Journal of Hydrology 514 (2014) 123–130

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Multivariate design in the presence of non-stationarity

Jens Bender [a,b,*], Thomas Wahl [b,c], Jürgen Jensen [a,b]

[a] Institute for Water and Environment, University of Siegen, Paul-Bonatz-Str. 9-11, 57076 Siegen, Germany
[b] Research Centre Siegen, University of Siegen, Weidenauer Str. 167, 57076 Siegen, Germany
[c] University of South Florida, College of Marine Science, 140 7th Avenue South, St. Petersburg, FL 33701, USA

ARTICLE INFO

Article history:
Received 22 January 2014
Received in revised form 1 April 2014
Accepted 8 April 2014
Available online 18 April 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Niko Verhoest, Associate Editor

Keywords:
Copulas
Non-stationarity
Bivariate model

SUMMARY

Over the last decade the number of applications of copula functions for multidimensional modeling of hydrological parameters has significantly increased. However, most of the studies assume stationarity in the marginal distribution parameters as well as in the dependence structure of the variables. This is because the available time series are often too short for using a non-stationary multivariate model. In this study we analyze the joint probability of flood peak and volume based on a discharge time series of the Rhine River providing 191 years of data. We find significant positive trends in the marginal distribution parameters as well as in the dependence measure from analyzing 50-year moving time windows. Fitting time dependent marginal distributions and time dependent copulas to the data sets, and comparing the results with the stationary approach, shows the influence of the non-stationary behavior of the variables. The results are illustrated by calculating the joint probability of the flood peak and volume for four cases: i. considering all parameters as time dependent, i.e. the location, scale and shape parameter of the marginals and the copula parameter, ii. considering the location and scale parameter of the marginals and the copula parameter as time dependent, iii. considering the location parameter of the marginals and the copula parameter as time dependent, and iv. considering only the copula parameter as time dependent. The results highlight that the joint probability, illustrated by the isoline of a given exceedance probability, varies significantly over time when non-stationary models are applied.

1. Introduction

Over the last years copula functions have been increasingly used for various multivariate hydrological analyzes. They were applied for rainfall frequency analysis (e.g. De Michele and Salvadori, 2003; Grimaldi and Serinaldi, 2006), flood frequency analysis considering peak flow and flood volume (e.g. Favre et al., 2004; Karmakar and Simonovic, 2009), drought frequency analysis (e.g. Shiau, 2006; Kao and Govindaraju, 2010), storm surge modeling (e.g. Wahl et al., 2012; Corbella and Stretch, 2013; Zhong et al., 2013), and for several other multivariate problems. The main advantage of copulas over other multivariate models is related to the fact that the dependency between variables can be modeled separately from their marginal distributions. Furthermore, copulas, especially those belonging to the Archimedean family, are relatively easy to construct and capable of modeling a broad range of dependence structures. To the authors' knowledge most of the hydrological studies where copulas were applied assumed stationarity of the parameters of the marginal distributions as well as the dependence structure. This assumption, however, might not hold and non-stationary behavior of one or more marginal parameters and/or the interdependence between the variables of interest may influence the results of the multivariate statistical analysis. In the univariate case, non-stationary extreme value models have been widely applied. A general and comprehensive introduction to the topic was given by Coles (2001). Katz et al. (2002), for example, applied non-stationary extreme value models to precipitation and discharge time series. Khaliq et al. (2006) reviewed several methods of modeling non-stationary hydro-meteorologic time series, and Mudersbach and Jensen (2010) derived future coastal design water levels using a non-stationary approach.

The necessity of applying non-stationary extreme value approaches is primarily due to trends in the time series that have to be modeled. Several authors already reported significant trends in the rainfall pattern over Europe (e.g. Moberg and Jones, 2005) and in the annual maximum discharges of major German rivers (e.g. Bormann et al., 2011), one of which (the Rhine River) is investigated here. Whether the trends reported here are related to climate change, land use changes within the catchments or

* Corresponding author at: Institute for Water and Environment, University of Siegen, Paul-Bonatz-Str. 9-11, 57076 Siegen, Germany. Tel.: +49 271 740 3401.
E-mail address: jens.bender@uni-siegen.de (J. Bender).

http://dx.doi.org/10.1016/j.jhydrol.2014.04.017
0022-1694/

in-stream river engineering is not further investigated as it is not relevant for the purpose of the study.

Although various authors (e.g. Chebana et al., 2013) assessed the dependence of different hydrologic parameters over time, the number of papers dealing with non-stationary multivariate extreme value models is rather sparse. Zhang (2005) first considered non-stationary marginal distributions in a conditional multivariate model based on copulas. The dependence structure was assumed to be constant over time. Corbella and Stretch (2013) also applied non-stationary marginal distributions combined with conditional copula functions with constant dependence measures. Chebana et al. (2013) first mentioned the idea of using copula functions with time dependent parameters in case of a changing dependence structure between the investigated variables.

There are several possible explanations that only few studies up to now were concerned with multivariate non-stationary statistics. From a practical point of view, time series are often too short to set up a non-stationary multivariate model. Furthermore, there are also still some mathematical issues that need to be solved; how can we handle for example the problem that the marginal distribution family may vary over time? However, when designing hydraulic structures often more than one parameter has to be taken into account, and having the anticipated lifetime of the structure in mind one is not only interested in reliable present-day design values, but also in estimating potential future changes in these values.

Here, we combine the univariate non-stationary extreme value analysis with a novel approach to model changes in the dependence of two variables over the time. The main objective is to outline how bivariate design parameters may change if non-stationarities are existent and taken into account in the univariate and bivariate statistical models. We use a time series of mean daily discharge of the Rhine River comprising 191 years of data measured at gauge Worms in Germany. We model the annual maxima of the flood discharge and the corresponding flood volumes of the direct runoff.

The remainder of the paper is organized as follows: In Section 2 we give a short introduction to the copula theory and to univariate and bivariate non-stationary extreme value analysis. In Section 3 the data set for the case study is introduced, before the results are presented in Section 4 and discussed in Section 5.

2. Methods

2.1. General copula theory

Since copula functions have been widely used over the last years in hydrology and a corresponding number of papers has been published, we only provide a short and basic introduction to the theoretical background of copula functions. More information can be found for example in Nelsen (2006) and Salvadori et al. (2007).

Copulas are flexible joint distributions for modeling the dependence structure of two or more random variables. First mentioned by Sklar (1959), the joint behavior of two (or more) random variables X and Y with continuous marginal distributions $u = F_X(x) = P(X \leqslant x)$ and $v = F_Y(y) = P(Y \leqslant y)$ can be described uniquely by an associated dependence function or copula function C. In the bivariate case, the relationship between all $(u,v) \in (0,1)^2$ can be written as:

$$F_{X,Y}(x,y) = C[F_X(x), F_Y(y)] = C(u, v) \tag{1}$$

where $F_{X,Y}(x,y)$ is the joint cumulative distribution function (cdf) of the random variables X and Y.

A copula function with a strictly monotonically decreasing generator function $\varphi: (0,1) \to (0,\infty)$ with $\varphi(1) = 0$ belongs to the Archimedean copula family. The general form of one-parametric Archimedean copulas is

$$C_\theta(u, v) = \varphi^{-1}[\varphi(u) + \varphi(v)] \tag{2}$$

where θ denotes the copula parameter. In this study we use three Archimedean copulas, namely the Clayton, Frank, and Gumbel copulas. They are relatively easy to construct, flexible, and capable of modeling the full range of tail dependencies. The Clayton copula has lower tail dependence, the Frank copula has no tail dependence, and the Gumbel copula has strong upper tail dependence (e.g. Nelsen, 2006).

In multivariate extreme value analysis one is often concerned with the problem of choosing an adequate design event out of a large number of possible parameter combinations, since all data couples (u,v) on the same probability level have the same bivariate probability of exceedance. However, some combinations of a given probability are, at least in theory, more likely than others. Therefore the relevant design event can be selected as the point with the largest joint probability density on the probability-isoline as outlined for example by Gräler et al. (2013):

$$(u, v) = \underset{C_{UV}(u,v)=k}{\operatorname{argmax}} f_{XY}(F_X^{-1}(u), F_Y^{-1}(v)). \tag{3}$$

The resulting design values (x,y) can then easily be calculated using the inverse of the cumulative distribution functions of the marginals:

$$x = F_X^{-1}(u) \text{ and } y = F_Y^{-1}(v) \tag{4}$$

For practical applications, such as reservoir design, there may exist more appropriate approaches for selecting the relevant design event. Process-based models can be applied to simulate the system's response to several combinations of the design relevant variables and the cost-benefit ratio can be maximized. Here, given the exemplary character of the study, we will outline the temporal changes of the bivariate design values of X and Y using the most likely event.

Furthermore, in hydrologic and hydraulic applications one is mostly interested in the mean interarrival time between two design events, usually given in years and also known as return period (RP). In a univariate context the return period is commonly defined as:

$$T = \frac{\mu_T}{1 - F_X(x)} \tag{5}$$

where μ_T denotes the mean interarrival time (typically given in years; when using annual maxima values, μ_T equals 1 year). $F_X(x)$ represents the cumulative distribution function of the univariate variable X.

In the multivariate domain, however, it is still discussed by the community which method is most suitable to transform the joint exceedance probability to a multivariate joint return period (JRP). Here, we follow the approach introduced by Salvadori et al. (2007). The JRP is based on the bivariate cdf $F_{XY}(x,y)$, expressed by the bivariate copula function $C_{UV}(u,v)$ (Eq. (1)) and the cdf's of the marginals. This JRP is often denoted as AND-joint return period ($T^\wedge$) since it is based on the probability that X and Y exceed the values x and y, respectively: $P(X > x \wedge Y > y)$. Hence, the bivariate AND-JRP $T^\wedge$ can be written as:

$$T^\wedge_{X,Y} = \frac{\mu_T}{P(X \geqslant x \wedge Y \geqslant y)} = \frac{\mu_T}{1 - F_X(x) - F_Y(y) + C[F_X(x), F_Y(y)]} = \frac{\mu_T}{1 - F_X(x) - F_Y(y) + C(u, v)} \tag{6}$$

A comprehensive overview about several other available methods to estimate the JRP is given e.g. in Volpi and Fiori (2014), De Michele et al. (2013), Gräler et al. (2013), and references therein.