**Nonlinear Processes
in Geophysics**

Comment on "Multivariate non-normally distributed random variables in climate research – introduction to the copula approach" by C. Schölzel and P. Friederichs, Nonlin. Processes Geophys., 15, 761–772, 2008

F. Serinaldi

Dipartimento di Idraulica Trasporti e Strade, "Sapienza" Università di Roma, Via Eudossiana 18, 00184 Rome, Italy

Honors Center of Italian Universities-H^{2}CU, "Sapienza" Università di Roma, Via Eudossiana 18, 00184 Rome, Italy

Schölzel and Friederichs (2008) provide an overview of the copula approach for use in meteorological and climate research. As stated by the authors, the field is vast and growing fast, with a comprehensive review rather difficult at the present stage. However, the work lacks some important references which would provide methods and tools to tackle some methodological problems pointed out by the authors, as well as some applications to meteorological and geophysical variables. Some of these studies had already been published or were in preview at the time of submission of the paper by Schölzel and Friederichs (2008). These references may be very useful to researchers who want to introduce copulas in their analyses exploiting the most recent results.

In Sect. 2, the authors cite six works dealing with copula in hydrology and meteorology. However, in the last three years, about 40 papers on multivariate environmental analyses using copulas have been published. An extensive bibliography up to 2007 can be found in Salvadori et al. (2007), and an update to 2008 is given by Villarini et al. (2008). Since copulas are already quite well introduced in geosciences it can be helpful to consult such literature as well as financial and econometric ones.

In Sect. 3, the authors state that copula approach can be easily extended to a mixture of discrete and continuous random variables. This extension requires caution, which is addressed by Genest and Nešlehová (2007). The authors then discuss the paucity of canonical ways to define correlations among random variables. In-depth discussions on dependence concepts, copulas, and other issues on relations among random variables are provided by Joe (1997) and Drouet-Mari and Kotz (2001).

In Sect. 4, and then in Sect. 8, the authors say that there is not a general procedure for copula selection. As stated by the authors in the conclusion, this is also a common problem in univariate frequency analysis. Theoretical reasons (e.g., physical background) should guide the choice, but practical aims (agreement to data according to some fitting criterion, or reproduction of some important properties of the sample) should be accounted for as well. Nikoloulopoulos and Karlis (2008) provide an alternative frequentist model averaging method of selection, and dos Santos Silva and Freitas Lopes (2008) give an example of Bayesian approach.

Helpful references on elliptical copulas presented in Sect. 4 are Fang et al. (2002) and Genest et al. (2007). A possible solution (perhaps partial, but feasible) of the problem of radial symmetry pointed out by the authors is the v-copula proposed by Bárdossy (2006), which preserves the ability to model data with pairwise different degrees of association (such as meta-elliptical copulas), and also shows asymmetric tail behavior. At the end of Sect. 4.1.2, the authors discuss tail dependence and rate of convergence as measures of tail behavior. Poulin et al. (2007) describe the impact of tail dependence on extreme quantiles of conditional distributions that can arise in some environmental studies.

In Sect. 4.2, the authors state that the interpretation of the parameters of Archimedean copulas in terms of statistical properties is not straight forward. We do not completely agree as 2-copulas are related to scale-free rank-based measures of association, such as Kendall's τ_K and Spearman's ρ_S, leading to relationships between such indices (and lower

(upper) tail dependence coefficient λ_L (λ_U)) and parameters of several Archimedean 2-copulas (e.g., Joe, 1997; Nikoloulopoulos and Karlis, 2008).

Section 5 deals with estimation and goodness-of-fit tests. A simple way (not mentioned in the paper) to estimate copula parameters for several families is the moment-like method based on rank-based measures of association (e.g., Genest et al., 2007). Among goodness-of-fit procedures, tests proposed by Genest and Rémillard (2008) and Genest et al. (2009) should be considered. These tests are based on Kolmogorov-Smirnov and Cramer-von Mises statistics computed on values of copulas or on their distribution functions. They do not require probability integral transformations and resort to bootstrap approach to compute critical values.

As far as copulas and extreme value theory are concerned (Sect. 6), Salvadori et al. (2007) thoroughly describe the methodology and inference procedures. They give several environmental and geophysical examples by deeply discussing univariate and multivariate extreme value theory, copula approach, and the relationships between these two research areas.

In the first example presented in Sect. 7, the authors fit a copula to non-zero five-day averages of precipitation recorded at two sites. Similar application is described by Serinaldi (2008, 2009) and Villarini et al. (2008), where a copula-based mixed model for zero-inflated data (such as rainfall records from two stations) as well as a study on λ_U and τ_K for rainfall series at daily and finer time scales are discussed.

In the same section, the authors state that "The asymmetric tail behavior also excludes regular correlation coefficients as a proper measurement of dependence". However, the rationale of τ_K or ρ_S is to describe the overall strength of monotonic association, and not to explicitly account for asymmetric tail behavior. Other measures such as tail dependence and quadrant dependence (e.g., Joe, 1997) can be used when other association aspects are of interest.

Furthermore, we do not completely agree on the statement: "copula does not provide a practical measurement of tail behavior of the original random variables, but upper/lower tail dependence of the transformed variables has to be considered when fitting a copula model". Pairwise upper (lower) tail dependence can be visualized as the accumulation of points in the upper-right (lower-left) corner of the plot of generic pairs of pseudo-observations $(\hat{F}_{X_i}(x_i), \hat{F}_{X_j}(x_j))$, where $\hat{F}_X$ denotes the empirical distribution function, without introducing any further transformation, and λ_L, λ_U, τ_K, and ρ_S are insensitive to monotonic marginal transformations (as stated by the authors at the end of Sect. 4.1.2). Hence a first selection among proper parametric copula families is possible in spite of the marginal choice (e.g., Michiels and De Schepper, 2008; Serinaldi, 2008; Villarini et al., 2008; Serinaldi et al., 2009). Then, a formal test can be applied to refine the selection. These procedures do not require marginal transformations, as only the original variables and the pseudo-observations obtained by empirical marginal distributions are needed.

References

Bárdossy, A.: Copula-based geostatistical models for groundwater quality parameters, Water Resour. Res., 42, W11416, doi: 10.1029/2005WR004754, 2006.

dos Santos Silva, R. and Freitas Lopes, H.: Copula, marginal distributions and model selection, Stat. Comput., 18, 313–320, 2008.

Drouet-Mari, D. and Kotz, S.: Correlation and dependence, Imperial College Press, London, 2001.

Fang, H. B., Fang, K. T., and Kotz, S.: The meta-elliptical distributions with given marginals, J. Multivariate Anal., 82, 1–16, 2002.

Genest, C. and Nešlehová, J.: A primer on copulas for count data, The Astin Bulletin, 37, 475–515, 2007.

Genest, C. and Rémillard, B.: Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, Ann. I. H. Poincare B, 44, 1096–1127, 2008.

Genest, C., Favre, A.-C., Béliveau, J., and Jacques, C.: Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data, Water Resour. Res., 43, W09401, doi: 10.1029/2006WR005275, 2007.

Genest, C., Rémillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study, Insur. Math. Econ., 44, 199–213, 2009.

Joe, H.: Multivariate models and dependence concepts, Chapman & Hall, New York, 1997.

Michiels, F. and De Schepper, A.: A copula test space model how to avoid the wrong copula choice, Kybernetika, 44, 864–878, 2008.

Nikoloulopoulos, A. K. and Karlis, D.: Fitting copulas to bivariate earthquake data: The seismic gap hypothesis revisited, Environmetrics, 19, 251–269, 2008.

Poulin, A., Huard, D., Favre, A.-C., and Pugin, S.: Importance of tail dependence in bivariate frequency analysis, J. Hydrol. Eng., 12, 394–403, 2007.

Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R.: Extremes in nature: An approach using copulas, Springer, 2007.

Schölzel, C. and Friederichs, P.: Multivariate non-normally distributed random variables in climate research – introduction to the copula approach, Nonlin. Processes Geophys., 15, 761–772, 2008, http://www.nonlin-processes-geophys.net/15/761/2008/.

Serinaldi, F.: Analysis of inter-gauge dependence by Kendall's τ_K, upper tail dependence coefficient, and 2-copulas with application to rainfall fields, Stoch. Env. Res. Risk A., 22, 671–688, 2008.

Serinaldi, F.: Copula-based mixed models for bivariate rainfall data: An empirical study in regression perspective, Stoch. Env. Res. Risk A., doi:10.1007/s00477-008-0249-z, in press, 2009.

Serinaldi, F., Bonaccorso, B., Cancelliere, A., and Grimaldi, S.: Probabilistic characterization of drought properties through copulas, Phys. Chem. Earth, doi:10.1016/j.pce.2008.09.004, in press, 2009.

Villarini, G., Serinaldi, F., and Krajewski, W. F.: Modeling radar-rainfall estimation uncertainties using parametric and nonparametric approaches, Adv. Water Resour., 31, 1674–1686, 2008.