)	54 ⁺ (29)	33 ⁺ (11)	70 (88)	93 (93)	47 ⁺ (35)	17 (12)
Test value	111.09 *				192.66 *				30.35 *			

States of 0.0, 0.1–10.0, 10.1–35.0, and at least 35.1 mm of rainfall are written as 0, 0–10, 10–35, and >35, respectively. The observed counts are shown outside the parentheses and the expected counts under the null hypothesis of the chi-square test of independence are shown inside the parentheses

⁺The count is significantly higher than the expected value according to the residual analysis under the 0.05 significance level

⁻The count is significantly lower than the expected value according to the residual analysis under the 0.05 significance level

*The test value suggests the rejection of the null hypothesis of the Chi-square test under the 0.05 significance level

Table 7 Counts of transitions from 10.1–35.0 and at least 35.1 mm rainfall days to 0.0, 0.1–10.0, 10.1–35.0, and at least 35.1 mm rainfall days, and test values from the chi-square test of independence at stations over the joint area (351201, 376203, and 376201)

Previous day rainfall (mm)	Latter day rainfall (mm)											
	Station 351201				Station 376203				Station 376201			
	0	0–10	10–35	>35	0	0–10	10–35	>35	0	0–10	10–35	>35
10–35	209 (168)	317 (320)	188 (194)	72 (84)	153 (136)	229 (228)	131 (138)	57 (67)	175 (164)	258 (253)	144 (154)	55 (62)
>35	47 ⁻ (68)	120 (117)	77 (71)	42 ⁺ (30)	29 ⁻ (46)	76 (77)	54 (47)	33 ⁺ (23)	36 (47)	68 (73)	54 (44)	25 (18)
Test value	15.98 *				16.14 *				10.20 *			

States of 0.0, 0.1–10.0, 10.1–35.0, and at least 35.1 mm of rainfall are written as 0, 0–10, 10–35, and >35, respectively. The observed counts are shown outside the parentheses and the expected counts under the null hypothesis of the chi-square test of independence are shown inside the parentheses

⁺The count is significantly higher than the expected value according to the residual analysis under the 0.05 significance level

⁻The count is significantly lower than the expected value according to the residual analysis under the 0.05 significance level

*The test value suggests the rejection of the null hypothesis of the Chi-square test under the 0.05 significance level

Table 8 Same as Table 7 but for sample stations over the mountainous region (327501, 331201, and 328201)

Previous day rainfall (mm)	Latter day rainfall (mm)											
	Station 327501				Station 331201				Station 328201			
	0	0–10	10–35	>35	0	0–10	10–35	>35	0	0–10	10–35	>35
10–35	202 (202)	362 (350)	160 (167)	31 (36)	182 (171)	346 (345)	186 (196)	56 (58)	212 (205)	292 (292)	146 (148)	34 (39)
>35	42 (42)	60 (72)	42 (35)	12 (7)	36 (47)	93 (94)	63 (53)	18 (16)	39 (46)	66 (66)	36 (34)	14 (9)
Test value	7.95 *				5.73				5.29			

*The test value suggests the rejection of the null hypothesis of the chi-square test under the 0.05 significance level

Table 9 Same as Table 7 but for stations over the plain area (i.e., 378201, 380201, and 400201)

Previous day rainfall (mm)	Latter day rainfall (mm)											
	Station 378201				Station 380201				Station 400201			
	0	0–10	10–35	>35	0	0–10	10–35	>35	0	0–10	10–35	>35
10–35	240 (232)	365 (358)	164 (176)	59 (63)	267 (251)	329 (326)	144 (156)	56 (63)	229 (223)	275 (275)	120 (125)	46 (47)
>35	59 (67)	97 (104)	63 (51)	22 (18)	62 (78)	98 (101)	60 (48)	26 (19)	70 (76)	93 (93)	47 (42)	17 (16)
Test value	6.50				11.03 *				1.37			

*The test value suggests the rejection of the null hypothesis of the chi-square test under the 0.05 significance level

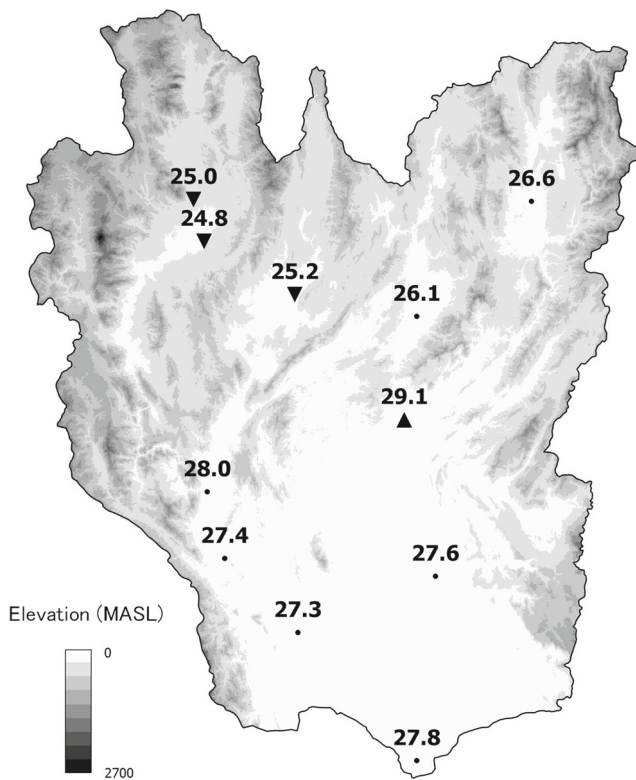


Fig. 5 Average rainfall amount on days with at least 10.1 mm of rainfall. An upward arrow represents a station where the rainfall amount is higher than the average, and a downward arrow represents a station where the rainfall amount is lower than the average, according to the Student's t test with the 0.05 significance level

in the mountainous region (i.e., 327501, 331201, 329201, 328201, and 330201), there are three stations with significantly low rainfall amounts (327501, 329201, and 328201), while the other two stations (331201 and 330201) have rainfall amounts close to the average. For the plain area, the rainfall amount is close to the average at each station (378201, 380201, and 400201). However, the average rainfall amount over the mountainous region is less than 27.0 mm for each station, while it lies between 27.0 and 28.0 mm at each station over the plain area. Hence, it can be concluded that rainfall amounts on days with at least 10.1 mm of rainfall are lower over the mountainous region than over the plain area.

5 Conclusions

Many flood events in the Chao Phraya basin coincide with consecutive rainfall days, and many previous studies have suggested that the variations of rainfall characteristics are due to topography. Therefore, this study performed a statistical analysis to investigate the variations of the characteristics of consecutive rainfall days over the upstream

part of the Chao Phraya basin, which encompasses a mountainous region in the north and plain area in the south. The Markov chain probability model has proved capable of describing the probability of consecutive rainfall days accurately. Using the Markov chain probability model in conjunction with the chi-square test of independence, the consecutiveness of days with at least 10.1 mm of rainfall was determined for the entire area. This showed that consecutiveness of days with at least 35.1 mm of rainfall was evident over the joint area, where the mountainous region meets the plain area. Furthermore, the Student's t test indicated that the mountainous region has a lower amount of rainfall than the plain area on days with at least 10.1 mm of rainfall. Hence, we can describe the characteristics of variation of consecutive rainfall days as follows.

1. Over the plain area, days with at least 10.1 mm of rainfall are consecutive.
2. Over the mountainous region, days with at least 10.1 mm of rainfall are also consecutive, but the rainfall amount on those days is not as high as over the plain area.
3. Over the joint area, days with at least 10.1 mm of rainfall are also consecutive, but days with at least 35.1 mm of rainfall are consecutive more evidently than over other areas.

These findings are useful for flood management because the different characteristics of consecutive rainfall days can contribute differently to the flood waters in the Chao Phraya basin depending on location. This study suggested that rainfall over the joint area should be monitored closely and studied in greater detail because of the high probability of consecutive days of heavy rainfall in this area. Apart from flood management, many others hazards such as landslides, which also occur frequently in this area, can also be triggered by days of consecutive rainfall. Hence, the results of this study could also be applied to mitigate other rainfall-related geohazards.

Acknowledgements The authors are thankful to the Thai Meteorological Department for the rainfall data and to the Hydro and Agro Informatics Institute for the information on flooding in Thailand. We are also grateful for the financial support from a Japanese Government Scholarship.

References

- Dahale SD, Panchawagh N, Singh SV, Ranatunge ER, Brikshavana M (1994) Persistence in rainfall occurrence over tropical South-east Asia and equatorial Pacific. *Theor Appl Climatol* 49(1):27–39
- Dastidar AG, Ghosh D, Dasgupta S, De UK (2010) Higher order Markov chain models for monsoon rainfall over West Bengal, India. *Indian J Radio Space Phys* 39(1):39–44
- Hiscott RN (1981) Chi-square tests for Markov chain analysis. *Math Geol* 13(1):69–80

- Hossain MM, Anam S (2012) Identifying the dependency pattern of daily rainfall of Dhaka station in Bangladesh using Markov chain and logistic regression model. *Agric Sci* 3(3):385–391
- Hydro and Agro Informatics Institute (2007) Studying of rainfall amount for the determination of an appropriate criteria for flood warning system. Hydro and Agro Informatics Institute, Thailand (in Thai)
- Hydro and Agro Informatics Institute (2008a) Summary of Thailand water situation in 2548 and 2549 BE. Hydro and Agro Informatics Institute, Thailand (in Thai)
- Hydro and Agro Informatics Institute (2008b) Summary of Thailand water situation in 2550 BE. Hydro and Agro Informatics Institute, Thailand (in Thai)
- Hydro and Agro Informatics Institute (2009) Summary of Thailand water situation in 2551 BE. Hydro and Agro Informatics Institute, Thailand (in Thai)
- Kuraji K, Gomyo M, Punyatrong K (2009) Inter-annual and spatial variation of altitudinal increase in rainfall over Mount Inthanon and Mae Chaem Watershed, Northern Thailand. *Hydrol Res Lett* 3:18–21
- Lee SK, Mapes BE, Wang C, Enfield DB, Weaver SJ (2014) Spring-time ENSO phase evolution and its relation to rainfall in the continental US. *Geophys Res Lett* 41(5):1673–1680
- Mahavik N, Satomura T, Shige S, Sysouphanthavong B, Phonevilay S, Wakabayashi M, Baimoung S (2014) Rainfall pattern over the middle of Indochina Peninsula during 2009–2010 summer monsoon. *Hydrol Res Lett* 8(1):57–63
- Moon SE, Ryoo SB, Kwon JG (1994) A Markov chain model for daily precipitation occurrence in South Korea. *Int J Climatol* 14(9):1009–1016
- Okumura K, Satomura T, Oki T, Khantiyanan W (2003) Diurnal variation of precipitation by moving mesoscale systems: radar observations in northern Thailand. *Geophys Res Lett* 30(20):2073
- Prajamwong S, Suppataratarn P (2007) Integrated flood mitigation management in the lower Chao Phraya river basin. Thailand Development Research Institute, Thailand
- Raju PVS, Mohanty UC, Rao PLS, Bhatla R (2002) The contrasting features of Asian summer monsoon during surplus and deficient rainfall over India. *Int J Climatol* 22(15):1897–1914
- Sheskin DJ (1996) Handbook of parametric and nonparametric statistical procedures. CRC Press, USA
- Sonnadara DUJ, Jayewardene DR (2015) A Markov chain probability model to describe wet and dry patterns of weather at Colombo. *Theor Appl Climatol* 119(1–2):333–340
- Szyniszewska AM, Waylen PR (2012) Determining the daily rainfall characteristics from the monthly rainfall totals in Central and Northeastern Thailand. *Appl Geog* 35(1–2):377–393
- Takahashi HG (2010) Seasonal changes in diurnal rainfall cycle over and around the IndoChina Peninsula observed by TRMM-PR. *Adv Geosci* 25:23–28
- Takahashi HG, Fujinami H, Yasunari T, Matsumoto J (2010) Diurnal rainfall pattern observed by tropical rainfall measuring mission precipitation radar (TRMM-PR) around the indochina peninsula. *J Geophys Res* 115(D7):D07109
- Thai Meteorological Department (2012) The climate of Thailand. Thai Meteorological Department, Thailand
- The World Bank (2012) Thai flood 2011: rapid assessment for resilient recovery and reconstruction planning. The World Bank, Thailand
- Weiss LL (1964) Sequences of wet and dry days described by a Markov chain probability model. *Mon Weather Rev* 92(4):169–176
- Yokoi S, Satomura T (2008) Geographical distribution of variance of intraseasonal variations in western IndoChina as revealed from radar reflectivity data. *J Climate* 21(19):5154–5161