Interactive comment on “Copulas for hydroclimatic applications – A practical note on common misconceptions and pitfalls” by Faranak Tootoonchi, Jan Olaf Haerter, Olle Räty, Thomas Grabs, Mojtaba Sadegh, and Claudia Teutschbein

Ioannis Tsoukalas (e-mail: itsoukal@mail.ntua.gr)

Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, 15780 Zographou, Greece

Introduction

This is a particularly interesting work that concerns a very active topic of research in the hydrological domain (and beyond). Below there are a few comments that I hope the Authors might find useful, aiming to improve the quality of the manuscript, as well as better highlight some common misconceptions and pitfalls that regard particularly the case of Gaussian copula.

Comments

1. **L88-89**. The Authors write: “*Since the early 2000’s, copula methods have been adopted in hydrological modeling, which was triggered by the study of Salvadori and De Michele (2010).*”

   With above sentence in mind I would like to bring to the Authors attention the works of Favre et al. (2004) and Salvadori and De Michele (2004), which if I am not mistaken are the first applications of copulas in hydrological domain (chronologically preceding the one already mentioned in the manuscript).

2. **L90-101**. In this paragraph the Authors mention numerous works that have used the notion of copulas for the development of various methods in hydrological domain. In this extent I think that it is useful to mention that copulas have also been used for the generation of synthetic hydroclicmatic data, such as synthetic time series of rainfall, runoff, etc. (an important task required by many uncertainty-aware methods/models driven by stochastically-generated data). As in the case of random variables and multivariate distributions, also in this case copulas offer the necessary flexibility for modelling/simulation of non-Gaussian processes. For instance see the works of Lee and Salas (2011), Chen et al. (2015) and Hao and Singh (2013), as well as recent approaches in hydrological domain, based on the Gaussian copula (a construct related with the Nataf’s joint distribution; see Lebrun and Dutfoy (2009), and references below for a discussion in a hydrological context) that allow the parsimonious simulation of multivariate stationary and cyclostationary processes with any marginal distribution and correlation structure (Kossieris et al., 2019; Tsoukalas et al., 2018a, 2018b, 2020) - also in a multi-scale/disaggregation context (Tsoukalas et al., 2019). In addition, you might also find useful another recent work ours (i.e., Tsoukalas et al., 2018c), with emphasis on section 4, where we highlight an unwanted effect on the established dependence patterns of classical synthetic data generation methods (e.g., the Thomas and Fiering approach) that lack an explicit assumption about joint distribution, hence copula, of the process.

3. **L204-205**. With reference to Elliptical copula (i.e., the Gaussian and Student-t copula), the Authors write: “*Later, Aas (2004) showed that the co-dependence structure for Elliptical copulas can be presented by linear (Pearson) correlation. Correspondingly, their copula parameter θ can*

either be estimated as being equal to the linear (Pearson) correlation or derived from Kendall' τ or Spearman's S. For more details on the corresponding equations, we refer the readers to Aas (2004)."

Indeed there are relationships that *link* the Pearson's correlation coefficient with Kendall's and Spearman's rank-based correlation coefficients, yet as highlighted in Tsoukalas et al. (2018b), section 3.2.3, these are valid if and only if both the marginals, and the copula are Gaussian (see also references therein).

When the copula is Gaussian, and the marginals are not (which is typical in hydrology), these relationships are no longer valid. In fact, in such cases, the Pearson's correlation coefficient depends on the marginals; since it involves the first cross-product moment among the variables (i.e., it involves the term E[X1, X2]), while the Kendall's and Spearman's correlations do not (since they are rank-based measures of dependence). In the case of Gaussian copula and non-Gaussian marginal, there is a non-analytical relationship that links the Pearson's correlation coefficient in Gaussian (in the manuscript's notation, the Gaussian copula parameter θ) and target domain that has to be found by resolving of a double infinite integral. In particular, and with reference to hydrological domain, see Tsoukalas et al. (2018a, 2018b, 2019, 2020) and references therein.

In my view, the above are delicate, often neglected, points that concern the Gaussian copula, and therefore should be made clear in the manuscript, since they are both (very) common misconceptions/pitfalls that concerns the later (widely-used) copula.

4. **L310-318**. In this paragraph, as well as in other parts of the manuscript, the Authors discuss the debate between stationarity and non-stationarity. On this topic, and beyond the work of Lins and Cohn (2011), already cited in the manuscript, my suggestion to the Authors would be to review, (and cite if it is considered appropriate), the recent works of, Serinaldi et al. (2018), with emphasis on section 4.2, Koutsoyiannis and Montanari, (2007), (2015), Lins and Cohn (2011), Matalas (2012), and Montanari and Koutsoyiannis (2014). All these works discuss the importance of the assumption of stationarity, highlighting that it is an essential tool for inferencing from data (e.g., model fitting). See also the very interesting, note of Harry F. Lins[1], which concludes as follows:

$$\text{Stationarity} \neq \text{static}$$

$$\text{Non-stationarity} \neq \text{change (or trend)}$$

In my view, stationarity should not be viewed as a shortcoming, nor considered dead. It is recalled that non-stationarity implies non-ergodicity, which in turn makes inference from observed data impossible, unless of course the deterministic dynamics of the process (and hence potential *change*) are known; which in my understanding, is never the case in hydrological sciences.

Regards,

Ioannis Tsoukalas

[1] http://www.wmo.int/pages/prog/hwrp/chy/chy14/documents/ms/Stationarity_and_Nonstationarity.pdf

References

Chen, L., Singh, V.P., Guo, S., Zhou, J., Zhang, J., 2015. Copula-based method for multisite monthly and daily streamflow simulation. J. Hydrol. 528, 369–384. https://doi.org/10.1016/j.jhydrol.2015.05.018

Favre, A., El Adlouni, S., Perreault, L., Thiémonge, N., Bobée, B., 2004. Multivariate hydrological frequency analysis using copulas. Water Resour. Res. 40. https://doi.org/10.1029/2003WR002456

Hao, Z., Singh, V.P., 2013. Modeling multisite streamflow dependence with maximum entropy copula. Water Resour. Res. 49, 7139–7143. https://doi.org/10.1002/wrcr.20523

Kossieris, P., Tsoukalas, I., Makropoulos, C., Savic, D., 2019. Simulating Marginal and Dependence Behaviour of Water Demand Processes at Any Fine Time Scale. Water 11, 885. https://doi.org/10.3390/w11050885

Koutsoyiannis, D., Montanari, A., 2015. Negligent killing of scientific concepts: the stationarity case. Hydrol. Sci. J. 60, 1174–1183. https://doi.org/10.1080/02626667.2014.959959

Koutsoyiannis, D., Montanari, A., 2007. Statistical analysis of hydroclimatic time series: Uncertainty and insights. Water Resour. Res. 43, 1–9. https://doi.org/10.1029/2006WR005592

Lebrun, R., Dutfoy, A., 2009. An innovating analysis of the Nataf transformation from the copula viewpoint. Probabilistic Eng. Mech. 24, 312–320. https://doi.org/10.1016/j.probengmech.2008.08.001

Lee, T., Salas, J.D., 2011. Copula-based stochastic simulation of hydrological data applied to Nile River flows. Hydrol. Res. 42, 318–330. https://doi.org/10.2166/nh.2011.085

Lins, H.F., Cohn, T.A., 2011. Stationarity: Wanted dead or alive? J. Am. Water Resour. Assoc. https://doi.org/10.1111/j.1752-1688.2011.00542.x

Matalas, N.C., 2012. Comment on the Announced Death of Stationarity. J. Water Resour. Plan. Manag. 138, 311–312. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000215

Montanari, A., Koutsoyiannis, D., 2014. Modeling and mitigating natural hazards: Stationarity is immortal! Water Resour. Res. 50, 9748–9756. https://doi.org/10.1002/2014WR016092

Salvadori, G., De Michele, C., 2004. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. Water Resour. Res. 40. https://doi.org/10.1029/2004WR003133

Serinaldi, F., Kilsby, C.G., Lombardo, F., 2018. Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. Adv. Water Resour. https://doi.org/10.1016/j.advwatres.2017.10.015

Tsoukalas, I., Efstratiadis, A., Makropoulos, C., 2019. Building a puzzle to solve a riddle: A multi-scale disaggregation approach for multivariate stochastic processes with any marginal distribution and correlation structure. J. Hydrol. 575, 354–380. https://doi.org/10.1016/j.jhydrol.2019.05.017

Tsoukalas, I., Efstratiadis, A., Makropoulos, C., 2018a. Stochastic Periodic Autoregressive to Anything (SPARTA): Modeling and Simulation of Cyclostationary Processes With Arbitrary Marginal Distributions. Water Resour. Res. 54, 161–185. https://doi.org/10.1002/2017WR021394

Tsoukalas, I., Kossieris, P., Makropoulos, C., 2020. Simulation of Non-Gaussian Correlated Random Variables, Stochastic Processes and Random Fields: Introducing the anySim R-Package for Environmental Applications and Beyond. Water 12, 1645. https://doi.org/10.3390/w12061645

Tsoukalas, I., Makropoulos, C., Koutsoyiannis, D., 2018b. Simulation of Stochastic Processes Exhibiting Any-Range Dependence and Arbitrary Marginal Distributions. Water Resour. Res. 54, 9484–9513. https://doi.org/10.1029/2017WR022462

Tsoukalas, I., Papalexiou, S., Efstratiadis, A., Makropoulos, C., 2018c. A Cautionary Note on the Reproduction of Dependencies through Linear Stochastic Models with Non-Gaussian White Noise. Water 10, 771. https://doi.org/10.3390/w10060771