Journal of Computational Mathematics, Vol.12, No.4, 1994, 330–338.

THE TRANSFORMED NONPARAMETRIC FLOOD FREQUENCY ANALYSIS *

Kaz Adamowski
(Department of Civil Engineering, University of Ottawa, Ottawa, Canada)
Wojciech Feluch
(Institute of Environmental Engineering, Technical University of Warsaw, Warsaw, Poland)

Abstract

The nonparametric kernel estimation of probability density function (PDF) provides a uniform and accurate estimate of flood frequency-magnitude relationship. However, the kernel estimate has the disadvantage that the smoothing factor h is estimate empirically and is not locally adjusted, thus possibly resulting in deterioration of density estimate when PDF is not smooth and is heavy-tailed. Such a problem can be alleviate by estimating the density of a transformed random variable, and then taking the inverse transform. A new and efficient circular transform is proposed and investigated in this paper.

1. Introduction

Flood magnitude and the corresponding frequency can be estimated from the available date sample by the parametric method whereby various theoretical distributions (i.e., Log-Pearson Type III) are employed. During the past several years parametric modeling has been a subject of intensive investigations by many researchers (Singh, 1986). It is now well recognized that the main problems of parametric procedures are due to the presence of asymmetrical and multimodal densities in the observed flood data. The data also might be of such a type that there is no suitable parametric family that gives a good fit (i.e., the separation effect, Beran et al., 1986) and subsequently will lead to erroneous conclusions.

Many parametric distributions have been recommended for use in hydrology. However, there is no general consensus among hydrologists as to the "best"theoretical frequency distribution for use in flood frequency analysis (Wallis et al., 1985). In order to obtain some degree of uniformity when performing flood analysis (Thomas, 1985), several countries imposed a choice of the procedure (i.e. Log-Peason Type III in A, Generalized Extreme Value in U.K.). In other countries the choice of a distribution

* Received January 14, 1987.

is suggested to be limited to several methods (i.e. Log-Normal Type III, Log-Pearson Type III, Generalized Extreme Value, Wakeby and Weibull Distribution) as is the situation in Canada (Pilon et al., 1985). It is then up to the designer to make the decision as to which method is the most appropriate for a given circumstance. Both of the above administrative recommendations (i.e., imposition of a bass method, or limiting the choice) illustrate a need for a new method which would be uniform, give accurate results, and be suitable for asymmetrical and multimodal distributions.

Under such circumstances, a new nonparametric method was investigated by Adamowski (1985). The nonparametric density estimation does not require assumption of any functional form of density. In fact, very little is assumed, and the assumptions made are mild (Rao, 1984). Adamowski (1985) compared the performance of several parametric and nonparametric estimators and concluded that the nonparametric method is accurate, uniform, and particularly suitable for multimodal data.

The nonparametric method requires a selection of a kernel function $K(.)$, and a smoothing factor h. The choice of a kernel has little effect on the efficiency of the method. Nevertheless, there exists an optimal kernel of Epanechnikov which is in the form of a circular function (Rao, 1983, p.66).

However, the choice of a smoothing factor h plays a crucial role because it affects the bias and variance of the estimator. The optimal choice of a smoothing factor depends on the unknown a priori density and the derivatives of that density. In practical situations since density is unknown, therefore h has to be estimated empirically by various methods (Devroye et al, 1985, p. 191, Adamowski, 1985).

The potential of the nonparametric density estimation is not fully realized in hydrology primarily because of the following two difficulties : a) the value of the smoothing factor h is constant and empirically derived, and b) the nonparametric method places small probability value in the tails of a distribution (thus the extrapolation for return periods exceeding the record length might be influenced too much by the highest observation in the sample).

When the value of h is constant and is not locally adjusted, then the performance of the kernel estimate might deteriorate especially for a density which is not smooth and heavy tailed (skewed). Such problems can be alleviate by the following two modifications, namely a) using the transformation, and b) employing a variable kernel method (Breiman, et al., 1977).

The problem of placing low probability values in the tails of a distribution can be resolved by the introduction of a mixture of parametric and nonparametric methods (Schuster and Yakowitz, 1985).

The purpose of this paper is to introduce the transformed kernel method for estimation of flood frequency and magnitude relationship.