Turk J Engin Environ Sci
24 (2000) , 419 – 428.

Use of Gamma Distribution in Hydrological Analysis

Hafzullah AKSOY
Istanbul Technical University,
Civil Engineering Faculty, Hydraulics Division,
80626 Ayazaga, Istanbul-TURKEY
e-mail: haksoy@itu.edu.tr

Received 17.09.1999

Abstract

In this study, amounts of daily rainfall and the ascension curve of the hydrograph are investigated. In both cases, the 2-parameter gamma distribution is used. The distribution is fitted to the amounts of daily rainfall and to the differences between the flows of successive days on the ascension curve of the hydrograph. The shape and scale parameters of the distribution, in both cases, are estimated in a monthly time interval. A 30-year daily rainfall series and a 35-year daily runoff series are used for the application. It may be seen that the distribution fits very well to the rainfall data and also that the ascension curve of the hydrograph can easily be represented by the distribution.

Key Words: Gamma distribution, daily rainfall, daily runoff, hydrograph, ascension curve

Gama Dağılımının Hidrolojik Analizlerde Kullanımı

Özet

Bu çalışmada günlük yağış verileri ve hidrografın yükselme eğrisi incelenmektedir. Hem yağış verileri hem de yükselme eğrisi için 2-parametreli gama dağılımı kullanılmaktadır. Dağılım günlük yağış verilerine ve yükselme eğrisi üzerindeki ardışık günlerin akımları arasındaki farklara uydurulmaktadır. Dağılımın şekil ve ölçek parametreleri aylık zaman aralığında belirlenmektedir. Uygulamada 30 yıllık bir yağış serisi ile 35 yıllık bir akım serisi kullanılmış, sonuçta dağılımın hem günlük yağış verilerini hem de hidrografın yükselme eğrisini iyi temsil ettiği görülmüştür.

Anahtar Sözcükler: Gama dağılımı, günlük yağış, günlük akım, hidrograf, yükselme eğrisi

Introduction

It is very hard for hydrologists to measure, collect and store hydrological data such as rainfall and runoff. A great deal of hydrological data is required in the design of water-related structures. In most cases, the available amount of data is limited and it may also contain some gaps in the series. When it is decided to carry out a water resources project in a hydrological region, it is first necessary to collect all the information related to the region and then to analyze the collected data. A frequency analysis of the data is the most commonly applied method. The planning engineer then searches for a mathematical equation characterizing the available data in hand to fill the gaps in the observations and/or to extend it to a longer period. The gaps in the data can be filled using correlated data obtained from hydrological areas that are close by geographically and hence similar to the area under consideration. It is possible to fill gaps in a data series with historical mean values. The observed data can also be extended by a mathematical equation (model).

Although hydrological variables are of the continuous type, they are discretized and used as a discrete series. It is generally assumed that a hydrological variable has a certain distribution type. Some of the most common and important probability distributions used in hydrology are the normal, log-normal, gamma, Gumbel, and Weibull. The normal distribution generally fits to the annual flows of rivers. The log-normal distribution is also used for the same purpose. In hydrology, the gamma distribution has the advantage of having only positive values, since hydrological variables such as rainfall and runoff are always positive (greater than zero) or equal to zero as a lower limit value (Markovic, 1965). The Gumbel and Weibull distributions are used for extreme values of hydrological variables. The Gumbel distribution is used in the frequency analysis of floods (Gumbel, 1954) and the Weibull distribution in the analysis of low flow values observed in rivers (Bulu and Aksoy, 1998).

This study gives results obtained from two applications of the gamma distribution in hydrology. The amounts of daily rainfall are represented by the gamma distribution in the first application, while the ascension curve of the hydrograph is simulated by the same distribution in the second. In the following parts of the study, the gamma-distribution is introduced together with parameter estimation techniques and gamma-distributed random number generation mechanisms.

Gamma Distributions

The gamma distribution function has three different types, 1-, 2- and 3-parameter gamma distributions. If the continuous random variable x fits to the probability density function of

$$f(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} \; ; \; x \geq 0 \tag{1}$$

it is said that the variable **x** is 1-parameter gamma distributed, with the shape parameter α. In Equation (1), $\Gamma(\alpha)$, the incomplete gamma function, is given by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \tag{2}$$

The distribution function has a form of the simple exponential distribution in the case of $\alpha = 1$. If x in Equation (1) is replaced by x/β, where β is the scale parameter, then the 2-parameter gamma distribution is obtained as

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \; ; \; x \geq 0 \tag{3}$$

which returns to the 1-parameter gamma distribution for $\beta = 1$. If x is replaced by $(x-\gamma)/\beta$, where γ is the location parameter, then the 3-parameter gamma distribution is obtained as

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} (x-\gamma)^{\alpha-1} e^{-(x-\gamma)/\beta}; \;\; x \geq \gamma \tag{4}$$

The gamma distribution, on which more detailed information can be found in Yevjevich (1972) and Bobee and Ashkar (1991), can also be given in the Pearson-type distribution.

Use of the 1-parameter gamma distribution in hydrology is limited due to its relative inflexibility in fitting to frequency distributions of hydrologic variables, but the 2- and 3-parameter gamma distributions are commonly used in place of the log-normal distribution function with 2 and 3 parameters. The disadvantage of the gamma distribution is that the cumulative distribution function can not be plotted as a straight line on the probability graph paper.

Markovic (1965) used five probability distribution functions, the normal, log-normal with 2 parameters, log-normal with 3 parameters, gamma with 2 parameters and gamma with 3 parameters to find the best-fitting distribution function of the annual precipitation and river flow at 2506 selected precipitation and river gauging stations in the USA and Canada. At the end of the study, it was found that there was no difference between the log-normal and gamma distributions and also that there was no significant advantage in the 3-parameter gamma distribution when compared to the 2-parameter gamma distribution. The gamma distribution was also found to be the probability distribution of monthly rainfall in arid regions (Sen and Eljadid, 1999).

Parameter Estimation

The characteristics of a population are given by parameters. A population of hydrological variables is not always exactly known. Therefore, statisticians use samples taken from the population. Parameters estimated from a sample are called statistics, which are functions of the elements of the sample and reflect the properties of the population.

Different techniques are used in estimating the parameters. These are listed here in ascending order of efficiency, from the least efficient to the most efficient: the graphical method, the least-squares method, the method of moments, and the maximum likelihood method. The recently developed method of probability-weighted moments (L-moments) can also be used.

The graphical method for parameter estimation is particularly used for distributions which can be plotted as a straight line on a probability graph paper. The normal distribution is the most common example. It is not easy to use this method to estimate the parameters of distributions which cannot be plotted as a straight line on the probability graph paper, as in the case for gamma distribution. The sum of squares of differences between the coordinates of the observed values and their corresponding values on the fitted distribution should be the smallest in the least-squares method. Being one of the oldest and the most useful methods of parameter estimation, the method of moments uses relations between the central moments and parameters of the distribution. Parameter estimation is given for the 2-parameter gamma distribution as

$$E(x) = \alpha\beta \ \ \mathrm{Var}(x) = \alpha\beta^2 \tag{5}$$

where $E(.)$ denotes the expected value of the variable and Var(.) the variance. In the maximum likelihood method, values maximizing the maximum likelihood function of the distribution are taken as estimates of the parameters.

Developed by Greenwood et al. (1979) originally for distributions expressable in inverse form, the method of probability-weighted moments (Hosking and Wallis, 1997) can also be used for parameter estimation of the 2-parameter gamma distribution, a distribution which is non-expressable in inverse form. The method is based on relations between the probability-weighted moments or L-moments and parameters of the distribution.

Generation of Gamma-Distributed Random Numbers

The generation of gamma-distributed random numbers is harder than the generation of random numbers having distributions such as the uniform, normal or exponential. The complexity of the generation of gamma random numbers increases with the number of parameters included in the distribution. Several methods have been developed to generate gamma-distributed random numbers (Bobbee and Ashkar, 1991). In this study, an algorithm given in Kottegoda (1980) is used in generating the gamma distributed random numbers.

Application I: Daily Rainfall Amounts

When the generation of a daily rainfall series is required, it is first necessary to determine the state of the day. A day is defined as either wet or dry. A Markov chain-based algorithm is one of the best ways of determining the state of a day (Aksoy, 1998).

If it is definitely known that a day is wet (rainy), it is necessary to determine the amount of rainfall on such a day. Daily rainfall data can be characterized by a probability distribution function known from the statistical literature. In the present study, the 2-parameter gamma distribution is considered the probability distribution function of the data, as used by Katz (1977) and Wilks (1989).

The model is applied to a daily rainfall data series taken from the Goztepe rain-gauge station of the State Meteorological Works (Devlet Meteoroloji Işleri, DMI), in the Asian part of Istanbul. The data series includes the daily total amount of rainfall in mm for a 30-year period from 1961 to 1990.

The shape and scale parameters of the gamma distribution (α and β) are determined from the daily rainfall data of the gauging station in a monthly time interval, by Equation (5) and given in Table 1. It may be seen that α varies between 0.341 (July) and 0.569 (December) and β, between 6.892 (June) and 19.94 (September).

A series of 30 years' duration is generated using a Markov chain-based algorithm (Aksoy et al., 1998) in determining the state of a day (wet/dry days). Generated gamma-distributed random numbers with parameters of α and β are used in calculating the amount of rainfall on wet days. The results obtained from the generated series are compared to the ones from the observed series. Comparisons were made according to average (Figure 1), standard deviation (Figure 2), skewness (Figure 3), and lag-one correlation coefficient (Figure 4) in a daily time interval. In addition, the monthly (Figure 5) and annual averages (Figure 6) are compared. It can be seen from Figures 1-6 that all the characteristics are well preserved. A comparison of the shape and scale parameters (Figures 7 and 8) also shows agreement between the amounts of observed and simulated rainfall.

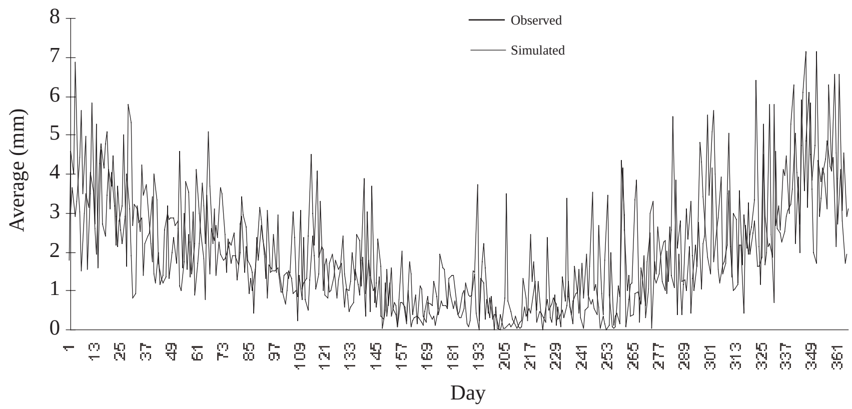

Figure 1. Average Daily Rainfall

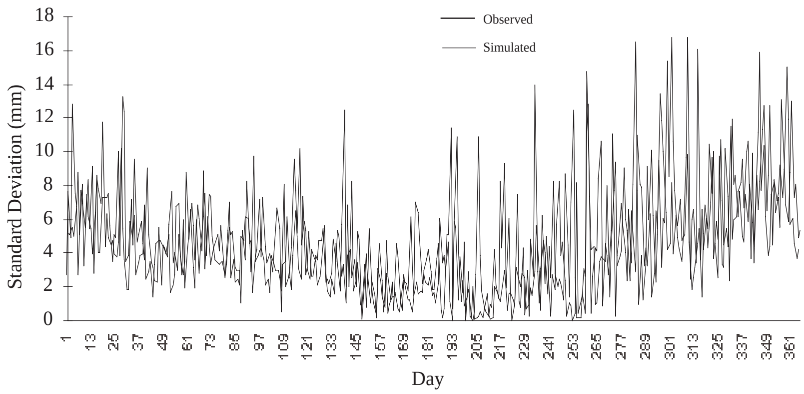

Figure 2. Standard Deviation of Daily Rainfall

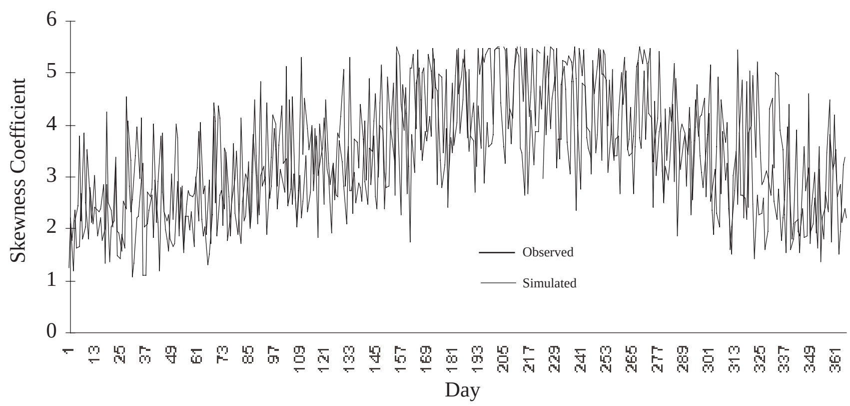

Figure 3. Skewness of Daily Rainfall

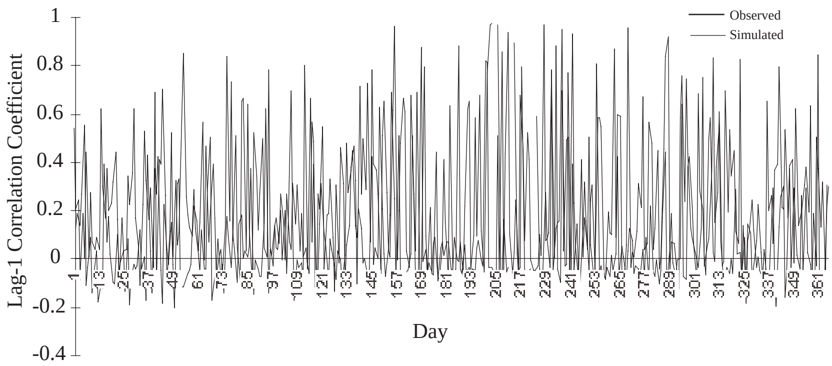

Figure 4. Lag-one Correlation Coefficient of Daily Rainfall

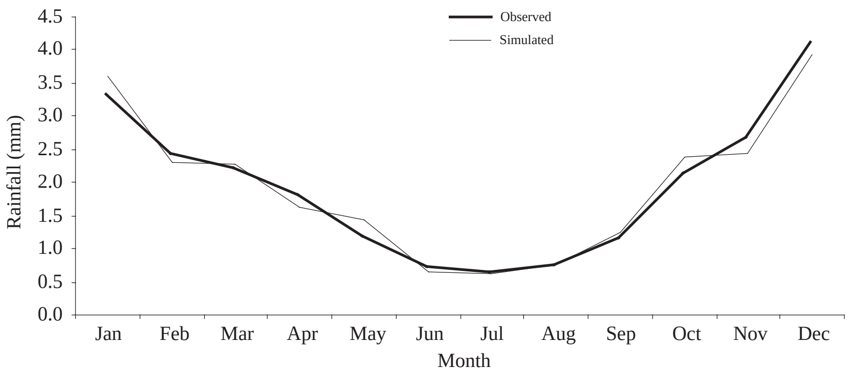

Figure 5. Average Monthly Rainfall

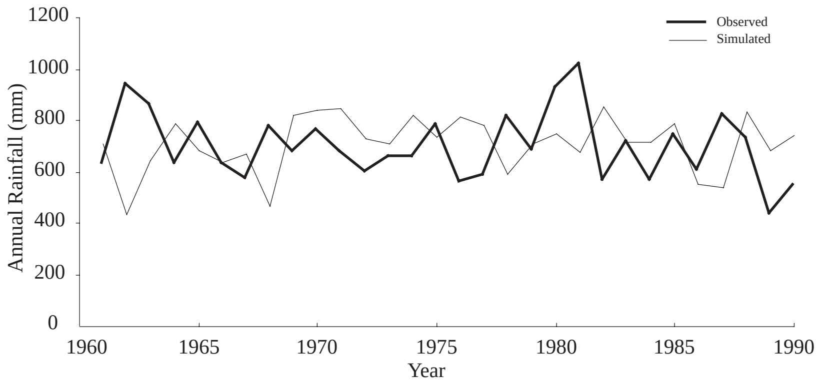

Figure 6. Annual Rainfall

Figure 7. Observed and simulated values of the shape parameter of the 2-parameter gamma distribution for rainfall data

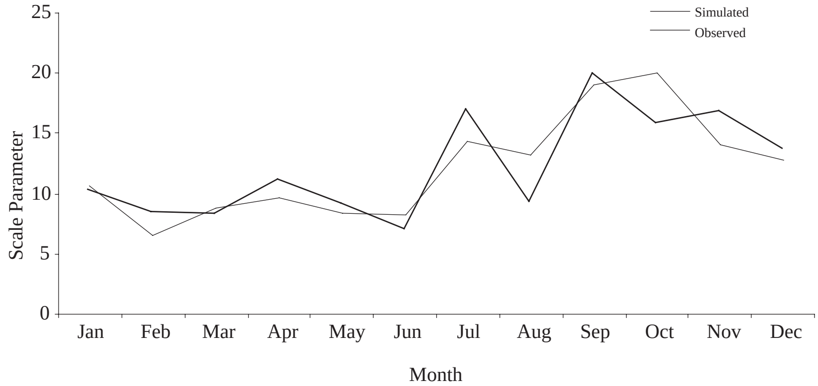

Figure 8. Observed and simulated values of the scale parameter of the 2-parameter gamma distribution for rainfall data

Table 1. Shape and scale parameters of the 2-parameter gamma distribution determined for rainfall data

Month	α	β
January	0.561	10.23
February	0.504	8.386
March	0.536	8.241
April	0.436	11.05
May	0.453	9.057
June	0.495	6.892
July	0.341	16.87
August	0.432	9.216
September	0.355	19.94
October	0.441	15.77
November	0.375	16.72
December	0.569	13.65

Application II: Ascension Curve of the Hydrograph

A typical hydrograph has three components: the ascension curve (rising limb), peak point and recession curve (falling limb). It can be said that a day with a flow value greater than the flow of the previous day is on the ascension curve of the hydrograph (which corresponds to a flow increment). Otherwise, the day is assumed to be on the recession curve (which corresponds to a flow decrement).

Many studies on the recession curve of a hydrograph were found in the literature (Hall, 1968; Kavvas and Delleur, 1984; Clausen, 1992; Wittenberg, 1994, 1999; Tallaksen, 1995; Wittenberg and Sivapalan, 1999). However, no studies on the ascension curve were found in the literature. Therefore, the aim in this paper is to propose the random structure of the ascension curve of a hydrograph using gamma distribution. It is assumed that differences between the flows of successive days (flow increments) on the ascension curve of the hydrograph are gamma distributed.

Parameters which will be used for the investigation of the flow increment are determined according to the method explained below: along the ascension curve of the hydrograph, the flow increment may have a value of zero due to a very small unmeasurable increment. Such a day is considered a day with an increment of zero. If the same phenomenon is observed during a time period on the recession curve, it is considered a day with a very small unmeasurable decrement. These zero-flow increments are included in the estimation of the shape and scale parameters of the ascension curve. Parameters are determined for each month of the year, as in the case of the previous application, which results in a total of 24 parameters.

Treiber and Plate (1977) and Sargent (1979) suggested the use of the 2-parameter gamma distribution for simulation of the ascension curve of the hydrograph, but they did not use this distribution due to constraints related to software for the generation of gamma variates. In the present study, the flow increments on each ascension curve are successively generated by 2-parameter gamma-distributed random numbers. The generated random numbers on the ascension curve are then ranked, so that the higher the increment, the closer it is to the peak of the hydrograph (Aksoy, 1998). The shape and scale parameters of the ascension curve are determined by the method of moment, using Equation (5).

Testing of the technique is carried out using daily flow data from Seytan Deresi, an intermittent stream at Babaeski in the Thrace region, for the 35-year period 1958-1992. The gauging station is operated by the Electrical Power Resources Survey and Development Administration (Elektrik Işleri Etüt Idaresi, EIEI) and the number of the station is 101. The station is 50 m above the mean sea level. The area above the gauging station is 478.4 km^2 and the long-term daily mean flow is 2.44 m^3/s.

The shape (α) and scale (β) parameters of the 2-parameter gamma distribution are calculated for each month of the year (Table 2).

Table 2. Shape and scale parameters of the 2-parameter gamma distribution determined for runoff data

Month	α	β
October	0.023	14.760
November	0.056	14.215
December	0.075	31.765
January	0.083	44.304
February	0.187	20.017
March	0.251	13.527
April	0.231	5.749
May	0.084	12.160
June	0.148	5.422
July	0.105	4.142
August	0.134	1.907
September	0.100	1.053

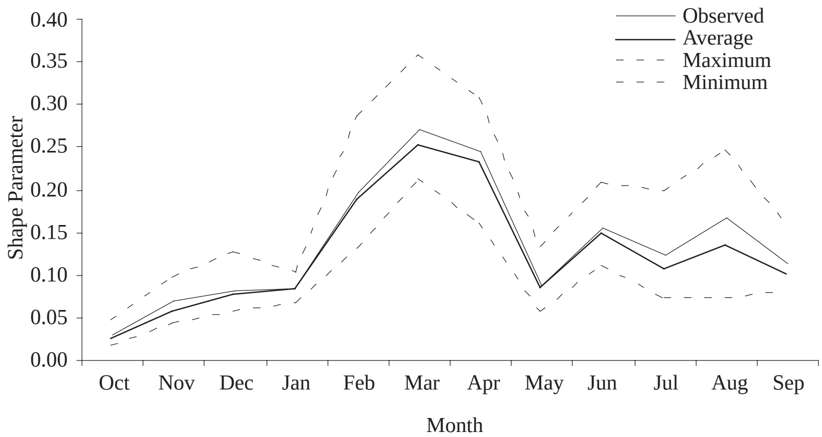

Figure 9. Observed and simulated values of the shape parameter of the 2-parameter gamma distribution for runoff data

The results are given as the average of ten simulations, each 35 years long. As in the previous case, using a Markov chain, it is first determined whether the flow has an increment or a decrement (Aksoy, 1998). An increment is generated by using the 2-parameter gamma-distributed random numbers, if the flow in a day has an increment. Figures 9 and 10 show a graphical comparison of the shape and scale parameters of the distribution. It is seen that both parameters are well preserved. The minimum and maximum values obtained from ten simulations are also given in the figures. The monthly mean, standard deviation and skewness coefficient of the increments in the flow obtained as an average of ten simulations are given in Figures 11-13 together with the observed values. The mean and standard deviation of the flow increments are well preserved. Although a 2-parameter gamma probability distribution is used, the skewness coefficient is generally close to the observed value with the exception of months (from October to January) with very high skew.

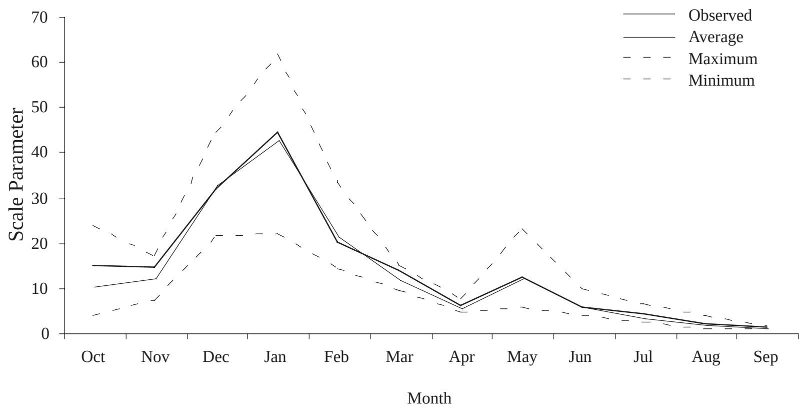

Figure 10. Observed and simulated values of the scale parameter of the 2-parameter gamma distribution for runoff data

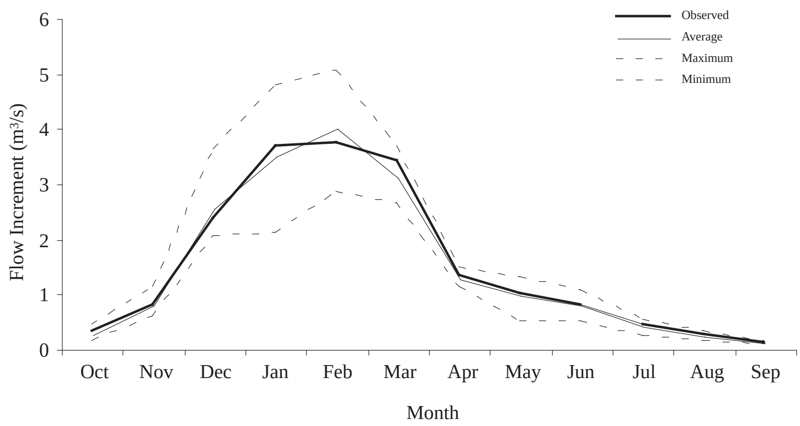

Figure 11. Observed and simulated values of monthly mean of increments in the flow

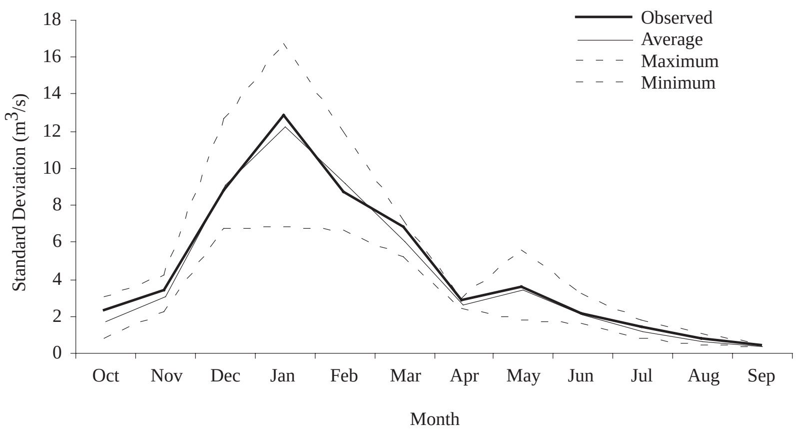

Figure 12. Observed and simulated values of standard deviation of increments in the flow

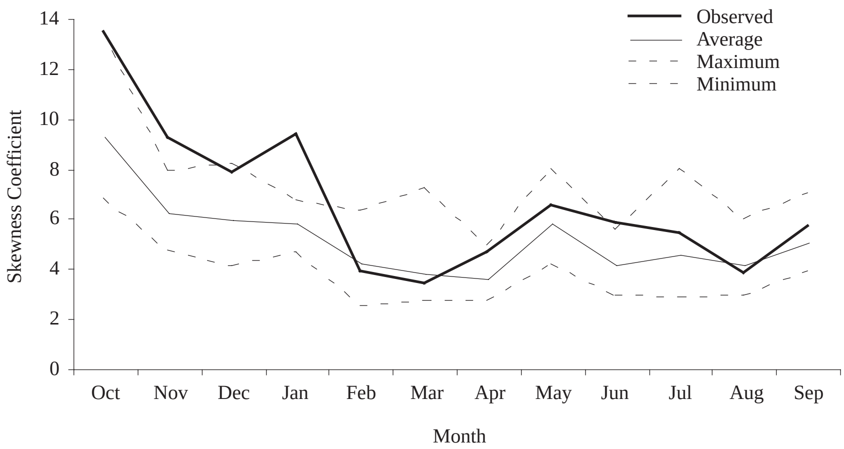

Figure 13. Observed and simulated values of skewness coefficient of increments in the flow

Conclusions

The generation of long sequences of daily rainfall data can be performed by means of an assumed probability distribution function. The proposed method suggests that the amount of rainfall on wet days can be very easily generated by a 2-parameter gamma-distributed random number generation algorithm once wet and dry days are determined. It may be seen, as was noted in previous studies by Katz (1977) and Wilks (1989), that the 2-parameter gamma distribution fits to the daily rainfall data. Short- and long-term characteristics such as daily average, standard deviation skewness, and monthly and annual averages are well preserved.

The ascension curve of the hydrograph was also analysed. It is known that the ascension curve of the hydrograph is a probabilistic process rather than a deterministic process and hence that it follows a probability distribution function. Differences between the flow values of successive days (increments in the flow for each day) are used. The 2-parameter gamma-distribution function is fitted to these increments. At the end of the application, it is found that the gamma family of distribution functions is a very useful tool in investigating the ascension curve of the hydrograph.

Symbols

$E(.)$	:	Expected value
x	:	Variable
Var(.)	:	Variance
α	:	Shape parameter
β	:	Scale parameter
γ	:	Location parameter
$\Gamma(.)$	:	Incomplete gamma function

References

Aksoy, H., Kuruyan Akarsularin Gunluk Akimlarinin Modellenmesi, Doktora Tezi, ITU Fen Bilimleri Enstitusu, Istanbul, 1998.

Aksoy, H., Ozdemir, A.D. and Kirmizigul, H., Markov Zincirleri Kullanilarak Eksik Yagis Verilerinin Tamamlanmasi, II. Ulusal Hidrometeoroloji Sempozyumu, 18-20 Kasim 1998, Ankara, 177-185, 1998.

Bobee, B., and Ashkar, F., The Gamma Family and Derived Distributions Applied in Hydrology, Water Resources Publications, Littleton, Colorado, 1991.

Bulu, A., and Aksoy, H., Low Flow and Drought Studies in Turkey, Proc. Low Flows Expert Meeting, 10-12 June 1998, University of Belgrade, Belgrade, 1998.

Clausen, B., "Modelling Streamflow Recession in Two Danish Streams", Nordic Hydrology, 23, 73-88, 1992.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C., and Wallis, J.R., "Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form", Water Resour. Res., 15(5), 1049-1054, 1979.

Gumbel, E.J., Statistics of Extremes, Columbia University Press, New York, 1954.

Hall, F.R., "Base Flow Recessions - A Review"., Water Resour. Res., 4(5), 973-983, 1968.

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

Katz, R.W., "Precipitation as a Chain-Dependent Process", Journal of Applied Meteorology, 16, 671-676, 1977.

Kavvas, M.L., and Delleur, J.W., "A Statistical Analysis of the Daily Streamflow Hydrograph", J. Hydrol., 71, 253-275, 1984.

Kottegoda, N.T., Stochastic Water Resources Technology, John Wiley & Sons, New York, 1980.

Markovic, R.D., Probability Functions of Best Fit to Distributions of Annual Precipitation and Runoff, Hydrology Paper, No.8, Colorado State University, Fort Collins, Colorado, 1965.

Sargent, D.M., "A Simplified Model for the Generation of Daily Streamflows", Hydrological Sciences Bulletin, 24(4), 509-527, 1979.

Sen, Z. and Eljadid, A.G., "Rainfall Distribution Function for Libya and Rainfall Prediction", Hydrol. Sci. J., 44(5), 665-680, 1999.

Tallaksen, L.M., "A Review of Baseflow Recession Analysis.", J. Hydrol., 165, 349-370, 1995.

Treiber, B., and Plate, E.J., "A Stochastic Model for the Simulation of Daily Flows", Hydrological Sciences Bulletin, 22(1), 175-192, 1977.

Wilks, D.S., "Conditioning Stochastic Daily Precipitation Models on Total Monthly Precipitation", Water Resources Research, 25 (6), 1429-1439, 1989.

Wittenberg, H., Nonlinear Analysis of Flow Recession Curves. In FRIEND: Flow Regimes from International Experimental and Network Data, ed. by P. Seuna, A. Gustard, N.W. Arnell and G.A. Cole, IAHS Publication, 221, 61-67, 1994.

Wittenberg, H., "Baseflow Recession and Recharge as Nonlinear Storage Processes", Hydrological Processes, 13, 715-726, 1999.

Wittenberg, H., and Sivapalan, M., "Watershed Groundwater Balance Estimation using Streamflow Recession Analysis and Baseflow Separation", J. Hydrology, 219, 20-33, 1999.

Yevjevich, V., Probability and Statistics in Hydrology, Water Resources Publ., Fort Collins, Colorado, 1972.

TECHNICAL NOTE