Bivariate Copula Models for Dependence: Application to Wind Speed Data

Lídia M. André
Faculdade de Ciências da Universidade de Lisboa, lidiamandre@gmail.com

P. de Zea Bermudez
Faculdade de Ciências da Universidade de Lisboa and CEAUL, pcbermudez@fc.ul.pt

Palavras–chave: Dependence, Copulas, Wind speed, Bayesian approach

Resumo: Pearson's correlation coefficient is used for quantifying the intensity of linear associations between variables. Alternatively, non-parametric correlation coefficients, such as Spearman's and Kendall's, are also commonly used. However, in real data applications, the associations between variables are frequently non-linear. In these situations, it would be most interesting to fit a bivariate model to the random variables which describes other types of dependence structures. The traditional Gaussian bivariate model does not generally describe the complexity of the association between the variables. Gaussian models can hardly be adequate to model data which show, e.g., strong asymmetries or heavy tails.

Copulas allow the description of the joint dependence of several variables. In a bivariate framework, they enable to create the joint distribution for the pair of random variables independently of their marginal distributions ([5]). Copulas have been widely used in several areas, such as insurance, banking, finance and climate ([1], [2] and [5]). Basically, in a bivariate context, a copula connects the joint distribution function of the random pair (X, Y) to the two univariate distribution functions. That is, if X and Y are random variables with joint distribution function $H(x, y) = P(X \leq x, Y \leq y)$ and marginal distribution functions $F(x) = P(X \leq x)$ and $G(y) = P(Y \leq y)$, then there exists a copula C such that

$$H(x, y) = C(F(x), G(y)).$$

Moreover, C is unique if the random variables are continuous. This result is the basis of the copula theory and is due to Sklar ([6]).

This work uses copulas to analyse the dependence between daily maximum wind speeds (Km/h) observed in 40 stations spread out in the continental part of Portugal from 2000 up to 2012, X, and simulated wind speeds produced by a simulator, at a regular grid of $81km^2$ grid cell size, Y. One of the most relevant features of the observed data is the extremely high proportion of missing observations, which reaches 90% in some stations. Consequently, only 40 stations, out of a total of 117, were considered (the ones with less than 30% of NAs). One of the major benefits of using the simulated data is that it has no missing values. The problem is that the simulated and the observed daily maximum wind speeds, in some stations, do not match well and tend to differ, mostly in the right tail. Consequently it is very

important to understand the dependence between X and Y.

The observed and simulated data will be analysed by season. In every season, only every 5 observations will be considered in order to overcome the short-term dependence that exists within the series. Several copulas, namely Gumbel and t-Student ([3] and [4]), will be fitted to the observed and to the simulated data. Observations will be simulated from the fitted copulas and compared to the real and simulated wind speeds. Moreover, estimates of the dependence of joint, mild, strong and very strong wind speeds will be presented. Finally, the classical modelling will be briefly compared with a Bayesian approach.

Agradecimentos: The first author was partially supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the project PTDC/MAT-STA/28649/2017.
The second author thanks FCT through the projects UID/MAT/00006/2019 and PTDC/MAT-STA/28649/2017 for financial support.

Both authors thank IPMA for allowing the use of the observed data and Instituto Dom Luiz, Universidade de Lisboa, for providing the simulated data.

Bibliografia

[1] Díaz, M. G. Bayesian prediction of glacial discharge in Antartida using copulas. *PhD Thesis*, Universidad Carlos III, Madrid, Spain, 2017.

[2] Embrechts, P., Lindskog, F. and McNeil, A. Modelling Dependence with Copulas and Applications to Risk Management. *Handbook of Heavy Tailed Distribution in Finance.*, S. T. Rachev, Elsevier/North Holland, Amsterdam, 2013.

[3] Joe, H. *Dependence Modeling with Copulas*. 1st Edition, Chapman & Hall/CRC Monographs on Statistics and Applied Probability, 2014.

[4] Nelson, R .B. *An Introduction to Copulas*. Springer Science+Business Media, New York, NY, 2nd Edition, 2006.

[5] Shemyakin, A. and Kniazev, A. *Introduction to Bayesian Estimation and Copula Models of Dependence*. John Wiley & Sons, New York, NY, 2017.

[6] Sklar, A. Fonctions de répartition n dimensions et leurs merges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229–231, 1959.