Probability Distribution of Summer Daily Precipitation in the Huaihe Basin of China Based on Gamma Distribution

LIANG Li[1,2](梁 莉), ZHAO Linna[1,2*](赵琳娜), GONG Yuanfa[3](巩远发), TIAN Fuyou[2](田付友), and WANG Zhi[1,2](王 志)

1 *Public Meteorological Service Center, China Meteorological Administration, Beijing* 100081

2 *National Meteorological Center, China Meteorological Administration, Beijing* 100081

3 *Chengdu University of Information Technology, Chengdu* 610225

(Received September 30, 2010; in final form December 8, 2011)

ABSTRACT

The probability distribution of precipitation in the Huaihe basin (HB) is analyzed with the shape and scale parameters of a Gamma distribution. The summer daily precipitation records of 158 meteorological raingauges are applied over the HB during the period of 1980–2007, and the precipitation samples are classified into unconditional rainy days and conditional rainy days which have a dry or wet preceding day over the years. The Kolmogorov-Smirnov (K-S) test and the comparison between the Gamma distribution probability density function and the sample frequency of daily precipitation records of five representative stations are conducted and analyzed.

The results show that the HB is a “scale-dominated” region characterized with large scale parameters of the Gamma distribution, where rainfall is likely to exhibit large variability leading to extreme wet or dry conditions. Further analysis shows that the confluence area of Sha River and Ying River within the stream between Wangjiaba dam and Bengbu station, the eastern branch of the Huaihe River (HR) between Bengbu station and the Hongze Lake, and the downstream area below the Hongze Lake, are all the areas with a high probability of large precipitation under the condition that the rainy day has a dry preceding day. The eastern part of the Yishu River watershed and the region near Wangjiaba dam are the center of a high probability of large precipitation under the condition that the rainy day has a wet preceding day.

Moreover, the day following a dry preceding day has a greater probability of small rainfall. The probability distribution of summer daily precipitation of the HB is significantly skewed. The probability distribution could be more applicable if the rainy days are preceded by a dry or wet day.

Key words: daily precipitation, wet/dry preceding day, Gamma distribution, probability distribution, Huaihe basin

Citation: Liang Li, Zhao Linna, Gong Yuanfa, et al., 2012: Probability distribution of summer daily precipitation in the Huaihe basin of China based on Gamma distribution. *Acta Meteor. Sinica*, **26**(1), 72–84, doi: 10.1007/s13351-012-0107-2.

1. Introduction

The Huaihe basin (HB) is located in the eastern part of China between the Yellow River and the Yangtze River, composed of the Huaihe River (HR), Yihe River, Shuhe River, and Sihe River. The HR covers the area 31°–35°N, 112°–121°E. It flows through four provinces, namely, Henan, Anhui, Jiangsu, and Shandong, and its watershed area is about 2.7×10^5 km^2. As the weather and climate conditions are complex in the basin, both drought and flood disasters happen frequently there, causing serious damages not only to human life and property but also to the sustainable development of the region. Since the 1950s, major floods have happened in the HB, such as those in 1954, 1991, and 2003. The rainstorms that occurred in June and July 2007 over the HB gave rise to severe floods in the region (Zhao et al., 2007).

Supported by the China Meteorological Administration Special Public Welfare Research Fund (GYHY201006037, GYHY200906007, and GYHY(QX)2007-6-1)

*Corresponding author: zhaoln@cma.gov.cn.

Though floods always occur in summer and autumn there, rainfalls mainly gather in summer. The precipitation from June to August in the HB is about 53% of the annual total precipitation (Cheng, 2004). The mean annual total precipitation and runoff depth are approximately 888 and 240 mm, respectively. The precipitation spatial and temporal distributions are very irregular and change from year to year. This is attributed to the fact that the basin is located in the transitional area between the monsoon climate to its south and the continental climate to its north. The precipitation from June to August in the HB determines whether drought or flood will occur and their severity degrees. It is necessary to obtain further understanding of the characteristic of summer precipitation distribution over the HB so as to provide climatic background information for Quantitative Precipitation Forecast (QPF) and water and agriculture resources management.

In the statistical studies of precipitation distribution, many distribution models are often used, such as the exponential, Gamma, Kappa, and Weibull distributions. Much research has been carried out related to the spatial distribution of precipitation. Todorovic (1975) assumed an exponential distribution of daily precipitation, but the results showed that the exponential distribution does not fit very well the probability distribution of daily precipitation. Katz (1977) proposed a model that considers precipitation as a chain-dependent process, which is more agreeable to the realistic situation than the independent random process, and this model can reproduce the probability distribution of maximum daily precipitation in a number of days. Woolhiser (1992) reviewed the studies applying various normalizing transforms, as well as the Kappa and Weibull distributions. Gregory et al. (2007) used the Gamma distribution to represent monthly rainfall in Africa. They found that the Gamma distribution is well suited to the fitting of precipitation data and its function formulation is simpler than that of the Weibull distribution, which leads to reduced computation time.

In China, the research on temporal and spatial distributions of precipitation probability using different methods has also progressed. Zhang (1990) obtained the theoretical expressions for the distribution function of the total amount of precipitation and the daily maximum precipitation based on the Gamma probability distribution. Cheng et al. (2003) carried out a model fitting test, which tested the probability of precipitation at different time intervals, such as 10 days, a month or a year, in Henan Province by using various theoretical probability distribution models, such as normal, quasi-normal, and skewed Gamma distributions. Wu et al. (2004) used the Gamma distribution to estimate the conditional probability density functions of the daily precipitation on the rainy day after a dry or wet preceding day in Guangxi Autonomous Region. The results show that the Gamma distribution is well suited to the sampling of daily precipitation frequency distribution given a wet or a dry preceding day. Su et al. (2007) analyzed the probability distribution pattern of the sequence of extreme precipitation over the Yangtze River and proved that the Wakeby distribution function can simulate the probability distribution of extreme precipitation. Zhang and Wei (2009) used the Generalized Extreme Value (GEV) distribution and other statistical diagnoses to study the probability distribution of extreme precipitation of southern China in the recent 46 years.

The above studies focused more on application of various statistical methods to analysis of precipitation probability distribution over some areas of China, excluding the HB. The statistical characteristics of the spatial and temporal distributions of precipitation probability in the HB have not been well studied. Though the precipitation distribution characteristics over the HB have been studied by Zhu (1998), Liang et al. (2001), and Cheng (2004), these studies mostly focused on the spatial and temporal variability of precipitation (rather than precipitation probability) in this region. Considering the important role that the HB plays in controlling flood and relieving drought, a better understanding of the features of spatial and temporal variability in precipitation probability could provide decision support to water and agriculture resources management (Zhao et al., 2010) over this region.

There are many probability distributions that could be successfully utilized to parameterize rainfall distributions. This paper makes use of the Gamma distribution to study the summer daily precipitation probability in the HB over the past 28 years.

The study intends to reveal the feature of the precipitation probability distribution in the flood season of the HB. It will help to achieve a better understanding of the precipitation pattern over the basin especially in the context of climate change. It will also serve as a scientific reference for daily precipitation forecasts by providing the necessary climate background.

2. Data and methodology

2.1 Data

A total of 158 raingauge stations with daily precipitation records are available in the HB. Figure 1 illustrates the raingauge locations. This study focuses on summer precipitation in the basin. The summer season is defined from 31 May to 31 August each year. The daily precipitation is the accumulative precipitation for 24 h from 1200 UTC of the previous day to 1200 UTC of the present day. The summer precipitation data from 1980 to 2007 are used. A number of gauges with incomplete records are excluded.

2.2 Methodology

Unlike other meteorological variables such as temperature and sea level pressure, precipitation does not obey the normal distribution. In the variety of stochastic variables in hydro-meteorological prediction, precipitation is the most critical variable. Although it exhibits obvious skewed distributions, the specific types of the overall distribution are impossible to obtain. It is practical to understand stochastic processes of precipitation by using statistical models. Many researchers (Gregory et al., 2007; Zhang, 1990; Wu et al., 2004) showed that it is feasible to estimate stochastic processes of daily precipitation using the Gamma distribution. Therefore, we also use density function of the Gamma distribution to estimate the probability distribution of daily precipitation. Specifically, we adopt the model proposed by Katz (1977),

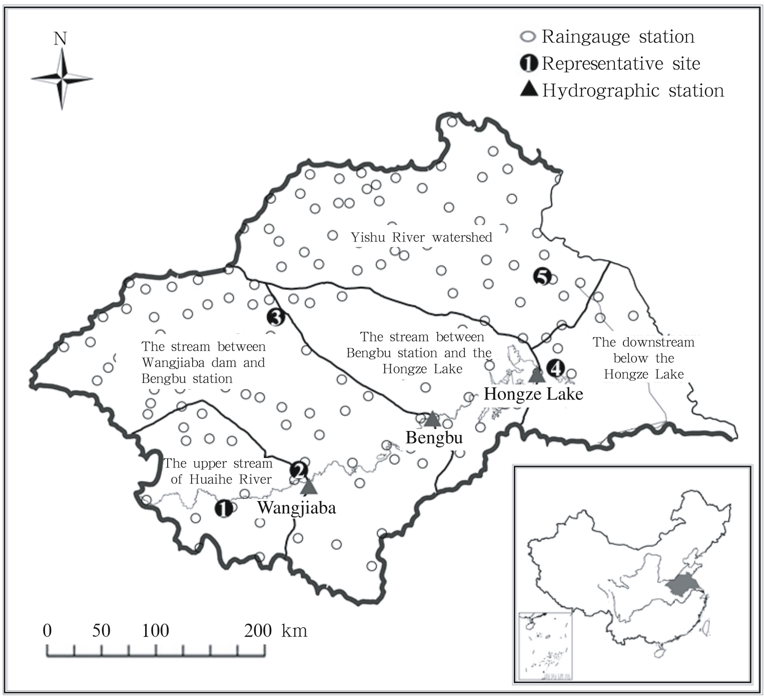

Fig. 1. Illustration of the catchments of the Huaihe basin and locations of the five raingauge stations in the test area (①, ②, ③, ④, and ⑤ represent Xixian, Fuyang, Shangqiu, Huaian, and Lianyungang raingauge sites, respectively).

which assumes that the probability distribution of daily precipitation is related to whether the day before the rainy day is dry or wet. The evolution of wet and dry day obeys the one order Markov chain model (sometimes called the one day-related links). This model is employed to estimate the probability distribution of precipitation on unconditional and conditional rainy days, given a dry or wet preceding day.

The Gamma distribution density function is shown in Eq. (1). The complete Gamma function is shown in Eq. (2) (Gregory et al., 2007).

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}\exp\Big(-\frac{x}{\beta}\Big), \quad x>0, \tag{1}$$

$$\Gamma(\alpha) = \int_0^{\infty} \mathrm{e}^{-1}t^{\alpha-1}\mathrm{d}t, \tag{2}$$

where α and β are the shape and scale parameters ($\alpha > 0$, $\beta > 0$), and x represents the daily precipitation. The shape parameter is a measure of the skewness of the distribution. As the shape value increases, the distribution curve becomes more symmetrical. When $\alpha < 1$, the shape of the curve is like a reverse "J" and the value of probability density is the maximum when x is near zero. The scale parameter is a measure of the steepness of the distribution. When α varies little and β varies a lot, the smaller the β is, the more positively skewed the curve becomes.

The Gamma distribution parameters are estimated through the maximum likelihood estimation as follows (Gregory et al., 2007):

$$A = \ln(\overline{x}) - \frac{\sum_{i=1}^{n_p}\ln(x_i)}{n_p}, \tag{3}$$

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

$$\hat{\beta} = \frac{\overline{x}}{\hat{\alpha}}, \tag{5}$$

$$\hat{\alpha}\hat{\beta}^2 \approx s^2. \tag{6}$$

In Eq. (3), A is a measure of the skewness of the distribution, where x_i holds non-zero values, and $\overline{x}$ is the arithmetic mean of all the values of x_i. In Eq. (4), $\hat{\alpha}$ represents the value of A. Then, $\hat{\alpha}$ and $\hat{\beta}$ could be derived by Eqs. (5) and (6), that is, $\hat{\alpha} = \frac{\overline{x}^2}{s^2}$, $\hat{\beta} \approx \frac{s^2}{\overline{x}}$. In general terms, it means a wet area (i.e., one with a large mean) with a relatively small variance has a large estimated shape parameter, while a dry area with a relatively high variance results in a small one. Similarly, large scale values may indicate relatively high variance in the dry region (i.e., one with a low mean), while small scale values indicate relatively little variance in the wet region.

Compared with the normal distribution in which a single parameter can directly provide an intuitive understanding of some aspects of the distribution, the Gamma distribution requires that both the shape and scale parameters are interpreted together. In regions with the same shape parameter but different scale parameters, the probability density function will differ. The mean or standard deviation can describe intuitive characteristics of some aspects of the distribution.

Gregory et al. (2007) evaluated the parameter fields and found that areas receiving a minimal amount of rainfall are described by either a large shape parameter or a large scale parameter, but not large values for both parameters. The term "shape-dominated" is used to refer to the locations with a larger shape parameter, and "scale-dominated" is used to refer to the locations with a larger scale parameter. The "shape-dominated" regime describes a region where it is likely to have consistent rains but fewer extreme events. The "scale-dominated" regime describes a region where it is likely to have variable rains and more extreme rainfall events.

Since the probability distribution of daily precipitation is related to whether the preceding day is dry or wet, the stochastic process of precipitation could be described by a dual variable, that is, $((J_{n-1}, X_n)$, $n = 1, 2, 3 \ldots\ldots)$, where $J = 1$ represents a rainy day while $J = 0$ represents a rainless day. X_n is the precipitation of the nth day. The dual variable (J_{n-1}, X_n) expresses that the probability distribution of the nth day precipitation is related to whether the (n–1)th day is rainy. The start numbers of n for X_n and J_n are 0 and 1, respectively. Assuming the precipitation of rainy days under the condition that $J_{n-1} = 0$ or $J_{n-1} = 1$ obeys the Gamma distribution, the parameters could be divided into two groups in order to represent these two conditions, i.e., α_1 and β_1, α_2 and β_2. The probability density function under these two kinds of conditions could be derived from Eq. (1).

$$f_i(x) = \frac{1}{\beta_i^{\alpha}\Gamma(\alpha_i)}(\frac{x}{\beta_i})^{\alpha_i-1}\exp\Big(-\frac{x}{\beta_i}\Big), \tag{7}$$

$$\Gamma(\alpha_i) = \int_0^{\infty} \mathrm{e}^{-1}t^{\alpha_i-1}\mathrm{d}t. \tag{8}$$

The samples of rainy days of 158 sites in the HB are selected and classified into three types including unconditional rainy days and conditional rainy days which have a dry or wet preceding day. The parameters α_i and β_i as well as $\overline{x}_i$ and s_i^2 are estimated by the maximum likelihood estimation, that is, by solving Eqs. (5) and (6). Finally, the probability density function $f_i(x)$ is calculated within a certain interval according to Eqs. (7) and (8).

2.3 Goodness-of-fit test

Once the distribution parameters are estimated, their accuracy in approximating the real precipitation distribution must be confirmed. This can be done with the Kolmogorov-Smirnov (K-S) goodness-of-fit test. The K-S test compares the cumulative distribution functions of the theoretical distribution with the observed values and returns the maximum difference between these two cumulative distributions (Wilks, 1995). The maximum difference in these two cumulative distribution functions is referred to as the K-S statistic. In this paper, the K-S test is employed to examine the goodness-of-fit of the Gamma distribution. Specific steps could be found in Sun et al. (2001).

3. Results

3.1 Parameter analysis of the unconditional and conditional rainy days

Sixteen stations are selected as representative stations from five catchments respectively, namely, the upper stream of the HR, the stream between Wangjiaba dam and Bengbu station, the stream between Bengbu station and the Hongze Lake, the Yishu River watershed, and the downstream area of the HR below the Hongze Lake (Fig. 1). Table 1 lists the samples of the representative stations and their shape and scale parameters.

As shown in Table 1, the sample size of unconditional rainy days is about twice of that of the conditional rainy days in all the five catchments. Among the conditional samples of rainy days, the samples with a wet preceding day are obviously more than the ones

Table 1. Shape (α_i) and scale (β_i) parameters of the representative stations in the five catchments of the Huaihe basin

Catchment	Representative station	Unconditional rainy days			Conditional rainy days: Rainy days after a dry preceding day			Conditional rainy days: Rainy days after a wet preceding day		
		Total	α_0	β_0	Total	α_1	β_1	Total	α_2	β_2
Upper stream	Xinyang	1022	0.303	49.13	421	0.433	25.34	601	0.302	58.47
	Xixian	924	0.401	37.91	408	0.427	29.42	516	0.411	42.19
	Zhumadian	947	0.302	53.13	416	0.422	29.65	531	0.293	64.41
The stream between Wangjiaba dam and Bengbu station	Zhengzhou	854	0.398	28.58	405	0.318	30.86	449	0.48	26.67
	Zhoukou	909	0.383	37.74	420	0.363	35.95	489	0.404	38.82
	Fuyang	991	0.329	44.81	420	0.351	34.69	571	0.335	49.73
	Liuan	1060	0.395	33.98	424	0.494	20.37	636	0.401	39.03
The stream between Bengbu station and the Hongze Lake	Bengbu	989	0.383	37.81	435	0.348	30.33	554	0.437	40.12
	Shangqiu	841	0.381	34.6	417	0.356	35.07	424	0.406	34.12
	Suqian	949	0.379	39.98	436	0.376	36.12	513	0.388	42.45
The downstream area below the Hongze Lake	Huaian	1001	0.363	42.11	435	0.312	43.21	566	0.404	41.16
	Sheyang	994	0.438	35.6	440	0.462	30.47	554	0.432	38.83
Yishu River watershed	Zaozhuang	911	0.418	38.33	472	0.402	35.72	439	0.442	40.25
	Xuzhou	876	0.383	43.23	426	0.433	33.93	450	0.366	49.96
	Lianyungang	923	0.421	38.65	439	0.516	30.50	484	0.368	45.57
	Rizhao	935	0.38	35.37	452	0.46	26.05	483	0.351	42.17

with a dry preceding day. The values of the shape parameter of the representative stations in the basin are all less than one. This means that $f(x)$ is monotonically decreasing as x increases. In other words, heavy precipitation events are less likely to happen than light precipitation events.

Figure 2 illustrates that the shape parameter is one order of magnitude less than the scale parameter, indicating that the basin is a "scale-dominated" region, that is to say, this region is likely to have variable rains and more extreme events. During the period 1980–2007, flood or drought occurred about once every two years in the HB. The periods 1985–1988 and 1992–1999 were relatively dry. Severe drought occurred in 1999, and flood occurred in 1991, 2003, and 2007 (Liang et al., 2001; Bi et al., 2004b; Zhao et al., 2007). For example, the precipitation from 1 June to 22 July in 2003 exceeded the mean annual precipitation by about 50–100 mm in many parts of the basin. Extreme heavy precipitation (100–300 mm higher than the annual mean) was observed in northern Anhui, most of Henan, and southern Shandong provinces. Continuous heavy rain occurred in this basin in 2007, heavier than that in 2003, which caused severe flooding (Bi et al., 2004a, b).

In Fig. 3, values of the shape parameters of three rainy day types are all below 0.5. This indicates that the HB is not a "shape-dominated" region. The relationship between the shape parameters of the three rainy day types is investigated. Analysis of the differences between α_0 and α_1, α_0 and α_2, α_1 and α_2 (figures omitted) shows that these shape parameters of the three rainy day types have little difference. It can be seen clearly from Fig. 3 that the distribution of the shape parameter of conditional rainy days (Figs. 3b and 3c) presents more high-value areas. In Fig. 3b, the high value center of the shape parameter is located at the boundary of the stream between Bengbu station and the Hongze Lake, and the biggest shape parameter reaches 0.48. In Fig. 3c, the large shape parameter is centered on the boundary of the stream between Bengbu station and the downstream area below the Hongze Lake, and the center value is 0.46.

It is clear from Fig. 4 that the values of scale parameters of the HB are above 30. As the value is large, the distribution of precipitation probability shows large variability. The relationship between the scale parameters of the three rainy day types is also investigated. Analysis of the differences between β_0 and β_1, β_0 and β_2, β_1 and β_2 (figures omitted) shows that the scale parameters of such three types have large differences, with $\beta_1 < \beta_0 < \beta_2$. This means that the

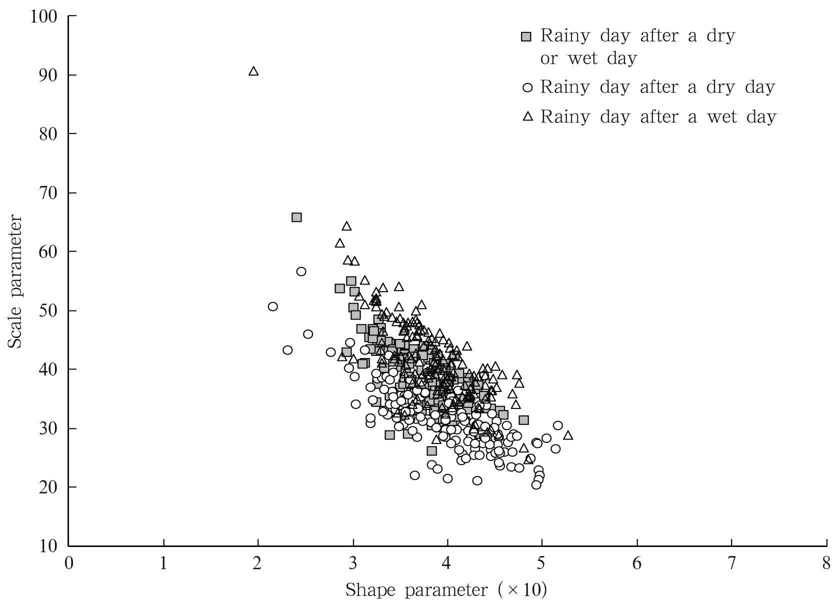

Fig. 2. Scatter plot of shape and scale parameters of the Gamma distribution used to fit the summer daily precipitation in the Huaihe basin.

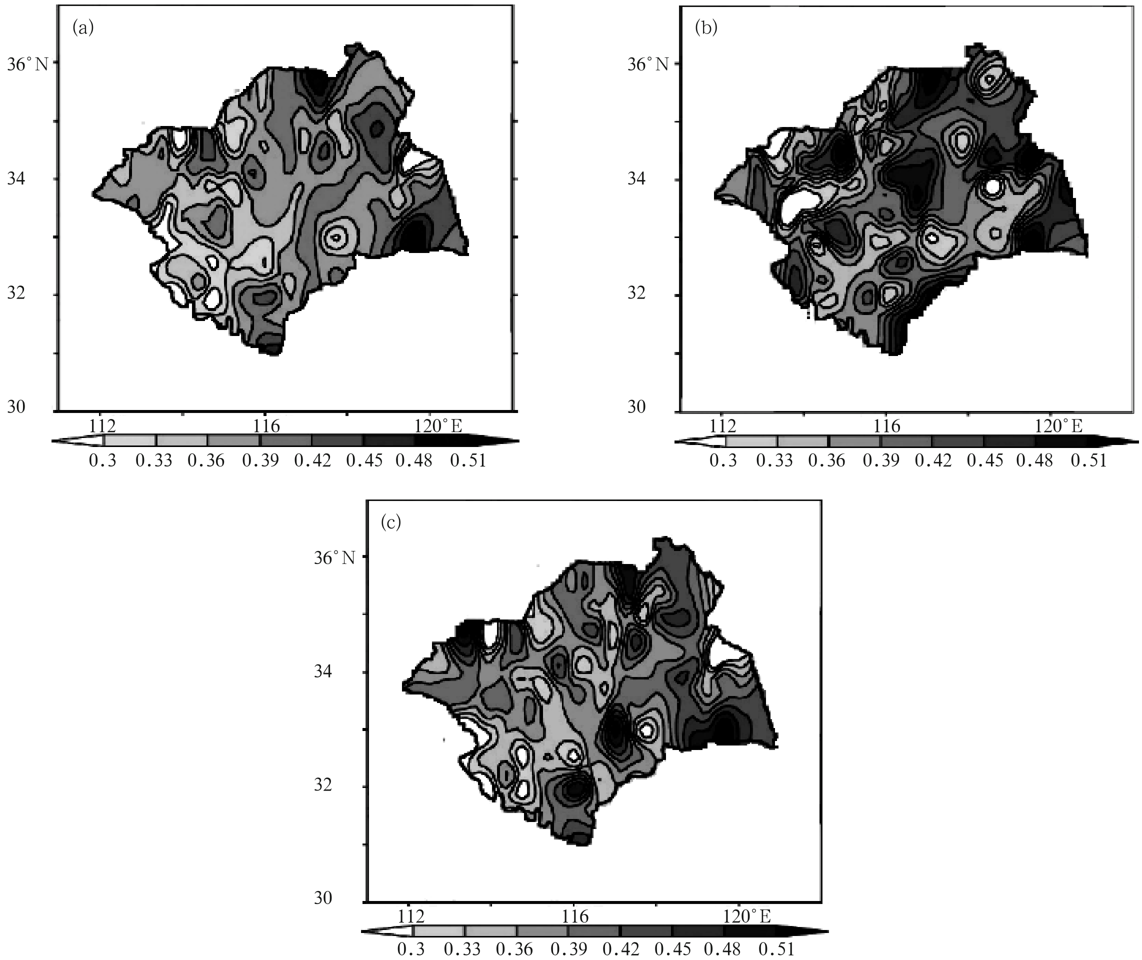

Fig. 3. Spatial distributions of shape parameters on (a) rainy days after a dry or wet preceding day, (b) rainy days after a dry preceding day, and (c) rainy days after a wet preceding day, respectively.

amount of precipitation of the rainy day with a dry preceding day is larger than that with a wet preceding day. The confluence area of Sha River and Ying River within the stream between Wangjiaba dam and Bengbu station, the eastern branch of the HR between Bengbu station and the Hongze Lake, and the downstream area below the Hongze Lake, all have a high probability of large precipitation under the condition that the rainy day has a dry preceding day. The eastern part of the Yishu River watershed and the region near the Wangjiaba dam are the center of high probability of large precipitation under the condition that the rainy day has a wet preceding day.

As continuous heavy rain is the main reason that causes the flood in the HB (Ding, 1993), the area with a high probability of large precipitation under the condition that the rainy day has a wet preceding day also becomes a potential flooding zone.

The variation in the distribution of large daily rainfall results from the variation of the atmospheric circulation. For example, according to the comparative analysis of large-scale atmospheric circulation features of the floods in 1991 and 2003 in the HB (Ding, 1993; Jiao et al., 2004), the main rainfall areas were concentrated between the Yangtze River and the HR (i.e., the area between Wangjiaba dam and Bengbu station and the area between Bengbu station and the Hongze Lake) in 1991, while the heavy rain was concentrated along the HR (i.e., the Yishu River watershed and the downstream area below the Hongze Lake) in 2003. This was mainly caused by the northward shift of the subtropical high ridge line compared with

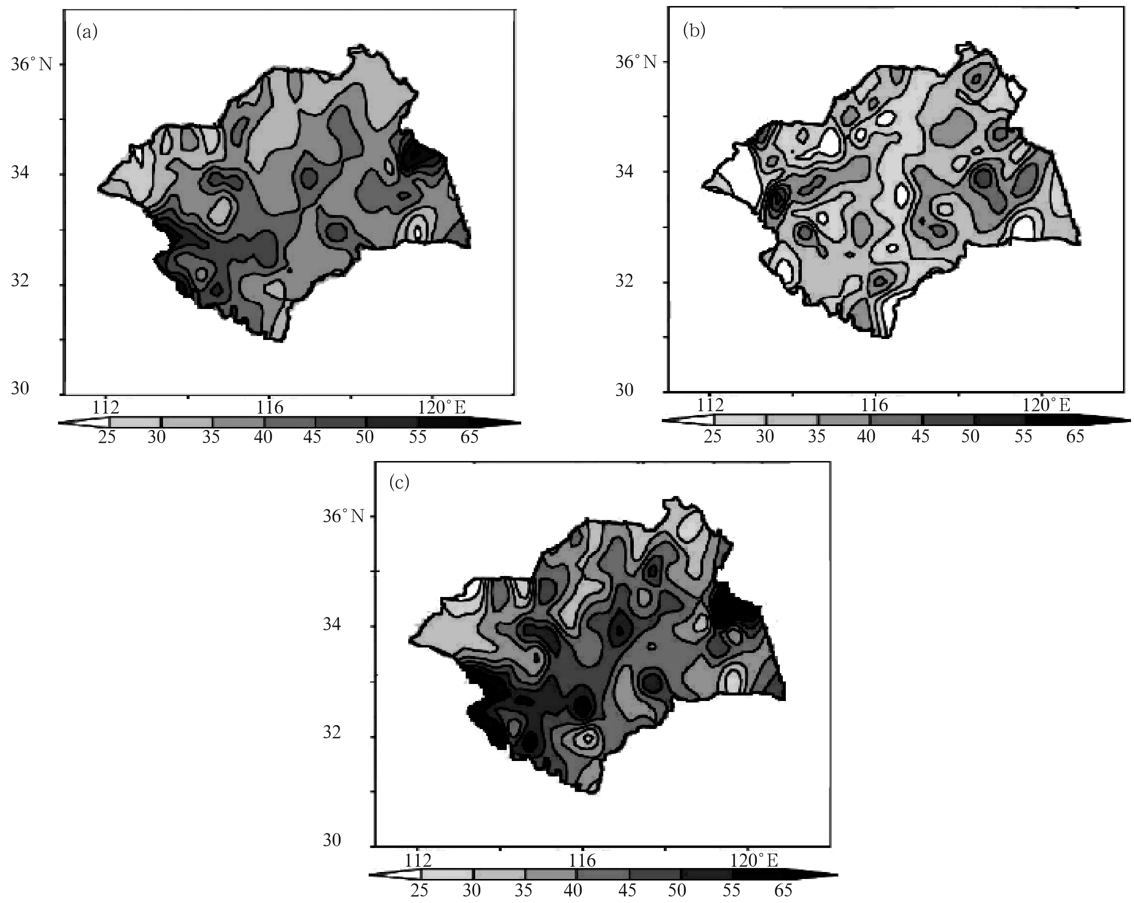

Fig. 4. As in Fig. 3, but for scale parameters.

its climatological average position. Many factors, such as the intensity and position of the western Pacific subtropical high (being relatively stable and southward, or actively advancing eastward or retreating westward with strong intensity), may cause the change of the distribution of large precipitation in the HB.

Moreover, Eqs. (5) and (6) indicate that the mean and the variance of daily precipitation are mainly determined by the scale parameters. Examination of the mean and variance distributions of the three types of rainy days (figures omitted) finds that the spatial distributions of mean and variance are consistent with those of the scale parameters of the Gamma distribution.

The analysis above gives a more useful statistical understanding on the pattern of the short-duration and high intensity precipitation in the HB, which may help the authorities to make better decision during flood seasons.

3.2 Probability distributions of daily precipitation of unconditional and conditional rainy days

A fundamental objective of this study is to establish the link between statistical probability distribution of precipitation and extreme precipitation events, which could help with rainfall prediction. The probability distribution of daily precipitation of certain type of rainy days can be created by using the estimated shape and scale parameters of the Gamma distribution. Five raingauge sites, i.e., Xixian, Fuyang, Shangqiu, Huaian, and Lianyungang, are selected as representative stations from the five catchments (mentioned in Section 3.1), respectively.

To evaluate how accurate the probability density function of Gamma distribution is, as compared with the real situation, here we derive the sample frequency curve of daily precipitation records under the corresponding condition. First, calculating the frequency that the precipitation of the conditional rainy days falls into the interval (0, 5], (5, 10], …, (145, 150], respectively. Then, the frequency divided by the interval width 5.0 is taken as the sample frequency density. The results are compared with the probability density function of Gamma distribution in quantity. Accordingly, the results of $f_1(x)$ and $f_2(x)$ are calculated, where $x = 2.5, 7.5, \ldots, 147.5$, corresponding to the range of the sample frequency (Wu et al., 2004).

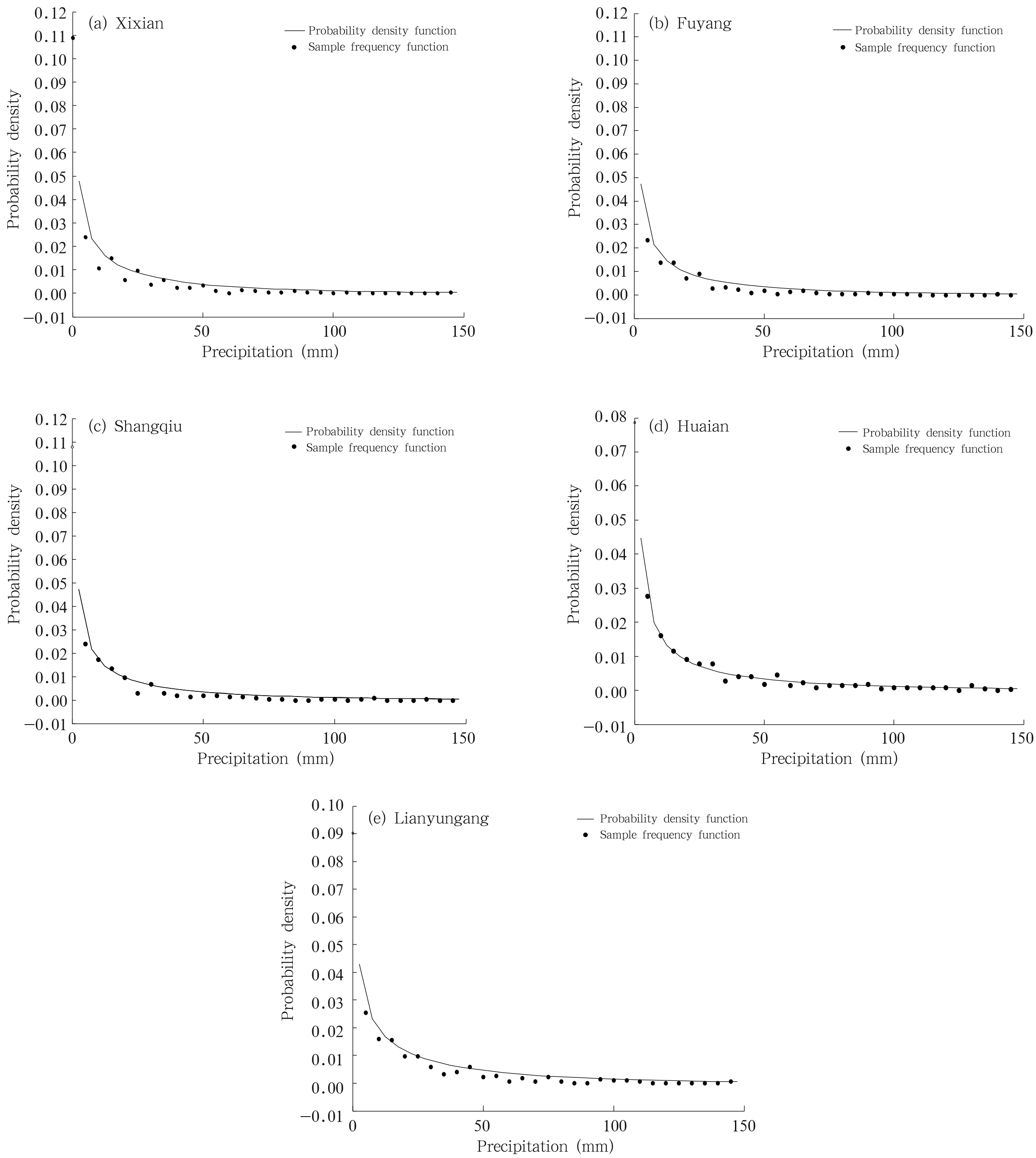

Fig. 5. Gamma distribution density function (solid line) and sample frequency function (dots) of daily precipitation under the condition of a dry preceding day at the representative stations of Xixian (a), Fuyang (b), Shangqiu (c), Huaian (d), and Lianyungang (e), respectively.

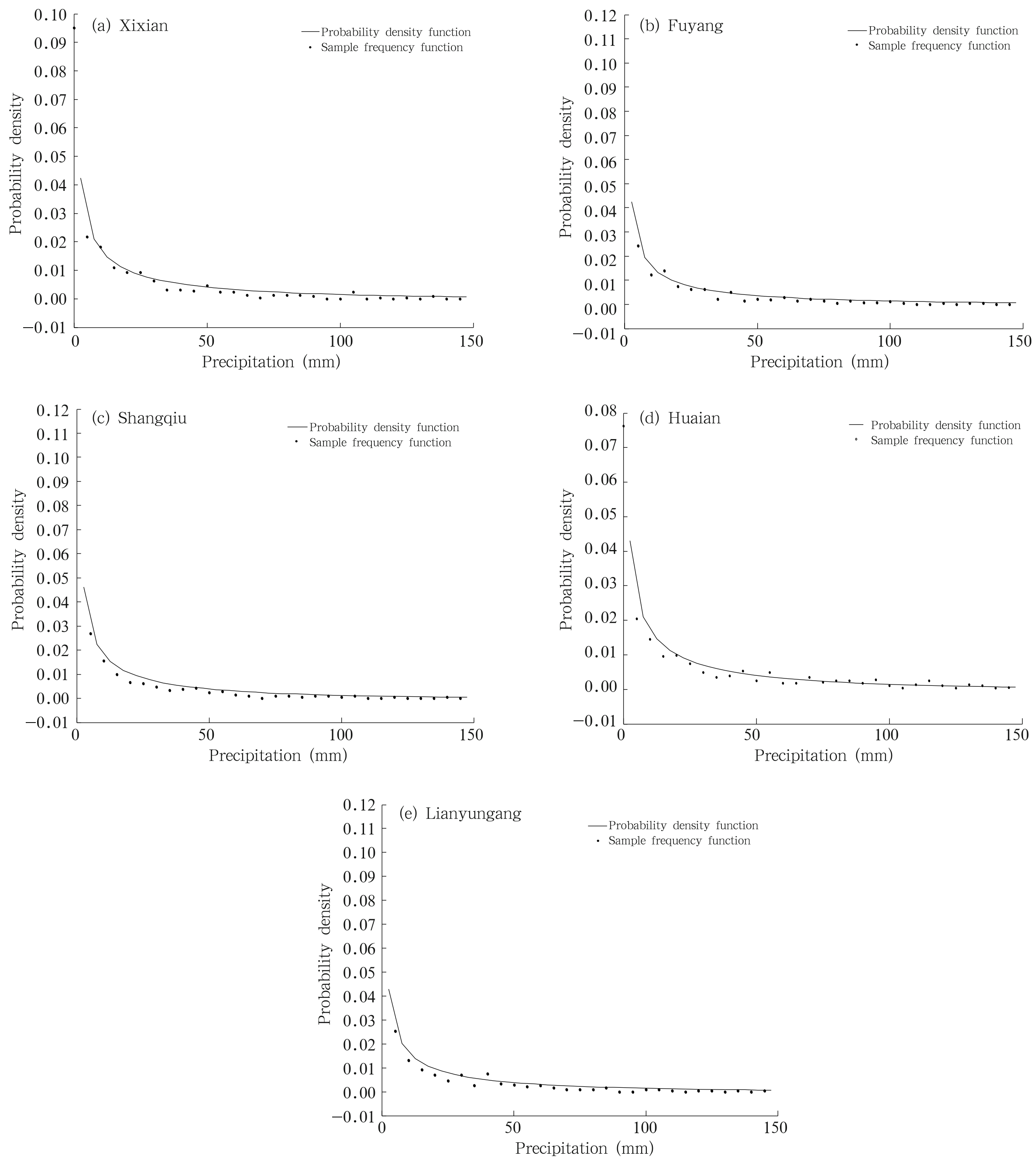

Fig. 6. As in Fig. 5, but for a wet preceding day.

Figures 5 and 6 show the Gamma distribution density function and sample frequency of daily precipitation at the representative stations under the condition that the rainy day has a dry or wet preceding day.

Furthermore, both the probability density curve of the Gamma distribution and the sample frequency curve present the shape of a reverse"J" (Figs. 5 and 6). The fitness of the two curves increases with the daily precipitation. In Figs. 5c and 6b, the two curves almost overlap. As a result, it illustrates that the Gamma distribution function can well describe the probability distribution of the daily precipitation under the two conditions. Meanwhile, this has reduced the influence that the random oscillation of the samples causes on the estimate of the probability distribution of daily precipitation to some extent. The curves of both the probability density function and the sample frequency reflect that the daily precipitation occurs more often in the interval less than 20 mm. The fitness is better on the range of more than 60 mm. The largest deviation of the two curves exists in the range less than 30 mm. The probability density function is almost highly estimated. This means that the Gamma distribution function can describe not only the large probability of light and moderate rains, but also the small probability of heavy rains.

Although the estimation of the Gamma distribution parameters is influenced by the number of samples and random vibration (Wu et al., 2004), the probability density function of Gamma distribution can better fit the frequency of the samples by using the daily precipitation records of 28 yr in the HB. The small vibrations, as indicated in the curves, are the random vibrations of the samples, and they can be reduced when the number of samples is large enough.

Although differences exist in the Gamma distribution parameters of the five HB catchments, the probability density functions at the five representative stations do not differ significantly. This may be because the representative stations are selected from each catchment to enable a wide spread as far as possible. This, however, may result in the under-representative probability density function of the five stations.

In addition, the K-S test is performed to examine the goodness-of-fit of the Gamma distribution of the five representative stations. The K-S statistics of the five sites are all less than the threshold. The test results are acceptable at the confidence level of 0.05. The test shows that, overall, the Gamma distribution appears to be a satisfactory precipitation distribution function for the HB.

3.3 Influence of sample size on the estimates of Gamma distribution parameters

Since the dataset covers a period of 28 yr, which is long enough for statistical analysis, the statistical results exhibit certain interdecadal variations. A question arises here: if the data used start before 1980 (e.g., from 1960) with a longer period of coverage, will there be any change in the results?

Liu (2005) found that when summer precipitation is fitted with the Gamma distribution in the periods 1961–1975 and 1976–2000 separately, the shape parameters of the Gamma distribution in the two periods are both less than 1 under the unconditional and conditional rainy days. And its geographical distributions are consistent with each other in the two periods. The geographical distributions of the scale parameters in the two periods are also consistent except that their numerical values are different.

To verify the above conclusion, the shape and scale parameters of the five representative stations, namely, Xixian, Fuyang, Shangqiu, Huaian, and Lianyungang, are estimated for the period of 1960–2007. It is difficult to compare all stations over the HB as the new stations were established and the old ones were removed during the period of 1960–2007, during which the number of stations and the dataset are not consistent with those in the period of 1980–2007.

From Table 2, it is evident that there are little differences in the shape and scale parameters under the unconditional rainy days between the period 1960–2007 and the period 1980–2007. However, the shape and scale parameters under the unconditional rainy days vary a lot between the two periods.

The shape parameters of the five representative stations during 1960–2007 are smaller than those during 1980–2007, while the scale parameters are larger than those during 1980–2007.

It is necessary to classify the rainy days into conditional rainy days which have a dry or wet preceding day over the years. The longer the years we studied, the greater the differences existed in the samples of conditional rainy days.

In conclusion, the sample size has little effect on parameter estimation of unconditional rainy days when fitting the summer daily precipitation in the HB with the Gamma function. This means that it would have no influence when figuring out the climatology type of precipitation in this region. However, it would

Table 2. Comparison of shape parameter α_i and scale parameter β_i at the five representative stations in the HB between the period 1980–2007 and the period 1960–2007

Representative station	Investigation period	Unconditional rainy days		Conditional rainy days: Rainy days after a dry preceding day		Conditional rainy days: Rainy days after a wet preceding day	
		α_0	β_0	α_1	β_1	α_2	β_2
Xixian	1980–2007	0.401	37.91	0.427	29.42	0.411	42.19
	1960–2007	0.355	42.53	0.111	112.96	0.248	69.81
Fuyang	1980–2007	0.329	44.81	0.351	34.69	0.335	49.73
	1960–2007	0.345	40.35	0.107	109.46	0.242	64.91
Shangqiu	1980–2007	0.381	34.60	0.356	35.07	0.406	34.12
	1960–2007	0.372	34.57	0.150	77.49	0.224	62.90
Huaian	1980–2007	0.363	42.11	0.312	43.21	0.404	41.16
	1960–2007	0.378	38.30	0.130	99.36	0.250	62.80
Lianyungang	1980–2007	0.421	38.65	0.516	30.50	0.368	45.57
	1960–2007	0.407	39.83	0.188	85.98	0.219	74.36

have an influence on the parameter estimation of conditional rainy days, meaning that the probability distribution under the condition that the rainy days have a dry or wet preceding day may vary.

4. Summary and discussion

This study calculates precipitation parameters for the Gamma distribution by using the maximum likelihood estimates.

Analysis of the estimated parameters of Gamma distribution shows that the Huaihe basin is a "scale-dominated" region where it is likely to have variable rains and more extreme events. Extremely high or low rainfall events (i.e., drought or flood) often occur there.

By combining the probability density function of Gamma distribution with spatial distribution of scale parameter, it has been found that the chance of precipitation on the rainy days with a dry preceding day is larger than that with a wet preceding day. The confluence area of Sha River and Ying River between Wangjiaba dam and Bengbu station, the eastern area of the HB between Bengbu station and the Hongze Lake, and the downstream area of the HB below the Hongze Lake, have a high probability of large precipitation under the condition that the rainy day has a dry preceding day. The eastern part of the Yishu River watershed and the region near the Wangjiaba dam are the center of a high probability of large precipitation under the condition that the rainy day has a wet preceding day.

The probability density distribution function of precipitation in the HB overcomes to a certain extent the influence that the random oscillation of the samples causes on the estimate of the probability distribution of daily precipitation.

Through the Kolmogorov-Smirnov test and the comparison between the Gamma distribution probability density functions at five representative stations, i.e., Xixian, Fuyang, Shangqiu, Huaian, and Lianyungang, and the sample frequency of the daily precipitation records, it is proved that the Gamma distribution function can be an adequate fitting to the probability distribution of the precipitation in summer on conditional rainy days, given a wet or dry preceding day.

The probability distribution of summer daily precipitation over the HB is significantly skewed. The probability distribution could be more applicable if the rainy days were given the condition that it has a dry or wet preceding day.

The sample size has little effect on parameter estimation of unconditional rainy days when fitting the summer daily precipitation of the HB with the Gamma function. That is to say, the sample size would have no influence on figuring out the climatology type of precipitation in the HB. However, it would have an influence on the parameter estimation of conditional rainy days, meaning that the probability distribution under the condition that the rainy days has a dry or a wet preceding day may vary.

This paper has analyzed the statistical precipitation probability distribution. The large-scale atmospheric circulation background causing the variation of large daily rainfall has only been briefly discussed. The correlation between the variability of large-scale atmospheric circulation and the large daily rainfall over the HB, as well as the differences of precipitation distribution between regions or sites in the HB, all are worth of future investigations.

REFERENCES

Bi Baogui, Jiao Meiyan, and Li Zechun, 2004a: Contrast analysis of meteorological and hydrological features of extremely heavy rainfall causing severe floods in Huai River valley. *Journal of Nanjing Institute of Meteorology*, **27**(5), 577–585. (in Chinese)

—–, —–, and Liao Yaoming, 2004b: Analysis of precipitation and river flow in the Huaihe basin during the summer of 2003. *J. Appl. Meteor. Sci.*, **15**(6), 681–687. (in Chinese)

Cheng Binyan, Xie Xiaoli, Zhu Yeyu, et al., 2003: A study on application of precipitation probability model in monitoring and assessment of drought and water logging. *Journal of Catastrophology*, **18**(1), 20–25. (in Chinese)

Cheng Huaqiong, 2004: The space-time change feature and physical factor of the rainstorm in Huaihe basin and the statistical forecast model. M.S. Thesis, Dept. of Atmospheric Science, Chinese Academy of Meteorological Sciences, Beijing, 60 pp. (in Chinese)

Ding Yihui, 1993: *The Research of the Persistent Heavy Rainfall in 1991 in Huaihe basin.* Meteorological Press, Beijing, 47–65. (in Chinese)

Gregory, J. H., J. Michaelsen, and C. Funk, 2007: Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *Int. J. Climatol.*, **7**, 935–944.

Jiao Meiyan, Yao Xuexiang, Zhou Bing, et al., 2004: *The Weather Analysis and Research of the Flood in 2003 in Huaihe River.* Meteorological Press, Beijing, 29–42. (in Chinese)

Katz, R., 1977: Precipitation as a chain-dependent process. *J. Appl. Meteor.*, **16**, 671–676.

Liang Shuxian, Yang Yaqun, and Xu Ming, 2001: The distribution of drought and flood in Huaihe basin from June to August, *J. Hydrology*, **21**(2), 54–56. (in Chinese)

Liu Xuehua, 2005: Analysis of the interdecadal variations of the probability distribution of temperature and precipitation in China in recent 40 years. Master thesis, Nanjing University of Information Science & Technology, Nanjing, 54 pp. (in Chinese)

Su Buda, Marco Gemmer, Jiang Tong, et al., 2007: Probability distribution of precipitation extremes over the Yangtze River basin during 1960–2005. *Advances in Climate Change Research*, **3**(4), 208–213. (in Chinese)

Sun Jiliang, Qin Dayong, and Sun Hangguang, 2001: *The General Model of the Hydrometeorological Statistics.* The Waterpub Press, Beijing, 20–21. (in Chinese)

Todorovic, W., 1975: A stochastic model of n-day precipitation. *J. Appl. Meteor.*, **14**, 17–24.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction.* Academic Press, San Diego, CA, 676 pp.

Woolhiser, D. A., 1992: Modeling Daily Precipitation-Process and Problems. *Statistics in the Environmental & Earth Sciences*, Walden AT, Guttorp P. Halsted Press, London, 71–89.

Wu Hongbao, Wang Panxing, and Lin Kaiping, 2004: Probability distribution of the maximum amount of daily precipitation in certain days in June and July for Guangxi. *J. Trop. Meteor.*, **20**(5), 586–592. (in Chinese)

Zhang Ting and Wei Fengying, 2009: Probability distribution of precipitation extremes during raining seasons in South China. *Acta Meteor. Sinica*, **67**(3), 442–451. (in Chinese)

Zhang Yaocun, 1990: A stochastic distribution model of N-day precipitation amount. *Journal of Nanjing Institute of Meteorology*, **13**(1), 23–31. (in Chinese)

Zhao Linna, Yang Xiaodan, and Qi Dan, 2007: The analysis of precipitation and flooding features in the Huaihe basin during the summer of 2007. *Climatic and Environmental Research*, **12**(6), 729–736. (in Chinese)

—–, Wu Hao, Tian Fuyou, et al., 2010: Assessment of probabilistic precipitation forecasts for Huaihe basin using TIGGE data. *Meteor. Mon.*, **36**(7), 133–142. (in Chinese)

Zhu Guoren, 1998: The analysis of precipitation of the Huaihe basin. *J. Hydrology*, **1**, 53–55. (in Chinese)