Fuzzy awakening in rainfall-runoff modeling

Zekai Şen and Abdüsselam Altunkaynak

Istanbul Technical University, Civil Engineering Faculty, Hydraulics Division, Maslak 34469, Istanbul, Turkey.
E-mail: *altunkay@itu.edu.tr*

Received 29 April 2003; accepted in revised form 27 August 2003

Abstract Rainfall-runoff relationships are widely used in many engineering hydrologic designs in urban and rural areas. Such relationships are obtained through the application of regression analysis in many studies. Unfortunately, in the classical regression approach to determine rainfall-runoff relationships, internal uncertainties are not taken explicitly into consideration. In this paper, an alternative to the classical regression approach is proposed through fuzzy system modeling. It is concluded that the fuzzy systems approach yields comparatively less relative error than a regression approach and, therefore, it is recommended for use in future. The application is presented for rainfall-runoff records at two sites near Istanbul, Turkey.

Keywords Drainage; fuzzy logic; hydrology; prediction; rainfall-runoff; regression analysis

Nordic Hydrology **Vol 35 No 1 pp 31–43** © IWA Publishing 2003

Introduction

The common and simple approach in the assessment of rainfall-runoff relationships is the use of regression analysis. The simplicity of this method lies in its basic data requirements, specifically rainfall and runoff records. On the other hand, it is well known that the use of regression equations is a remedy for the very complicated rainfall-runoff processes that exist (at preliminary stages of the runoff estimation for design). However, prior to its use, the following imbedded hydrologic assumptions, comments, and simplifications should be considered:

1. In simple regression relationships rainfall is assumed as uniformly distributed over the drainage area. Such an assumption might be valid for small areas, but as the area increases, the validity of this approach must be questioned. Consequently, more uncertainties become included in the overall rainfall-runoff transformation process.
2. Depending on the antecedent soil and surface conditions of the drainage area, the portion of the rainfall that appears as direct runoff will be different even when the peak rainfall amounts are the same. This indicates that the transformation to runoff is not a static, but rather a dynamic process according to the environmental conditions. For instance, during wet periods its value is comparatively larger than during dry spells. It should be noted here that the words ‘wet’ and ‘dry’ are linguistically fuzzy in content.
3. Logically, rainfall depth is greater than the generated runoff depth from the same storm, and consequently, the proportionality factor assumes values between 0 and 1. However, such a proportionality coefficient is not a constant throughout the year and it also depends on the antecedent conditions. It is not possible to consider such variations in the coefficients through regression analysis where the calculations are based on all data considered. Although multiple regression analysis can be used, still the problem of parameter estimation remains. On the other hand, the fuzzy approach furnishes a basis for considering such uncertainties through vaguely defined membership functions. Again, phrases such as ‘greater than’ and ‘assumes values between 0 and 1’ are fuzzy statements.

In addition to these hydrological requirements, most often given the measurements of rainfall and runoff, the determination of the rainfall-runoff relationship is a first-stage analysis, which requires a suitable methodology. Since the rainfall and runoff measurements show haphazard fluctuations around an average value, the appropriate methodology should be based on uncertainty techniques such as conventional statistical or modern fuzzy approaches. Most often a simple linear relationship between the dependent rainfall and independent runoff variables is preferred. For any regression application in finding the rainfall-runoff relationship, the following steps are necessary in the calculations.

(1) In practice, the rainfall and runoff measurements are plotted on a rectangular coordinate system and the result is a scatter of points. In fact, each one of these points corresponds to different antecedent or environmental conditions. However, in the regression line fitting such a distinction is not considered, and each point in the scatter is treated equally leading to the best straight line. Use of the regression approach brings into view additional restrictive assumptions such as the equivalence of variances, *i.e.* uniform variance throughout the data set, independence of deviations of each scatter point from the fitted regression line, their Gaussian distribution. If these procedural assumptions are not satisfied then the regression approach leads to biased rainfall-runoff relationships.

(2) The regression methodology yields single values of parameters. This implies that irrespective of seasonality as an important factor, its influence on vegetation cover and infiltration rates, the parameters are considered as having the same values for different seasons or months of the water year.

(3) Although the scatter diagram shows the random behavior of the drainage basin, in the sense that the rainfall-runoff transformation does not change significantly with the physical characteristics of the basin, the regression parameters are expected to vary with the duration of rainfall and antecedent conditions, which is not evident in the scatter diagram on the regression.

The main purpose of this paper is to propose a fuzzy rainfall-runoff modeling procedure as an alternative to the classical regression approach which requires a set of restrictive assumptions. However, in the fuzzy approach none of these assumptions are required. The application of fuzzy and regression methods are presented for two different drainage basins within the metropolitan city boundary of Istanbul, Turkey.

Rainfall-runoff scatter diagrams

In general, the linear regression is fitted to the scatter of rainfall-runoff points, conventionally with single and constant model parameter values. However, such an approach ignores the dynamic behavior of the rainfall-runoff process, and consequently the variations of the data are rendered to a completely deterministic world. It is well known that the rainfall-runoff process is dynamic and nonlinear in its nature where proportionality and superposition principles do not apply (Kundzewicz and Napiorkowski 1986). The statistical approach adopted herein is to group runoff coefficients in terms of months, and calculation of average rainfall and runoff calculations from given data. Thus, the scatter of 12 monthly average rainfall and runoff values appear on rectangular coordinate systems. The connection of the rainfall-runoff points in the logical monthly sequence leads to irregular polygons on the coordinate systems (Kadıoğlu and Şen 2001). Such a hypothetical rainfall and runoff scatter diagram with 12 points and their successive corrections are presented in Figure 1.

Since the 12-sided shape is in the form of an irregular polygon, it is also referred to as the rainfall-runoff polygon (Hoyt 1936). The following interpretive features are evident from such polygons.

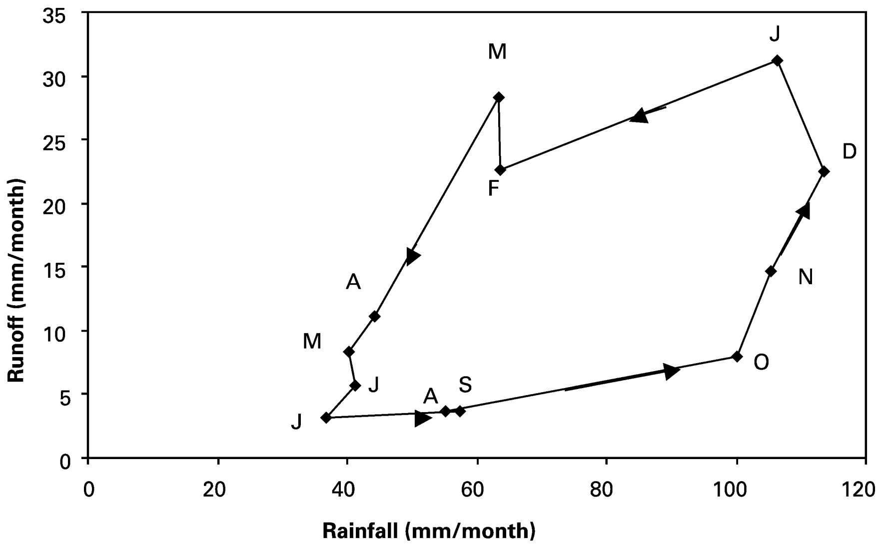

Figure 1 Monthly rainfall-runoff polygon

(a) The lengths of polygon sides indicate the change in average values of precipitation or runoff for consecutive months.
(b) The length of each polygon side indicates the value of the proportionality coefficient between rainfall and runoff between consecutive months.
(c) The closeness of the slope of each side to the vertical or horizontal indicates the relative proportions of the rainfall and runoff. Similar interpretations for all the months during one year provide a basis for qualitative interpretations about the rainfall-runoff occurrences in catchments. Unfortunately, in any regression approach these differences in the proportionalities during one year are not taken into consideration at all. However, the fuzzy approach accounts for such differences.
(d) Along each side of the polygon, runoff is assumed to change linearly with rainfall. Such a linearity assumption during time intervals smaller than one year yields more reliable results in the runoff volume calculations. The polygon constitutes finite straight-line portions for the validity of a linearity assumption on a monthly basis. Practically, if all of the sides fall along a single direction within 5% or 10% deviations, then the corners in the polygon diagram might be considered as scattered along a straight line which represents the monthly rainfall-runoff relationship. The narrower the polygon, the more representative will be the regression approach for rainfall-runoff modeling. In contrast, wide polygons imply heterogeneous temporal variations, dynamism and non-linearity in rainfall-runoff relationships for the catchment area considered.
(e) The smaller the area of the polygon, the more consistent the monthly rainfall and the more reliable is the regression estimation of the resulting runoff. Otherwise, the results are not reliable and instead, the fuzzy approach must be employed in finding the rainfall-runoff transformation relationship.

Fuzzy logic approach

The fuzzy approach is based on the linguistic uncertain expressions rather than numerical uncertainty measures. A detailed account of fuzzy logic and systems is presented by Zadeh (1965). Many researchers have applied the fuzzy approach to various engineering problems (Mamdani 1974; Pappis and Mamdani 1977; Sugeno 1985; Şen 1998; Şen 2001). The basis of fuzzy logic is to consider hydrological variables in a linguistically uncertain manner, in the forms of subgroups, each of which is labeled with successive fuzzy word attachments such as "low", "medium", "high", etc. In this way, the variable is considered not as a global and numerical quantity but in partial groups which provide better room for the justification of sub-relationships between two or more variables on the basis of fuzzy words. For instance, in this paper rainfall and runoff variables are considered as five partial subgroups, namely, "low", "medium low", "medium", "medium high" and "high". A small number of fuzzy subgroups selection leads to unrepresentative predictions whereas a large number imply unnecessary calculations. In practical studies, in the preliminary stage the number of subgroups is selected as four or five (Şen 2001). Five subgroups in each variable imply that there are $5 \times 5 = 25$ different partial relationship pairs that may be considered between the rainfall and runoff variables. However, many of these relationships are not physically plausible. For instance, if the rainfall is "high" it is not possible to state that the runoff is "low" or ever "medium". Figure 2 shows the relative positions of the fuzzy words employed in this paper. Each one of the middle fuzzy words is shown as a triangle with the maximum membership degree at its apex. The most left and right fuzzy words, namely, "low" and "high" are represented by trapeziums.

It is significant to consider that neighboring fuzzy subsets interfere with each other providing the fuzziness in the modeling. In the case of systems that lie outside the preconceived triangle, these triangles change shape depending on the climatic regions and soil types, *e.g.* sandy desert can take 150 mm of rainfall and 20 mm runoff.

Since the rainfall-runoff relationship, in general, has a direct proportionality feature, it is possible to write the following five rule-bases for the description of fuzzy rainfall-runoff modeling. These rules are simply

R1: IF rainfall is L THEN runoff is L or
R2: IF rainfall is ML THEN runoff is ML or
R3: IF rainfall is M THEN runoff is M or
R4: IF rainfall is MH THEN runoff is MH or
R5: IF rainfall is H THEN runoff is H

where L, ML, M, MH and H are abbreviations for fuzzy subgroups of "low", "medium low", "medium", "medium high" and "high", respectively. The general appearance of such a model is given in Figure 3 where, rather than a regression model of the classical calculations, a fuzzy region of rainfall-runoff relationships is considered with uncertainty domain.

In order to show two applications of fuzzy inference, in Figure 4 two rules of the rainfall-runoff relationship are shown with membership degree, μ.

Given the rainfall intensity, $i = 22.5$ mm both ML and M fuzzy subgroups of rainfall variable are triggered. The consequent part of each runoff variable appears as a truncated

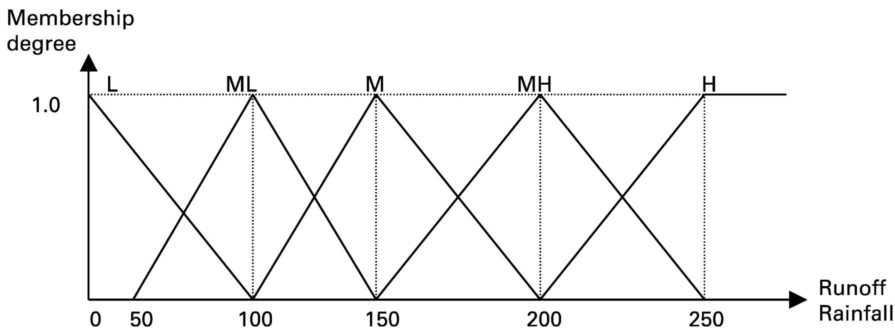

Figure 2 Hypothetical fuzzy subgroups of rainfall and runoff

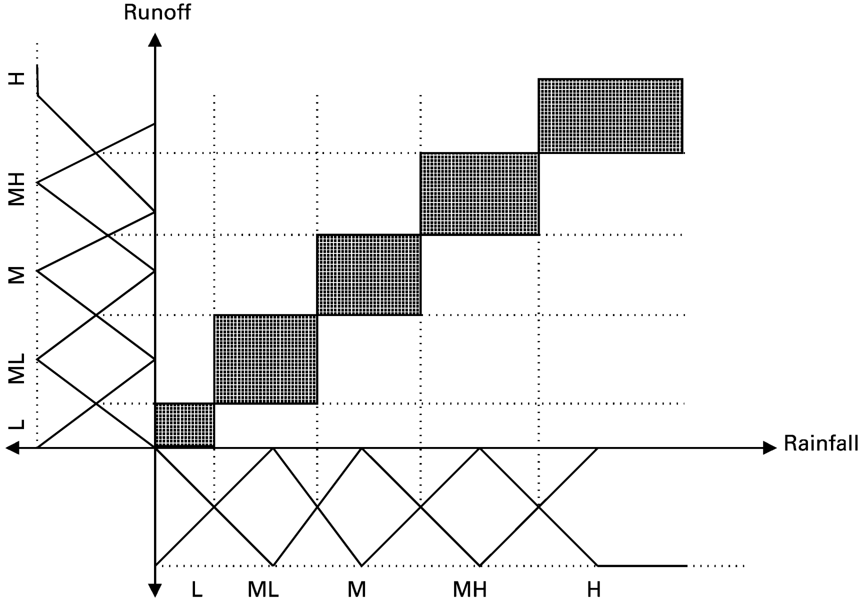

Figure 3 Fuzzy rule-base rainfall-runoff relationship domain

trapezium for each rule on the right hand side in Figure 4. The overlapping of these two truncated trapeziums indicates the combined inference from these two rules as in the lower part of the same figure which is represented in Figure 5 with relevant numbers. In this figure A1, A2, A3, A4, A5 and A6 indicate triangular and rectangular subareas in the fuzzy inference.

For hydrologic design purposes, it is necessary to deduce from these combined fuzzy subgroups a single value which is referred to as "defuzzification" in the fuzzy systems terminology. The purpose of defuzzification is to convert the final fuzzy set representing the overall conclusion into a real number that, in some sense, best represents this fuzzy set. Although there are various defuzzification methods the most common method is centroid defuzzification (Ross 1995; Şen 2001). In general, given a fuzzy set with membership degree $\mu(x)$ defined on the interval [b, c] of variable x, the centroid defuzzification prediction, $\hat{x}$, is defined as

$$\hat{x} = \frac{\int_{b}^{c} x\mu(x)dx}{\int_{b}^{c} \mu(x)dx}$$

By applying this formula to the fuzzy inference set in Figure 5, it is possible to obtain a defuzzification value by numerical calculation as:

$$\hat{x} = \frac{\sum_{i=1}^{6} x_i A_i}{\sum_{i=1}^{6} A_i}$$

$$= \frac{0.05*6.67 + 0.90*7.88 + 0.05*9.02 + 0.10*9.25 + 0.25*10.13 + 0.03*10.83}{0.05 + 0.90 + 0.05 + 0.10 + 0.25 + 0.03} = 8.4$$

which is shown in the same figure.

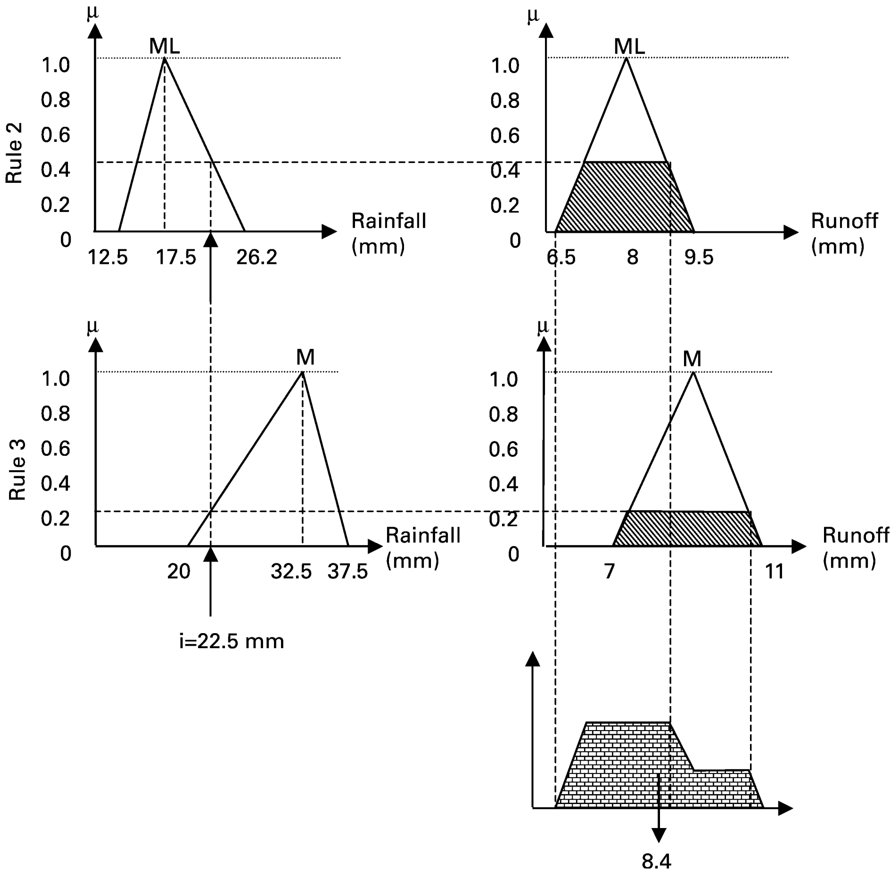

Figure 4 Rainfall-runoff relationship rules

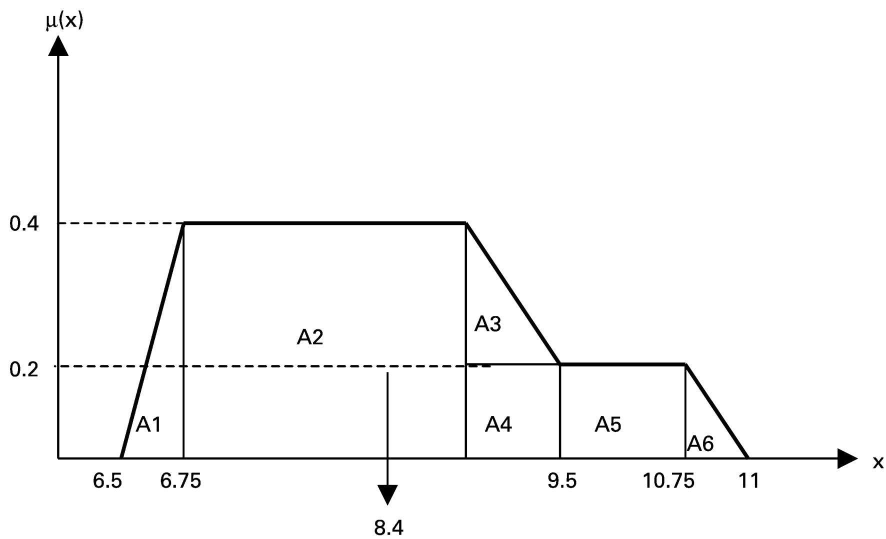

Figure 5 Fuzzy inference subset

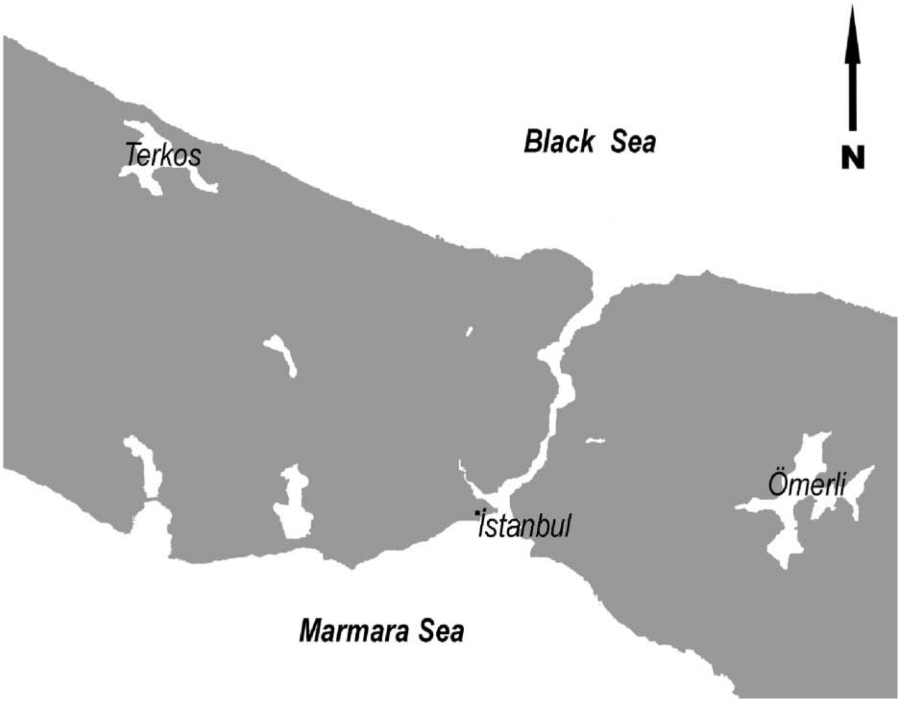

Figure 6 Catchment location map

Application and discussion

The application of the two methodologies (regression and fuzzy) is performed for two different drainage basins on the European and Asian sides of Istanbul (see Figure 6). For better water supply to Istanbul, these catchments are connected jointly through a submarine pipe between the two continents. Both catchments have been measured for almost 20 years in order to depict the rainfall-runoff relationships. Hence, simultaneous measurements of monthly rainfall and runoff averages are available for each drainage basin.

Figure 7 shows the classical regression line fittings for the determination of runoff coefficient in each catchment for a monthly term. It is obvious that in both of these catchments, the scatter of rainfall-runoff data is considerable. This suggests instability in the relationship between rainfall and runoff. In addition, the prerequisite assumptions, especially the variance constancy, are not met, because in each scatter diagram, the variance is small for small rainfall-runoff values, but it increases as the rainfall-runoff values increase. Hence, the classical regression lines cannot provide reliable relationships.

Tables 1 and 2 give the monthly averages and standard deviations, respectively, of rainfall and runoff for each catchment. Figure 8 shows monthly rainfall-runoff scatter and consequent polygons are shown. In both drainage basins the scatter of the points are reduced to a significant extent when compared with scatter diagrams in Figure 7. For comparison purposes and obtaining more stable regression estimates, in the same figure, regression lines based on the new scatter diagrams are also shown.

The application of fuzzy system model as explained in this paper is performed for each catchment through a calculation program and the results are presented in Tables 3 and 4, respectively, for Terkos and Ömerli catchments. The relative error for each runoff prediction through the classical regression and fuzzy models is also presented. It is observed that the fuzzy approach invariably provides better estimates than the classical regression

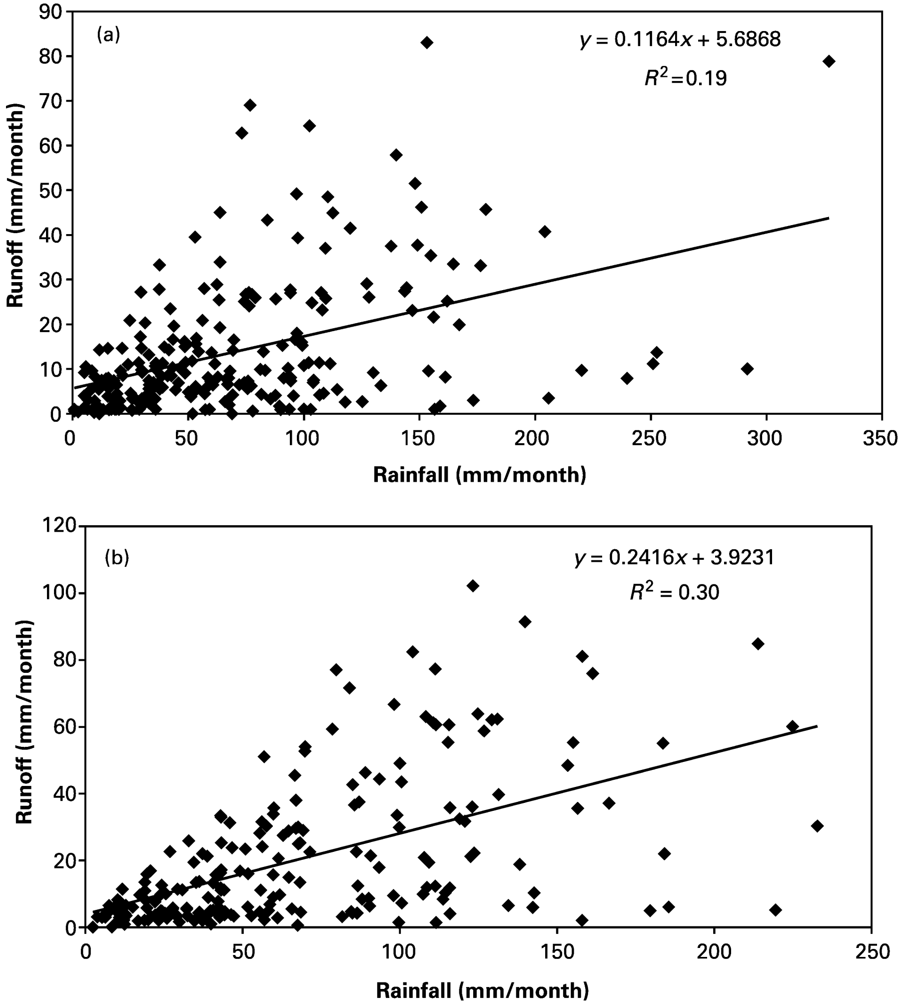

Figure 7 Regression method for monthly rainfall-runoff relations: (a) Terkos and (b) Ömerli catchments

Table 1 Monthly rainfall data statistics (mm)

Month	Terkos Ave.	Terkos St. Dev.	Ömerli Ave.	Ömerli St. Dev.
January	106.4	76.1	103.7	54.0
February	63.7	35.8	51.4	25.5
March	63.3	38.6	64.5	39.2
April	44.3	25.6	45.6	28.7
May	40.3	38.0	33.8	22.2
June	41.2	32.7	27.2	20.9
July	36.9	26.4	29.6	32.8
August	57.5	81.2	39.3	45.7
September	55.2	41.9	37.9	33.6
October	100.1	56.9	91.5	60.7
November	105.3	48.6	99.1	44.3
December	113.4	57.6	110.1	48.2
Average	69.0	46.6	61.1	38.0

Table 2 Monthly runoff data statistics (mm)

Month	Terkos		Ömerli	
	Ave.	St. Dev.	Ave.	St. Dev.
January	24.7	30.3	49.0	24.1
February	49.7	19.7	31.9	17.3
March	51.0	17.5	36.2	25.6
April	34.9	13.2	17.2	13.3
May	39.2	1.6	9.2	6.7
June	36.9	6.0	5.0	4.4
July	31.6	7.4	5.0	4.9
August	69.3	16.8	6.7	7.8
September	48.6	9.4	6.0	8.4
October	78.5	30.5	11.4	10.9
November	76.9	40.1	19.0	17.3
December	85.5	39.5	45.8	22.6
Average	52.2	19.3	20.2	13.6

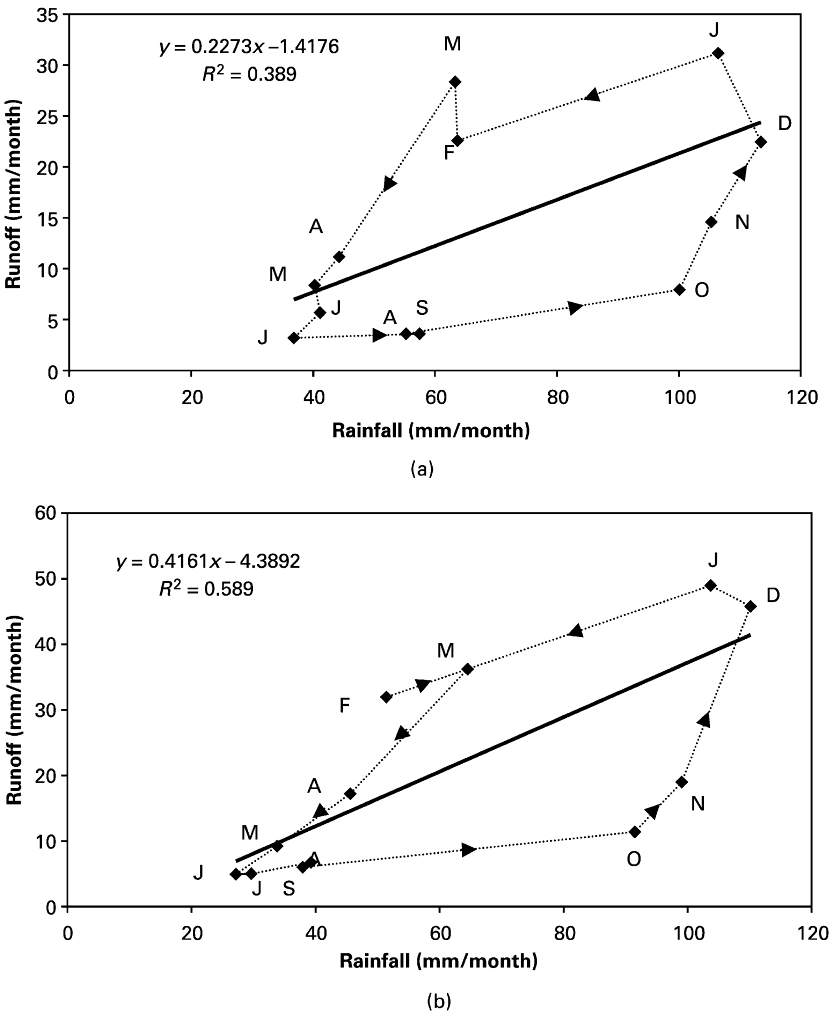

Figure 8 Rainfall-runoff polygons: (a) Terkos and (b) Ömerli catchments

Table 3 Terkos fuzzy rule-base system and regression prediction error

Observation		Runoff prediction		Relative error (%)	
Rainfall (mm)	Runoff (mm)	Fuzzy	Regression	Fuzzy	Regression
5.0	9.2	7.4	6.3	19.2	31.9
9.6	8.3	8.3	6.8	0.2	18.0
12.5	7.5	8.8	7.1	14.8	4.8
14.4	7.8	9.1	7.4	14.4	5.6
15.5	8.0	9.3	7.5	13.9	6.4
20.7	9.6	10.1	8.1	5.0	15.7
24.0	11.0	10.6	8.5	3.6	22.9
28.6	11.4	11.1	9.0	2.6	20.9
29.7	9.8	11.3	9.1	13.3	6.7
32.9	13.2	11.8	9.5	10.6	27.9
35.7	10.7	12.1	9.8	11.6	8.0
36.0	11.3	12.2	9.9	7.4	12.6
39.7	14.9	12.6	10.3	15.4	30.8
41.6	14.3	12.8	10.5	10.5	26.4
43.4	16.6	13.0	10.7	21.7	35.3
48.3	16.1	13.4	11.3	16.8	29.8
51.7	11.8	13.6	11.7	13.2	0.8
54.4	13.9	13.8	12.0	0.7	13.5
60.1	13.7	14.1	12.7	2.8	7.4
69.3	14.2	14.5	13.8	2.1	3.2
90.7	15.3	16.8	16.2	8.9	5.8
96.2	16.6	19.6	16.9	15.3	1.7
96.9	18.0	20.0	17.0	10.0	5.7
103.4	24.8	22.5	17.7	9.3	28.5
107.9	23.2	24.0	18.3	3.3	21.4
109.4	25.8	24.4	18.4	5.4	28.6
127.2	29.1	28.6	20.5	1.7	29.6
137.7	37.5	30.7	21.7	18.1	42.1
149.1	37.7	33.1	23.0	12.2	38.9
154.9	35.4	34.4	23.7	2.8	33.0
164.6	33.5	36.8	24.9	9.0	25.8
178.6	45.7	41.2	26.5	9.9	42.1
327.0	78.8	78.3	43.8	0.6	44.5
			Average	**9.28**	**20.49**

rainfall-runoff relationship. In Table 4 the fuzzy logic model prediction yields less relative error as compared to the regression method. In order to present visual inspection of the results, Figure 9 presents actuaŞ data scatter with classical regression and fuzzy approach models. It is obvious that fuzzy model follows the general trend in the scatter more closely than the regression solution.

In two catchments the fuzzy model average relative errors are less than the practically acceptable 10%.

Conclusions

Runoff calculation based on rainfall has frequently been used in practical applications for storm drainage design by using regression approach. It is still being used for various purposes preferably in small drainage areas. However, there are pitfalls in its use if based especially on the regression approach. These pitfalls are due to the basic assumption requirements in the regression method and the uncertainty within the rainfall and runoff data. The same set of assumptions remains also in the application of multiple regression. In order

Table 4 Ömerli fuzzy rule-base system and regression prediction error

Observation		Runoff prediction		Relative error (%)	
Rainfall (mm)	Runoff (mm)	Fuzzy	Regression	Fuzzy	Regression
3.8	3.1	3.0	4.8	4.8	36.0
5.3	3.0	3.0	5.2	1.3	42.4
6.3	3.0	3.0	5.5	1.0	44.9
6.5	3.8	3.1	5.5	19.0	30.8
7.4	5.3	4.0	5.7	25.5	7.2
8.0	5.2	4.5	5.9	13.7	11.2
9.2	6.0	5.5	6.2	9.0	2.4
9.7	4.9	5.8	6.3	15.7	21.8
10.6	6.6	6.4	6.5	3.0	1.8
11.0	7.6	6.7	6.6	12.5	13.4
12.8	6.6	7.6	7.0	13.5	5.9
17.2	9.7	9.4	8.1	3.2	16.7
18.9	10.9	9.9	8.5	9.1	22.1
23.9	9.8	11.0	9.7	10.9	1.1
24.2	12.6	11.1	9.8	11.9	22.5
25.6	10.0	11.3	10.1	11.5	1.1
29.3	11.4	11.9	11.0	4.2	3.5
30.8	11.2	12.0	11.4	6.7	1.5
35.2	13.4	12.5	12.4	6.7	7.3
36.2	13.7	12.6	12.7	8.0	7.5
40.6	13.3	12.9	13.7	3.0	3.2
41.5	15.7	13.0	14.0	17.2	11.2
43.0	16.3	13.1	14.3	19.6	12.2
49.1	16.8	13.4	15.8	20.2	6.0
51.7	16.0	13.5	16.4	15.6	2.5
59.7	15.7	13.7	18.4	12.7	14.4
61.4	20.6	20.6	18.8	0.0	8.9
66.9	29.6	23.3	20.1	21.3	32.1
67.7	25.0	23.7	20.3	5.2	18.9
68.2	25.3	23.9	20.4	5.5	19.4
69.2	29.0	24.3	20.6	16.2	28.8
71.4	22.5	25.1	21.2	10.4	5.9
99.1	33.5	32.5	27.9	3.0	16.8
99.8	29.9	32.8	28.0	8.8	6.2
115.9	35.8	37.2	31.9	3.8	10.8
119.0	32.5	38.1	32.7	14.7	0.5
123.0	36.0	39.3	33.6	8.4	6.6
131.4	39.7	42.6	35.7	6.8	10.2
153.4	48.5	49.2	41.0	1.4	15.5
155.1	55.3	49.8	41.4	10.0	25.1
183.7	55.1	60.2	48.3	8.5	12.3
224.9	60.1	66.9	58.3	10.2	3.1
			Average	**9.85**	**13.61**

to get rid of such assumptions, it is proposed in this paper to apply a fuzzy modeling approach. The basis of this approach is explained and applications are performed for two drainage areas near Istanbul Turkey. Possible uncertainties are explained and finally, it is suggested that in the case of rainfall-runoff record existence, it is preferable to apply the fuzzy modeling for runoff estimations from given rainfall measurements. In this model,

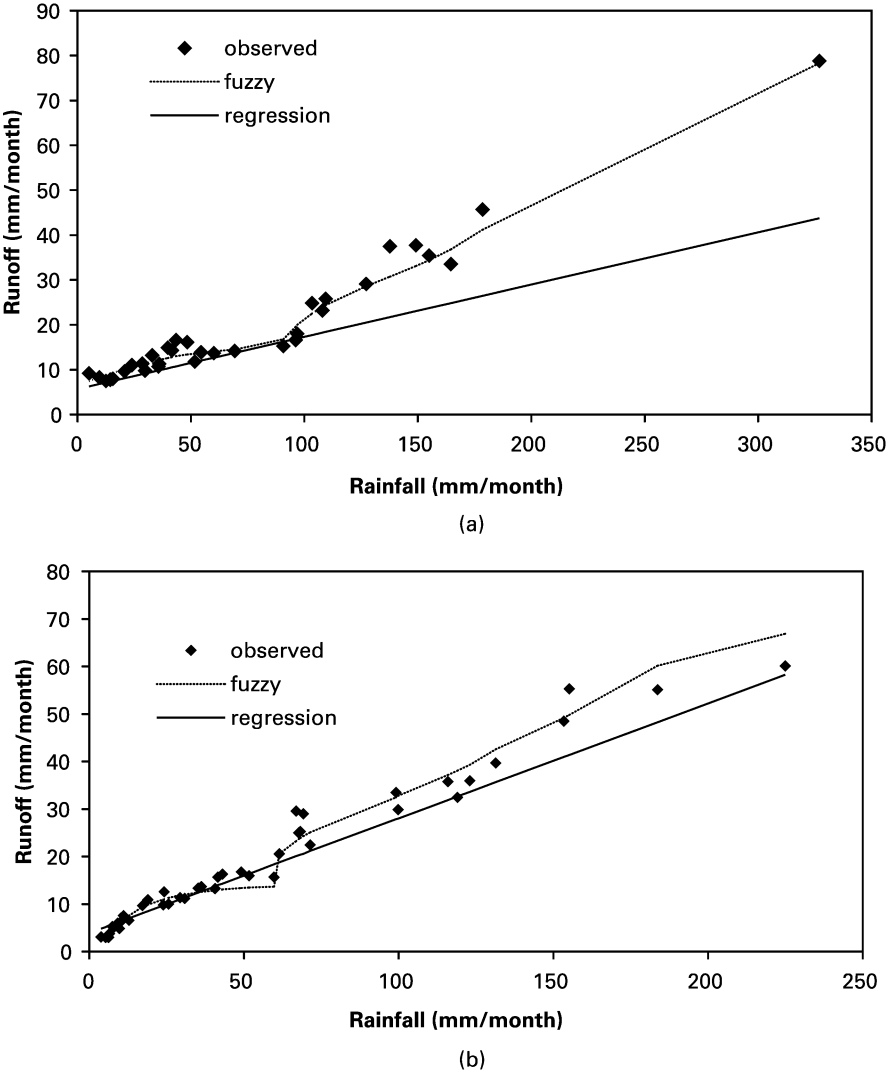

Figure 9 Regression and fuzzy models: (a) Terkos and (b) Ömerli catchments

although, the model parameters do not appear explicitly, the runoff estimations are achieved within an acceptable relative error of less than 10%.

References

Hoyt, W.G. (1936). Studies of Relations of Rainfall and Run-off in the United States. *U.S. Geological Survey Water-Supply Paper,* **772**, 301p.

Kadıoğlu, M. and Şen, Z. (2001). Monthly precipitation-runoff polygons and mean runoff coefficients. *Hydrol. Sci. J.*, **46**(1), 3–11.

Kundzewicz, Z.W. and Napiorkowski, J.J. (1986). Nonlinear models of dynamic hydrology. *Hydrol. Sci. J.*, **37**, 163–185.

Mamdani, E.H. (1974). Application of fuzzy algorithms for simple dynamic plant. *Proc. IEE*, **121**, 1585–1588.

Pappis, C.P. and Mamdani, E.H. (1977). A fuzzy controller for a traffic junction. *IEEE Trans. Syst. Ma Cybern.*, **7**(10), 707–717.

Ross, J.T. (1995). *Fuzzy logic with engineering applications*, McGraw-Hill Inc., New York, 593pp.
Sugeno, M. (Ed.), (1985). *Industrial Applications of Fuzzy Control*, *North-Holland*, New York.
Şen, Z. (1998). Fuzzy algorithm for estimation of solar irradiation from sunshine duration. *Solar Energy*, **63**(1), 39–49.
Şen, Z. (2001). *Bulanık mantık ve modelleme ilkeleri*, Bilgi Kültür Sanat Publication House (Fuzzy Logic and Modeling principles book in Turkish) 176pp.
Zadeh, L.A. (1965). Fuzzy sets. *Inform. and Control*, **12**(2), 94–102.