Geophysical Research Abstracts
Vol. 12, EGU2010-10967-2, 2010
EGU General Assembly 2010

Analysis of hydrological data in temporal scale using copulas

Takayuki Sugimoto and András Bárdossy

Universtät Stuttgart, Institute of Hydraulic Engineering, Hydrology and Hydrogeology, Stuttgart,, Germany
(takayuki.sugimoto@iws.uni-stuttgart.de)

Global climate change can have the impact on characteristics of rainfall-runoff event and also on the water regime in the catchment. In this context, it is motivated to investigate the existing data to detect the impact of structural modification of the river, change of land use and hydrological regime. The objective of the research is to develop the statistical methodology to identify hydrological characteristics in the catchment, which is transferable for the different catchments. This research is sponsored by The Global Runoff Data Center(GRDC) in Federal Institute of Hydrology in Germany , which fosters multinational and global long-term hydrological studies and offered world wide runoff data collected since 1988. In general, this research is motivated to develop further hydrological methodology and apply it to the global runoff data in order to detect the noticeable information.

For this aim, statistical dependency should be analysed for a large number of hydrological data, which are located different continents. In the application of multivariate statistics, each correlated data are differently skewed and its marginal distributions are not uniformed. However, the statistical measure such as covariance or coefficient can't describe all the details of the dependence structure, subsequently could lose the crucial information.

Copula is a statistical concept, which describes the statistical dependence of correlated data by transforming the original data on the uniformed domain with uniformed marginal distribution. The advantage of copulas is the capability of measuring more detailed dependency and possibly be able to reveal the significant information about dependency from variety of datasets.

A copula is defined as a distribution function on the n dimensional unit cube. All marginal distributions are uniformed on [0,1]. Formally

$$C : [0,1]^n \rightarrow [0,1]$$
$$C\left(\mathrm{u}^{(i)}\right) = u_i \;\; if \;\; \mathrm{u}^{(i)} = (1, \ldots 1, u_i, 1, \ldots 1)$$

Copulas and multivariate distribution are linked to each other by Sklar's theorem[*Skalar*,1959].

$$F\left(\mathrm{x}\right) = C\left(F_{x_1}\left(x_1\right), \ldots, F_{x_n}\left(x_n\right)\right)$$

Where $F_{x_i}\left(x_i\right)$ represents the i-th one dimensional marginal distribution of the multivariate distribution. Copulas can be constructed from distribution functions, as described by [*Nelsen* 1999]:

$$C\left(\mathrm{u}\right) = F\left(F_{x_1}^{-1}\left(u_1\right), \ldots, F_{x_n}^{-1}\left(u_n\right)\right)$$

This copula was applied to the given time series data and its dependence structure as well as its applicability was discussed.

Similar to the concept in time series analysis like autocovariance, the copula can be calculated from the data in the same location with time lag k. This could differentiate the characteristics of the catchment, thus it was tested and further discussed.

Based on the copula, which is calculated by transforming original on the copula domain, measure of asymmetry can be defined for bivariate case [*Jing* 2008]:

$$A\left(t\right) = E\left[\left(F_x\left(x_t\right) - 0.5\right)^2 \cdot \left(F_x\left(x_{t+k}\right) - 0.5\right) + \left(F_x\left(x_t\right) - 0.5\right) \cdot \left(F_x\left(x_{t+k}\right) - 0.5\right)^2\right]$$

Such a measure is useful to compare the copulas constructed from different pairs of data and practical to interpret the meaning of the copula objectively. The measure of asymmetry in copula corresponding to different lag $k$was

computed and plotted on the graph. There seems to be similar behaviour with different intensity. The meaning and its usability was further discussed.