Nordic Hydrology, 31 (3), 2000, 187-206

A Method Based on Taking the Average of Probabilities to Compute the Flow Duration Curve

Hikmet Kerem Cigizoglu

Civ. Eng. Faculty, Istanbul Technical University,
Maslak 80626 Istanbul, Turkey

In this study a method based on taking the average of the probabilities is presented to obtain flow duration curve. In this method the exceedance probability for each flow value is computed repeatedly for all time periods within a year. The final representing exceedance is just simply the average of all these probabilities. The applicability of the method to daily mean flows is tested assuming various marginal probability distributions like normal, Pearson type III, log-Pearson type III, 2-parameter lognormal and 3-parameter lognormal distributions. It is seen that the observed flow duration curves were quite well approximated by the 2-parameter lognormal average of probabilities curves. In that case the method requires the computation of the daily mean and standard deviation values of the observed flow data. The method curve enables extrapolation of the available data providing the exceedance probabilities for the flows higher than the observed maximum flow. The method is applied to the missing data and ungauged site problems and the results are quite satisfactory.

Introduction

Flow duration curves are one of the three graphical tools most familiar to the hydrologist together with the hydrograph and mass curve (Foster 1934). A flow-duration curve (*FDC*) represents the relationship between the magnitude and the frequency of daily, weekly, monthly (or some other time interval of) streamflow for a particular river basin, providing an estimate of the percentage of time a given streamflow was equaled or exceeded over a historical period. Though *FDC* is widely used in water resources engineering the number of the studies related to this subject is small compared with its importance. The majority of the realized studies are about the utilization domains of duration curves in water resources engineering and some of them cover the determination of the regional flow duration curves. *FDC*'s have been advocated for use in hydrologic studies such as hydropower, water-supply, and irrigation planning (Chow 1964; Warnick 1984; Maidment 1992; Vogel and Fennessey 1995; Cigizoglu 1997b). Male and Ogawa (1984) showed how *FDC*'s can be used to illustrate and evaluate the trade-offs among the variables involved in the selection of a waste-water treatment-plant capacity. Hughes and Smakhtin (1996) developed a method based on the *FDC* for patching and extending observed time series of daily streamflow.

As many water resource projects are located at ungauged sites, for which flow duration curves are required but unavailable, the information spatially transferred from gauged sites may be useful for hydrologic design at ungauged sites. Regional analysis is used to facilitate extrapolation from a site at which records have been collected to others at which data are required but are unavailable. Hence studies of regional flow duration curves are useful for hydrological analysis and have received much attention. Regional *FDC* procedures have been developed for ungauged sites by Singh (1971), Dingman (1978), Quimpo *et al.* (1983), Mimikou and Kaemaki (1985), Fennessey and Vogel (1990), Yu and Yang (1996) and Smakhtin *et al.* (1997). Mimikou and Kaemaki regionalized flow duration curve by using morphoclimatic characteristics of the drainage basin. In their study Fennessey and Vogel approximated the lower half of daily flow duration curves using a 2-parameter lognormal probability density function and employed a conjugate gradient algorithm to fit log normal density functions to the lower half of observed flow-duration curves at various basins. Yu and Yang applied multivariate statistical analysis (principal component and cluster analysis) to daily flow data from gauged-stations in southern Taiwan.

The annual-based interpretation in the study of Vogel and Fenessey (1994) enables confidence intervals and recurrence intervals to be associated with *FDC*'s in a nonparametric framework. Boutillier and Waylen (1993) define the *FDC* with the ranking statistics. Cigizoglu (1997, a) and Cigizoglu and Bayazit (1999) developed a method based on the convolution theorem for two random variables to obtain the flow duration curve and the method is applied to some daily flow data successfully.

Each part of the flow duration curve carries significance for different type of water resources applications. Low flow part is required to assure the maintenance of water quality standards. This part is also used in establishing the value of the compensation flow released to downstream of the river by the reservoir. The upper part of the *FDC*, on the other hand, is significant for storage volume regulation in a reservoir where spillways are used for flood control purpose. Therefore robust estimation procedures are needed for all parts of the flow duration curve.

In this study an approach based on taking the average of probabilities is used to obtain *FDC*. The method is straightforward and requires only the computation of the statistical parameters of the daily mean flows.

The correct choice of a marginal probability distribution for the flow duration curve is critical as it governs both the general form or shape of the flow duration curve. Once a probability distribution is fitted successfully to the period of record curve then it is possible to make an extrapolation by simply reading the exceedance probabilities corresponding to the flow values higher than the observed maximum flow from the fitted probability curve.

In this study five different probability distributions are tested using the method presented. Having decided upon the most adequate distribution, application to the missing data and ungauged site problems has been carried out. The method curves produced satisfactory results in both applications.

The Definition of the Method

If the number of the time intervals considered through a T period is represented with n and $\Delta t/T$ is the exceedance probability ($\Delta t = 1$ for a month $n = 12$, $\Delta t = 1$ for a day $n = 365$), it can be assumed that the flow values at different time intervals are homogeneous, that is their distributions and statistical parameters are constant. In that case for each Δt_i time interval (i=1,2, ...n) the probability that the flow x_i exceeds an x_0 value, $P(x_i > x_0)$, is determined by considering the distribution function and parameters of the flow in that time interval. The average of these probabilities gives the probability, that the X flow value exceeds the mentioned x_0 through a year (Cigizoglu 1997a). This result can be obtained using the total probability theorem. If time intervals are mutually exclusive and exhaustive events then

$$P(x_0 \text{ is exceeded given an arbitrary } \Delta t \text{ time interval}) = \sum_{i=1}^{n} P(x_0 \cap \Delta t = i) \tag{1}$$

Using multiplication rule we have

$$\sum_{i=1}^{n} P(x_0 \cap \Delta t = i) = \sum_{i=1}^{n} P(x_0 \text{ is exceeded} \mid \Delta t = i) P(\Delta t = i) \tag{2}$$

and total probability equation can be obtained by replacing Eq.(2) in Eq. (1),

$$P(x_0 \text{ is exceeded given an arbitrary } \Delta t) = \sum_{i=1}^{n} P(x_0 \text{ is exceeded} \mid \Delta t = i) P(\Delta t = i) \tag{3}$$

If '$P(\Delta t=i) \equiv 1/n$' for all time intervals where n is the total number of time intervals and if we define x_0 is exceeded given an arbitrary time interval with $X>x_0$ where X represents all the flows observed in the flow record then

$$P(X>x_0)=\frac{\sum_{i=1}^{n}P(x_i>x_0)}{n} \tag{4}$$

for the daily flows

$$P(X>x_0)=\frac{\sum_{i=1}^{365}P(x_i>x_0)}{365} \tag{5}$$

where x_i represents the flows observed only in that day i throughout the whole flow record and i=1, ..., 365.

The flow duration curve of the daily streamflows can be computed using the formula above. Daily streamflows are so highly skewed that ordinary product moment ratios such as the coefficient of variation and skewness are remarkably biased and should be avoided, even with samples with tens of thousands of flow observations (Vogel and Fennessey 1993). The authors concluded in their study that product moment estimates of coefficient of variation, c_v, and skewness, c_{sx}, should be replaced by L moment estimators for most goodness of fit applications in hydrology including both small- sample, low-c_v applications and both small-and large-sample, high-c_v applications. The method of L moments is an exact analogue to the method of moments. However, L moment estimators are nearly unbiased for all sample sizes and all distributions. Therefore in this study the probability distribution parameters are computed using L moments rather than the classsical method of moments method. The computation formulas for L moments using probability weighted moments (PWMs) are introduced by Hosking (1990). Accordingly the unbiased sample estimates of the PWMs, for any distribution can be computed from

$$b_0=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \quad ; \quad b_1\equiv\sum_{i=1}^{n-1}\frac{(n-i)}{n(n-1)}x_i \quad ; \quad b_2\equiv\sum_{i=1}^{n-2}\frac{(n-i)(n-i-1)}{n(n-1)(n-2)}x_i \tag{6}$$

where x_i represents the ordered streamflows with x_1 being the largest observation and x_n the smallest. For any distribution the first three L moments are easily computed from the PWMs using

$$\lambda_1=b_0 \quad ; \quad \lambda_2=2b_1-b_0 \quad ; \quad \lambda_3=6b_2-6b_1+b_0 \tag{7}$$

Analogous to the product moment ratios, coefficient of variation, $c_v=S_x/\bar{x}$, and skewness, c_{sx}, Hosking (1990) defines the L moment ratios

$$\tau_2=\frac{\lambda_2}{\lambda_1}\equiv L \text{ coefficient of variation}; \quad \tau_3=\frac{\lambda_3}{\lambda_2}\equiv L \text{ skewness} \tag{8}$$

The number of daily parameters (365 for each parameter) looks to be high violating the statistical parsimony principle in number of parameters. Fitting Fourier series to the statistical parameters is considered as a solution to eliminate this problem (Salas *et al.* 1980). In this study Fourier series values are used for the parameters. The application procedure of the method for different probability distributions is presented below.

Normal Distribution

In the case of normal distribution assumption following steps are followed to compute the exceedance probability:

1) Mean, $\bar{x}_i$, and standard deviation, s_{xi}, values for each day, *i*, of the year are computed from the observed flow record using *L* moments (Hosking and Wallis 1997)

$$\bar{x}_i = \lambda_1; \quad s_{xi} = \pi^{\frac{1}{2}}\lambda_2 \tag{9}$$

2) An interval is selected initially by just dividing a selected maximum flow value to the total number of intervals, *M*. In this study maximum flow value is selected to be equal to $1.5x_{max}$, where x_{max} is the maximum observed flow value. The selection of a maximum value higher than x_{max} enables the examination of the behavior of the probability distribution in extrapolation. The total number of intervals, *M*, is taken to be equal to 1,000. The consecutive x_0's are computed just by adding this interval value to the previous one.

3) The standard normal variable is found for each x_i=x_0 for all time periods. Since daily flows are investigated in this study 365 different z_i values are computed as follows

$$z_i = \frac{x_0 - \bar{x}_i}{s_{xi}} \tag{10}$$

where s_{xi} and $\bar{x}_i$ represent the standard deviation and the mean of that particular day, respectively, and *i* takes values between 1 and 365.

4) The exceedance probability for each z_i is computed using the approximation formulas given by Abromowitz and Stegun (1972) for normal distribution.$P(x>x_0)$ is then just the average of these 365 values as given in Eq. (5).

Pearson Type III Distribution (P3)

The Pearson type III distribution is one of the probability distributions considered in the literature for skewed hydrologic random variables. It has three parameters α, β and γ. If $\gamma \neq 0$ then

$$\alpha = \frac{4}{c_{xsi}^{\,2}} \;; \quad \beta = s_{xi} \,|\, c_{sxi} \quad \text{and} \quad \xi = \bar{x}_i - \frac{2s_{xi}}{c_{sxi}} \tag{11}$$

$\overline{x}_i$, s_{xi} and c_{sxi} can be computed using L moments as given by Hosking and Wallis (1997)

$$c_{sxi} = 2\alpha^{-\frac{1}{2}}\,\mathrm{sign}(\tau_3) \;;\quad s_{xi} = \lambda_2 \pi^{\frac{1}{2}} \alpha^{\frac{1}{2}} \frac{\Gamma(\alpha)}{\Gamma(\alpha+1/2)} \quad \text{and} \quad \overline{x}_i = \lambda_1 \tag{12}$$

where

$$\alpha \approx \frac{1+0.2906t}{t+0.1882t^2+0.04442t^3} \quad \text{with} \quad t = 3\pi\tau_3^2 \quad \text{if} \quad 0<|\tau_3|<\frac{1}{3} \quad \text{and}$$

$$\alpha \approx \frac{0.36067t-0.59567t^2+0.25361t^3}{1-2.78861t+2.56096t^2-0.77045t^3} \quad \text{with} \quad t = 1-|\tau_3| \quad \text{if} \quad \frac{1}{3}<|\tau_3|<1 \tag{13}$$

The exceedance probability for the x_i=x_0 value can be computed via Modified Wilson Hilferty Transformation (Kirby 1972; Cigizoglu and Bayazit 1998; Cigizoglu 2000). Accordingly there is such a relation between the ε_i independent variable and the z_i standard Normal variable

$$\varepsilon_i = A\left(\max(H, 1 - (\frac{G}{6})^2 + (\frac{G}{6})z_i)^3 - B\right) \tag{14}$$

where

$$H \equiv (B - \frac{2/c_{sxi}}{A})^{1/3} \quad ; \quad A \equiv \max(\frac{2}{c_{sxi}}, 0.40)$$

$$B = 1+0.0144\max(0, c_{sxi}-2.25)^2 \quad ; \quad G = c_{sxi} - 0.063\max(0, c_{sxi}-1)^{1.85} \tag{15}$$

Since for the majority of the c_{sxi} values

$$H < \left(1 - (\frac{G}{6})^2 + (\frac{G}{6})z_i\right) \tag{16}$$

Eq. (14) becomes the following form

$$\varepsilon_i = A\left((1 - (\frac{G}{6})^2 + (\frac{G}{6})z_i\right)^3 - B \tag{17}$$

In this case the z_i variable is computed in the following way

$$z_i = \left((\frac{\varepsilon_i}{A} + B)^{1/3} - 1 + (\frac{G}{6})^2 \right) \frac{6}{G} \tag{18}$$

The ε_i value in this equation is obtained with the $\varepsilon_i = x_o - \overline{x_i}/S_{xi}$ formula. Finally the exceedance probability for the z_i value obtained using the Eq. (18) is computed again using the approximation formulas.

Log Pearson Type III Distribution (LP3)

All the computations are as described for the Pearson type III distribution except that x_i is replaced with the y_i=log (x_i).

Two-Parameter Lognormal Distribution (LN2)

The procedure for 2-parameter lognormal distribution is the same as followed for the normal distribution case except that the daily mean and standard deviations are computed for the y_i=ln(x_i) values. Hence, Eq. (3) is rewritten as

$$z_i = \frac{y_0 - \bar{y}_i}{s_{yi}} \tag{19}$$

where s_{yi} and $\bar{y}_i$ are the two parameters of the 2-parameter lognormal distribution computed using L moments as described in Normal distribution case.

Three-Parameter Lognormal Distribution (LN3)

Three-parameter lognormal distribution is one of the probability distributions in the literature proposed to be adequate for daily river flows (Limbrunner *et al.* in press). It has the parameters ξ (location), α (scale) and k (shape). They are computed using L moments as presented by Hosking and Wallis (1997)

$$k \approx -\tau_3 \frac{E_0+E_1\tau_3^2+E_2\tau_3^4+E_3\tau_3^6}{1+F_1\tau_3^2+F_2\tau_3^4+F_3\tau_3^6} \; ; \quad \alpha = \frac{\lambda_2 k e^{-k^2/2}}{1-\Phi(-k/\sqrt{2})} \; , \quad \xi = \lambda_1 - \frac{\alpha}{k}(1-e^{k^2/2}) \tag{20}$$

where E_0, E_1, E_2 , E_3, F_1, F_2 and F_3 are the coefficients and their values are presented in Table 1. $\Phi(-k/\sqrt{2})$ is the Normal distribution cumulative frequency value, $F(-k/\sqrt{2})$.

Table 1 – Coefficients used in Eq. (20)

Coefficient	E_0	E_1	E_2	E_3	F_1	F_2	F_3
	2.046653	-3.6544371	1.8396733	-0.2036024	-2.0182173	1.2420401	-0.2174180

$\bar{y}_i$, s_{yi} and ζ_i are then found consecutively

$$s_{yi} = -k \; ; \quad \bar{y}_i = \log\left(\frac{\alpha}{s_{yi}}\right) \quad \text{and} \quad \zeta_i = \xi - \exp(\bar{y}_i) \tag{21}$$

Once the parameters are determined the procedure explained in Normal distribution case is followed. Here however $\bar{y}_i$, s_{yi} and ζ_i are computed for y_i=log(x_i) whereas

$$z_i = \frac{\log(x_i - \zeta_i) - \bar{y}_i}{s_{yi}} \tag{22}$$

Table 2 – Statistical parameters of the daily flows in Turkish rivers (the values in parentheses represent the corresponding values computed using method of moments)

Station No.	Basin Area (km2)	Years	$\bar{x}$ (m^3/s)	$\frac{\sum_{i=1}^{365} s_{xi}}{365}$ (m^3/s)	$\frac{\sum_{i=1}^{365} c_{sxi}}{365}$	X_{min} (m^3/s)	X_{max} (m^3/s)
2145	5822	1963-1994	24.51	9.32 (10.17)	1.40 (1.20)	10.10	251
701	948	1963-1989	6.25	6.25 (6.22)	1.47 (1.53)	0.02	181
1708	1416	1963-1993	40.22	19.10 (21.02)	1.10 (1.14)	0.68	110
1714	10065	1963-1994	120.07	57.76 (65.29)	1.24 (1.30)	1.24	1119
1717	1055	1963-1994	6.11	2.82 (3.04)	1.22 (1.22)	1.22	107
1719	3450	1970-1994	53.19	28.78 (33.65)	1.19 (2.18)	1.19	760
1720	4304	1970-1994	45.27	19.96 (21.97)	1.18 (1.18)	1.18	579

Application of the Method to Probability Distributions

The method is applied to the daily flow data of three different gauged stations in different parts of Turkey. These are the stations with Nos. 701, 1714 and 2145. The related information for the selected stations is presented in Table 2. The daily means and standard deviations, of which the annual averages are presented in Table 2, are the ones computed using the Eq. (9) whereas the daily skewnesses, again represented with annual average in Table 2, are the ones computed using the Eq. (12). The x_{max} and x_{min} values are the maximum and minimum flows observed in the stations throughout the full flow record (Table 2). The values in parentheses represent the corresponding values computed using the classical method of moments.

Observed flow duration curve is computed for each station. For that the whole data for a station is sorted in increasing order and the exceedance probability is found using the well known Weibull formula

$$P(X>x_0) \equiv 1 - \frac{i}{N+1} \equiv \frac{N-i+1}{N+1} \tag{23}$$

where N is the total number of days in observed record, found by multiplying the total number of years, m, in the record with 365 and i varies between 1 and N. The utilization of different plotting formulas other than the one above does not affect the flow duration curve because the sample size N is high (Vogel and Fenessey 1994).

The observed flow duration curves are plotted together with the curves obtained for five different probability distributions, *i.e.* normal, Pearson type III, log-Pearson type III, 2-parameter lognormal and 3-parameter lognormal distributions for the selected three stations (Fig. 1 a, b, and c). In the low flow part, *i.e.* in the part where the exceedance probability is larger than 0.90, the LN2 distribution and LP3 distribution

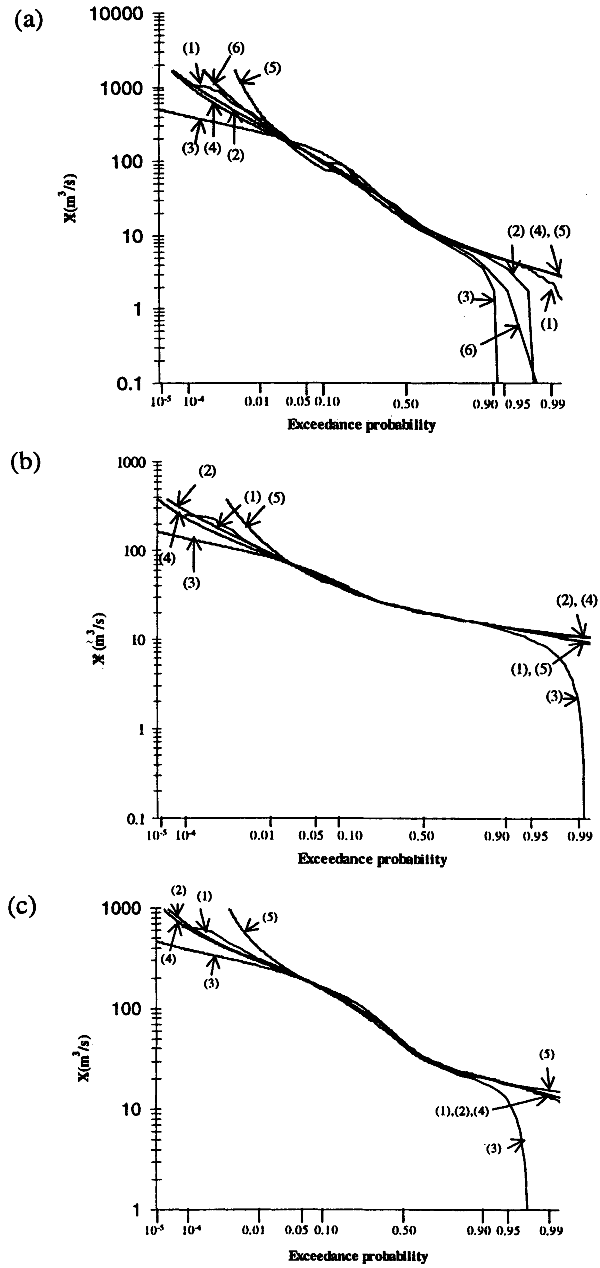

Fig. 1. *FDC*'s for the observed flow, 1), together with the Pearson type III distribution, 2), normal distribution, 3), and 2-parameter lognormal distribution, 4), log-Pearson type III distribution, 5), and 3-parameter lognormal distribution, 6), average of probabilities curves, for the stations 701, a), 2145, b), and 1714, c). (The LN3 distribution could be applied only to station 701)

curves were the ones closest to the observed *FDC*. The P3 distribution curve was close to the observed *FDC* only in the case of station 701. The LN2 and P3 distribution curves were the curves closest to the observed one for the upper part of the *FDC*, where the exceedance is less than 0.01. The Normal distribution curve showed a significant deviation from the observed one in both the low and high flow parts. During the application of LN3 distribution some problems appeared in the parameter computation stage. The ζ_i parameter had positive values for a significant number of days of the year creating negative 'x_i-ζ_i' differences making $\log(x_i-\zeta_i)$ computation impossible. However, the days with skewness very close to zero the ζ_i parameter had extremely high negative values so that even $\log(x_i-\zeta_i)$ values were very high. The application of this distribution was therefore limited to station 701 where only a few days had low positive ζ_i values. In these days the negative 'x_i-ζ_i' differences were replaced with x_i. The resulting LN3 curve for station 701 approximated the high flow part of the observed one well whereas it showed deviation in low flow part. Similar problems were mentioned by Yevjevich (1984) while considering LN3 distribution for daily streamflows. As a conclusion the P3 and LN2 distributions are found to be superior to the other three distributions for the extrapolation study where the exceedances for the flows higher than the maximum observed flow are to be investigated. In this study both distributions provided satisfactory exceedance approximations for the flow values in the $x_{\max}$-$1.5x_{\max}$ interval. This result is important especially for the flood studies. In the following part of the study however the LN2 distribution is preferred to P3 distribution since the latter provides a poor approximation in the low flow part. The LN2 distribution was also found adequate for daily flows by Beard (1943) and Fennessey and Vogel (1990). Beard suggested in his study to use of a lognormal probability density function to approximate flow duration curves. He proposed plotting flow *versus* exceedance probability on lognormal probability paper and drawing a best fit line through the data. Fennessey and Vogel employed an unconstrained optimization algorithm known as the conjugate gradient method to estimate the optimum values of the lognormal probability distribution parameters.

Comparison of the Method with the Mean and Median Flow Duration Curves

Vogel and Fennessey (1994) brought the median and the mean annual *FDC* concept to the literature. In this case an *FDC* is computed for each year of the period of the record. For each exceedance probability the median and mean values of discharge are computed using the *m* individual annual *FDC*'s, where *m* represents the duration of the historical record in terms of years. For example 27 separate annual *FDC*'s have to be established for Station 701. Considering 27 discharge values for each exceedance probability the mean and median are computed and the obtained values

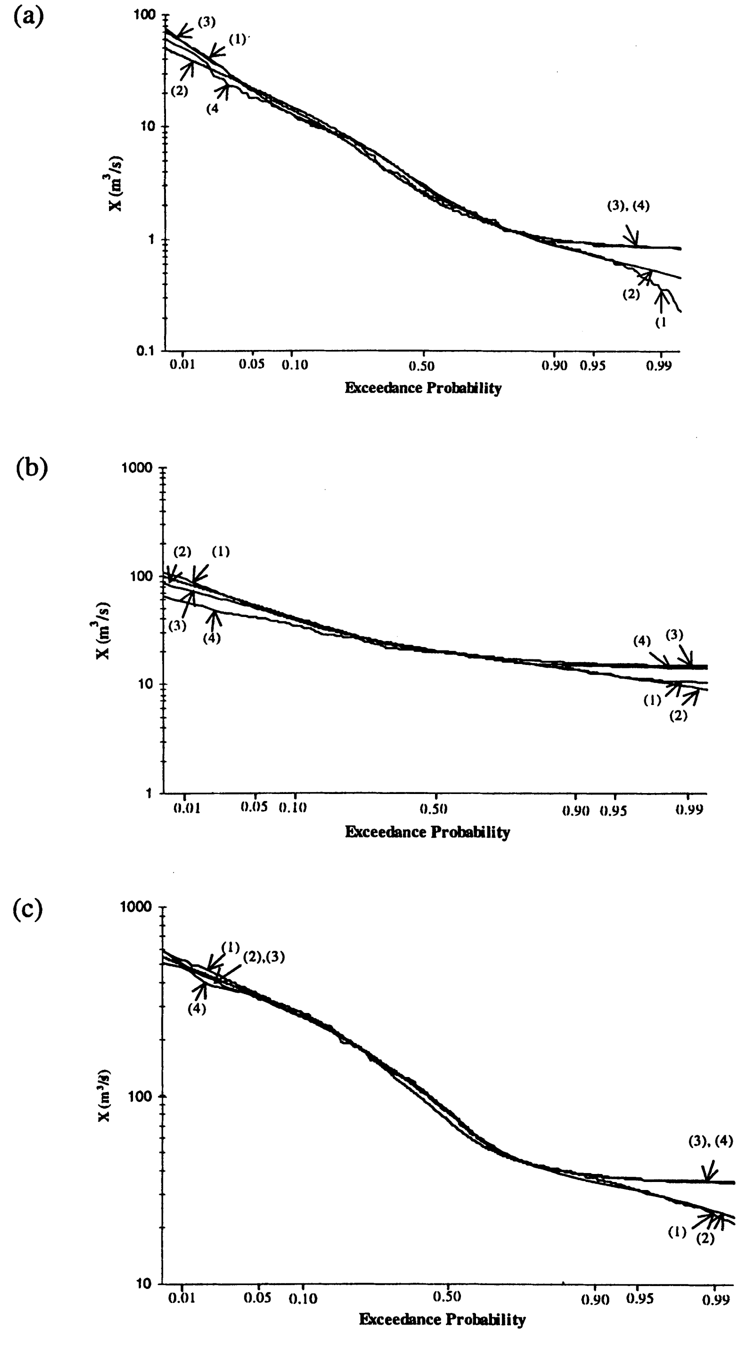

Fig. 2. Observed *FDC*, 1), together with the LN2 average of probabilities *FDC*, 2), mean *FDC*, 3), and median *FDC*, 4), for the stations 701, a), 2145, b), and 1714, c)

form the annual median and mean *FDC*'s. The mean and median *FDC*'s are compared with the observed ones and the *FDC* computed using the average of probabilities method for lognormal distribution for 3 stations (Fig. 2 a, b, and c). Average of probabilities method provided closer estimations for all quantiles compared with the other two curves. It is obvious that the mean and median annual *FDC*'s deviate from the observed curve significantly especially in the low flow part where exceedance is higher than 0.90. On the high flow part the mean *FDC* is closer to the original one compared with the corresponding median curve. This result is expected because the observed flow duration curve is highly sensitive to the hydrologic extremes, whereas the mean and median annual *FDC*'s are not as nearly sensitive. The median and mean annual *FDC* is supposed to deviate from the observed *FDC*, because the observed *FDC* gives steady state probability distribution, whereas the median annual and mean annual *FDC* represent 'typical' years. If ones interest is in steady state low flow response, then one should not be using the median or mean annual *FDC*. Application of the mean and median *FDC*'s produced the same results in the study of Vogel and Fennessey (1994) for the region with exceedance probability higher than 0.80. In their study high flow part was well approximated by both curves which is not the case in this study especially for the median *FDC*. This may bring an advantage for some water resources problems but in some cases like the computation of the compensation flow, which will be released to the downstream of a reservoir, the close approximation of the low flow part of *FDC* is significant. An overestimation of low flows may result in a high compensation flow estimation reducing the water quantity supplied to the region by the reservoir. When the flow record of the observed *FDC* contains years having missing seasons the low flow part of the curve may differ from the real one necessitating close approximation procedures. This issue is further investigated in the following part of the study further.

Application of the Method to the Missing Data Problem

In this part of the study the sensitivity of the average of probabilities curve to the missing flow record is investigated. It is assumed that there exists a historical record shorter than the original series. The statistical parameters, daily mean and standard deviation, are computed using the missing data and *FDC* is obtained using Eq. (5) for the three flow stations. As an example two possible scenarios are considered for station 701, *i.e.* it is assumed that the historical data is 25 years or 20 years long (the last two or seven years of the original data are missing respectively). The obtained average of probabilities curves are compared with the *FDC*'s of original duration (*e.g.* 27 years for station 701) for 3 stations (Fig. 3 a, b and c).

The shape of the average of probabilities curves look to be nearly unaffected for the 2 years data missing scenario and slightly affected for the 7 years data missing scenario. To be able to analyze the sensitivity of the average of probabilities method

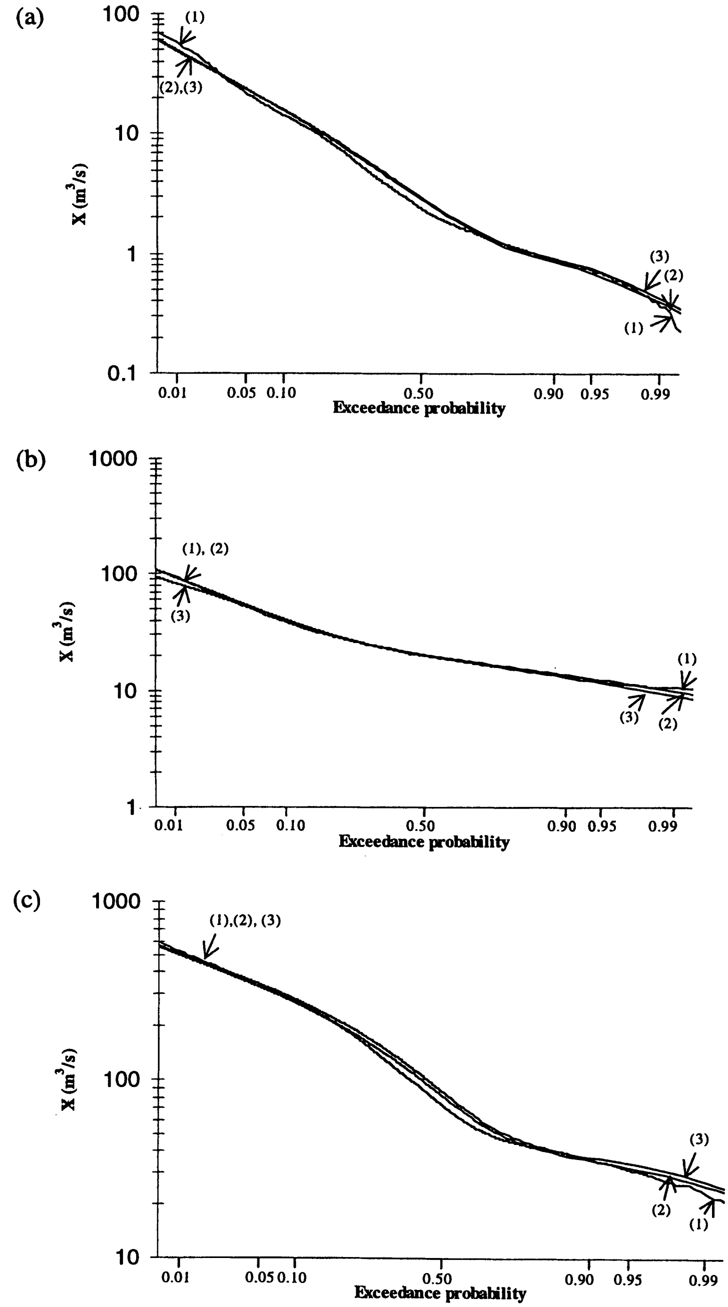

Fig. 3. Observed *FDC* for the whole flow record, 1), together with the LN2 average of probabilities *FDC*'s for the flow record with last two years missing , 2), and last seven years missing, 3), for the stations 701, a), 2145, b), and 1714, c).

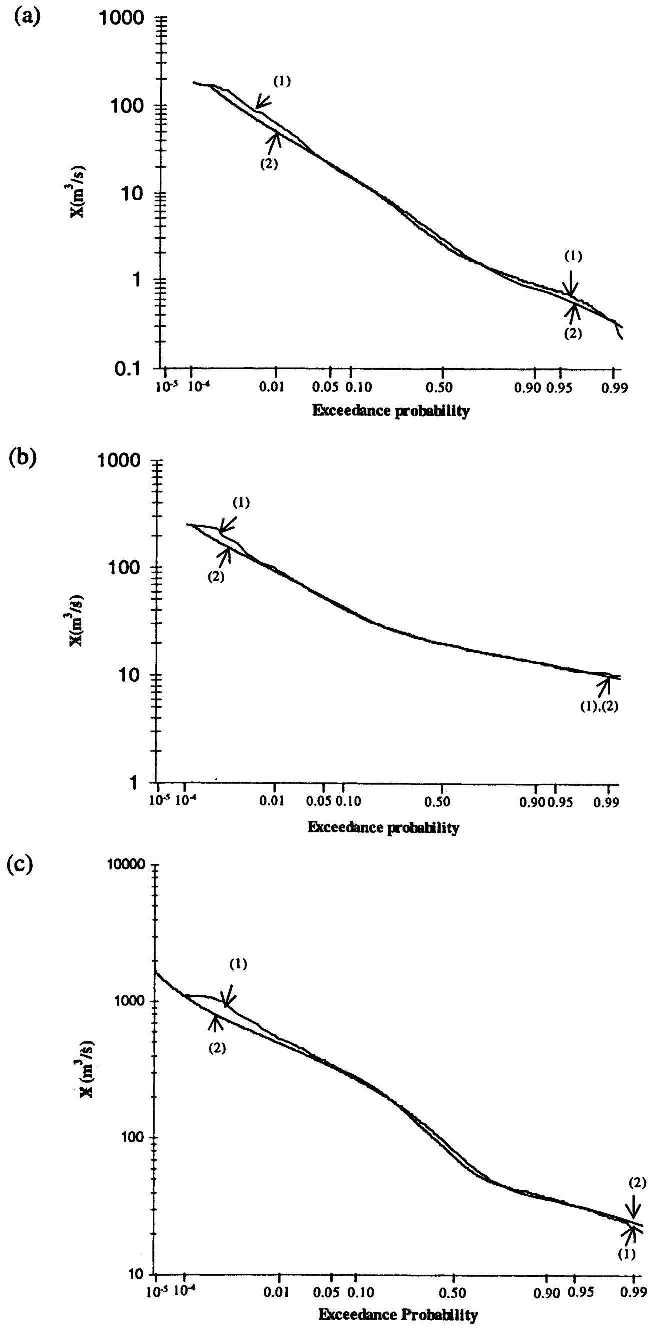

Fig. 4. *FDC*'s for the whole observed flow record, 1), compared with the LN2 average of probabilities *FDC*, 2), for the flow record with wettest and driest years missing for the stations 701, a), 2145 ,b), and 1714, c).

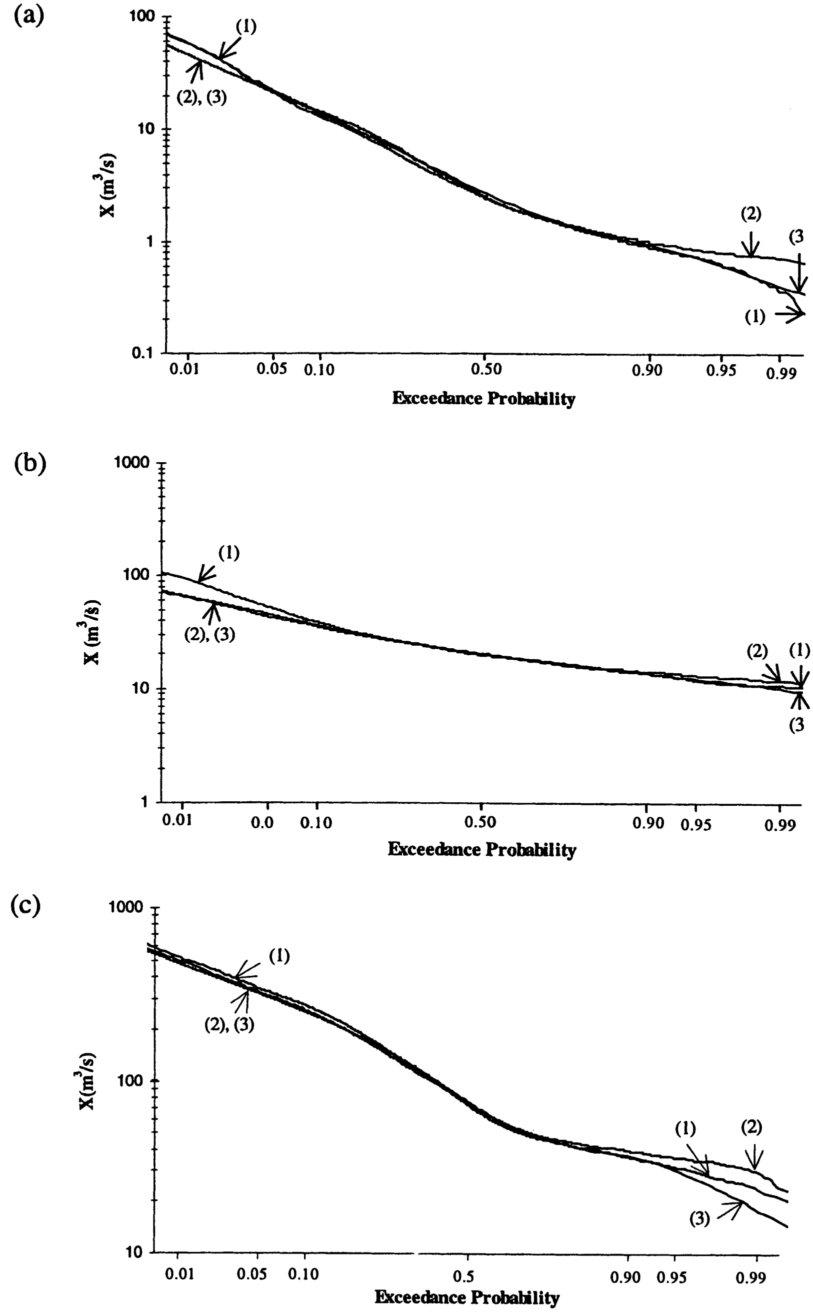

Fig. 5. The LN2 average of probabilities curve, 3), obtained for missing seasons, together with the observed *FDC* with missing seasons, 2), and for full record, 1), for the stations 701, a), 2145, b), and 1714, c).

to the wet and dry years application is realized to the observed record omitting the wettest and driest two years, *i.e.* totally neglecting four years. The average of probabilities curves are computed and plotted for this case together with the corresponding observed full record *FDC*'s for 3 stations in Fig. 4 a,b and c. The results are practically identical with the ones obtained for the previous two scenarios. The method curves are close to the observed one. This is important because it shows that the average of probabilities curve can be used in reliability even in the absence of the data having the maximum and minimum years missing. As a result it can be said that the method curves can be used for extrapolation even in the absence of the data covering years. The maximum observed flows in the threee stations with flow records having wettest and driest two years missing were 113 m^3/s, 167 m^3/s and 1118 m^3/s for stations 701, 1245 and 1714, respectively. The average of probability curves enabled the extrapolation of these values to 170 m^3/s, 251 m^3/s and 1677 m^3/s ,respectively, as illustrated in Fig. 4a, b and c.

As mentioned before the missing seasons in the flow record of the observed *FDC* may cause deviation from the full record curve. To be able to analyze the applicability of the method curve to this problem the following scenario is considered. The wettest months (February, March, April and May) of the wettest five years and the driest months (July, August and September) of the driest five years are assumed to be missing. The observed and method *FDC*'s for the flow record with missing seasons are compared with the full record *FDC* for three stations in Fig. 5a, b and c. In all the three stations the low flow part of the method curve is closer to the full record curve compared with the observed *FDC* for flow record with missing data. The high flow part of average of probability curve was indistinguishable from the corresponding missing record observed *FDC* in all three stations. Both curves showed some deviation in this part of the *FDC* from the full record observed *FDC*.

Application of the Method to the Ungauged Sites

There are various studies related to the computation of the flow duration curves for the ungauged sites in the literature. Quimpo *et al.* (1983) derived an exponential relation between the exceedance probability and the specific discharge using the data of 35 stations. In this equation two parameters, Q_A and c, were employed. The former parameter, Q_A, was obtained with the help of two other coefficients, p and m, and the catchment area, A, using a power type of relation. Similarly Mimikou and Kaemaki (1985) found a regressional relation between the exceedance probability and the monthly mean flow using four coefficients which were computed using the morphological characteristics like the catchment area, the hypsometric fall and the mean annual areal precipitation. Their study was based on the flow data of 11 stations. Fennessey and Vogel (1990) established regional regressions based on the basin parameters drainage area (A) and basin relief (H) to find the parameters μ' and σ' using the flow records of 23 stations in the region. These parameters obtained for an ungauged site were used to obtain the exceedance probabilities with an exponential equation. The regression *FDC*'s were compared well with the observed curves for the exceedance probabilities equal and higher than 0.50. Smakhtin *et al.* (1997) constructed non-dimensional *FDC*'s for 17 stations by dividing discharges by the long-term mean daily flow. Then all individual *FDC*'s in the region were superposed on one plot and a composite regional non-dimensional *FDC* was obtained. They used a regression equation to estimate the long-term mean of an ungauged site using the catchment area (A) and mean annual precipitation (MAP) values.

It is obvious that the studies in the literature to derive the regional flow duration curve for an ungauged site are based on the flow data of the numerous gauged stations in the region and are applicable in the availability of various morphoclimatic characteristics. These methods can not be employed if there are limited number of gauged station data in the region. The problem becomes more complicated if one has the catchment area, A, of the ungauged site as the unique available morphoclimatic characteristic. That is why Cigizoglu and Bayazit (1999) tried to compute the coefficient C_1 with a regression equation having the catchment area and the average of the daily coefficient of variation values, c_v', as the independent variables, using the daily flow data available in various regions of Turkey. Here C_1 is the coefficient of first harmonic for the daily mean flow.

The commonly used method for this case in practice is taking the average of the specific discharges of the stations in the region for each exceedance probability. The *FDC* for the ungauged site is then computed simply by multiplying this average value with the corresponding catchment area and this is repeated for all exceedances. In this part of the study the average of probabilities method is applied to derive the *FDC* in an ungauged site. The station 1720 was selected for the application. There is daily flow data available for four stations (1708, 1714, 1717 and 1719) in the region of station 1720. The related information for this data is presented in Table 2. The daily flow data for each station is non-dimensionalized by dividing with the corresponding catchment area. The exceedance probabilities for selected specific discharge values are computed as explained previously for the LN2 distribution case for each station separately. Then the average of these exceedances is taken for each selected specific discharge. The *FDC*'s for the station assumed to be ungauged is obtained then simply by multiplying the specific discharges with the corresponding catchment area value. The resulting curve is compared with those of the observed *FDC* and the classical specific discharge averaging method in Fig. 6. The average of probabilities method curve was closer to the observed *FDC* in all regions of the *FDC*. In the low flow part the classical method curve showed significant deviation from the observed one whereas the method curve was very close. This result is significant because the average of method curve carries also the advantage of extrapolation which is not possible with the classical method.

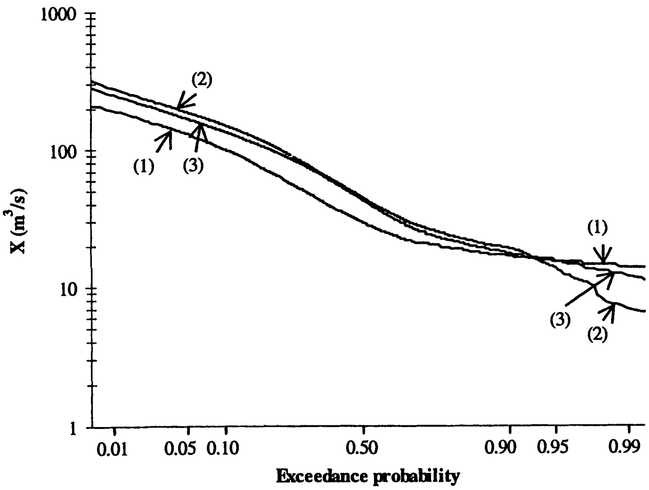

Fig. 6. The average of probabilities curve, 3), together with the average of specific discharges curve, 2), and observed *FDC*, 1), for station 1720.

Summary and Conclusions

A method based on the average of probabilities is developed to obtain the flow duration curve of the daily river flows. The method is applied to the present data using different probability distributions including the normal, Pearson Type 3, Log Pearson Type 3, 2-parameter and 3-parameter lognormal distributions. It is seen that the method produced the best results when 2-parameter lognormal distribution is employed. In that case the daily mean and standard deviation values are required to compute the *FDC*. *L*-moments method is preferred to the classical method of moments to obtain unbiased estimates for these parameters. The method enables the extrapolation of observed *FDC* providing some further information for the flood studies.

The method was then applied to the frequently encountered missing data and ungauged site problems. The method curves compared very well with the available alternative methods in the literature providing reliable estimates for the observed flow duration curve. In the event that one has a complete record of streamflow, the presented approach is identical to fitting a probability distribution to the observations directly, which are easier than the approach suggested. However, the developed method is really useful when one has either missing observations, an ungaged stream, or one is missing key flood or drought observations. Studies about the applicability of the method should be broadened with the availability of further data. It is aimed that the developed method will shed a light in the related hydrological studies.

References

Abramowitz, M., and Stegun, I. A. (1972) *Handbook of Mathematical Functions,* 10th Edition, General Publishing Company Ltd., Canada, p. 932.

Beard, L. R. (1943) Statistical analysis in hydrology, *Trans., ASCE, Vol. 108,* pp. 1110-1160.

Chow, W. T. (1964) *Handbook of Applied Hydrology,* Mc-Graw Hill Book Co., New York, N.Y.

Cigizoglu, H. K. (1997a) Mathematical models of the flow duration curves, PhD Thesis, Istanbul Technical University, Civil Engineering Faculty, Division of Hydraulics, pp. 120.

Cigizoglu, H. K. (1997b) Flow duration *curve* and its applications, Friend-AMHY Conference, Slovenia, October 1997.

Cigizoglu, H. K. (2000) Closure for the discussion by Srinivasan and Srinivas, *J. Hydrologic Engr. Div. Am. Soc. Civ. Eng., Vol. 5* (1), pp. 107-111.

Cigizoglu, H. K., and Bayazit, M. (1998) Application of the gamma autoregressive model in the analysis of the dry periods, *J. Hydrologic Engr. Div. Am. Soc. Civ. Eng., Vol. 3* (3), pp. 218-221.

Cigizoglu, H. K., and Bayazit, M. (2000) A generalized seasonal model for flow duration curve, *Hydrologic Pocesses, Vol. 14* (9).

Dingman, S. L. (1978), Synthesis of flow duration curves for unregulated streams in New Hampshire, *Water Resources Bulletin, Vol. 14* (6), pp. 1481-1502.

Fennessey , N. M., and Vogel, R. M. (1990) Regional flow duration curves for ungauged sites in Massachusets, *Journal of Water Resources Planning and Management, ASCE, Vol. 116* (4), pp. 530-549.

Foster, H. A. (1934) Duration curves, *Trans. ASCE, Vol. 99,* pp. 1213-1267.

Hosking, J. R. M., (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics, *J.R. Stat. Soc., Ser. B, Vol. 52* (2), pp. 105-124.

Hosking, J. R. M., and Wallis, J. R. (1997) *Regional frequency analysis* (An approach based on L-moments), Cambridge University Press, 224 pages.

Hughes, D. A., and Smakhtin, V. (1996) Daily flow time series patching or extension: A spatial interpolation approach based on flow duration curves, *Hydrological Sciences Journal, Vol. 41* (6), pp. 851-871.

Kirby, W. (1972) Computer-oriented Wilson Hilferty Transformation that preserves the first three moments and the lower bound of the Pearson Type 3 Distribution, *Water Resources Research, Vol. 8* (5), pp. 1251-1254.

LeBoutillier, D. W., and Waylen, P. R. (1993) A stochastic model of flow duration curves, *Water Resources Research, Vol. 29* (10), pp. 3535-3541.

Limbrunner, J. F., and Vogel, R. M. (2000) Estimation of the harmonic mean of a lognormal variable, *J. Hydrologic Engr. Div. Am. Soc. Civ. Eng., Vol. 5* (1), pp. 59-66.

Maidment, D. R. (1992) *Handbook of Hydrology*, Mc Graw Hill.

Male, J. W., and Ogawa, H. (1984) Tradeoffs in water quality management, *Journal of Water Resources Planning and Management*, ASCE, Vol. 110 (4), pp. 434-444.

Mimikou, M., and Kaemaki, S. (1985) Regionalization of flow duration characteristics, *Journal of Hydrology, Vol. 82,* pp. 77-91.

Quimpo, R. G., Alejandrino, A. A., and McNally, T. A. (1983) Regionalized flow duration curves for Philippines, *J. Wat. Resour. Planning Management*, 109, 320-330.

Salas, J. D., Delleur, J. W., Yevjevich, V., and Lane, W .L. (1980) *Applied modelling of hy-*

drologic time series, Water Resources Publications.
Singh, K. P. (1971) Model Flow Duration and Streamflow Variability, *Water Resources Research, Vol. 7* (4), pp. 1031-1036.
Smakhtin, V. Y., Hughes, D. A. and Creuse-Naudin, E. (1997) Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa, *Hydrological Sciences Journal, Vol. 42* (6), pp. 919-936.
Vogel, M., and Fennessey, N. M. (1993), L moment diagrams should replace product moment diagrams, *Water Resources Research, Vol. 29* (6), pp. 1745-1752.
Vogel, M., and Fennessey, N. M. (1994), Flow-duration Curves. I: New interpretation and confidence intervals, *Journal of Water Resources Planning and Management, ASCE, Vol. 120* (4), pp. 485-504.
Vogel, M., and Fennessey, N. M. (1995) Flow-duration Curves II: A review of applications in Water Resources Planning, *Water Resources Bulletin, Vol. 31* (6), pp. 1029-1039.
Yevjevich, V. M. (1984) *Structure of daily hydrologic series*, Water Resources Publications.
Yu, P. and Yang, T. (1996) Synthetic regional flow duration curves for Southern Taiwan, *Hydrological Processes, Vol. 10,* pp. 373-391.
Warnick, C. C. (1984) *Hydropower Engineering,* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, pp. 59-73.

Received: 16 April, 1999
Revised: 29 October, 1999
Accepted: 24 March, 2000

Address:
Current
School of Engineering,
University of Durham,
DH1 3LE Durham,
U.K.

Permanent
Division of Hydraulics, Civil. Eng. Faculty,
Istanbul Tehnical University,
Maslak 80626,
Istanbul,
Turkey.