Incorporating Daily Rainfall to Derive At-Site Hourly Depth-Duration-Frequency Relationships

Shiang-Jen Wu[1]; Yeou-Koung Tung[2]; and Jinn-Chuang Yang[3]

Abstract: Two models are herein presented to derive at-site hourly based rainfall depth-duration-frequency (DDF) relationships by incorporating daily rainfall information. The model development is motivated by the desire to utilize daily rainfall data that are in existence prior to the conversion from a manned gauge to an automatic one for improving the establishment of hourly based rainfall DDF relation at the site. The performance of the proposed models is assessed by comparing the rainfall DDF relationships derived from the conventional method using solely annual maximum rainfalls of various durations. Hourly rainfall data at Hong Kong Observatory over the period of 1884–1990 are used to demonstrate the application of the two proposed models and to examine their performance. Results from the numerical experiments show that the two proposed models, which incorporate daily rainfall data, are capable of producing more accurate and reliable rainfall DDF relationships than the conventional method solely on the basis of annual maximum hourly rainfalls.

DOI: 10.1061/(ASCE)HE.1943-5584.0000065

CE Database subject headings: Frequency analysis; Rainfall frequency; Poisson ratio.

Introduction

Rainfall depth of a storm event with a specified return period and duration, often required in conventional hydrosystem infrastructural designs, is estimated by rainfall frequency analysis. In general, the quality and record length of available rainfall data affect the reliable determinations of underlying probability distribution and the associated rainfall depth-duration-frequency (DDF) relationships. However, due to the evolution of measurement technology over time, the record length of available hourly (or shorter time resolution) rainfall data from automatic rain gauges is generally significantly shorter than that of nonrecording daily rainfall records. At a given site, daily rainfall data might have been in existence for some time before the rain gauge is upgraded or converted to an automatic one. In general, daily rainfall depth and the corresponding maximum rainfall for a specified duration, to a certain extent, are correlated. It is reasonable to hypothesize that incorporating concurrent or extended daily rainfall record could potentially enhance the reliability and accuracy in establishing hourly based rainfall DDF relationships.

As the availability of hourly rainfall data are limited by the operation period of automatic gauges, one approach to extend the length of hourly based rainfall record at a site by incorporating additional daily rainfall record is to disaggregate daily rainfall records into hourly data. Many studies have been made on rainfall disaggregation (Hershenhorn and Woolhiser 1987; Connolly et al. 1998; Koutsoyiannis and Onof 2001; Koutsoyiannis et al. 2004). The disaggregation approach is indirect in that daily rainfalls are first decomposed into hourly rainfalls through an appropriate mechanism. The annual maximum hourly rainfalls from disaggregated data are then used to establish the hourly based rainfall DDF relationships

In case hourly rainfall records are available, Nguyen and his coworkers (Nguyen and Chaleeraktrgkoon 1990; Nguyen and Pandey 1994; Nguyen et al. 1998) introduced a practical idea to relate rainfall characteristics of different durations through the establishment of quantile relationship. However, their quantile relationships are derived on all available hourly and daily rainfall data rather than the extreme type of rainfalls. This makes the assessment of protection level, in terms of annual exceedance probability, for a hydrosystem infrastructure difficult, if not impossible. For this reason, Wu et al. (2005) modified the idea of Nguyen and his coworkers to develop daily hourly rainfall quantile relationships by using the maximum hourly rainfalls corresponding to the annual maximum daily rainfall events. The rainfall quantile relationships established in such a way could facilitate the derivation of various hourly based rainfall DDF curves. The idea has been shown to be more accurate than the conventional method of exclusively using annual maximum hourly based data when rainfall duration is shorter than 6 h (Wu et al. 2005). This is primarily because, for a longer duration rainfall (say, 12 h), the events consisting of annual maximum 12-h rainfall may not occur inclusively within the day of annual maximum daily rainfall. The likelihood of hourly based annual t-h maximum rainfall that concurrently occurs with the annual maximum daily event decreases as the rainfall duration increases.

This paper modifies the earlier annual-maximum-event (AME) model proposed by Wu et al. (2005) and further develops an all-event (AE) model for establishing at-site annual maximum hourly based rainfall DDF relationships by incorporating concurrent and/or extended daily rainfall data. Through numerical appli-

[1]Associate Researcher, National Center for High-Performance Computing, 7 R and D 6th Rd., Hsinchu, 30076 Taiwan, Republic of China. E-mail: sjwu@nchc.org.tw

[2]Professor, Dept. of Civil and Environmental Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. E-mail: cetung@ust.hk

[3]Professor, Dept. of Civil Engineering, National Chiao-Tung Univ., 1001 University Rd., Hsinchu, Taiwan 300, Republic of China. E-mail: jcyang@mail.nctu.edu.tw

Note. This manuscript was submitted on July 17, 2007; approved on December 16, 2008; published online on February 19, 2009. Discussion period open until February 1, 2010; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, Vol. 14, No. 9, September 1, 2009. ©ASCE, ISSN 1084-0699/2009/9-992–1001/$25.00.

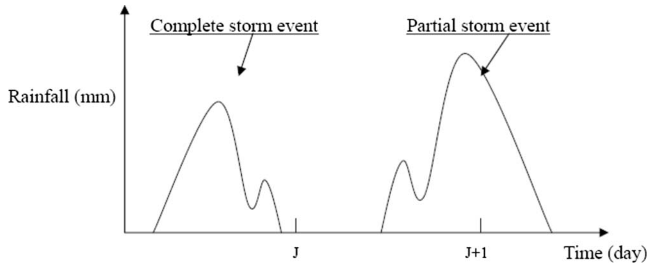

Fig. 1. Definition of complete and partial storm events

cations, the performance and capability of the two proposed models in establishing hourly based rainfall DDF relationships are examined.

Methodology

This section describes the three major tasks involved in the proposed models. They are (1) classification of rainstorm events; (2) establishment of quantiles relation between maximum hourly and daily based rainfall depths; and (3) derivation of probability distribution of hourly based annual maximum rainfall from available daily rainfall data.

Classification of Rainstorm Events

According to the occurrence time of rain, rainfall events can be classified into complete events and partial events as shown in Fig. 1. A complete rainfall event herein is defined as the one that begins and ends on the same day while a partial event begins on one day and ends on a different day. A day is designated as a rainy day if at least one rainfall event (complete or partial) occurs within that day. There are several ways to define a rainfall event in the literature and one can refer to Yen et al. (1993). In deriving an at-site rainfall DDF relationship, the rolling-time annual maximum rainfall depth for a specified duration is commonly retrieved from the recorded rainfall sequence consisting of both complete and partial events. A rainfall event may occur in the form of a complete event or a partial event that straddles across 2- or more consecutive rainy days. To extract maximum rainfall for storm durations between 2–24 h strictly within an individual rainy day may result in underestimating the annual maximum rainfalls in frequency analysis (Wu et al. 2005). Table 1 (a and b) lists the number of events with different periods of consecutive rainy days that contain the annual maximum rainfall for various durations at Hong Kong Observatory (HKO) over a 100-year record period. The table shows that most of the annual maximum rainfalls of different durations occur in an event straddling 2-consecutive rainy days. It is clear that the data series associated with the annual maximum events (AMEs) of varying periods of consecutive rainy day could not capture all annual maximum rainfall events of a fixed duration, especially when the rainstorm duration under consideration is short.

To circumvent the shortcomings of dealing with individual rainy day mentioned above and to better capture the occurrences of rolling-time annual maxima of hourly based rainfalls, the proposed models classify rainy days into different types of events according to the period of consecutive rainy days (see Fig. 2). For illustration, a single rainy day event is an isolated day which may contain one or more complete rainfall events within a single day.

Table 1. Frequencies of Varying Consecutive Rainy Day Events Containing Annual Maximum Values of Different Storm Durations

	Event duration				
	1 h	2 h	6 h	12 h	24 h
(a) All consecutive rainy days					
1	33	32	26	18	13
2	42	38	45	48	53
3	15	18	16	19	19
4	10	12	13	15	15
Total	100	100	100	100	100
(b) Annual maximum consecutive rainy days					
1	11	15	23	18	13
2	25	25	37	44	52
3	14	16	15	18	19
4	9	11	13	15	15
Total	59	67	88	95	99

A 2-consecutive rainy day event may contain multiple complete rainfall events or a mixture of several complete events and one single partial event that begins on one day and ends on the next day.

Quantile Relations of Hourly Based Maximum Rainfall and Total Rainfall

The proposed models utilize quantile relation between the total rainfall depth of the k-consecutive rainy day event and the associated t-h maximum rainfall defined as

$$h^t_{k,p} = g^t_k(d_{k,p}) \tag{1}$$

in which $g^t_k(\bullet)$=general expression of a function and $d_{k,p}$ and $h^t_{k,p}$=quantile values of the random total rainfall depth of the k-consecutive rainy day event (D_k), respectively, and the associated t-h maximum rainfall depth (H^t_k) at the same probability level p, that is, $F_{D_k}(d_{k,p})=\Pr(D_k \le d_{k,p})=\Pr(H^t_k \le h^t_{k,p})=F_{H^t_k}(h^t_{k,p})$

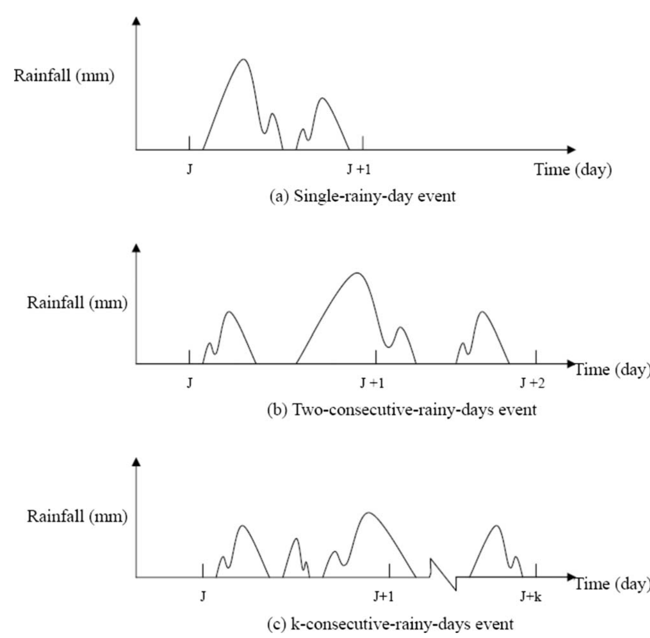

Fig. 2. Different types of consecutive rainy days events

J. Hydrol. Eng. 2009.14:992-1001.

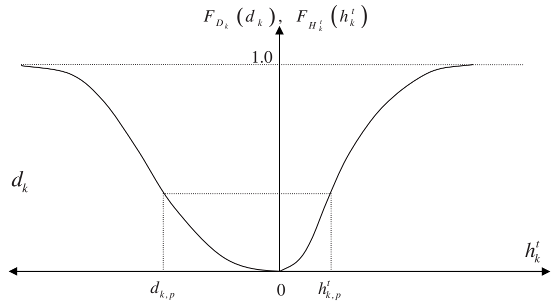

Fig. 3. Marginal distribution functions for D_k and H_k^t

$=p$ with $F_x(\bullet)$ representing the cumulative distribution function (CDF) of random variable X. Eq. (1) can be determined through the use of marginal frequency relationship for D_k and H_k^t, as shown in Fig. 3. Adopting a commonly used unimodal distribution for D_k and H_k^t, the quantile relationship between $h_{k,p}^t$ and $d_{k,p}$ will be monotonically increasing.

The idea behind Eq. (1) is from Nguyen and Chaleeraktrgkoon (1990) who used all available hourly and daily rainfall data to empirically establish such quantile functional relationship in several locations in Canada. As Nguyen's study used the complete hourly and daily rainfall data, the estimated rainfall quantiles are not easily usable to derive the annual exceedance probability associated with an event. Hence, a modification to consider daily based rainfall events and their corresponding maximum t-h rainfall depths, such as Eq. (1), would facilitate the derivation of probability distribution of maximum t-h rainfall on the annual basis.

Probability Distribution of Annual Maximum Hourly Based Rainfall

Quantile values of random maximum t-h rainfall associated with the k-consecutive rainy day event can be estimated from the quantiles of total rainfall depth of the corresponding event by Eq. (1). Since $h_{k,p}^t$ is a monotonically increasing function of $d_{k,p}$, the probability p can be calculated from the marginal probability distribution of the total rainfall depth associated with the k-consecutive rainy day event.

According to the total rainfall depths of selected k-consecutive rainy day events, two models are developed herein: (1) the annual–maximum-event (AME) model and (2) the AE model. The former model uses only the annual maximum total rainfall depths of k-consecutive rainy day events while the latter is based on the total rainfall depths of all such k-consecutive rainy day events in the record.

AME Model

Suppose that rainy days are classified into K varying periods of consecutive rainy day events and such events are statistically independent. Let $D_{\max,k}$ and $H_{\max,k}^t$, respectively, be the random annual maximum total rainfall depth associated with the k-consecutive rainy day event and the maximum t-h rainfall of the corresponding event. Since the overall annual maximum t-h rainfall depth $H_{\max}^t$ is equal to the maximum value of $\{H_{\max,k}^t\}_{k=1,2,\ldots,K}$, i.e., $H_{\max}^t = \max(H_{\max,1}^t, H_{\max,2}^t, \ldots, H_{\max,K}^t)$, the CDF of the annual maximum t-h rainfall depth can be expressed as

Table 2. Number of Years at the HKO Having Zero Occurrences of Events with Varying Number of Consecutive Rainy Days

Consecutive rainy days	Number of years with zero events
1	0
2	4
3	53
4	76

$$F_{H_{\max}^t}(h) = \Pr(H_{\max}^t \le h) = \Pr[(H_{\max,1}^t \le h) \cap (H_{\max,2}^t \le h) \cap \cdots \cap (H_{\max,K}^t \le h)] = \Pr[\bigcap_{k=1}^{K} (H_{\max,k}^t \le h)] \quad (2)$$

Under the condition of statistical independence for K rainfall events with varying period of consecutive rainy days, Eq. (2) can be rewritten as

$$F_{H_{\max}^t}(h) = \prod_{k=1}^{K} \Pr[(H_{\max,k}^t \le h)] \quad (3)$$

As the quantiles of $D_{\max,k}$ and $H_{\max,k}^t$ are related in the form of Eq. (1) as $h_{\max,k,p}^t = g_{\max,k}^t(d_{\max,k,p})$, substituting this relation into Eq. (3) results in

$$F_{H_{\max}^t}(h) = \prod_{k=1}^{K} \Pr[(H_{\max,k}^t \le h)] = \prod_{k=1}^{K} \Pr[(D_{\max} \le d_k)] = \prod_{k=1}^{K} F_{D_{\max}}(d_k) \quad (4)$$

where d_k represents a specific value of total rainfall depth associated with the annual maximum k-consecutive rainy day event obtainable from solving $h = g_{\max,k}^t(d_k)$. From Eq. (4), the return period (T-year) of the annual maximum t-h rainfall h_T can be calculated as $T = [1 - F_{H_{\max}^t}(h_T)]^{-1}$.

When extracting k-consecutive rainy day events used in the AME model, it is possible that such events might not occur in every single year. Hence, the annual maximum value of the total rainfall depth for the corresponding event is zero. This situation is becoming more likely to occur as the period of consecutive rainy day event increases. As can be seen in Table 2, based on 100 years of hourly rainfall record at HKO, except for the case of single rainy day, the number of years having zero occurrences of events with k ($k \ge 2$) consecutive rainy days is positive and it increases with k. Hence, the probability distribution used to fit the annual maximum total rainfall depth series for k-consecutive rainy day events with zeros should be treated.

Hosking and Wallis (1997) proposed a mixed CDF to fit rainfall data including nonzero and zero observations as

$$F_x(x) = \begin{cases} p_0, & x = 0 \\ p_0 + (1 - p_0)G_x(x), & x > 0 \end{cases} \quad (5)$$

in which $F_x(\cdot)$=CDF of rainfall depth; p_0=probability that rainfall depth is zero; and $G_x(\cdot)$=CDF for the positive-valued rainfall data. The parameters in $G_x(.)$ are calibrated with the positive-valued data. In this paper, Eq. (5) is applied in the AME model to establish the quantile relationship between the annual maximum total rainfall depth of a certain period of consecutive rainy days and its corresponding maximum t-h rainfall.

J. Hydrol. Eng. 2009.14:992-1001.

AE Model

Since the AE model considers all k-consecutive rainy day events, the probability of the overall annual maximum t-h rainfall should also depend on the annual number of such events. Suppose that the number of the k-consecutive rainy day event in a year is n_k. The total rainfall depth of the k-consecutive rainy day event and the associated maximum t-h rainfall depth are denoted as D_k and H_k^t, respectively. Referring to Eq. (2), the CDF for the annual maximum t-h rainfall depth $H_{\max,k}^t$, conditioned on n_k occurrences of k-consecutive rainy day events in a year, can be expressed as

$$F_{H_{\max}^t}(h|n_1,n_2,\ldots,n_K) = \Pr[(H_{\max}^t \le h|n_1,n_2,\ldots,n_K)] = \Pr[\bigcap_{k=1}^{K}(H_{\max,k}^t \le h|n_k)] \quad (6)$$

Assuming that the maximum t-h rainfall depths of different consecutive rainy day events are statistically independent, i.e., $\mathrm{Cov}(H_{\max,k}^t, H_{\max,k'}^t)=0$ for $k \ne k'$, Eq. (6) can be rewritten as

$$F_{H_{\max}^t}(h|n_1,n_2,\ldots,n_K) = \prod_{k=1}^{K}\Pr[(H_{\max,k}^t \le h|n_k)] \quad (7)$$

Furthermore, assume that the maximum t-h rainfall depths of n_k k-consecutive rainy day rainstorm events, $H_{k,i}^t(i=1,2,\ldots,n_k)$, are statistically independent from an identical probability distribution, that is, $H_{k,i}^t \overset{i.i.d.}{\sim} F_{H_k^t}(h)$. Then, Eq. (7) can be rewritten as

$$\begin{aligned} F_{H_{\max}^t}&(h|n_1,n_2,\ldots,n_K) \\ &= \prod_{k=1}^{K}\Pr[(H_{\max,k}^t \le h|n_k)] \\ &= \prod_{k=1}^{K}\Pr[(H_{k,1}^t \le h)\cap(H_{k,2}^t \le h)\cap \ldots \cap(H_{k,n_k}^t \le h)] \\ &= \prod_{k=1}^{K}\left[\prod_{i=1}^{n_k}\Pr[(H_{k,i}^t \le h)]\right] \\ &= \prod_{k=1}^{K}[\Pr(H_k^t \le h)]^{n_k} = \prod_{k=1}^{K}[F_{H_k^t}(h)]^{n_k} \end{aligned} \quad (8)$$

Through the quantile relationship of $d_{k,p}$ and $h_{k,p}^t$ defined in Eq. (1), the CDF for $H_{\max}^t$ conditioned on $(n_1,n_2,\ldots,n_K)$ can be expressed in terms of daily rainfall depth quantiles associated with the k-consecutive rainy day event as

$$F_{H_{\max}^t}(h|n_1,n_2,\ldots,n_K) = \prod_{k=1}^{K}[F_{H_k^t}(h)]^{n_k} = \prod_{k=1}^{K}[F_{D_k}(d_k)]^{n_k} \quad (9)$$

in which h and d_k are related in the form of Eq. (1).

Note that the number of the k-consecutive rainy day event in a year is random. Hence, let N_k be a discrete random variable representing the number of the k-consecutive rainy day event occurring yearly and $\Pr(N_k=n_k)$, for $n_k=0,1,2,\ldots,\infty$, be the probability of having n_k occurrences of the k-consecutive rainy day event in a year. Combining Eq. (9) with $\Pr(N_k=n_k)$, the unconditional CDF of the annual maximum t-h rainfall depth can be obtained as

$$F_{H_{\max}^t}(h) = \prod_{k=1}^{K}\left\{\sum_{n_k=0}^{\infty}\{[F_{D_k}(d_k)]^{n_k}\times\Pr(N_k=n_k)\}\right\} \quad (10)$$

Modeling Annual Occurrence of k-Consecutive Rainy Day Events

To calculate the probability of annual maximum t-h rainfall depth $H_{\max}^t$ by Eq. (10), the probability distribution of N_k must be specified. Poisson distribution is often adopted to describe the random number of occurrence of hydrologic events. It is defined by a single parameter, that is, the mean number of occurrence of hydrologic events of interest within a specified time interval (or the average rate of event occurrence per unit time). Although Poisson distribution is often used to describe the random number of occurrences of hydrologic events, it is not necessarily appropriate for all situations due to its requirement about the equality of the mean and variance.

The generalized Poisson distribution (GPD) introduced by Consul and Jain (1973) has two parameters θ and λ with the probability mass function for the random number of occurrence (N) as

$$P_r(N=n) = P_N(n|\theta,\lambda) = \begin{cases} \dfrac{\theta(\theta+n\lambda)^{n-1}e^{-(\theta+n\lambda)}}{n!}, & n=0,1,2,\ldots;\lambda \ge 0 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

The parameters (θ,λ) can be determined by the first two moments (Consul 1989) of N as

$$E(N) = \frac{\theta}{1-\lambda}; \quad \mathrm{Var}(N) = \frac{\theta}{(1-\lambda)^3} \quad (12)$$

The variance of the GPD model can be greater than, equal to, or less than the mean depending on whether the second parameter λ is positive, zero, or negative. The values of mean and variance of a GPD random variable tend to increase as θ increases. Joe and Zhu (2005) proved that the GPD model is a mixture of Poisson distribution and it fits counting data with large zero fractions better than negative binomial distributions. Therefore, The GPD model has greater flexibility to fit various types of random counting processes, such as binomial, negative binomial, or Poisson distribution, and many other observed data. As the condition of equality of mean and variance is not required, the GPD is used in this paper to describe the random annual number of occurrences of k-consecutive rainy day events.

Determining T-Year, t-h Maximum Rainfall Depth

In the process of determining the quantile value of annual maximum t-h rainfall depth for the AME and AE models, one-dimensional golden-section search technique is used to solve Eqs. (4) and (10) for the unknown h associated with a specified return period T by using the following objective function

$$\mathrm{Min}[\varepsilon(h_T^t)] = \left|\frac{F_{H_{\max}^t}(h_T^t) - \hat{F}_{H_{\max}^t}(h_T^t)}{F_{H_{\max}^t}(h_T^t)}\right| = \left|1 - \frac{T \times F_{H_{\max}^t}(h_T^t)}{T-1}\right| \quad (13)$$

where $F_{H_{\max}^t}(h_T^t)$=cumulative probability associated with a return period T-year specified in advance for the unknown rainfall depth h_T^t and $\hat{F}_{H_{\max}^t}(h_T^t)$=estimated cumulative probability computed by the AME and AE models using Eqs. (4) and (10), respectively. The T-year, t-h rainfall depth h_T^t is found when the objection function $\varepsilon(h_T^t)$ is less than the specified tolerance.

Table 3. *p*-Values of Two-Sample KS Test on the Equality of Annual Maximum Total Rainfall Amount Distributions for Events with Varying Consecutive Rainy Days

Record length	Consecutive rainy days		1 day	2 days	3 days
10 years	2 day	Mean	$3.75e-02$ [a]		
		Pr(α≤5%)	75.05%		
	3 day	Mean	$1.29e-02$ [a]	$6.55e-02$ [a]	
		Pr(α≤5%)	93.23%	56.44%	
	4 day	Mean	$2.28e-03$	$1.36e-02$ [a]	$2.15e-01$ [b]
		Pr(α≤5%)	99.99%	91.20%	18.95%
20 years	2 day	Mean	$2.06e-03$		
		Pr(α≤5%)	99.99%		
	3 day	Mean	$7.44e-04$	$6.93e-03$	
		Pr(α≤5%)	99.99%	96.52%	
	4 day	Mean	$2.00e-06$	$4.10e-05$	$2.41e-01$ [b]
		Pr(α≤5%)	99.99%	99.99%	14.37%
30 years	2 day	Mean	$3.22e-04$		
		Pr(α≤5%)	99.99%		
	3 day	Mean	$7.60e-05$	$2.71e-04$	
		Pr(α≤5%)	99.99%	99.99%	
	4 day	Mean	$0.00e+00$	$0.00e+00$	$2.16e-01$ [b]
		Pr(α≤5%)	99.99%	99.99%	20.96%
100 years	2 day	$0.00e+00$			
	3 day	$0.00e+00$	$0.00e+00$		
	4 day	$0.00e+00$	$0.00e+00$	$1.82e-02$ [a]	

Note: Pr(α≤5%) stands for the probability of *p* value less than the significant level of 5%.
[a]Significant at 1%.
[b]Significant at 5%.

Model Application

Description of Data

For the development, application, and performance evaluation of the proposed AME and AE models, hourly rainfall data in 1884–1990 at HKO, with an interruption in 1940–1946 due to the World War II, are used.

Since the model development focuses on the rainfall events which could potentially produce significant surface runoff, the criteria adopted for defining rainstorm events and their retrieval are the following: (1) rainstorm interevent time: ≥1 h; (2) event rainfall depth: ≥30 mm/event; and (3) hourly rainfall depth within an event: ≥10 mm/h. A day is considered to be a rainy day when at least one rainstorm event satisfying the above criteria exists. Accordingly, events with varying periods of consecutive rainy days are identified and the corresponding event total rainfall depth and maximum *t*-h rainfall depths are extracted and used to develop the proposed AME and AE models.

Classification of Events for Rainy Days

To develop the proposed AME and AE models, events with varying periods of consecutive rainy days are determined in advance. Referring to Eqs. (4) and (10), the computational burden hinges on the number of categories of such events. Therefore, it is desirable and practical to examine if the total number of categories of varying periods of consecutive rainy day events can be reduced by combining some of them. For this, two-sample Kolmogorov-Smirnov (KS) goodness-of-fit test is used to examine the equality of the probability distributions for the total rainfall depth associated with the different period of consecutive rainy days. In doing so, fewer but suitable number of event categories with the varying periods of consecutive rainy days can be determined.

Initially, events with different period of consecutive rainy days are classified into four categories, as shown in Table 1. To consider the sampling uncertainty in the two-sample KS test, 50 different subsamples of 10, 20, and 30 years of rainfall data following chronological sequence are randomly extracted from the complete record of 100 years. Tables 3 and 4 show the *p* values of the two sample KS test on the equality of probability distributions of the annual maximum total rainfall depth (for the AME model) and the total rainfall depth (for the AE model) between rainfall event with different period of consecutive rainy days based on 10, 20, 30, and 100 years of rainfall record. It is observed that, when the rainfall record length is longer than 10 years, the mean *p* values among events with 1-, 2-, and 3-consecutive rainy days are less than a significant level of 5%. While for a 10-year record length, the mean *p* values among events with 1- and 2-consecutive rainy days are smaller than 5%. Also shown in the parentheses below the mean *p* value are the percentages that the *p* value is less than the 5% significant level in the 50 random trials. The Pr(α≤5%) between 1- and 2-consecutive rainy day events (for 10 years) and those among 1-, 2-, and 3-consecutive rainy day events (for 20 years and longer) are greater than 90%. This indicates that the probability distribution functions of the total rainfall depth (for AE model) and the corresponding annual maximum value (for AME model) at HKO for 2-, 3-, and 4-consecutive rainy days are statistically similar to a 10-year record length and, hence, can be combined into one group. By the same token, rainfall depths for events with 3- and 4-consecutive rainy days can be combined into one single group when a rainfall record period is longer than 10 years.

J. Hydrol. Eng. 2009.14:992-1001.

Table 4. *p* Values of Two-Sample KS Test on the Equality of Total Rainfall Amount Distributions for Events with Varying Consecutive Rainy Days

Record length	Consecutive rainy days		1 day	2 day	3 day
10 years	2 day	Mean	$1.51e-04$		
		Pr($\alpha \leq 5\%$)	99.99%		
	3 day	Mean	$1.00e-08$	$2.35e-04$	
		Pr($\alpha \leq 5\%$)	99.99%	99.99%	
	4 day	Mean	$1.00e-08$	$3.82e-05$	$3.91e-01$ [a]
		Pr($\alpha \leq 5\%$)	99.99%	99.99%	1.92%
20 years	2 day	Mean	$7.00e-08$		
		Pr($\alpha \leq 5\%$)	99.99%		
	3 day	Mean	$0.00e+00$	$0.00e+00$	
		Pr($\alpha \leq 5\%$)	99.99%	99.99%	
	4 day	Mean	$0.00e+00$	$0.00e+00$	$3.88e-01$ [a]
		Pr($\alpha \leq 5\%$)	99.99%	99.99%	0.001%
30 years	2 day	Mean	$0.00e+00$		
		Pr($\alpha \leq 5\%$)	99.99%		
	3 day	Mean	$0.00e+00$	$0.00e+00$	
		Pr($\alpha \leq 5\%$)	99.99%	99.99%	
	4 day	Mean	$0.00e+00$	$0.00e+00$	$3.60e-01$ [a]
		Pr($\alpha \leq 5\%$)	99.99%	99.99%	0.001%
100 years	2 day	$0.00e+00$			
	3 day	$0.00e+00$	$0.00e+00$		
	4 day	$0.00e+00$	$0.00e+00$	$2.05e-03$	

Note: Pr($\alpha \leq 5\%$) stands for the probability of *p* value less than the significant level of 5%.
[a]Significant at 5%.

In summary, rainstorm events at HKO are classified into two groups (single- and more-than-1 consecutive rainy day events) for a 10-year rainfall record or shorter and into three groups (single, two-, and more-than-2 consecutive rainy day events) for a record length longer than 10 years to develop and verify the proposed AME and AE models.

Establishing h_k^t-d_k and $h_{max,k}^t$-$d_{max,k}$ Frequency-Quantile Relationships

As the magnitude of T-year, t-h rainfall depth is calculated by Eqs. (4) and (10), the frequency-quantile relationships for $(h_{k,p}^t, d_{k,p})$ and $(h_{\max,k,p}^t, d_{\max,k,p})$ should be established in advance. In the study, rainfall depth quantiles of different data series are determined through the frequency analysis. Two moment-based methods are commonly adopted in hydrologic frequency analysis: product-moment and L-moment methods. L moments are linear combination of ordered statistics and have theoretical advantages over the product-moment method of being able to characterize a wider range of distributions and, when estimated from a sample, of being more robust in the presence of outlier in the data (Hosking et al. 1997). The L-moment method is used in rainfall frequency analysis to estimate rainfall quantiles in this study.

Referring to Eq. (1), an often used parametric functional form for $g_{\max,k}^t(d_{\max,k})$ or $g_k^t(d_k)$ is

$$h_k^t = \alpha(d_k)^{\beta} \tag{14}$$

Fig. 4(a) shows $h_{\max,k,p}^t$ and estimated $\hat{h}_{\max,k,p}^t$ by Eq. (14) for the 2-consecutive rainy day event extracted from 100 years of hourly rainfall record at HKO. It is observed that $\hat{h}_{\max,k,p}$ starts to deviate from $h_{\max,k,p}^t$ as $d_{\max,k,p}$ increases, meaning that the discrepancy increases with return period. Furthermore, Fig. 4(b) shows that the relation between $h_{\max,k,p}^t$ and $d_{\max,k,p}$ could vary with the period of consecutive rainy days k. It is, therefore, impractical to use

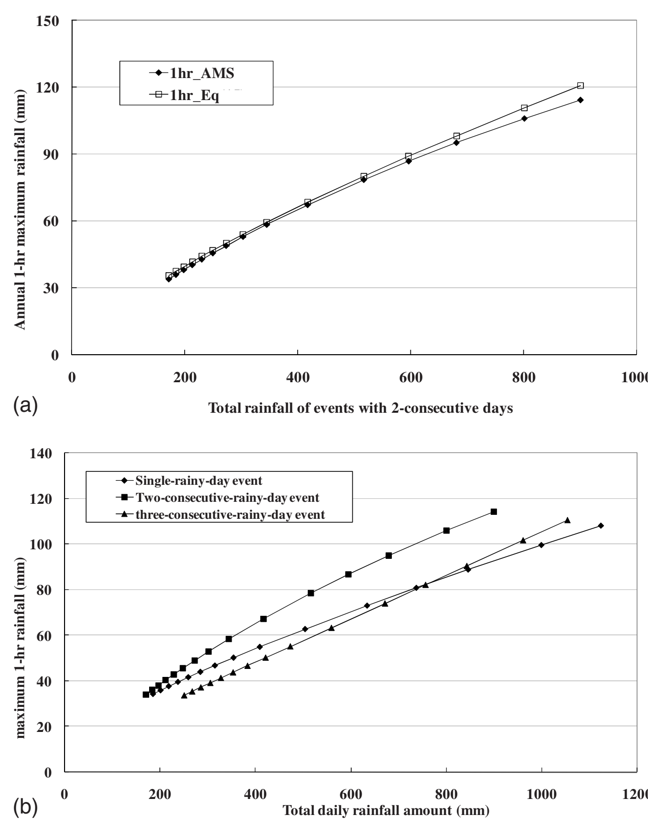

Fig. 4. (a) Comparison of quantile relationship between 1-h maximum rainfall derived by AMS (1-h AMS) and Eq. (14) and total rainfall amount (1-h_Eq) for the 2-consecutive rainy day rainfall events; (b) quantile relationships between 1-h maximum and total rainfall amount for 1-, 2-, and 3-consecutive rainy day rainfall events

J. Hydrol. Eng. 2009.14:992-1001.

a single parametric function form to describe frequency-quantile relationships for all durations. To circumvent the potential inaccuracy resulting from using inadequate parametric functions, interpolation approach is adopted in this study to estimate h_k^t (or $h_{\max,k}^t$) for a given d_k (or $d_{\max,k}$) based on the available quantile pairs of $(h_{k,p}^t, d_{k,p})$ and from the L-moment-based frequency analysis.

Performance Evaluation and Verification

To evaluate the proposed AME and AE models, rainfall frequency-quantile (rainfall DDF) relationships estimated by the proposed AME and AE models are compared with those obtained by the conventional method using solely observed annual maximum rainfall data of various durations. Specifically, the performance evaluation aims at (1) examining the ability of proposed AME and AE models to estimate quantiles of annual maximum rainfall depths of varying durations by incorporating concurrent daily rainfall data and (2) investigating the ability of two proposed models to establish more reliable and accurate quantiles estimation by incorporating extended daily rainfall record. The conventional method using exclusively the annual maximum t-h rainfall series is denoted herein as the annual-maximum-series (AMS) model. The procedure for performance evaluation is outlined below.

Step 1. Extract annual maximum rainfall depths of various durations from the complete n years of hourly rainfall record. Perform frequency analysis by using L-moment method to establish the rainfall DDF relationships, denoted as $\theta_{n,\mathrm{AMS}}$, and use it as the baseline for performance evaluation.

Step 2. Select m years out of n years of the complete record ($m \leq n$) and treat them as the "available" data. Then, from the m-year available rainfall record, extract k-consecutive rainy day events for various k values. The rainfall DDF relationships derived from m-year available hourly rainfall data by the AMS model and proposed AME and AE models are denoted as $\theta_{m,\mathrm{AMS}}$, $\theta_{m,\mathrm{AME}}$ and $\theta_{m,\mathrm{AE}}$, respectively. Additionally, using additional δ-year daily rainfall record ($m+\delta \leq n$), the rainfall DDF relationships obtained by the two proposed models are denoted as $\theta_{m+\delta,\mathrm{AME}}$ and $\theta_{m+\delta,\mathrm{AE}}$.

Step 3. Repeat Step 2 many times and calculate the statistical features of the rainfall DDF relationships $\theta_{m,\mathrm{AMS}}$, $\theta_{m,\mathrm{AME/AE}}$, and $\theta_{m+\delta,\mathrm{AME/AE}}$.

In the performance evaluation, three scenarios of hourly records with partial record lengths (m=10, 20, and 30 years) are used as the available data and four additional daily rainfall record of varying lengths (δ=10, 20, 30, and 40 years) are used for examining the performance of two proposed models. Fifty repetitions are carried out to calculate the statistical features of the rainfall DDF relationships by the AME, AE, and AMS models under various sample sizes for performance evaluation.

To compare the accuracy of estimated rainfall DDF relationships by the three models, the absolute relative error of derived rainfall DDF relationships by the three models with respect to the baseline ($\theta_{n=100,\mathrm{AMS}}$) under various available record length (m years) are computed by

$$\varepsilon_{m,\mathrm{model}} = \left| \frac{\theta_{m,\mathrm{model}} - \theta_{100,\mathrm{AMS}}}{\theta_{100,\mathrm{AMS}}} \right| \qquad (15)$$

where subscripts model=AMS, AME, AE and m=10, 20, 30. To evaluate the effect of additional daily rainfall data on the accuracy

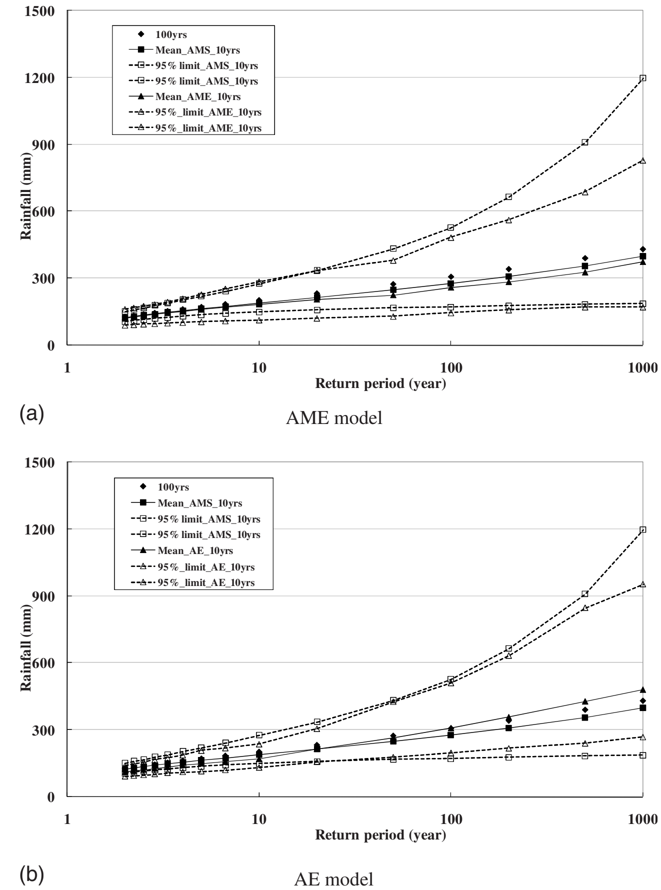

Fig. 5. Comparison of statistical features of 6-h rainfall DDF relationships of varying durations by the AMS and AME models under m=10 years: (a) AME model; (b) AE model

of derived rainfall DDF relationships, similar error measure is used

$$\varepsilon_{m+\delta,\mathrm{model}} = \left| \frac{\theta_{m+\delta,\mathrm{model}} - \theta_{100,\mathrm{AMS}}}{\theta_{100,\mathrm{AMS}}} \right| \qquad (16)$$

where model=AME, AE; m=10, 20, 30; and δ=10, 20, 30, and 40.

Comparison of Proposed Models with the Conventional Model

The relative performance of the AME and AE models under different conditions are evaluated by comparing their estimated rainfall DDF relationships with those conventionally derived solely from the same available m-year annual maximum hourly rainfall data $\theta_{m,\mathrm{AMS}}$. For the purpose of brevity, a typical result with rainfall duration t=6 h is chosen for illustration. For other durations, similar behavior can be observed and readers are referred to Wu (2006) for more detailed information. Figs. 5(a and b), respectively, show the comparison of the mean and 95% probability bounds for 6-h rainfall DDF relationships derived under the available record m=10 years by the conventional AMS model against the AME and AE models. As can be seen, the mean rainfall DDF curves derived by the AME and AE models are close to the baseline curves ($\theta_{n=100,\mathrm{AMS}}$). Also, the 95% probability bound of estimated rainfall DDF relationships, which indicates the model output variability can capture the baseline DDF relationship. For

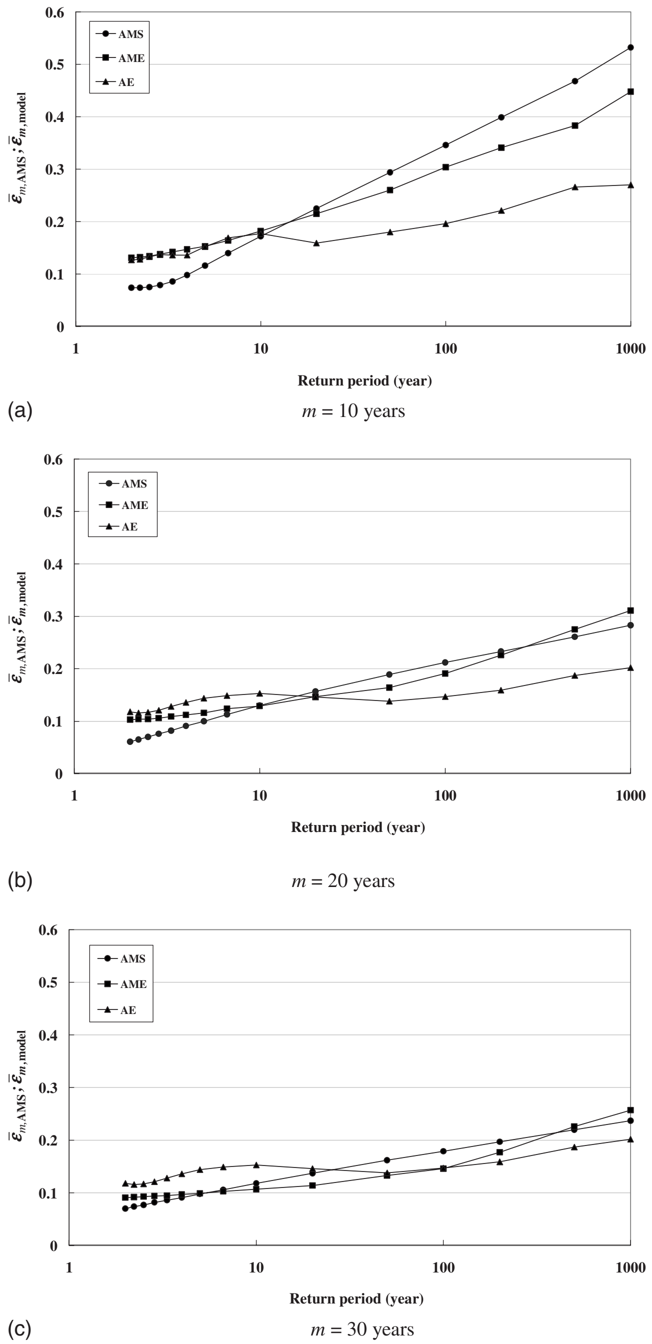

Fig. 6. Mean absolute relative error of 6-h rainfall DDF relationships for different "available" record lengths: (a) $m=10$ years; (b) $m=20$ years; and (c) $m=30$ years

great majority of the rainfall durations considered 1–18 h the width of 95% probability bounds for both AME and AE models are narrower than those of the AMS model under the same available record length, especially for high return periods. This is because a lower standard error is associated with the quantile estimation by the proposed models. A plausible explanation is that the incorporation of concurrent daily rainfall data enriches the information content in the process of deriving rainfall DDF relations.

Figs. 6(a–c) show the mean absolute relative error ($\bar{\varepsilon}_{m,\text{AME}}$, $\bar{\varepsilon}_{m,\text{AE}}$, and $\bar{\varepsilon}_{m,\text{AMS}}$) of 6-h rainfall DDF relationships from the 50 randomly drawn samples under different available rainfall record lengths. It is interesting to observe that the proposed models yield a larger error than the conventional AMS model when the return period is smaller than the available record length. However, as the return period increases, the proposed AME and AE models yield smaller error than the conventional AMS model.

As expected, the magnitude of error decreases as the available record length increases. Furthermore, the difference in error curves associated with the three models diminishes as the available record length increases. A general observation can be drawn from this study: when the return period of interest is significantly smaller than the available record length (e.g., $T<5$ years when $m=10$ years; $T<10$ years when $m=20$ years; and $T<20$ yr when $m=30$ years), the AME and AE models do not produce improved estimation of rainfall quantiles, if not making the estimation less accurate. Based on the results obtained, the proposed models, particularly the AE model, are capable of yielding more accurate and reliable at-site rainfall DDF relations than the AMS model when the return period is higher than the available record length.

Effect of Additional Daily Rainfall Record

In some situations, the rain gauge at a gauging location is upgraded from a previously manned gauge to an automatic recording gauge. Then, daily rainfall data exist prior to the availability of hourly rainfall data. This section examines the performance of two proposed models by incorporating additional daily rainfall records beyond the period of automatic gauging.

By incorporating additional daily rainfall record, the corresponding marginal frequency-quantile relationship of D_k (left-hand curve in Fig. 3) would be affected. Therefore, the frequency-quantile relations of D_k and $D_{\max,k}$ and, consequently, the $h_k^t \sim d_k$ and $h_{\max,k}^t \sim d_{\max,k}$ relations in Eq. (1) would have to be updated. Figs. 7(a and b) show the mean absolute relative error ($\bar{\varepsilon}_{m+\delta,\text{AME}}$, $\bar{\varepsilon}_{m+\delta,\text{AE}}$, and $\bar{\varepsilon}_{m,\text{AMS}}$) of the 6-h rainfall DDF relationships established by the AME and AE models, respectively, under available record length ($m=10$ years) along with different additional daily rainfall record of δ years. The figures clearly reveal that $\bar{\varepsilon}_{m+\delta,\text{model}}$ for the proposed models decrease as the record length of additional daily rainfall (δ) increases. This behavior has been observed for all other storm durations and available record length (m) considered (Wu 2006).

To consider the effect of sampling uncertainty of rainfall data on the rainfall DDF relationships, the probability of $\varepsilon_{m+\delta,\text{model}}$ (with model=AME, AE) being less than or equal to $\varepsilon_{m,\text{AMS}}$, $\Pr(\varepsilon_{m+\delta,\text{model}} \leq \varepsilon_{m,\text{AMS}})$, is calculated from the 50 absolute relative errors of estimated 6-h rainfall DDF relationships. Fig. 8 shows the relation between $\Pr(\varepsilon_{m+\delta,\text{model}} \leq \varepsilon_{m,\text{AMS}})$ and δ year at HKO for some selected return periods ($T=50$, 100, and 200 years) under available record $m=10$ years. It is observed that the value of $\Pr(\varepsilon_{m+\delta,\text{model}} \leq \varepsilon_{m,\text{AMS}})$ increases with additional daily rainfall record length (δ). The value of $\Pr(\varepsilon_{m+\delta,\text{model}} \leq \varepsilon_{m,\text{AMS}})$ is about 70% for all selected return periods and it could reach as high as 90% at higher return periods for some storm durations considered. Therefore, the proposed AME and AE models demonstrate their ability to enhance the accuracy of at-site rainfall DDF relationship by incorporating additional daily rainfall records, if such data exist.

Comparison between Proposed Models

It has been shown above that both AME and AE models have the ability to enhance more reliable and accurate estimation of at-site

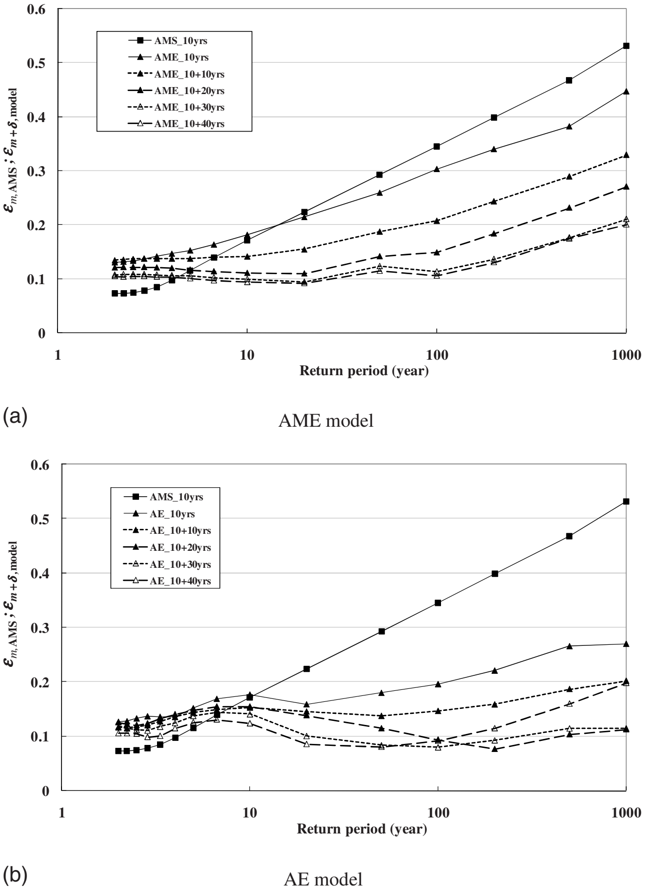

Fig. 7. Mean absolute relative error of 6-h rainfall DDF relationships different additional daily record lengths ($m=10$ years and $\delta=10,20,30,40$ years): (a) AME model; (b) AE model

rainfall DDF relationships by incorporating concurrent and/or additional daily rainfall data, if available. This section focuses on the comparison of two proposed models.

Referring to Fig. 6, without additional daily rainfall record ($\delta=0$), the mean absolute relative error values of rainfall quantiles calculated by the AE model are smaller than those of the AME model. Fig. 7 also shows that the rate of accuracy improvement decreases with additional daily record length. The rate of accuracy improvement by the AME model is significantly higher than the AE model, especially for the smaller number of additional daily record length, say, $\delta=10$ years. From Fig. 8, the value of $\Pr(\varepsilon_{m+\delta,\text{AE}} \leq \varepsilon_{m,\text{AMS}})$ is significantly higher than those of $\Pr(\varepsilon_{m+\delta,\text{AME}} \leq \varepsilon_{m,\text{AMS}})$, indicating that the AE model is more likely to yield a more accurate rainfall DDF relation than the AME model.

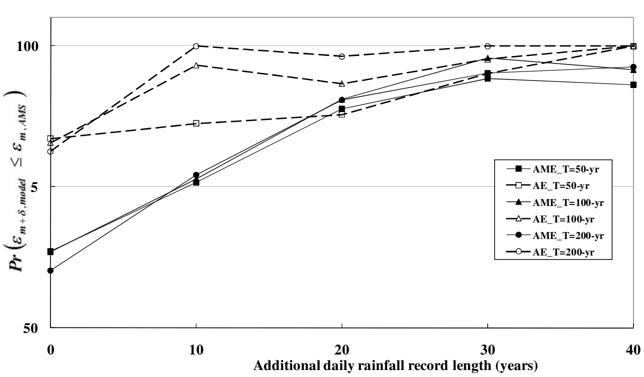

Fig. 8. $\Pr(\varepsilon_{m+\delta,\text{model}} \leq \varepsilon_{m,\text{AMS}})$ for AME and AE models of varying return periods under $m=10$ years

By incorporating concurrent daily rainfall data, the AE model, in general, produces more accurate rainfall DDF relationships than the AME model under 10–30 years available hourly rainfall record at HKO. However, the accuracy level between the two proposed models is compatible when the record length is about 30 years or possibly longer by extrapolation of the trend. From the viewpoint of data processing and computational complexity, the AME model is simpler than the AE model. Hence, it is reasonable to suggest the use of AME model rather than the AE model, when the available record length for hourly rainfall data are relatively long. Otherwise, the AE model is recommended.

Conclusions

This paper presents two models, namely, the AME model and the AE model, which incorporate daily rainfall information to establish at-site hourly rainfall DDF relationships. The application of proposed AME and AE models requires (1) classification of rainy days into events with varying periods of consecutive-rainy days; (2) establishment of the rainfall quantiles relation between hourly and daily rainfall depths; and (3) derivation of the probability distribution for annual maximum hourly rainfall depth in terms of the daily rainfall probability distribution.

For illustration and performance evaluation, the proposed AME and AE models are applied to estimate at-site rainfall DDF relation by incorporating daily rainfalls using the data at HKO. From the numerical experiments, the AME and AE models demonstrate their ability to yield more accurate DDF relations than the conventional model based solely on the basis of available annual maximum hourly rainfall record when the return period of interest is larger than the available record length. Moreover, the proposed AME and AE models show promising potential to further enhance the accuracy of rainfall DDF relationships by incorporating additional daily rainfall records at the site, if available. As a result, the proposed models can be viable for improving hourly based rainfall DDF relationships by using daily rainfall data available at the site prior to the conversion from a manned gauge to automatic recording.

To summarize the above results, the proposed AME and AE models would be carried out for the design of rain gauge network. In detail, the AME and AE models can evaluate "adequate" hourly rainfall record length at an automatic rain gauge from which the reliable and stable rainfall DDF relationships are to be established. As the hourly rainfall record for an automatic rain gauge reaches this adequate length, it can be redeployed to other strategic locations and replaced by a less expensive gauge for collecting daily rainfalls. Hence, a more cost effective rain gauge network, consisting of automatic and man gauges, can be established to derive reasonably accurate and reliable DDF relationships for hydrological applications.

In the development of AME and AE models, an important assumption was made. That is, consecutive rainy day events are statistically independent. In fact, there probably exists correlation among these events. As a result, a future work could be made to take a correlation among consecutive rainy day events into account to derive the AME and AE models.

Acknowledgments

This study is conducted under the auspice of research project "Investigating issues in rainfall intensity-duration and timescale relations in Hong Kong" (Project No. 620505) funded by the Research Grant Council of Hong Kong Special Administration Region, China.

Notation

The following symbols are used in this paper:

D_k, H_k^t = total rainfall depth of the k-consecutive rainy day event and the associated t-h maximum rainfall depth;

$D_{\max,k}, H_{\max,k}^t$ = annual maximum total rainfall depth of the k-consecutive rainy day event and the associated t-h maximum rainfall depth;

$F_{H_{\max}^t}(h)$ = CDF of t-h annual maximum rainfall depth;

n_k = number of k-consecutive rainy day event in a year;

$\varepsilon_{m,\mathrm{AMS}}, \varepsilon_{m,\mathrm{model}}$ = absolute relative errors of the rainfall DDF relationships $\theta_{m,\mathrm{AMS}}$, $\theta_{m,\mathrm{mode}}$ with respect to $\theta_{n=100,\mathrm{AM}}$;

$\varepsilon_{m+\delta,\mathrm{model}}$ = absolute relative error of the rainfall DDF relationships $\theta_{m+\delta,\mathrm{model}}$ with respect to $\theta_{n=100,\mathrm{AMS}}$;

$\theta_{m,\mathrm{model}}$ = rainfall DDF relationships derived from m-year available hourly rainfall data by "model";

$\theta_{n=100,\mathrm{AMS}}$ = rainfall DDF relationships derived by the AMS model using 100-year rainfall data; and

$\theta_{m+\delta,\mathrm{model}}$ = rainfall DDF relationships obtained by model using m-year available hourly rainfall data with additional δ-year daily rainfall record.

References

Consul, P. C., and Jain, G. C. (1973). "A generalization of Poisson distribution: Properties and application." *Technometrics*, 15, 791–799.

Consul, P. C. (1989). *Generalized Poisson distributions: Properties and application*, Marcel Dekker, New York.

Connolly, R. D., Schirmer, J., and Dunn, P. K. (1998). "A daily rainfall disaggregation model." *Agric. Forest Meteorol.*, 92, 105–117.

Hershenhorn, J., and Woolhiser, D. A. (1987). "Disaggregation of daily rainfall." *J. Hydrol.*, 95, 299–322.

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: An approach based on L-Moments*, Cambridge University Press, New York.

Joe, H., and Zhu, R. (2005). "Generalized Poisson distribution: the property of mixture of Poisson and comparison with negative binomial distribution." *Biom. J.*, 47(2), 219–229.

Koutsoyiannis, D., and Onof, C. (2001). "Rainfall disaggregation using adjusting procedures on a Poisson cluster model." *J. Hydrol.*, 246, 109–122.

Koutsoyiannis, D., Onof, C., and Wheater, H. S. (2004). "Multivariate rainfall disaggregation at a fine timescale." *Water Resour. Res.*, 39(7), 1–18.

Nguyen, V. T. V., and Chaleeraktrgkoon, C. (1990). "Estimation of hourly rainfall distribution based on available daily rainfall data." *J. Am. Water Resour. Assoc.*, 11, 67–76.

Nguyen, V. T. V., and Pandey, G. R. (1994). "Estimation of short-duration rainfall distribution using data measured at longer time scales." *Water Sci. Technol.*, 29(1), 39–45.

Nguyen, V. T. V., Nguyen, T. D., and Wang, H. (1998). "Regional estimation of short duration rainfall extremes." *Water Sci. Technol.*, 37(11), 15–19.

Wu, S. J. (2006). "Development and application of stochastic generation model for rainfall series with time scale." Ph.D. thesis, National Chiao-Tung Univ., Hsinchu, Taiwan.

Wu, S. J., Tung, Y. K., and Yang, J. C. (2005). "A simple model to estimate hourly rainfall DDF relationships by incorporating daily rainfall." *Proc., 31st IAHR Congress*, Korea Water Resources Association, Seoul, Korea, 3698–3708.

Yen, B. C., Riggins, R., and Ellerbroek, J. W., III. (1993). "Probabilistic characteristics of elapsed time between rainfalls." *Management of irrigation and drainage systems: Integrated perspectives*, ASCE, New York, 424–431.