GOÜ. Ziraat Fakültesi Dergisi, 2005, 22 (2), 51-56

Regional Analysis of Monthly Rainfalls over Amasya Province via L-Moments Method

Kadri Yürekli
Gaziosmanpasa University, Faculty of Agriculture, Department of Farm Structure and Irrigation, 60240, Tokat

Abstract: The study aims to perform regional frequency analysis of monthly rainfalls measured over Amasya province by using L-moment approach. Initially, Amasya province was formed two groups (as first and second) of rain gauge stations (sites) in the province to satisfy the homogeneity condition by using discordancy measure. Thereafter, heterogeneity (H) test was applied to assess whether the regions proposed as homogeneous according to discordancy measure of site characteristics are reasonably treated as a homogeneous region. This test confirmed the delineated regions as homogeneous. Choosing the best fit frequency distribution for the data from the sites of the selected regions was based on the Z-statistic. According to this statistic, the best fit distributions were estimated, generalized extreme value type I (GEV) for the first region, and Pearson type three (P3) or generalized extreme value type I for the second region.
Key Words: Monthly rainfall, L-moment, homogeneous region, heterogeneity measure, Z-statistic

L-Moment Yöntemi ile Amasya İlindeki Aylık Yağmurların Bölgesel Analizi

Özet: Bu çalışma, Amasya ilinde ölçülen aylık yağmurların bölgesel frekans analizini yapmayı amaçlamaktadır. Öncelikle, Amasya ili, homojenlik (discordancy) ölçüsü kullanılarak, homojenlik koşulunu yerine getirmek amacıyla ildeki yağmur istasyonları iki gruba (birinci ve ikinci olarak) ayrılmıştır. Daha sonra, homojenlik ölçüsüne göre homojen olarak önerilen bölgelerin, gerçekten homojen olup olmadığını değerlendirmek için heterojenlik testi uygulanmıştır. Bu test seçilen bölgelerin homojen olduğunu göstermiştir. Seçilen bölgelerin istasyonlarından elde edilen veriler için en uygun frekans dağılımının seçimi Z-istatistiğine göre belirlenmiştir. Bu istatistiğe göre, birinci bölge için genelleştirilmiş ekstrem tip I (GEV) , ikinci bölge için ise Person tip 3 (P3) yada genelleştirilmiş ekstrem tip I frekans dağılımları en uygun dağılımlar olarak belirlenmiştir.
Anahtar Kelimeler: Aylık yağmur, L-moment, homojen bölge, heterojenlik ölçüsü, Z-istatistiği

1. Introduction

Having information about distributions of precipitation depths is very important for the design of water-related structure, which protects agricultural land and downstream cities from flood and drought and supply agricultural water demand. But, a reliable design quantile estimate is commonly impossible. The selected quantile of under-or over design criterion concerning with hydraulic structures is exposed to risk as the return period is determined according to cost and economic-strategic significance of the structure. Selecting a reliable design quantile, which affect on design, operation, management and maintain of a hydraulic structure, considerably depends on statistical methods used in parameter estimation belonging to probability distribution (Hosking and Wallis, 1993). Therefore, defining a true distribution concerning with hydrological and meteorological events keeps on being major problem for researchers. Additionally, both the identification of appropriate statistical distribution for describing the observations and the estimation of the parameters of a selected distribution are complicated as many hydrologic and meteorological time series are too short for a reliable design quantile estimation (Hosking, 1990).

In the recent, researchers interested in hydrology and meteorology fields have focused on L-moment approach introduced Hosking (1990) and increasingly used in regional frequency analysis. The advantages of this method over conventional moments are that they are relatively insensitive to outliers and do not have sample size related bounds. Moreover, the parameter estimations are more reliable than the conventional method of moment estimates, particularly from small samples, and are usually computationally more tractable than maximum likelihood estimates. On the other hand, estimators of L-moments are virtually unbiased (Hosking, 1990; Park et al., 2001).

The overall objective of this study is to establish a monthly rainfall magnitude with any return period of occurrence. In order to achieve this by using monthly rainfalls over Amasya Province, delineating homogeneous regions based on discordancy measure of site characteristics and identification of suitable regional frequency distribution were included in the study.

2. Material and Method

Amasya region, selected as the study region, is bounded by latitudes 40º N and 41º 15' N, and longitudes 35º E and 36º 15' E, covering 551993 ha. Cropland, grassland and forest occupy about 57.2%, 12.6% and 40%, of the region, respectively. Due to abundance of dry farming in the region, Wheat is the major food crop. The major sources of irrigation are rainfall, canals and groundwater (Anonymous, 1991). Monthly rainfall amounts over Amasya province were used as a material in the study. There are ten rainfall gauge stations, which belong to Turkish State Meteorological Service (Figure 1). The activity of some stations on rainfall measurement has been stopped. The record lengths and elevations of the stations over the study area vary, from 67 to 16 years, and from 200 to 800m, respectively.

2.1 The Method of L-Moments

L-moments, as defined by Hosking (1990), are linear combinations of probability weighted moments (PWM). Greenwood et al. (1979) summarizes the theory of PWM and defined as

$$\beta_r = E\{X[F_X(x)]^r\} \qquad (1)$$

Where β_r is the r^{th} order PWM and $F_X(x)$ is the cumulative distribution function (cdf) of X. Hosking and Wallis (1997) defined unbiased sample estimators of PWMs as (b_i) and, obtained unbiased sample estimators of the first four L-moments by PWM sample estimators. Unbiased sample estimates of the PWM for any distribution can be computed from;

$$b_r = n^{-1}\sum_{j=1}^{n-r}\frac{\binom{n-j}{r}}{\binom{n-1}{r}}x_j \qquad (2)$$

Where x_j is an ordered set of observations $x_1 \le x_2 \le x_3 \le \ldots x_n$. For any distribution the first four L-moments are easily computed from PWM using;

$\lambda_1 = b_1$,
$\lambda_2 = 2b_2 - b_1$,
$\lambda_3 = 6b_3 - 6b_2 + b_1$,
$\lambda_4 = 20b_4 - 30b_3 + 12b_2 - b_1$ (3)

Sankarasubramanian and Srinivasan (1999) define the L-moment ratios (L-coefficient of variation, L-skewness and L-kurtosis, respectively)

$\tau_2 = \lambda_2/\lambda_1$,
$\tau_3 = \lambda_3/\lambda_2$,
$\tau_4 = \lambda_4/\lambda_2$ (4)

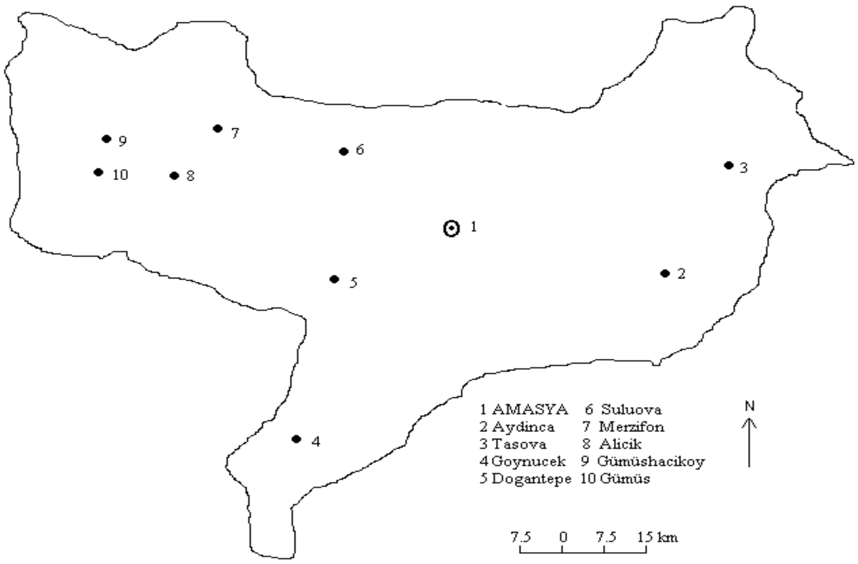

Figure 1. Rainfall Gauge Stations over Amasya Province

Rainfall amounts vary spatially within the region covered by a given storm. Therefore, the study region should be partitioned into hydrologically homogeneous regions in which rainfall amounts recorded at the rainfall gauge stations are assumed to be identical to obtain reliable results in hydrologic studies related to rainfall (Okman, 1994).

2.2 Screening of Data

The aim of this stage is to form groups of stations that satisfy the homogeneity condition, that stations with frequency distributions that are identical apart from a station–specific scale factors. This is usually carried out by dividing the sites into disjoint groups. Hosking and Wallis (1997) present a discordancy measure. In this approach, the L-moments ratio (L-coefficient of variation, L-skewness and L-kurtosis) of a site is used to describe that site as a point in three-dimensional space. A group of homogeneous sites will form a cluster of such points. If any point does not appear to belong to the cluster of such points on the L-moment diagram, that is, is far from the center of the cluster, the site related to that point should be removed from the region due to non-homogeneity condition. Discordancy measure (D_i) of a site can be calculated by

$$\bar{u} = N^{-1}\sum_{i=1}^{N} u_i \qquad (5)$$

$$S = (N-1)^{-1}\sum_{i=1}^{N}(u_i - \bar{u})(u_i - \bar{u})^T \qquad (6)$$

$$D_i = \frac{1}{3}(u_i - \bar{u})^T S^{-1}(u_i - \bar{u}) \qquad (7)$$

Let $u_i = \left[\tau_2^i, \tau_3^i, \tau_4^i\right]^T$ be a vector related to L-moment ratios of site i. Where N is the number of sites. Generally, any site with $D_i > 3$ is considered as discordant. In such a case, the site may properly belong to another region.

2.3 Heterogeneity Test for Regions

Heterogeneity (H) test by Hosking and Wallis (1993), which compares the inter-site variation (dispersion) in sample L-moments for the group of sites, is used to assess whether the regions proposed as homogeneous according to discordancy measure of site characteristics are reasonably treated as a homogeneous region. For this reason, the method fit the four-parameter Kappa distribution to the regional average L-moment ratios to generate 500 homogeneous regions with population parameters equal to the regional average sample L-moment ratios. The properties of the actual region are compared to the simulated homogeneous region. The heterogeneity (H) statistic and V statistic for the sample and simulated regions take the form, respectively:

$$H = (V_{obs} - \mu_V)/\sigma_V \qquad (8)$$

$$V = \left\{ \frac{\sum_{i=1}^{N} n_i (\tau_2^i - \tau_2^R)^2}{\sum_{i=1}^{N} n_i} \right\}^{1/2} \qquad (9)$$

n_i is record length at site i, τ_2^i is the sample L-coefficient of variation (L-Cv), τ_2^R is the regional average sample L-Cv, μ_V is the mean of simulated V values, σ_V is the standard deviation of simulated V values.

The value of H-statistic indicate that the region under consideration is acceptably homogeneous when H<1, possibly heterogeneous when 1≤ H <2, and definitely heterogeneous when H ≥2 (Hosking and Wallis, 1997).

2.4 Choosing the Best Fit Frequency Distribution

In regional frequency analysis, a single frequency distribution is fit to the data from several sites in a homogeneous region. Hosking and Wallis (1997) proposed an appropriate method for goodness of fit criterion based on L-kurtosis. This statistic is termed as the Z-statistic:

$$Z^{DIST} = (\tau_4^{DIST} - \bar{\tau}_4 + \beta_4)/\sigma_4 \qquad (10)$$

$$\beta_4 = N_{sim}^{-1}\sum_{m=1}^{N_{sim}}(\bar{\tau}_{4m} - \bar{\tau}_4) \qquad (11)$$

$$\sigma_4 = \left\{ (N_{sim} - 1)^{-1}\sum_{m=1}^{N_{sim}}(\bar{\tau}_{4m} - \bar{\tau}_4)^2 - N_{sim}\beta_4^2 \right\}^{1/2}$$
(12)

Where DIST refers to a candidate statistical distribution, τ_4^{DIST} is the population

L-kurtosis of selected distribution, $\bar{\tau}_4$ is the regional average sample L-kurtosis, β_4 is the bias of regional average sample L-kurtosis, σ_4 is the standard deviation of regional average sample L-kurtosis, and N_{sim} is realizations of a region with N sites. The four parameter Kappa distribution is used to simulate 500 regions similar to the actual region to estimate β_4 and σ_4. The $|Z^{DIST}|$ ≤ 1.64 should be for an appropriate regional distribution. But, the distribution giving the minimum $|Z^{DIST}|$ is considered as the best-fit distribution for the region.

2.5 Regional Estimates Based on Growth Curves

Data from different sites should be combined to appraise the parameters of the related distribution. Although alternative approaches exit for this reason, the index-flood method supported by Hosking and Wallis (1997) was used in the study. The method may be written as following

$$Q_i(F) = \mu_i q(F) \tag{13}$$

μ_i, is the at-site mean, q(F) is regional growth curve.

The regional frequency analysis of monthly rainfall depths over Amasya province was achieved by using the FORTRAN routines developed by Hosking (1996).

3. Results and Discussion

In order to achieve regional frequency analysis of monthly rainfall from different rainfall gauge stations over Amasya province, some basic L-moment statistics, which are L-mean (λ_1), L-coefficient of variation (τ_2), L-skevness (τ_3) and L-kurtosis (τ_4), belonging to that stations were estimated (Table 1). Hosking (1990) imply that L-moment ratios of a series are bounded, L-coefficient of variation (L-CV), L-skewness and L-kurtosis satisfy 0 < τ_2 <1, -1< τ_3 < 1, and $\frac{1}{4}(5\tau_3^2 - 1) \le \tau_4 < 1$, respectively. As it can be seen in Table 1, these conditions have been fulfilled.

The results of discordancy measure (D_i) recommended to form groups of homogeneous stations were given in Table 1. The values of that measure were estimated between 0.27 and 2.06 for the stations in the first region. The value was 1.0 for the stations in the second region. The selected regions may be accepted as homogeneous, owing to the D_i values for the stations in the first and second regions < 3.

Table 1. Discordancy Analysis Results of Rainfall Gauge Stations over Amasya Province

Region	Rainfall Gauge Station	λ_1	τ_2	τ_3	τ_4	D_i
I	AMASYA,	35.640	0.1017	0.1068	0.1139	0.33
	Merzifon,	32.890	0.1096	0.0947	0.1278	0.27
	Gümüshaciköy,	37.560	0.1217	0.0297	0.1461	1.64
	Gümüş,	41.770	0.0811	-0.1823	0.2173	2.06
	Suluova,	31.360	0.1193	0.0840	-0.0202	2.06
	Doğantepe,	33.670	0.1120	0.1473	0.0939	0.32
	Göynücek	36.160	0.0920	0.0872	0.1090	0.94
	Aydınca,	92.310	0.1180	0.1682	0.0392	0.38
II	Alıcık,	68.480	0.1351	0.2527	0.1394	1.00
	Taşova	32.660	0.1699	-0.0294	0.0539	1.00

The results related to assessment of dispersion of the at-site L-moment ratios for the entire study area and for two delineated regions based on discordancy measure were in Table 2. These values of heterogeneity indicate that the entire study area is definitely heterogeneous since H-statistic is greater than 2. But, the estimates belonging to the H-statistic for two delineated regions indicate that the regions are acceptably homogeneous due to H-statistic < 1. The parameter estimations of regional kappa distribution from which homogeneous regions with sites having records lengths the same as those of the observed data are generated and group average L-moments required for calculation of H-statistic are also presented in Table 2.

The identification of an appropriate regional distribution for each of the two regions was based on the Z-statistic. The Z-statistics concerning with some statistical distributions for the delineated regions were shown in Table 3. According to the analysis of results of goodness of fit test (Z-statistic), generalized extreme value type I (GEV) for the first region, and Pearson type three (P3) or generalized extreme value type I for the second region should be considered as the best-fit distributions, respectively as these distributions give the minimum Z-statistic. But, except these distributions, the values of Z-statistic related to GNO (generalized normal) and P3 distributions for the first region, and GNO and GPA distributions for the second region were less than the critical Z-statistic value (1.64). Therefore, these distributions may be used in regional frequency analysis for the regions.

The regional growth curve estimations related to some nonexceedence probability levels based on the distributions selected for the two regions as the best-fit distributions were given in Table 4. As can be seen the table, the regional growth curve estimations based on P3 and GEV statistical distributions belonging to second region were almost identical. This advocates the Z-statistics related to statistical distributions (P3 and GEV) selected for second region. Table 5 shows the quantiles of some probability levels from the selected distributions for the regions.

Table 2. The Results Related to Heterogeneity of The Selected Regions

Region	Group Average L-Moments			Parameters of Regional Kappa Distribution				H-Statistic
	τ_2^R	τ_3^R	τ_4^R	ξ	α	k	h	
All Sites	0.1148	0.0852	0.1029	0.8894	0.2190	0.2221	0.2129	2.59*
I	0.1085	0.0835	0.1046	0.8986	0.2036	0.2155	0.1903	-0.05
II	0.1544	0.0960	0.0919	0.8209	0.3285	0.2660	0.3558	0.62

ξ, location parameter
α, scale parameter
k and h, shape parameters

Table 3. The Simulation Results for The Z-statistic

Region	Statistical Distribution	Z-value
I	GLO	3.87
	GEV	1.04**
	GNO	1.38
	P3	1.17
	GPA	-4.50
II	GLO	1.70
	GEV	0.73
	GNO	0.82
	P3	0.72**
	GPA	-1.22

Table 4. Estimations of Growth Curve for The Regions According to Some Probability Level

Region	Probability Level					
	0.01	0.02	0.05	0.1	0.2	0.5
I-GEV	0.622	0.657	0.712	0.765	0.834	0.983
II-GEV	0.473	0.520	0.596	0.669	0.764	0.972
II-P3	0.477	0.522	0.595	0.667	0.763	0.973

Table5.Quantile Estimations from The Selected Distribution for The Regions

Region	Probability Level					
	0.01	0.02	0.05	0.1	0.2	0.5
I-GEV	25.14	26.55	28.77	30.91	33.70	39.72
II-GEV	22.96	25.25	28.94	32.48	37.09	47.19
II-P3	23.16	25.34	28.89	32.38	37.04	47.24

4. Conclusion

Many studies reported in the literature indicate that there are the advantages of using a regional frequency analysis. Therefore, the regional L-moment algorithm was applied to monthly rainfall data sequences over Amasya province in the study. With the reason, Amasya province was divided in two sub-regions as first and second regions according to L-moment ratios belonging to rainfall gauge stations over the province. The generalized extreme value type I (GEV), and Pearson type three (P3) and GEV statistical distributions for the first and second sub-regions were selected as best fit regional distributions, respectively.

References

Anonymous, 1991. Amasya İli Arazi Varlığı. Köy Hizmetleri Genel Müdürlüğü Yayinları, 101p, Ankara.

Greenwood, J.A., J.M. Landwehr, N.C. Matalas and J.R. Wallis, 1979. Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form. Water Resources Research, 15, 1049-1054.

Hosking, J.R.M., 1990. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal of The Royal Statistical Society Series B, 52, 105-124.

Hosking, J.R.M. and Wallis, J.R. 1993. Some Statistics Useful in Regional Frequency Analysis. Water Resources Research, 29, 271-281.

Hosking, J.R.M., 1996. Fortran Routines for Use with the method of L-Moments, Version 3, Research Report RC 20525, 33p, New York-USA.

Hosking, J.R.M. and Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, 224 p., USA.

Okman, C., 1994. Hydrology. University of Ankara Faculty of Agriculture, Publication No, 1388, 359p, Ankara. (in Turkish).

Park, J.S., Jung, H.S., Kim, R.S. and Oh, J.H., 2001. Modelling Summer Extreme Rainfall over the Korean Penissula Using Wakeby Distribution. International Journal of Climatology, 21, 1371-1384.

Sankarasubramanian, A. and Sirinivasan, K., 1999. Investigation and Comparison of Sampling Properties of L-Moments and Conventional Moments. Journal of Hydrology, 218, 13-34.