*Journal of Engg. Research* **Vol. 2 - (1) March 2014** pp. 33-46

# تأثير الوقود المستخرج من زيت النخيل كمعدل لمادة البيتومين لأجل تحسين مقاومة التهالك

***محمد رسلي هاينن، *رمضان سياه بوتراجايا، *نور اسما علي أكبر، **ديوي سري جيانتي و ***نور ايزي ام دي يوسف**

*كلية أبحاث الهندسة المدنية والبناء، الجامعة التكنولوجية الماليزية، ماليزيا

**كلية الهندسة الزراعية، جامعة سيا كوالا، أندونيسيا

*** كلية الهندسة المدنية والبناء، جامعة كبانسجان، ماليزيا

## خلاصة

إن الزيادة المطردة للمخلفات الناتجة من مصانع زيوت النخيل سبب مشكلة تلوث في البيئة ويمكن التعامل مع هذه المشكلة بالتدوير وإعادة الاستخدام. في هذه الورقة تم استخدام رماد وقود زيت النخيل- وهو ناتج جانبي في مصانع زيوت النخيل - كمادة معدلة للبيتومين. تم دراسة خواص البيتومين المعدل من حيث إمكانية الإحتراق ونقطة الليونة، وجدول الاختراق، ورقم لزوجة الاختراق وفحوصات اللزوجة. تم تنفيذ طريقتين لتمثيل التهالك (aging) في هذه الدراسة. الاولى الطريقة التقليدية فحص الفلم الرقرق بالفرن. والثانية فحص وعاء الصغط. أظهرت النتائج المخبرية زيادة في نقطة الليونة وانخفاض في الاختراق لعينة البيتومين بعد خلطها برماد وقود النخيل. إن إضافة هذا الرماد- وخاصة بكميات كبيرة - زاد من تحمل البيتومين للحرارة.

إن استخدام هذا الرماد أدى بصورة عامة إلى نقصان أثر معدل التقادم على الخواص الفيزيائية والانسيابية للبيتومين كما ظهر من نقصان مؤشر العمر لدرجة اللزوجة.

# Regional flood frequency analysis for gauged and ungauged cathments of seyhan river basin in Turkey

NESLIHAN SECKIN*, RECEP YURTAL*, TEFARUK HAKTANIR**

**University of Cukurova, Department of Civil Engineering, Adana, TURKEY*

***University of Erciyes, Department of Civil Engineering, Kayseri, TURKEY*

## ABSTRACT

Regional flood frequency analysis for the Seyhan Basin in Turkey is done by two different approaches, and generalized growth curves to be used at un-gauged sites are developed by either method. First, the homogeneity of the annual flood peaks series recorded at 11 un-regulated gauging stations in the Seyhan Basin is analyzed by the conventional $C_v$ (coefficient of variation) test; and second, discordance, homogeneity, and goodness-of-fit by the Z tests by the method of L-moments are performed. The Seyhan Basin is subdivided into three homogeneous sub-basins from the aspect of annual flood peaks based on the $C_v$ test, which detects regions having recorded series with variation coefficients close in magnitudes. The probability distributions of Log-Logistic, Pearson-3, log-Pearson-3, Gumbel, and Wakeby, whose parameters are computed by the method of probability-weighted moments, are taken as candidate distributions. The goodness-of-fit tests of Kolmogorov-Smirnov and Crammer Von Misses are applied, and the Wakeby distribution is found to be superior to the others. According to the analyses by the L-moments method, the Seyhan Basin as a whole turned out to be a homogeneous region. By inspection of the L-moment ratio diagrams of L-variation versus L-skewness and L-skewness versus L-kurtosis of those 11 series, and by the Z-statistic analysis of the L-moments method, the Log-Logistic distribution is found to represent the recorded series of annual flood peaks more closely. By regression, an expression relating average peak discharge ($Q_{av}$) to the catchment area is obtained, and growth curves for the frequency of flood peaks ($Q_{max.}$) by both the $C_v$-based and the L-moments methods are developed using the data of those 11 gauged sites.

**Keywords:** $C_v$-based method; flood frequency analysis; l-moments.

## INTRODUCTION

The purpose of flood frequency analysis is to make a rational prediction for the relationship between flood magnitudes, $Q_T$, versus their average return periods, T, which, in diagrammatical form is called a growth curve. An accurate estimation of flood magnitude $Q_T$ having a return period of T, as stipulated in pertinent regulations, is essential in dimensioning many hydraulic structures. Magnitudes of hydrologic events such as annual maximum flow (the same as annual flood peak), or annual maximum precipitation are random variables, and deterministic prediction of such random variables is not possible. If hydrologic information and recorded data at the sites of concern are available, flood peaks can be predicted by the statistical method called flood frequency analysis, which is performed by two different methods, and these are:

1 - At-site flood frequency analysis

2 - Regional flood frequency analysis

Accurate estimates of flood peaks at a potential dam site are difficult because (1) usually there is not a gauging station at that cross-section of the stream and (2) the observed data record is often too short to make a statistically significant frequency analysis. Regional frequency analysis can resolve this problem by using data from several sites, which are judged to have frequency distributions similar to the site of interest. In this study, a regional flood frequency analysis is applied for the Seyhan River Basin in Turkey using the hydrometric and geographical data of 11 selected stations in the basin. The locations of these stations are shown in Figure. 1

If flood frequency characteristics of all stations in the same region are approximately equal, that region is called homogenous and it is different from other regions for the flood producing mechanism. The objective of this study has been to carry out a regional flood frequency analysis first by a conventional method, and next by the lately popular L-moments approach, and the Seyhan Basin is taken as the application area. The conventional method is the so-called $C_v$-based test, which identifies homogeneous regions containing sites whose recorded series exhibit histograms of close variances. Because the coefficient of variation ($C_v$) is an objective statistic for the true measure of wideness or narrowness of distribution of a random variable, the $C_v$-based test determines those homogeneous regions having $C_v$'s of magnitudes close to each other. The scheme proposed by Hosking and Wallis (1997) based on the method of L-moments has gained popularity in recent years, and accordingly regional flood frequency analysis is also applied by the L-moments method again for the Seyhan Basin.

First, a regression equation is obtained for the average unit flood peak versus the catchment area as suggested by Wiltshire (1986), which is to be used as the final step after having determined the $Q_{max.}$ out of the growth curve developed by either of the two methods. The Seyhan Basin is divided into sub-areas on the basis of magnitudes of variation coefficients of the recorded annual flood peaks series by the ($C_v$)-based test. Annual maximum flood records are normalized by dividing them with the mean value of annual maximum flood records. In this way, dimensionless floods are obtained (Wiltshire, 1986). The flood records at the homogeneous region are combined together. Seven probability distributions commonly used for flood frequency are applied to the combined data, and regional frequency curves (growth curves) are obtained. In order to decide on the better fitting ones out of these seven candidate distributions, Kolmogorov-Smirnov and Crammer Von Misses goodness-of-fit tests are applied. Next, the regional flood frequency analysis of the Seyhan Basin is also performed by the L-moments procedure as suggested by Hosing and Wallis (1997).

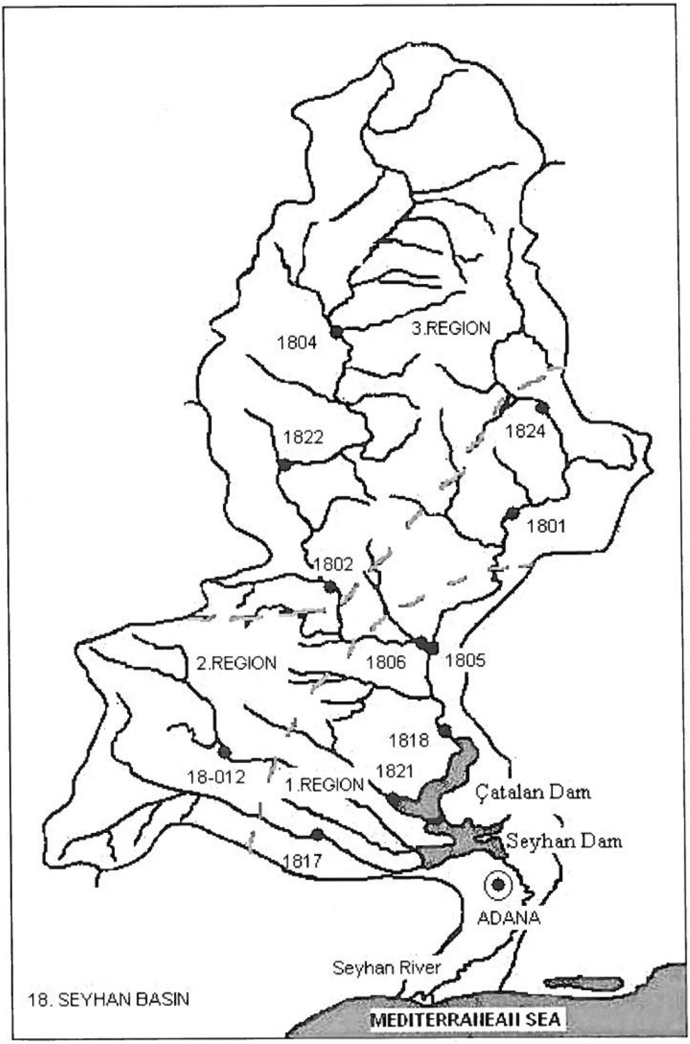


**Fig. 1.** Seyhan Basin and the borders of homogeneous regions

## REVIEW OF LITERATURE

Regional flood frequency analysis is one of the most active areas of research (e.g., Cunnane, 1988, 1989; National Research Council, 1988; Smithers & Schulze, 2001; Stedinger *et al.*, 1993). The homogeneity test and the index flood technique were first introduced by Dalrymple (1960). Wiltshire (1986) pointed out some drawbacks of the Dalrymple tests. Parameter estimates from small samples computed by the probability-weighted moments (PWM) method are sometimes more accurate than the maximum likelihood (ML) estimates, which was originally proposed by Greenwood *et al.* (1979). Since then it has been used widely in practice and in research studies also (e.g., Bobee, B. & Rasmussen, 1995; Cunnane, 1989; Rao & Hamed, 1997; Sankarasubramanian & Srinivasan, 1999). Hosking *et al.* (1984) investigated the properties of the parameters estimated by the PWM method for the generalized extreme values (GEV) distribution. Hosking *et al.* (1985) showed that the PWM method is superior to the ML method when the GEV distribution is used for longer return periods, i.e., $T > 100$. L-moments of a random variable were introduced by Hosking (1986). Hosking (1986, 1990) defined the L-moments as linear combination of the PWMs. Hosking (1990) used the L-moment ratio diagrams to identify the underlying parent distributions and the L-moment ratios for testing goodness-of-fit of different distributions. Hosking and Wallis (1993) suggested a homogeneity test based on the L-moment ratios. Hosking and Wallis (1993) extended the use of L-moments and developed statistics to measure discordance, regional homogeneity, and goodness-of-fit that can be used for regional frequency analysis. The details of the L-moments approach are available in current literature (Hosking & Wallis, 1997; Hosking, 1990). Jaiswal, *et al.* (2003), Kumar, *et al.* (2003), Rao & Hamed (1997), Parida, *et al.* (1998) and several others have investigated various issues involved in the regional flood frequency modeling with L-moments. Haktanir (1990, 1992) developed a computer code for the at-site flood frequency analysis which includes a few distributions and performs the Chi-square, Kolmogorov-Smirnov, and Cramer Von Misses goodness-of-fit tests. Based on these classical goodness-of-fit tests Haktanir found out that the Log-Logistic and Log-Pearson-3 distributions fit the annual flood peaks series of natural streams of Turkey better than the others.

## MATERIAL AND METHOD

### Identification of Homogeneous Regions

In this study, the $C_v$-based test is preferred for identification of homogeneous regions because of its simplicity and its being powerful in defining recorded series of similar variances. According to this test, any region in the basin should have two properties:

1 - Distinction from other groups

2 - Homogeneity of flood frequency characteristics to allow the group-average flood frequency curve to be accurately defined.

Mainly, the $C_v$ test delineates a homogeneous region including those sites having similar variation coefficients. Because the variance of a histogram of a recorded sample series or the variance of the density function of a distribution assumes magnitudes in parallel to the mean of the distribution, the coefficient of variation ($C_v$ = standard deviation / mean) is a true measure of wideness or narrowness of the distribution of the random variable within its range. Hence, this test aims to put those sites having similar variances in the same group. A narrow distribution of flood peaks should result from a hydrologically different flood producing mechanism than a wide distribution. And, the idea behind this test is: if streams in a particular geographical area are to be within the same homogeneous region, the variances of histograms of their flood peaks should exhibit similar appearances. A detailed description of the $C_v$-based test was given by Wiltshire (1986).

## $C_v$-based Test

In order to depict the variations in variation coefficients quantitatively, a variance ratio test, which is called F test is defined. The aim of this test is that variance of the site $c_v$'s within a group is to be minimized, and between groups variance of $c_v$'s is to be maximized. The F statistic is defined as:

$$F = \frac{between\ group\ var\ iance\ of\ Cv's}{within\ group\ var\ iance\ of\ Cv's} \tag{1}$$

This F statistic is formulated as follows:

$$A = \sum_k \sum_j \frac{(cv_{ik} - cv_{.k})^2}{u_{ik}} \tag{2}$$

$$B = \sum_k \left( \sum_j \frac{1}{u_{jk}} \right) (cv_{.k} - cv_{..}) \tag{3}$$

k and j in these equations denote groups and sites respectively. $m_k$ is the number of sites in group $k$, $n_{jk}$ is the length of record at $j^{th}$ site in group k, $cv_{jk}$ denotes sample coefficient of variation of site j in group $k$, $u_{jk}$ is the sampling variance of $cv_{jk}$, A is the within-group variation, B is the between-groups variation, $cv_{.k}$ is arithmetic average of the $c_v$'s in group $k$, $c_{v..}$ is arithmetic average of the $c_v$'s across all groups.

The coefficient of variation is defined as:

$$Cv = \frac{S_x}{\overline{x}} \tag{4}$$

where, $S_x$ denotes the sample standard deviation of the series, and $\overline{x}$ denotes the arithmetic average of the series. If there are $p$ number of groups in the basin, the F statistic is defined as:

$$F = \frac{B/(p-1)}{A/(\sum m_k - p)} \tag{5}$$

$S_k$ is total variation within groups. $S_k$ can be used as a test of within-group homogeneity.

$$S_k = \sum \frac{(cv_{jk} - cv_{.k})^2}{u_{jk}} \tag{6}$$

The F test quantitatively indicates the measure of the groups of basins being significantly different from each other. Degrees of freedom of $S_k$ is ($m_k$-1), which is a random variable obeying a $\chi^2$ distribution. The variance term of sampling can be determined as the average of the jack-knife estimation of variance computed for each site within a group. The jack-knife variance estimator was defined by Parr (1983).

## Regional flood frequency curve

The regional flood frequency curve represents the flood magnitude versus return period relationship ($Q_{max.} \leftrightarrow T$), which is a unique curve in a homogeneous region. The distribution of flood magnitudes is expressed as a function of the Gumbel-reduced variate $y$ ($y$ = -ln[-ln(1-1/T)], which is a common practice. The $y$ variate can be interpreted to be the average return period, T, expressed in a compact form. Annual maximum floods at sites in a homogeneous region are pooled. Then, these magnitudes are arranged in an order from minimum to maximum. In this study, the non-exceedance probability (F) of such a flood peak is estimated out of the recorded series by the Gringorten plotting position formula as also suggested by NERC (1975). The numerical values for F by the plotting position formulas of Cunnane (1989) and Chegodayev are actually fairly close to the Gringorten formula, which is:

$$F_{ij} = \frac{(i - 0.44)}{(n_j + 0.22)} \tag{7}$$

Next, the Gumbel reduced variate is calculated by:

$$y_{ij} = -Ln(-Ln(F_{ij})) \tag{8}$$

In these equations i denotes the rank of a random variable, j denotes site, $n_j$ is the number of random variable at $j^{th}$ site, $F_{ij}$ is non-exceedance probability of $i^{th}$ random variable at site j, and yij is gumbel-reduced variate.

According to the index flood method by the L-moments approach, the non-exceedance probability of a random variable is determined by the biased-Landwehr (also called Hosking, herein) plotting position formula, which is equation (9) below. The L-moment ratios diagrams, and the Z-statistics have been used for selecting the best-fit-distribution. The L-moment ratios diagrams are two graphs: one between L-skewness and L-variation, and the other between L-kurtosis and L-skewness, and they can be used to select a suitable probability distribution for a region. The idea is: if a probability distribution is a good fit to a sufficiently long recorded series of a random variable, then the plot of sample L-$C_v$ versus L-skewness should match the analytical plot of the L-$C_v$ versus L-skewness peculiar to that distribution (Hosking & Wallis, 1997). Similarly, the plot of sample L-skewness versus L-kurtosis should match the analytical plot of the L-skewness versus L-kurtosis peculiar to that distribution. There are formulas for computing the sample L-moments ratios using all the *n* number of elements of a recorded series arranged in ascending order (Hosking & Wallis, 1997).

$$F_{ij} = \frac{(i-0.35)}{(n_j)} \tag{9}$$

## RESULTS AND DISCUSSIONS

### Conventional Method

There have been some attempts to identify flood regions by methods other than geographical delimitation. These "alternative" techniques allocate basins to regions or groups by one of two general processes;

a. Classification by statistics of the basin flood frequency distribution;

b. Classification by basin characteristics.

For example, in New Zealand, Mosley (1981) used cluster analysis to form groups of basins characterized by specific mean annul flood and coefficient of variation.

A homogenous flood frequency region will contain annual maximum flood populations whose flood frequency relationships have similar slopes on a probability plot. Since the slope of a probability plot is related to the coefficient of variation ($C_v$) of the standardized annual maximum flood series it is reasonable to develop a homogeneity test based upon the regional variability in the site $C_v$'s (Wiltshire, 1986).

A homogeneity test was applied to the annual flood peaks series recorded at 11 stream-gauging stations of the Seyhan Basin, some characteristic information of which is given in Table 1. When the entire Seyhan basin is taken as a single region, the magnitude of the $S_k$ statistic computed by Eqn. (6) turns out to be 20.77. Using a degree of freedom of 10 (dof = 11-1 = 10), and the levels of significance of $\alpha=0.01$ and $\alpha=0.05$, the critical Chi-squared values taken from the table of Chi-square distribution are: 4.46 and 8.65. Since the calculated value 20.77 is greater than the critical values, the null hypothesis ($H_0$) of "there is no difference between the coefficients of variations in the Seyhan basin as a whole" was rejected at both significance levels. Namely, according to the magnitudes of variation coefficients of the recorded series, the Seyhan basin cannot be analyzed as a single homogeneous region. Both the magnitudes of the coefficients of variations of the 11 sites in the Seyhan basin (Table 1) and the geographical positions of the sites were analyzed, and based on the locations of the gauging stations, the Seyhan basin was divided in to three regions. Region one has five, region two has three, and region three has three gauging stations respectively. The $C_v$'s of five stations in region 1 are approximately in the interval of: $0.50 < C_v < 0.57$, those of three stations in region 2 are in: $0.60 < C_v < 0.65$, and those of three stations in region 3 are in: $0.36 < C_v < 0.41$. The homogeneity test was applied to each region and the results are presented in Table 2.

**Table 1.** some relevant information about the Seyhan Basin

| Station Name | Station Number | Name of the Stream | Record Length (year) | Drainage Area (km$^2$) | Elevation (m) | Coefficient of variation |
|---|---|---|---|---|---|---|
| GÖKDERE | 1805 | Göksu River | 51 | 4242.8 | 350 | 0.542 |
| ERGENUŞAĞI | 1806 | Zamanti River | 36 | 8698.4 | 347 | 0.569 |
| ARAPALİ | 1817 | Çakt Creek | 17 | 1582.4 | 150 | 0.497 |
| ÜÇTEPE | 1818 | Seyhan River | 24 | 13846 | 180 | 0.501 |
| SARIMEHMETLİ | 1821 | Eğlence Creek | 16 | 628.8 | 75 | 0.503 |
| HYMMETLİ | 1801 | Göksu River | 55 | 2596.8 | 665 | 0.610 |
| KAMIªLI | 18-012 | Körkün Creek | 19 | 1065 | 1109 | 0.621 |
| ÇUKURKIªLA | 1824 | Göksu River | 34 | 1526.4 | 1200 | 0.652 |
| FARASA | 1802 | Zamanti River | 19 | 7379.3 | 858 | 0.365 |
| SÖĞÜTLÜ | 1804 | Zamanti River | 20 | 4389.2 | 1345 | 0.407 |
| FIRAKTIN KÖP. | 1822 | Zamanti River | 21 | 6334.8 | 1270 | 0.368 |

**Table 2.** Results of the Regional Homogeneity Analysis Using Wiltshire's Method

| Region | n | S | $\chi^2_{0.90} - \chi^2_{0.95}$ | Comments |
|---|---|---|---|---|
| Region1(1805, 1806, 1817, 1818, 1821) | 5 | 7.331 | 9.49-13.28 | Homogeneous |
| Region2(1801, 18-012, 1824,) | 3 | 4.595 | 5.99-9.21 | Homogeneous |
| Region3(1802, 1804, 1822) | 3 | 3.124 | 5.99-9.21 | Homogeneous |
| **ALL** | 11 | **F** = 20.773 | 4.459-8.649 | Heterogeneous |

Lastly, the differences between these three homogeneous regions were analyzed. The value of the F statistic computed by Eqn. (5) was 20.773. Since the F value obtained from the F probability distribution table is smaller than the computed F value, all these three regions are said to be different from each other sufficiently in both significance levels.

In this study, the estimates of the non-exceedence probabilities of the elements of the recorded sample series, $P_{ij}$, were calculated with Gringorten formula, Eqn(7), and then, the $y_{ij}$ values were computed by Eqn.(8). To obtain a smooth curve free of unnecessary noises and congestion, the $y_{ij}$ values were divided into sub-intervals of slice widths of 0.5 and the average values of each sub-interval were used in calculations. These average values of $X_k$ and $y_k$ are presented graphically in Figures 2, 3, and 4. The abscissa coordinates of the plotted points in these figures representing the Gumbel reduced variate extends all the way up to a *y* of 5.296, which is the *y* value for the return period of 200 years.

The Wakeby, Gumbel, log-Pearson-3, Pearson-3, log-normal, and log-logistic (also known as Log-Logistic) probability distributions were applied to the Seyhan Basin and their parameters were computed by the method of PWM (Table 3).

**Table 3.** Regional Dimensionless Probability Weighted Moments (PWM) for the Seyhan Basin

| Region | $M_0$ | $M_1$ | $M_2$ | $M_3$ |
|---|---|---|---|---|
| 1 | 1.000 | 0.371 | 0.215 | 0.147 |
| 2 | 1.000 | 0.345 | 0.193 | 0.130 |
| 3 | 1.000 | 0.401 | 0.238 | 0.166 |

For the return periods of T = 2, 5, 10, 25, 50, 100, and 200 years, values of $Q_{max.}$ and their corresponding *y* values (reduced gumbel values) were calculated. Along with the flood frequency curves which were computed for each region by any one of the chosen distributions (Figures 2, 3, and 4), the averaged points computed by the Gringorten plotting position formula are also shown. Obviously, a good-fit probability distribution is more significant than any plotting position formula, and therefore, the frequency curve of the best-fit distribution is more meaningful than the plots given by the plotting position formula. It is a traditional practice however, to give both plots simultaneously, and somehow, the analyst expects them to be in conformity to some extent. The parameters of the distributions are given in Table 4, and $Q_{max}$ values are given in Table 5.

**Table 4.** Parameters of some probability distributions for the Seyhan basin

| Distributions | Parameters | Region 1 | Region 2 | Region 3 |
|---|---|---|---|---|
| LN3-PWM | Regional Skew | -0,618 | -0,301 | 0,057 |
| | Regional Shape | 0,648 | 0,665 | 0,349 |
| | Regional Scale | 376,856 | 95,596 | 78,615 |
| | Regional Loca. | 90,672 | 16,579 | -10,889 |
| GUMBEL-PWM | Regional Scale | 0,000 | 0,000 | 0,000 |
| | Regional Loca. | 420,569 | 100,108 | 59,311 |
| PEARSON 3 (P3-PWM) | Regional Skew | 1,702 | 1,846 | 1,079 |
| | Regional Shape | 1,502 | 1,174 | 6,166 |
| | Regional Scale | 237,539 | 75,508 | 16,785 |
| | Regional Loca. | 208,963 | 46,211 | 14,498 |
| LOG-PEARSON 3 (LP3-PWM) | Regional Skew | 0,094 | 0,294 | 0,101 |
| | Regional Shape | 62221,137 | 150,646 | 25,823 |
| | Regional Scale | 0,019 | 0,088 | 0,008 |
| | Regional Loca. | 4,203 | -1,346 | 3,542 |
| LOG-LOGISTIC-PWM | Regional Shape | 0,283 | 0,308 | 0,178 |
| | Regional Scale | 450,861 | 114,987 | 90,674 |
| | Regional Loca. | 25,417 | -0,342 | -23,211 |
| WAKEBY-PWM | a | 266,661 | 52,082 | 14,532 |
| | b | 8,534 | 4,158 | 20,851 |
| | c | 569,393 | -9,725 | -198,817 |
| | d | 0,254 | 0,154 | -0,247 |
| | m | 114,512 | 37,812 | 27,461 |

For all the three regions, the frequency curves of all of the distributions used are in close agreement up to the *y* values of 3, which corresponds to the return periods up to T = 20 years for region 1 (Figures 2, 3, and 4). The observed values plotted by the Gringorten plotting position formula also show close agreement to the frequency curves of the distributions. Because an appropriate probability distribution, like the ones considered in this study, whose parameters are computed by the method of PWM is more meaningful and powerful than a plotting position formula of simplistic expression, which is based on the rank numbers of a finite-length of *n* elements of the ordered sample series, conformity or disconformity of a frequency curve to the plotted points cannot be true guidance as to the goodness of that distribution, however.

In order to guide for the better-fitting ones of the distributions considered in this study, the performances of the distributions by the goodness-of-fit tests of Kolmogorov-Smirnov and Cramer Von Misses are assessed in a quantitative way by tallying the bests, second bests,..., and fifth bests the way done in previous studies (Haktanir 1990, 1992). And, the results of these tedious ranking performances were summarized in tables, which are not given here to prevent unnecessary congestion of numbers. The overall winner by these two goodness-of-fit tests turned out to be the Wakeby distribution for all the three regions (Table 6).

**Table 5.** $Q_{max}$. values for the return periods of T= 2,5,10,25,50,100,200 for the Seyhan Basin by all the distributions whose parameters are computed by the method of PWM

| Region | T Return Period | y | Log Normal (PWM) | Wakeby (PWM) | Gumbel (PWM) | Log Logistic (PWM) | Log Pearson Type 3 (PWM) | Gumbel (PWM) |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 0,367 | 467,537 | 480,420 | 496,793 | 476,270 | 465,864 | 480,331 |
| | 5 | 1,500 | 729,377 | 672,880 | 732,496 | 689,702 | 724,621 | 715,373 |
| | 10 | 2,250 | 934,627 | 842,566 | 888,548 | 859,318 | 910,002 | 884,032 |
| | 25 | 3,199 | 1229,198 | 1123,403 | 1085,737 | 1124,093 | 1148,753 | 1110,901 |
| | 50 | 3,902 | 1473,938 | 1391,916 | 1232,041 | 1368,832 | 1326,390 | 1289,580 |
| | 100 | 4,600 | 1740,165 | 1723,497 | 1377,251 | 1663,908 | 1502,293 | 1476,399 |
| | 200 | 5,296 | 2029,963 | 2135,833 | 1521,900 | 2021,679 | 1676,895 | 1672,737 |
| 2 | 2 | 0,367 | 122,599 | 126,513 | 133,456 | 125,298 | 121,534 | 124,410 |
| | 5 | 1,500 | 204,295 | 192,885 | 208,178 | 192,568 | 204,475 | 199,906 |
| | 10 | 2,250 | 269,886 | 246,889 | 257,648 | 247,445 | 265,439 | 258,762 |
| | 25 | 3,199 | 346,577 | 334,028 | 320,155 | 334,972 | 344,927 | 343,595 |
| | 50 | 3,902 | 445,836 | 416,310 | 366,523 | 417,465 | 404,543 | 414,697 |
| | 100 | 4,600 | 533,880 | 516,835 | 412,552 | 518,521 | 463,862 | 492,877 |
| | 200 | 5,296 | 630,287 | 640,145 | 458,413 | 642,947 | 522,955 | 579,094 |

**Cont. Table 5.** $Q_{max}$. values for the return periods of T= 2,5,10,25,50,100,200 for the Seyhan Basin by all the distributions whose parameters are computed by the method of PWM

| Region | T Return Period | y | Log Normal (PWM) | Wakeby (PWM) | Gumbel (PWM) | Log Logistic (PWM) | Log Pearson Type 3 (PWM) | Gumbel (PWM) |
|---|---|---|---|---|---|---|---|---|
| 3 | 2 | 0,367 | 67,728 | 66,605 | 67,516 | 67,464 | 66,811 | 67,019 |
| | 5 | 1,500 | 91,958 | 93,243 | 92,890 | 90,486 | 93,334 | 92,666 |
| | 10 | 2,250 | 108,086 | 110,650 | 109,691 | 107,118 | 110,648 | 110,465 |
| | 25 | 3,199 | 128,636 | 131,710 | 130,919 | 131,144 | 131,905 | 133,937 |
| | 50 | 3,902 | 144,103 | 146,965 | 146,667 | 151,853 | 147,219 | 152,184 |
| | 100 | 4,600 | 159,721 | 162,156 | 162,299 | 175,429 | 162,079 | 171,136 |
| | 200 | 5,296 | 175,612 | 177,700 | 176,700 | 202,465 | 175,600 | 190,973 |

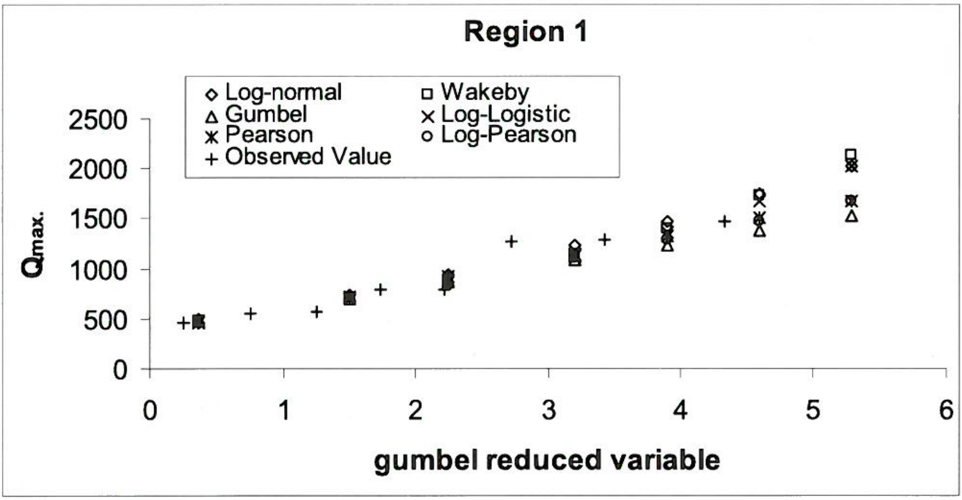


**Fig. 2**. The regional growth curves for region 1 in the Seyhan Basin

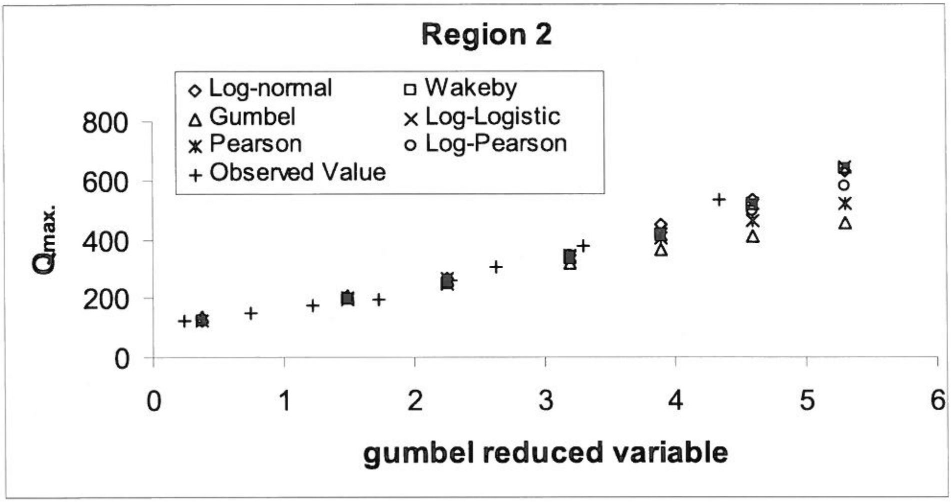


**Fig. 3.** The regional growth curves for region 2 in the Seyhan Basin

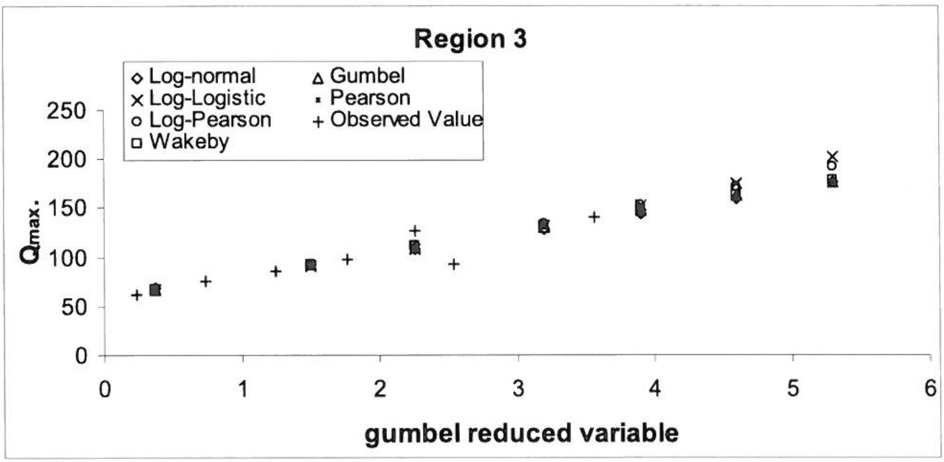


**Fig. 4.** The regional growth curves for region 3 in the Seyhan Basin

**Table 6.** Kolmogorov Smirnov and Cramer Von Mises test for the Seyhan basin

| Region | Station | Best Dist. According to the Kolmogorov-Smirnov test | Kolmogorov-Smirnov statistic D | Best Dist. According to the Cramer Von Mises test | Cramer Von Mises statistic C |
|---|---|---|---|---|---|
| 1 | Gökdere | Wakeby(PWM) | 0,055 | Wakeby(PWM) | 0,0179 |
| | Ergenuqad | Wakeby(PWM) | 0,096 | Wakeby(PWM) | 0,0565 |
| | Arapali | Log-Logistic(PWM) | 0,094 | LP3(PWM) | 0,023 |
| | Üçtepe | Gumbel(PWM) | 0,089 | Wakeby(PWM) | 0,0355 |
| | Sarmehmetli | Wakeby(PWM) | 0,087 | Wakeby(PWM) | 0,0211 |
| 2 | Himmetli | Gumbel(PWM) | 0,08 | Wakeby(PWM) | 0,0431 |
| | Kamql | P3 (PWM) | 0,093 | LP3(PWM) | 0,0359 |
| | Çukurkqla | Log-Logistic(PWM) | 0,066 | Wakeby(PWM) | 0,0325 |
| 3 | Farasa | Gumbel(PWM) | 0,104 | LP3(PWM) | 0,0314 |
| | Södütlü | Wakeby(PWM) | 0,093 | Wakeby(PWM) | 0,0243 |
| | Fraktn Köp. | LP3(PWM) | 0,088 | Log-Logistic(PWM) | 0,0263 |

## L-moments Method

Using the annual flood peaks series recorded at 11 unregulated sites in the Seyhan basin, the L-moments ratios of sample L-coefficients of L-variation coefficient (L-$C_v$), L-skewness coefficient (L-Cs) and L-kurtosis coefficient (L-Ck) are computed, which are given in Table 7. Values of the site discordance measure, Di, the heterogeneity measure, H, and the goodness-of-fit measure ($Z_{dist}$) were computed for the whole Seyhan basin using the Fortran computer program developed by Hosking (1991). Discordance Di, has been found to be 2.22, which suggests that no site is discordant because it is less than 3.0.

**Table 7.** L$C_v$, LCs, LCk and Di values at various gauging stations of the Seyhan Basin

| Site No | N | Gauging Stn. No | L-$C_v$ | L-Cs | L-Ck | Di |
|---|---|---|---|---|---|---|
| 1 | 51 | SN1805 | 0.2762 | 0.2836 | 0.2655 | 0.27 |
| 2 | 36 | SN1806 | 0.262 | 0.3426 | 0.3188 | 0.65 |
| 3 | 17 | SN1817 | 0.286 | 0.1915 | 0.1327 | 0.63 |
| 4 | 24 | SN1818 | 0.2563 | 0.2663 | 0.2683 | 0.34 |
| 5 | 16 | SN1821 | 0.266 | 0.2667 | 0.2853 | 0.88 |
| 6 | 55 | SN1801 | 0.2971 | 0.3085 | 0.2567 | 0.27 |
| 7 | 19 | SN18012 | 0.3407 | 0.3078 | 0.1199 | 2.22 |
| 8 | 34 | SN1824 | 0.3402 | 0.3042 | 0.1931 | 0.84 |
| 9 | 19 | SN1802 | 0.2122 | 0.2075 | 0.1548 | 1.16 |
| 10 | 20 | SN1804 | 0.2423 | 0.0889 | 0.0224 | 2.11 |
| 11 | 21 | SN1822 | 0.1914 | 0.2833 | 0.2895 | 1.62 |
| | | Weighted | 0.2759 | 0.2742 | 0.2266 | |

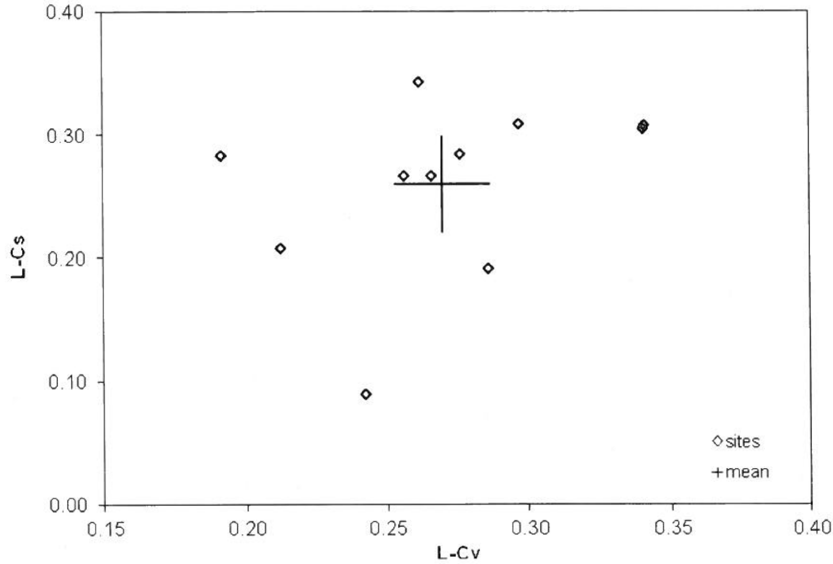


**Fig. 5**. LCv-LCs Moment ratio diagram for the Seyhan Basin.

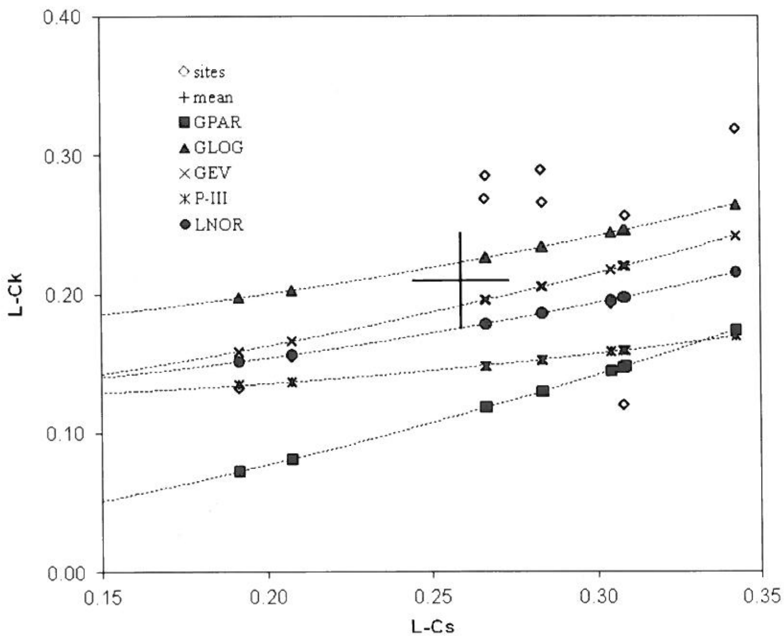


**Fig. 6**. LCs-LCk ratios diagrams for the Seyhan Basin.

The $LC_v \longleftrightarrow LCs$ moment ratio diagrams for different stations is shown in Figure 5. The $LCs \longleftrightarrow LCk$ moment ratio diagram for these 11 sample series as well as the theoretical $LCs \longleftrightarrow LCk$ relationships for several three-parameter distributions are shown in Figure 6, which shows that the Log-Logistic distribution is closer to the regional average value. Heterogeneity measures computed by carrying out 500 simulations using the data of 11 sites are given in Table 8. H(1), H(2) and H(3) are less than 1.0. So, the results of both tests reveal that the Seyhan Basin may be considered as a homogeneous region based on the L-moments approach.

**Table 8.** Heterogeneity measures for the Seyhan Basin

| HETEROGENEITY MEASURES | |
|---|---|
| OBSERVED S.D. OF GROUP L-CV | 0.0406 |
| SIM. MEAN OF S.D. OF GROUP L-CV | 0.0437 |
| SIM. S.D. OF S.D. OF GROUP L-CV | 0.0119 |
| STANDARDIZED TEST VALUE H(1) | -0.26 |
| OBSERVED AVE. OF L-CV / L-SKEW DISTANCE | 0.0591 |
| SIM. MEAN OF AVE. L-CV / L-SKEW DISTANCE | 0.1012 |
| SIM. S.D. OF AVE. L-CV / L-SKEW DISTANCE | 0.0239 |
| STANDARDIZED TEST VALUE H(2) | -1.76 |
| OBSERVED AVE. OF L-SKEW/L-KURT DISTANCE | 0.0806 |
| SIM. MEAN OF AVE. L-SKEW/L-KURT DISTANCE | 0.1278 |
| SIM. S.D. OF AVE. L-SKEW/L-KURT DISTANCE | 0.0292 |
| STANDARDIZED TEST VALUE H(3) | -1.62 |

The goodness-of-fit value ($Z_{dist}$) computed for the distributions of: Log-Logistic (GLO), general extreme values (GEV), log-normal-3 (also known as generalized normal), Pearson-3, and generalized Pareto are given in Table 9, which shows that the $Z_{dist}$ is lower than 1.64 for only two distributions, namely, GEV and GLO. Further, the $Z_{dist}$ for the GLO distribution is the smallest and much smaller than that of GEV, also. Thus, the GLO distribution is the robust distribution for the Seyhan basin by the L-moments approach advocated by Hosking and Wallis (1997).

**Table 9.** Zdist Statistics values for various distributions

| | | |
|---|---|---|
| **LOG-LOGISTIC** | L-KURTOSIS = 0.229 | **Z VALUE = -0.38** * |
| GENERAL EXTREME VALUES | L-KURTOSIS = 0.200 | Z VALUE = -1.29 * |
| GENERALIZED NORMAL | L-KURTOSIS = 0.182 | Z VALUE = -1.84 |
| PEARSON-3 | L-KURTOSIS = 0.150 | Z VALUE = -2.81 |
| GENERALIZED PARETO | L-KURTOSIS = 0.124 | Z VALUE = -3.63 |

The magnitudes of the regional parameters for the GEV and GLO distributions as well as the 5-parameter Wakeby distribution are given in Table 10. The regional parameters of the Wakeby distribution are included in Table 10 because the Wakeby distribution has five parameters, more than most of the well-known distributions, and it has a wider range of distributional shapes than the other distributions. This makes the Wakeby distribution particularly useful for simulating long synthetic data for use in studying the robustness (e.g., Cunnane, 1989; Hosking & Wallis, 1997; Houghton, 1978).

**Table 10.** Regional parameters for the frequency distributions chosen to be better fit by the L-moments approach

| | | | | | |
|---|---|---|---|---|---|
| LOG-LOGISTIC | $\xi=0.88$ | $\alpha=0.243$ | $k=-0.274$ | | |
| GENERAL EXTREME VALUES | $u=0.745$ | $\alpha=0.337$ | $k=-0.156$ | | |
| WAKEBY | $\xi=0.279$ | $\alpha=1.737$ | $\beta=4.751$ | $\gamma=0.342$ | $\delta=0.185$ |

Using the GEV, GLO, and Wakeby distributions, standardized quantiles have been computed at selected return periods of 1.01, 1.58, 2, 5, 10, 25, 50, 100, and 200 years and plotted against the respective return periods (Table 11). The regional curve thus developed is shown in Figure 8.

**Table 11.** Magnitudes of the growth factors ($Q_T/Q_{ave}$) for various return periods by the three distributions

| RETURN PERIODS (year) | 1.0101 | 1.58 | 2 | 5 | 10 | 20 | 50 | 100 | 200 |
|---|---|---|---|---|---|---|---|---|---|
| LOG-LOGISTIC | 0.245 | 0.757 | 0.880 | 1.290 | 1.613 | 2.112 | 2.570 | 3.118 | 3.777 |
| GENERAL EXTREME VALUES | 0.287 | 0.744 | 0.872 | 1.314 | 1.653 | 2.142 | 2.555 | 3.012 | 3.519 |
| WAKEBY | 0.299 | 0.766 | 0.883 | 1.285 | 1.626 | 2.148 | 2.607 | 3.128 | 3.720 |

The accuracy of the simulated quantiles is estimated by the monte carlo simulation as suggested by Hosking (1991). The root mean square error (RMSE) of quantile estimates from the simulated regional quantiles, and the 90% error bounds for the growth curve are determined via the computer program prepared by Hosking. These accuracy measures computed by the GLO distribution for the selected values of non-exceedance probability (F) are given in Table 12. The estimated growth curve together with its 90% error bounds are shown in Figure 7.

**Table 12.** Accuracy measures for the growth curve computed by the GLO distribution

| F (1-1/T) | qhat(F) | RMSE | Error bounds | |
|---|---|---|---|---|
| | | | 0.950 PT | 0.050 PT |
| 0.01 | 0.245 | 0.703 | 0.153 | 0.304 |
| 0.367 | 0.757 | 0.057 | 0.716 | 0.795 |
| 0.5 | 0.880 | 0.033 | 0.839 | 0.921 |
| 0.8 | 1.290 | 0.035 | 1.232 | 1.331 |
| 0.9 | 1.613 | 0.059 | 1.539 | 1.680 |
| 0.96 | 2.112 | 0.088 | 1.942 | 2.257 |
| 0.98 | 2.570 | 0.109 | 2.316 | 2.832 |
| 0.99 | 3.118 | 0.130 | 2.737 | 3.567 |
| 0.995 | 3.778 | 0.151 | 3.223 | 4.488 |

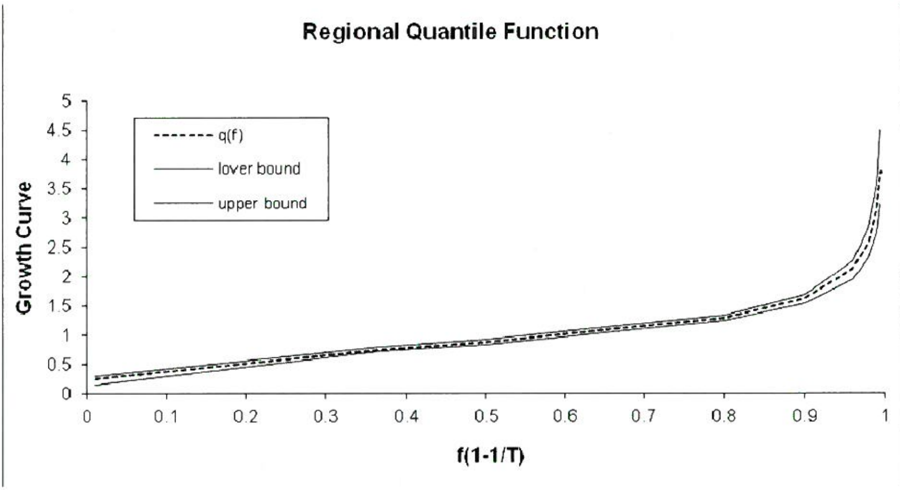


**Fig. 7.** Regional Quantile Function (GLO Dist. With 90% Error Bounds)

## Relationship between mean annual Peak Flood and Catchment Area

For ungauged cathments, at-site mean cannot be computed because of the absence of the observed flow data. In this study, a regional relationship has been developed for q (mean unit discharge) as a function of catchment area. For ungauged cathments, mean annual peak flood ($\overline{Q}$) can be computed by Eqn. (10) below, which is obtained by regression using the observed data of the 11 stations.

$$q = \frac{\overline{Q}}{A} = 0.0233e^{\left(\frac{1790.53}{A}\right)} \quad (10)$$

A graphical representation of the relationship between q (mean unit discharge) and catchment area is presented in Figure 9. The result of the Z test analysis indicates that the GLO distribution is the best fit distribution for the Seyhan basin. The regional parameters of the GLO distribution are $\xi = 0.88$, $\alpha = 0.243$, k = -0.274. Accordingly, magnitude of a flood peak for a given return period can be computed by Eqn. (11) below, which is the quantile function of the GLO distribution with these regional parameters.

$$Q = \left[0.88 - 0.887\left(1 - \left(\frac{1}{1-T}\right)^{-0.274}\right)\right] \times \overline{Q} \quad (11)$$

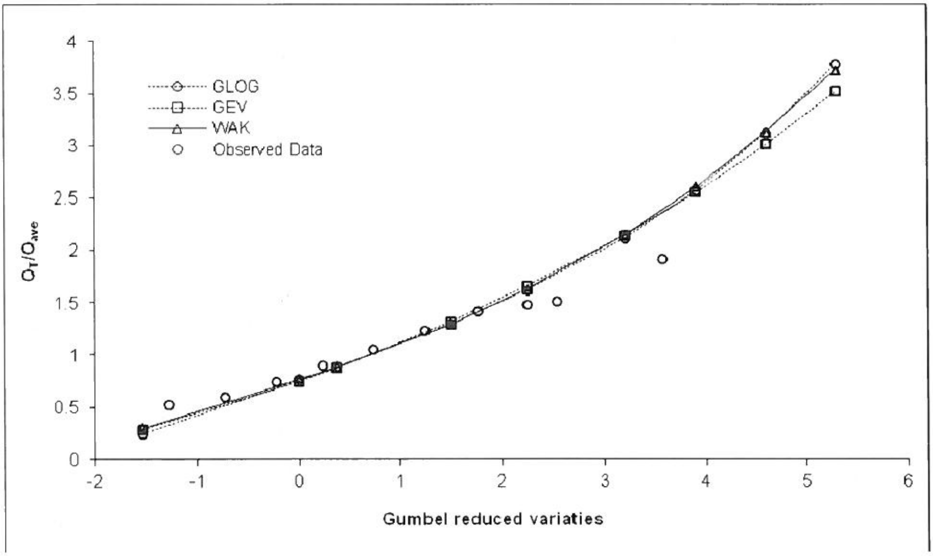


**Fig. 8**. Regionalized growth curves by the three distributions for the Seyhan Basin

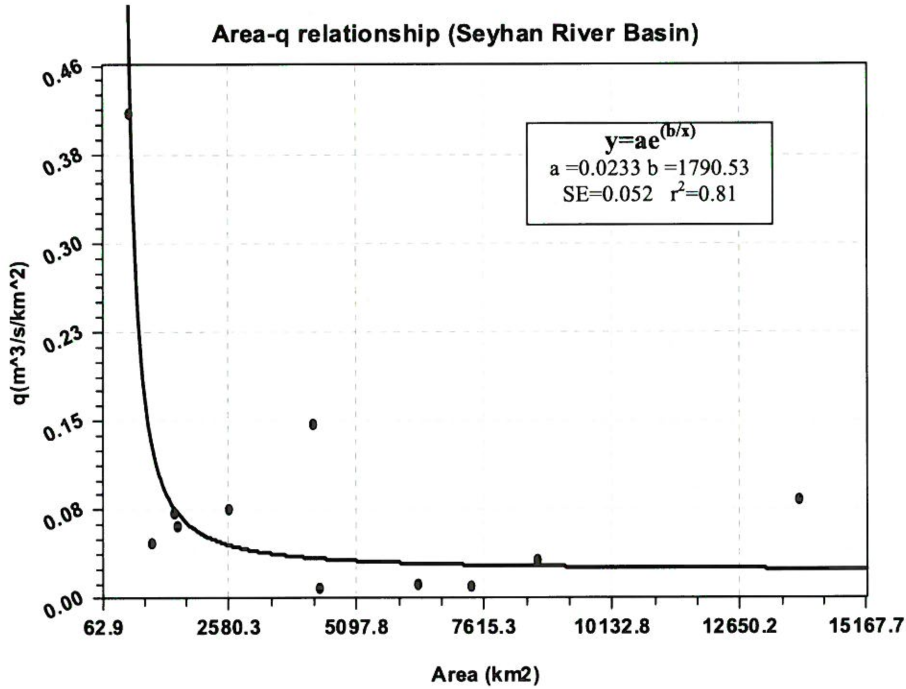


**Fig. 9**. Relationship between q (mean unit discharge) and catchment area with observed data of 11 gauged stations.

## CONCLUSIONS

In this study, the conventional method applied, which are calling the $C_v$-based method, and the L-moments method, which is gaining popularity over the last ten years especially, with the purpose of a regionalized flood frequency analysis in one of the major basins of Turkey. This is an important field of applied hydrology, and there is a serious need for regionalized flood frequency analysis, especially in developing countries where the number of stream-gauging stations is too few and the recorded data is direly lacking. Yet, water resources projects of all kinds, including dams, are being implemented in the developing countries because there is an urgent need for these projects, mainly by financial credits provided by developed countries. This study will stimulate similar studies especially in developing countries. Even in developed countries, there are many dam sites at which there are no recorded stream flow data In this study, firstly, the Seyhan Basin is divided into three homogeneous regions (sub-areas) determined by applying the coefficient of variation ($C_v$)-based test. The flood frequency analyses of annual flood peaks series recorded at 11 stream-gauging sites in the Seyhan Basin are done by seven different probability distributions, which, according to the pertinent literature, are known to be commonly used for flood frequency analyses. The better-fit ones of these distributions are determined by evaluations of goodness-of-fit tests of Kolmogorov-Smirnov and Cramer von Mises. In general, the Wakeby distribution, whose parameters are computed by the method of PWM, is the best-fit distribution in all the three regions of the Seyhan basin. In region 3, however, the Pearson type-3 distribution also performs closely to the Wakeby distribution. (Table 6)

The Wakeby distribution, whose parameters are computed by the method of PWM, is the best-fit distribution for four station according to kolmogorov simirnov test and for seven stations according to cramer von mises test.

Secondly, the hydrological homogeneity of the Seyhan Basin is also identified by the L-moments approach as suggested by Hosking and Wallis (1997). According to the L-moments method of analysis, the sample series recorded at 11 unregulated stream-gauging stations are all found to be concordant (not discordant), and the Seyhan Basin is found to be homogeneous as a whole. The magnitudes of the regional parameters of various distributions are computed by the L-moments approach. Based on the Z test, the Log-Logistic distribution is identified as the robust distribution for the Seyhan Basin.

Eqn. (10), which is obtained by regression using the data of the gauged sites, can be used first to predict the average flood peak. Next, the preferred one of the growth curves given in Figures 2, 3, 4, for sub-regions 1, 2, 3, respectively, will be used to estimate the flood peak versus a given return period if the conventional approach is to be used. Preferably the Wakeby distribution should be chosen. On the other hand, Eqn.(11) can be used directly for any location in the Seyhan basin to estimate the flood peak versus any given return period if the L-moments approach is preferred.

## REFERENCES

**Bobee, B. & Rassmussen, P.F. 1995.** Recent advances in flood frequency analysis. Reviews of Geophysics, Supplement, US national Report to International Union of Geodesy and Gephysics, 1991-1994, 1111-1116.

**Cunnane, C. 1988.** Methods and merits of regional flood frequency analysis. Journal of Hydrology, **100:** 269-290.

**Cunnane, C. 1989.** Statistical Distributions for Flood Frequency Analysis. World Meteorological Organization Operational Hydrology, Report No. 33, WMO, No.718, Geneva, Switzerland.

**Dalrymple, T., 1960.** 'Flood Frequency Methods' Water Supply Paper 1543, US Geological Survey, pp11-51.

**Greenwood, J.A., Landwehr, J.M., Matalas, N.C., & Wallis, J.R, 1979.** Probability weighted moments: definition and relation to parameters of several distribution exprensible in Inverse form, Water Resources Research, **15(5):** 1049-1054.

**Haktanir, T. 1990.** A few distribution compiled together for flood frequency analysis. DOGA, Turkish Journal of Engineering and Environmental Sciences. **14**(1), 146-165.

**Haktanir, T. 1992.** Comparison of various flood frequency distributions using annual flood peaks data of rivers in Anatolia. Journal of Hydrology. **136** (1-4), 1-31.

**Hosking, J.R.M., 1990.** L-Moments:Analysis and Estimation of Distributions Using Lineer Combinations of Order Statistics. J. Royal Statistical Society, 52(2): 105-124.

**Hosking, J.R.M., 1991**. FORTRAN Routines for use with the method of L-moments. Research Report RC-17097, IBM Research Division, T. J. Watson Research Center, Yorktown Heights, New York.

**Hosking, J.R.M., Wallis J.R., & Wood, E.F., 1984.** Estimation of the Generalized Extrem Value Distribution by the Method of Probability Weighted Moments. Institute of Hydrology. Report No 89. 25 pp.

**Hosking, J.R.M., Walls, J.R., & Wood, E.F., 1985b.** Estimation of The Generalized Extreme Value Distribution By The Method of Probability Weighted Moments. Technometrics, 27(3): 251-261.

**Hosking, J.R.M., 1986.** The Theory of Probability Weighted Moments. Research Rep. RC 12210, 160 pp., IBM Research Division, Yorktown Heights, NY

**Hosking, J.R.M., 1991.** Approximations for use in Constructing L-moments Ratio Diagrams. Res Report, RC-16635, vol 3, IBM Res Division, New York.

**Hosking, J.R.M., & Wallis J.R., 1997.** Regional Frequency Analysis-an Approach based on L-moments. Cambridge University Pres, New York.

**Hosking, J.R.M., & Wallis J.R, 1993.** Some Statistics useful in Regional Frequency Analysis. Water Resources Research, vol 29, no 2, pp 271-281.

**Houghton, J.C. 1978.** Birth of a parent: the Wakeby distribution for modeling flood flows. Water Resources research, 14(6), 1105-1110.

**Jaiswal, R.K., Goel, N.K., Singh, P. & Thomas, T., 2003.** L-moment based Flood Frequency Modelling, The Institution of Engineers (India), vol 84, pp 6-10.

**Kumar, R.,Chatterjee, C., Panigrihy N., Patwary, B.C., & Singh., R.D., 2003.** Development of Regional Flood Formulae using L-moments for Gauged and Ungauged Catchments of North Brahmaputra River System. The Institution of Engineers (India), **84**: 57-63.

**Mosley, M.P. 1981.** Delimination of New Zeland Hydrological Regions. J. Hydrology, **49**: 173-192.

**National Research Council 1988.** Estimating Probabilities of Extreme Floods, Methods and Recommended Research. Committe on Techniques for Estimating Probabilities of Extreme Floods, National Academy Press, Washington, D.C.

**NERC 1975.** Flood studies report. National Environment Research Council, London, **1**(5).

**Parida, B.P., Kachroo, R.K., & Shrestha, D.B., 1998.** Regional Flood Frequency Analysis of Mahi-Sabarmati Basin (Subzone 3-a) using Index Flood Procedure with L-moments. Water Resources Management. **12**:1-12.

**Parr, W.C., 1983**. A Note on The Jackknife, The Bootstrap and The Delta Method Estimators of Bias and Variance. Biometrica, **70(3):** 719-722.

**Rao R., & Hamed, K.H., 1997.** Regional Frequency Analysis of Wabash River Flood Data By L-moments. Journal of Hydrologic Engineering, pp 169-179.

**Sankarasubramanian, A & Srinivasan, K 1999.** Investigation and comparison of sampling properties of L-moments and conventional moments. Journal of Hydrology, **218**, 13-34.

**Smithers, J.C. & Schulze, R.E. 2001.** A methodology for the estimation of short duration design storms in Africa using a regional approach based on L-moments. Journal of Hydrology, **241**, 42-52.

**Stedinger, J.R.; Vogel, R.M.; Foufoula-Georgiou, M. 1993.** Frequency Analysis of Extreme Events. in Handbook of Hydrology ed. D.R. Maidment, Ch.18, McGraw-Hill.

**Wiltshire, S.W. 1986.** Identification of homogeneous regions for flood frequency analysis. Journal of Hydrology. **84**, 287-302.

***Submitted :*** *08/07/2013*
***Revised :*** *28/10/2013*
***Accepted :*** *17/09/2013*

## Notation

MOM : method of moments

PWM : probabilty weighted moments

$m_k$ : the number of sites in group k

F : variance ratio test (F statistics)

$n_{jk}$ : the length of record at $j^{th}$ site in group k,

$cv_{jk}$ : sample coefficient of variation of site j in group k

$u_{jk}$ : the sampling variance of $cv_{jk}$

A : the within groups variation

B : the between groups variation

$cv_{.k}$ : arithmetic average of the cv's in group k

cv.. : arithmetic average of the cv's across all groups.

Cv : the coefficient of variation

$S_x$ : the sample standard deviation of the series

$\overline{x}$ : number of groups in basin

$S_k$ : total variation within groups

$n_j$ : the number of random variable at $j^{th}$ site

$F_{ij}$ : non-exceedance probability of $i^{th}$ random variable at site j

yij : gumbel-reduced variate.