第Ⅱ部門

Suitability of sample size for identifying distribution function in regional frequency analysis

DPRI, Kyoto University, Student Member ○ Binay Kumar MISHRA
DPRI, Kyoto University, Member Yasuto TACHIKAWA
DPRI, Kyoto University, Fellow Kaoru TAKARA

Introduction

Reliable estimation of the hydro-metrological extremes for the given return period is of particular importance in planning and design of hydraulic structures. Frequency analysis relates the magnitude of extreme events to their frequency of occurrence through the use of probability distributions. It needs a large number of historical data at the place of interest. In practice, either there is no data or of very short length data at the point of interest. In such cases, regional frequency analysis can be an effective tool. In regional frequency analysis, delineation of homogeneous region is made with the help of subjective consideration of physiographic and climatic factors as well as with the help of observed historical hydrometric extreme data in the region and finally testing their homogeneity. Once the homogeneous region is delineated, best fitting distribution function representing the regional data is determined from the plot of conventional method of moments or recently developed L-moment skewness versus corresponding kurtosis against various frequency distribution functions. If the data used is of short length, misleading of best fitting distribution function takes place since skewness and kurtosis varies largely with the length of data or sample size. In this paper, it is intended to show the fluctuation pattern of skewness and kurtosis with the change in sample size in terms of root mean square error (RMSE). Annual maximum rainfall data of stations having a length of about 100 years lying in Yodo river basin, Japan have been used for analysis. It was found from the study that the root mean square error (RMSE) in skewness and kurtosis reduces largely as the sample size increases from 20 to 50 and then changes in very small way. Numerically, the average RMSE in skewness and kurtosis fluctuated about or more than 50 % for sample size 20 to less than or about 20% for sample size 50 when compared with that of large sample.

Methodology

Test for the independence and homogeneity of the observed historical annual maximum rainfall data was performed using Mann-Whitney test for further analysis. Next, since the moment coefficients (skewness and kurtosis) get highly deviated with the presence of outliers, Grubb's test was used to detect such outliers. There were no lower outlier found but only higher outlier was found in the considered data. 10 numbers of continuous dataset were selected randomly for each sample size from 20 to 80 at an increment of 10. For each dataset, skewness and kurtosis were determined using conventional method of moments and recently developed L-moment method. Root mean square errors (RMSE) were calculated for each dataset of different sample size by comparing with respective skewness and kurtosis for whole sample.

Discussion

Figure 1 and 2 shows that fluctuation in the kurtosis and skewness is very high for the data set of lower sample size. Reduction in the fluctuation is very rapid when the sample size varies from 20 to 50 and attains a less fluctuation in skewness and kurtosis for the sample size 50 upwards. Fluctuation pattern with the sample size in the case of conventional as well as L-moment approach is similar. Figure 3 shows the plot of skewness versus kurtosis for different data set of various size against the point for large sample size (94 years) at Hirakata, Yodo basin, Japan. For data set of small sample size the points are seen deviated largely from themselves as well as from the point of large sample size. For the data set of sample size 50 or more, the points are seen concentrated in between themselves as well as from the point of whole sample size.

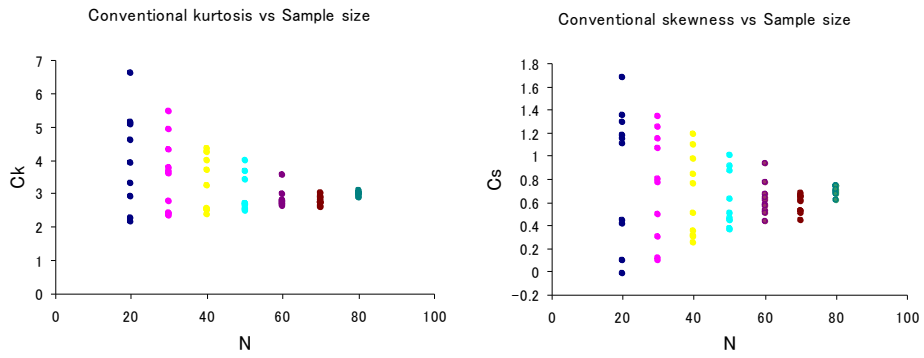

Figure 1: Conventional kurtosis and skewness variation with sample size for different dataset at Hirakata rain-gauge station, Yodo river basin, Japan

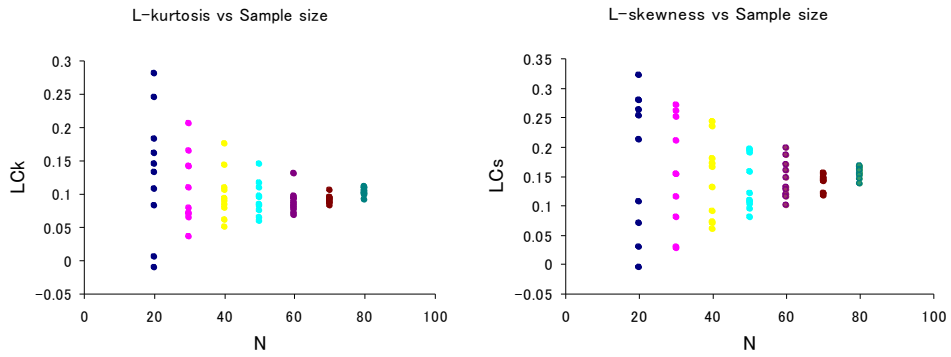

Figure 2: L-kurtosis and L-skewness variation with sample size for different dataset at Hirakata raingauge station, Yodo river basin, Japan

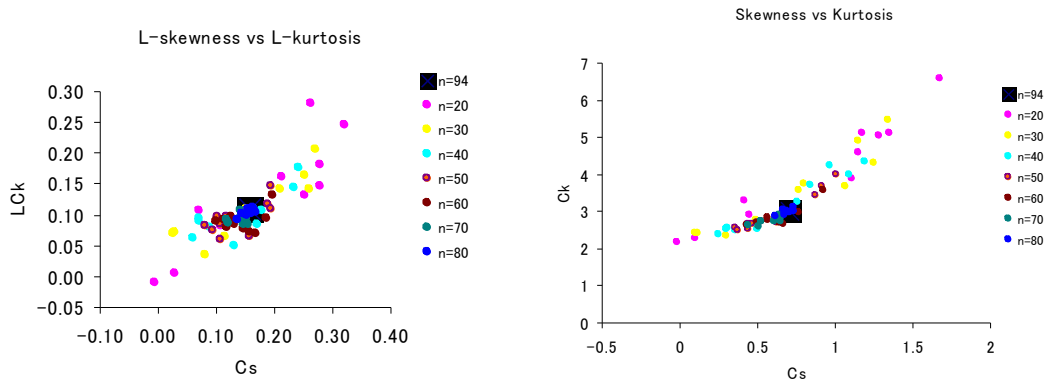

Figure 3: Plot of skewness and kurtosis for different dataset of various sample size at Hirakata raingauge station, Yodo river basin, Japan