2011 Symposium on Data-Driven Approaches to Droughts

Drought Research Initiative Network

5-3-2012

Bivariate drought analysis using entropy theory

Zengchao Hao
Texas A & M University - College Station, hzc07@tamu.edu

Vijay P. Singh
Texas A & M University - College Station, vsingh@tamu.edu

Bivariate drought analysis using entropy theory

Zengchao Hao[1], Vijay P. Singh[1,2]

1: Department of Biological & Agricultural Engineering,Texas A&M University

2: Department of Civil and Environmental Engineering,Texas A&M University

Introduction

Drought analysis is important for water resources planning and management. Drought duration and severity are two main characteristics that have often been used for drought analysis, which can be defined using run theory using hydrological variables (e. g., streamflow) [*Yevjevich*, 1967] as shown in Figure 1 .

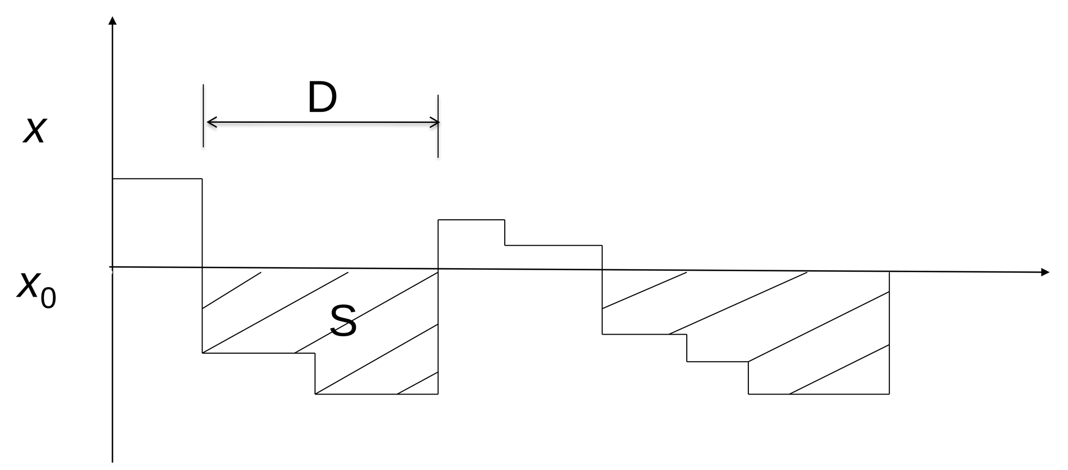

Figure 1 Defination of drought duration (D) and severity (S) using run theory for variables *x* and truction level x_0

A traditional way to characterize the drought duration or severity is based on fitting a probability density function. The drought duration can be modeled by a geometric distribution (discrete) or an exponential distribution (continuous). The gamma distribution is generally used to model drought severity.

Bivariate drought analysis is often needed to characterize the correlation between drought duration and severity. This article proposes a new method for constructing the joint density function of drought duration and severity with different marginal distributions based on the principle of maximum entropy. The proposed method is applied for drought analysis based on the monthly streamflow of Brazos River at Waco, Texas.

Methods

For a continuous random variable *X* with probability density function (PDF) *f*(*x*) defined on the interval [a, b], entropy is defined as a measure of uncertainty expressed as [*Shannon*, 1948]:

$$H_1 = -\int_a^b f(x)\ln f(x)\,dx$$

For continuous random variables *X* and *Y* with a PDF *f*(*x*, *y*) defined over the space [a, b]x[c, d], the Shannon entropy can be defined as:

$$H_2 = -\int_c^d \int_a^b f(x,y)\ln f(x,y)\,dxdy$$

The principle of maximum entropy is proposed by *Jaynes* [1957] which states that the probability density function should be selected among all the distributions that satisfy the constraints. It has been shown that many of the commonly used distributions can be derived from entropy theory with different constraints and the maximum entropy distribution incorporates these distributions as special cases [*Singh*, 1998].

To derive the joint density function *f*(*x*,*y*) of drought duration (*X*) and severity (*Y*), constraints for variables *X* and *Y* need to be specified separately and jointly. These constraints of variables X and Y can be specified as

$$\int_a^b f(x,y)x^2\,dx = \overline{x^2} \qquad \int_a^b f(x,y)y^2\,dx = \overline{y^2}$$

$$\int_a^b f(x,y)\ln y\,dy = \overline{\ln y} \qquad \int_c^d \int_a^b xyf(x,y)\,dxdy = \overline{XY}$$

With these constraints above , the joint PDF can be obtained by maximizing entropy H_2 as:

$$f(x,y) = \exp(-\lambda_0 - \lambda_1 x^2 - \lambda_2 y^2 - \lambda_3 \ln y - \lambda_4 xy)$$

where λ_i (*i*=0,1,2…,4) are the Lagrange multipliers . The univariate distribution for drought duration *X* can be obtained by integrating the joint PDF f(*x*, *y*) over *Y* . Similarly, the unviariate distribution for drought severity can be obtained accordingly.

Results

Monthly streamflow data of Brazos River at Waco, TX (USGS 08096500) for the period from January 1941 to December 2009 was used for drought analysis. The mean streamflow of each month is used as the truncation level to define the drought event. Significant correlation exists between drought duration and severity and the bivariate distribution was used to model them jointly.

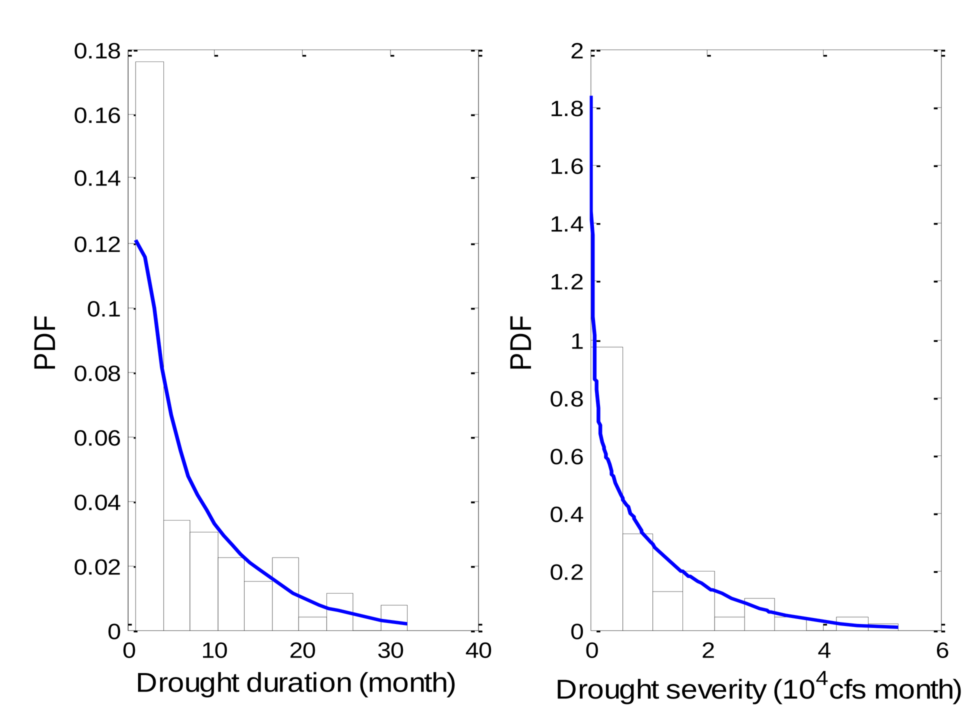

Figure 2 Comparison of the histograms and fitted PDFs

Histograms and fitted marginal distributions of drought duration and severity are shown in Figure 2. The fitted PDFs capture the general pattern of the histograms. The empirical probabilities estimated from the Gringorten's plotting position formula and theoretical cumulative probabilities are shown in Figure 3. It can be seen that generally theoretical probabilities fitted empirical probabilities well.

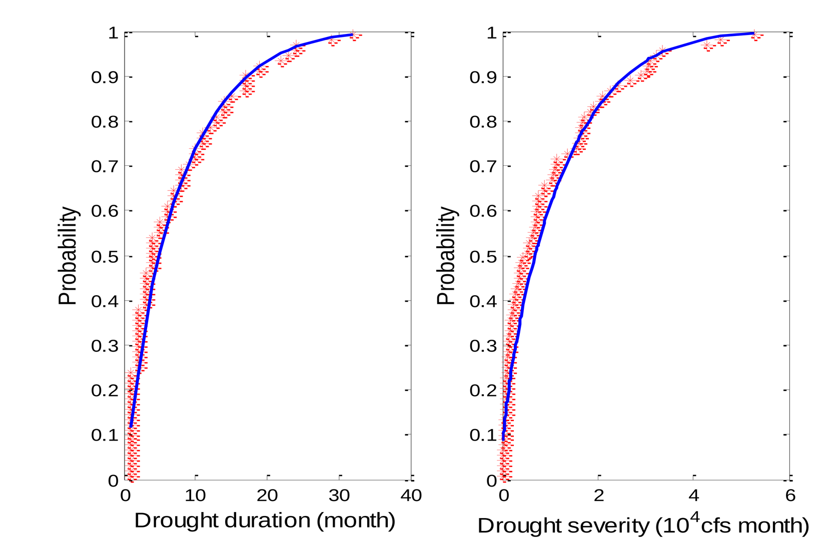

Figure 3 Comparison of the empirical and theoretical probability

The univariate return periods of 2,5,10,20,50 and 100 years defined by separate drought duration and severity are summarized in Table 1. The joint return period and conditional return period are also computed as shown in Figure 4 and Figure 5.

Table 1 Return period defined by drought duration and severity separately

Return Period (Year)	Drought duration (Month)	Drought severity (10^4cfs months)
2	6.9	1.0
5	14.1	2.1
10	19.0	2.8
20	23.4	3.5
50	28.5	4.3
100	32.0	4.8

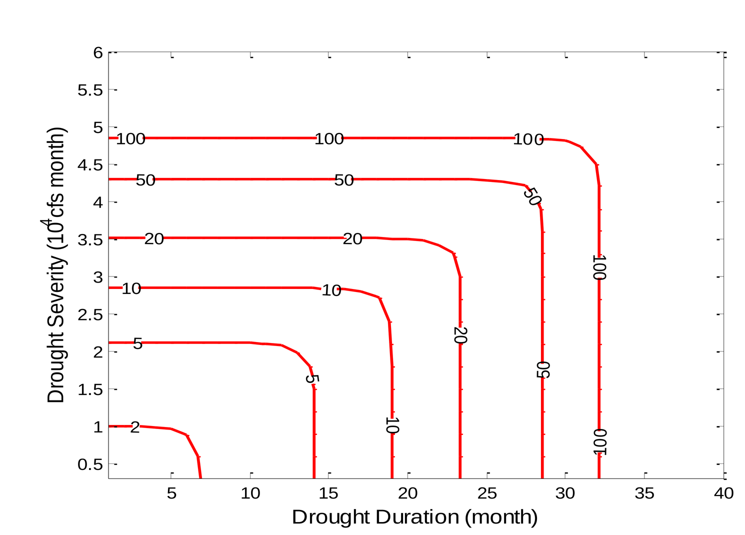

Figure 4 Bivariate drought duration and severity return periods

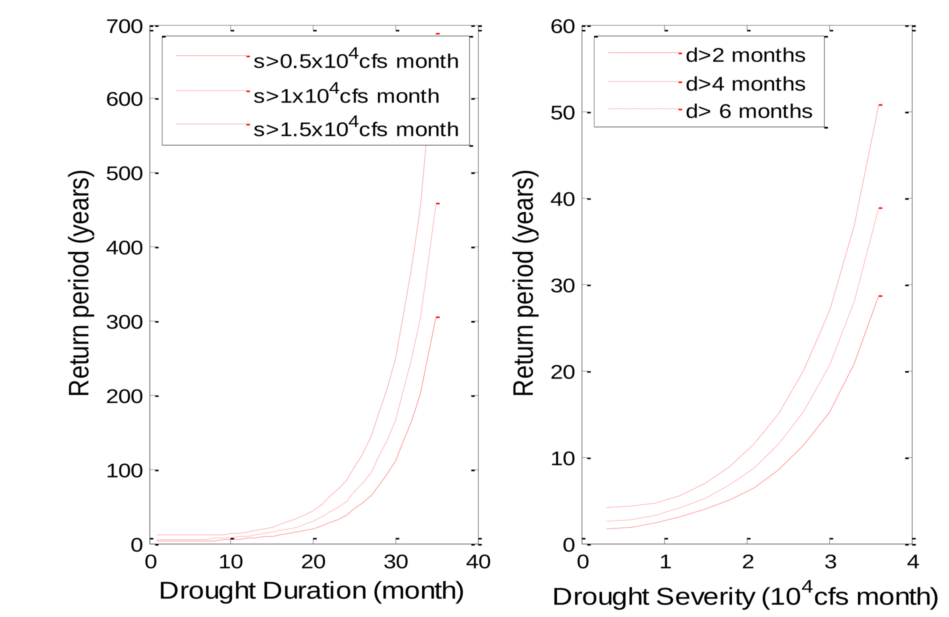

Figure 5 Conditional return periods of drought duration and severity

Conclusions

A bivariate distribution based on entropy theory is proposed to model the drought duration and severity. The advantage of the proposed method is that it is flexible to incorporate different forms of marginal distributions of drought duration and severity. Drought events defined by the monthly streamflow at Brazos River at Waco, Texas are used to illustrate the application of the proposed method for drought analysis. The results show that the proposed method is a useful tool to derive the joint distribution of drought duration and severity for drought analysis.

Future work

Only the mean and logarithmic mean are specified as constraints to derive the joint distribution of drought duration and severity in the current study. Different constraints will be used in the future to constructed the joint distribution and the one with better performance will be selected for the drought analysis.

References

Yevjevich, V. (1967), Objective approach to definitions and investigations of continental hydrologi droughts, Hydrology Paper 23, Colorado State U, Fort Collins.

Jaynes, E. (1957), Information Theory and Statistical Mechanics, Physical Review, 106(4), 620-630.

Shannon, C. E. (1948), A mathematical theory of communications, Bell Syst. Tech. J. , 27(7), 379-423.

Singh, V. P. (1998), Entropy-Based Parameter Estimation in Hydrology, Kluwer Academic Publishers. Netherlands.