Exploring the joint probability of precipitation and soil moisture over Europe using copulas

Carmelo Cammalleri[1], Carlo De Michele[1], Andrea Toreti[2]

[1]Dipartimento di Ingegneria Civile e Ambientale (DICA), Politecnico di Milano, Milan, 20133, Italy.
[2]European Commission, Joint Research Centre (JRC), Ispra, 21027, Italy.

Correspondence to: Carmelo Cammalleri (carmelo.cammalleri@polimi.it)

Abstract. The joint probability of precipitation and soil moisture is here investigated over Europe with the goal to extrapolate~~d~~ meaningful insights on the potential joint use of these variables for the detection of agricultural droughts within a multivariate probabilistic modeling framework. The use of copulas is explored, being the framework ~~as a parametric approach~~ often used in hydrological studies for the analysis of bivariate distributions. The analysis is performed for the period 1996-2020 on the empirical~~ly~~ frequencies derived from ERA5 precipitation and LISFLOOD soil moisture datasets, both available as part of the Copernicus European Drought Observatory. The results show an overall good correlation between the ~~empirical frequency series derived from the~~ two standardized ~~datasets~~ series (Kendall's $\tau = 0.42 \pm 0.1$), but also clear spatial patterns in the tail-dependence derived with both non-parametric and parametric approaches. About half of the domain shows symmetric tail-dependence~~s~~, well reproduced by the Student-t copula;~~,~~ whereas the rest of the domain is almost equally split between low and high tail-dependences (both modeled with the Gumbel family of copulas). These spatial patterns are reasonably reproduced by a random forest classifier, suggesting that this outcome is not driven by chance. This study stresses how a joint use of standardized precipitation and soil moisture for agriculture drought characterization may be~~ more~~ beneficial in areas with strong low tail-dependence, and how this behavior should be carefully considered in multivariate drought studies.

1. Introduction

Agricultural drought, defined as a condition of unusually high precipitation shortages and/or soil water deficits causing adverse effects on crop yields and production (Panu and Sharma, 2002), is probably the most recognized of the four main drought types or phases (Wilhite and Glantz, 1985). This is mainly due to the more direct and easier to understand impacts compared to the other types of droughts (Mishra and Singh, 2010). The scientific literature on agricultural drought ~~has produced~~provides a ~~very~~ large ~~number~~ variety of indices (WMO and GWP, 2016), with the aim of reproducing the temporal dynamics ~~and the effects~~ of crop water deficit through a combination of climatic observations, hydrological modeling, and remote sensing data (Zargar et al., 2011).

The difficulty in capturing the multi-facet nature of agricultural drought events across the world with a single ~~index~~ approach (Sivakumar et al., 2011) is confirmed by the absence of consensus in the scientific literature on the most reliable agricultural drought index. ~~However, d~~Despite the large range of available indices, some common characteristics can be identified, such as the focus on some proxy variables of plant water availability – through soil moisture (Dutra et al., 2008), actual evapotranspiration (Anderson et al., 2011) or basic meteorological information (Vicente-Serrano et al., 2010) – and the need to account for deviations from long-term conditions (i.e., use of standardized anomalies~~scores~~).

Meteorological drought indicators computed on appropriate aggregation time scales (McKee et al., 1993; Vicente-Serrano et al., 2010) have demonstrated a good capability to represent agricultural drought conditions in several case studies (e.g., Bachmair et al., 2018; Mohammed et al., 2022; Tian et al., 2018). They have been successfully integrated in a number of operational drought monitoring systems, thanks to their minimal input data requirements and ease of use. Among those indices, the Standardized Precipitation Index (SPI, McKee et al., 1993) computed on short-to-medium aggregation periods (i.e., SPI-3 and -6) is often adopted as a suitable proxy variable for agricultural droughts (WMO, 2012).

As highlighted by Sheffield and Wood (2007), simplified indices for drought monitoring, such as the Palmer Drought Severity index (PDSI; Palmer, 1965) or the previously mentioned meteorological indicators, have been slowly integrated with indices directly based on modeled soil moisture data.~~,~~ This transition is fostered by ~~thanks to~~ the increasing availability worldwide of process-based hydrological models. Soil moisture percentile, or similarly standardized quantities, are often used for this scope (Mo and Lettenmeier, 2013; Xia et al., 2014). The ever-growing records of remote sensing-based estimates of soil moisture are becoming an additional data source to support the development of dedicated soil moisture-~~-~~based drought indices (Cammalleri et al., 2017; Carrão et al., 2016).

In the context of agricultural drought, an overall good agreement between SPI and soil moisture indices has been demonstrated over a large range of agricultural practices, crop types and climatic conditions. Halwatura et al. (2017) showed how SPI-3 represents a good approximation of modeled soil moisture over three different climatic regions in eastern Australia. Sims et al. (2002) found high correlation between short-term precipitation deficit and soil moisture variations in North Carolina, while Ji and Peters (2003) highlighted the high correlation between SPI-3 and vegetation growth over croplands and grasslands in the U.S. Great Plains. Wang et al. (2015) observed a good matching between soil moisture dynamics and SPI at the scale of 1-3 months when testing various indices over China. In Europe, Manning et al. (2018) highlighted how precipitation is the main driver of soil moisture droughts for a set of both dry and wet sites.

In spite of the above-mentioned consistencies, the outcome of any~~of a~~ drought analysis is ~~will be~~ inevitably affected by the index selected to characterize drought conditions over a certain study region, ~~will inevitably affect the outcome of the drought analysis,~~ as also highlighted by Quiring and Papakryiakou (2003) in testing different indices over the Canadian prairies. These a~~A~~uthors suggest that a variety of drought indices should always be tested to determine the most appropriate one for ~~each specific~~a given application. It follows that the synergy between multiple indices can be exploited by the use of multivariate indicators (Hao and Singh, 2015), a family of approaches that encompasses a variety of merging strategies, including combined cascading indices (Cammalleri et al., 2021a; Rembold et al., 2019), composite and integrated approaches (Brown et al., 2008; Svoboda et al., 2002), and joint probability functions (Bateni et al., 2018; Hao and AghaKouchak, 2013; Kanthavel et al., 2022).

The latter category~~lass of approaches~~, in particular, aims at capturing the complex statistical ~~inter~~dependence among different drought-related variables (Hao and Singh, 2015), and it has seen a growing relevance in many hydrological applications thanks to the introduction of copula functions and their ability to model a wide range of dependence structures (Nelsen, 2006; Salvadori et al., 2007; Joe, 2015). In the field of drought indices, the approach proposed by Kao and Govindaraju (2010) for the computation of the Joint Deficit Index (JDI) has been applied to a variety of drought-related quantities over different regions, often including precipitation and soil moisture (i.e., Dash et al., 2019; Kwon et al., 2019).

A ~~One~~ key feature in using~~of~~ joint probability is the possibility to characterize the so-called tail-dependence (TD), namely the asymptotical dependence of~~at~~ the extremes (Frahm et al., 2005). While TD has received large attention in the scientific literature of hydrological extremes (e.g., Aghakouchak et al., 2010; Poulin et al., 2007; Serinaldi, 2008), its use is largely unexploited in ~~limited attention to this quantity is still given in the~~ studies focusing on combined drought indices.

Studies on the marginal distribution of either precipitation or soil moisture usually adopt ~~have somewhat converged on adopting~~ the Gamma distribution for precipitation and the Beta distribution for soil moisture. The use of the Gamma family for the implementation of the SPI at different accumulation periods has become~~en~~ a standard practice in many applications (e.g., Mo and Lyon, 2015; Yuan and Wood, 2013). While other distributions have also proven to be reliable, such as the exponentiated Weibull (Pieper et al., 2020) and the Person Type III (Ribeiro and Pires, 2016), fitting the Gamma is still the most adopted ~~one~~approach. Over Europe, Stagge et al. (2015) demonstrated how the Gamma outperformed the other tested distributions across all accumulation periods and regions.

A more limited number of applications based on soil moisture data are available in the scientific literature compared to SPI. The use of the Beta distribution for soil moisture data has been introduced as early as the late ‘70s, with the pioneer study of Ravelo and Decker (1979), following the consideration that soil moisture is a double-bounded quantity, ranging between residual and saturation. Sheffield et al. (2004) successfully applied this standardization for drought analyses over the US, while the same distribution has been adopted by Cammalleri et al. (2016) on modeled data over Europe. Most recently, the Beta distribution was also used to characterize the frequency of global satellite soil moisture data (Sadri et al., 2020).

Conversely, no standard approaches have been identified for the application of copulas to model the bivariate joint distribution of precipitation and soil moisture, mainly due to the large variety of probabilistic structures than may be observed between these two quantities. Common fitting strategies rely on the application of various copula families to identify the optimal ~~best fitting~~ for each specific site (e.g., Hao and AghaKouchak, 2013), or are based on an a-priori selection of a copula family following empirical evidence (e.g., Dixit and Jayakumar, 2021). Independently from the selection strategy, the adopted copula implicitly assumes~~carries with it~~ an underling TD behavior, which influence~~impacts~~ on extreme detection should be properly accounted.

A comprehensive study on the joint probabilistic dynamics of ~~these two quantities~~precipitation and soil moisture ~~, and on their bivariate distribution,~~ is currently lacking in the scientific literature of multivariate drought modeling. Hence, the main goal of this study is to fill this gap, by investigating the mutual relationship between the empirical frequencies of precipitation (cumulated over 3 months, as for SPI-3) and soil moisture datasets as available over Europe as part of the European Drought Observatory of the Copernicus Emergency Management Service (EDO, https://edo.jrc.ec.europa.eu).

A large set of copulas is tested for this purpose across the entire European domain, to identify an optimal modeling of the dependence especially in proximity of the tails (given its major role in extreme detection). The spatial distribution of the results is analyzed to infer evidence of common patterns and behavior, which may support future operational applications based on similar parametric approaches.

2. Materials and Methods

2.1 Precipitation and soil moisture datasets

The study focuses on Europe and makes use of the dataset of indicators available over the region as part of EDO. Precipitation data accumulated over consecutive 3-month periods are used here, as the quantity at the base of the SPI-3 index. Hourly total precipitation maps from the ECMWF ERA5 global atmospheric reanalysis model (https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5) are collected through the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/) and cumulated at monthly updates (no missing values are present in the reanalysis dataset). This dataset has proven to be quite reliable over Europe for drought analyses (e.g., Cammalleri et al., 2021b; van der Wiel et al., 2022), as it is currently employed in near-real time as part of the operational tools of EDO. Empirical frequencies of 3-month precipitation are derived from the rainfall records, in order to obtain a non-parametric calculation of the standardized anomaly, ~~dataset conceptually analogous to the commonly used~~ SPI-3, ~~but~~ without the possible artifact introduced by the fitting of a theoretical distribution (i.e., Gamma distribution) (see Soľáková et al., 2014). From here on, we will refer to this dataset as standardized precipitation.

Soil moisture records over the entire European domain are derived from the simulations of the LISFLOOD distributed hydrological rainfall–runoff model (de Roo et al., 2000). LISFLOOD runs in near-real time as part of the European Flood Awareness System (Thielen et al., 2009), and it provides daily soil moisture maps for the root zone at a spatial resolution of 5-km. Daily modeled data are averaged at monthly scale and converted into a Soil Moisture Index (SMI) as in Seneviratne et al. (2010). The model is calibrated and validated over an extensive network of river discharge stations following the procedure described in Arnal et al. (2019), and it has been successfully tested for drought analyses over Europe as part of EDO for the computation of the Soil Moisture Anomaly (SMA) index (Cammalleri et al., 2015). Similar to ~~As for the~~ precipitation, empirical frequencies are computed from the monthly soil moisture data in order to obtain a non-parametric calculation of the standardized anomaly, SMA, thus ~~but~~ independent from a theoretical fitting (i.e., Beta distribution). We will refer to this dataset as standardized soil moisture from hereafter.

In this study, data collected for~~on~~ the most recent 25 years (1996-2020) are used as a common period. This period is chosen to minimize the effects of non-stationarity in precipitation records and to avoid the inclusion of early LISFLOOD records that are affected by a lower number of ground meteorological stations in the forcing (Thieming et al., 2022). The time series of both standardized precipitation and soil moisture at grid-cell scale are preliminarily tested for auto-correlation using the partial auto-correlation function (PACF, Box and Jenkins, 1976). This analysis returned positive and statistically significant (95% confidence interval) values only at ~~for~~ lag = 1, suggesting a substantial absence of auto-correlation beyond what is expected for time series with smooth temporal dynamics such as ~~derived on a~~ 3-month cumulative ~~data~~precipitation and soil moisture.

The 300 maps (12 months × 25 years) for the two standardized datasets are then spatially interpolated on a common Lambert azimuthal equal-area (LAEA) projection on a regular grid of 5-km using the nearest neighbor algorithm. This is done to preserve the high-resolution information of the soil moisture and by considering the smooth spatial dynamics of precipitation accumulated over 3 months.

2.2 Copula families

The introduction of copulas in multivariate probability modeling has provided to hydrologists a flexible tool to reproduce the joint probability of multiple dependent variables characterized by a variety of marginal distributions (De Michele and Salvadori, 2003; Salvadori and De Michele, 2004).

Limiting the focus on bivariate variables, the joint probability distribution, F, of two random variables (X_1 and X_2) can be expressed, thanks to the Sklar's theorem, as:

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \quad (1)$$

where F_1 and F_2 are the marginal distribution of X_1 and X_2, respectively, and C is the copula function (Salvadori et al., 2007).

A large variety of parametric formulations has been introduced in the literature to explicitly link the marginal distributions to the joint probability ~~distributions~~, with some of the most common copula families used in hydrology belonging to the Elliptical and Archimedean copulas (Chen and Guo, 2019). Two ~~non-parametric~~ measures of dependence play a major role in parametric copula inference. The Kendall rank correlation coefficient (τ) is commonly used as a non-parametric measure of overall ordinal association, while the so-called Tail-Dependence (TD) coefficients~~,~~ (Salvadori et al., 2007) are~~is~~ used to estimate the asymptotical degree of dependence in the upper and lower extremes (upper tail-dependence, λ_U, and lower tail-dependence, λ_L, respectively). The estimation~~assessment~~ of ~~the non-parametric~~ TD non-parametrically is not an easy task, as highlighted by Serinaldi et al. (2015), as it aims at assessing an ~~reconstructing an~~ asymptotic behavior~~value starting~~ from a finite sample. Several formulations are proposed in the scientific literature (see Frahm et al., 2005), and ~~The non-parametric values of both TDs can be evaluated following~~ the method proposed by Schmidt and Stadtmueller (2006) is here used to ~~evaluate~~obtain ~~the~~ non-parametric ~~values~~estimates of both TD coefficients.

In this study, the parametric bivariate probability of standardized precipitation and soil moisture is assessed by using the R package "VineCopula" (Aas et al., 2009; Dissman et al., 2013). The Akaike Information Criterion (AIC, Stoica and Selen, 2004) is used to select, for each spatial grid cell, the best fitting copula among the wide range of families available in the package. The main properties of some relevant copulas are reported in Table 1, as they will be useful to interpret the successive results.

In particular, from the data in Table 1 it is important to highlight how the BB7 copula is a combination of ~~the~~ Joe and Clayton copulas, of which it inherits the tail-dependences, and how the TD behavior of a copula can be inverted (i.e., the upper tail-dependence can become the lower and *vice versa*) by simply considering the reciprocal marginals (commonly known as rotated forms, identified by the suffix 180). Information from both non-parametric and parametric approaches are here jointly used to discriminate between different TD behaviors.

Table 1. Main copulas analyzed in this study and their ~~coefficients for the~~ upper and lower tail-dependence coefficients (λ_L and λ_U, respectively).

Copula	λ_L	λ_U
Gaussian	0	0
Student-t	$2t_{v+1}\left(-\sqrt{v+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$	$2t_{v+1}\left(-\sqrt{v+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$
Gumbel	0	$2-2^{\frac{1}{\theta}}$
Clayton	$2^{\frac{-1}{\theta}}$	0
Joe	0	$2-2^{\frac{1}{\delta}}$
BB7	$2^{\frac{-1}{\theta}}$	$2-2^{\frac{1}{\delta}}$

Even if a copula is selected as the optimal based on the AIC, this does not necessarily exclude that other copulas may perform similarly. For this reason, we introduced a further test based on the relative likelihood criterion (Burnham and Anderson, 2002), $\exp\left(\frac{\mathrm{AIC}_{\min}-\mathrm{AIC}_i}{2}\right)$, to establish the likelihood that an AIC value of a given copula (AIC_i) is ~~statistically~~ significantly different tha~~t~~n the minimum value ($\mathrm{AIC}_{\min}$) obtained for the optimal solution.

2.3 Random forest classification of selected copulas

The interpretation of the selected copula functions may help highlighting the transferability of the observed results over different contexts. For this reason, the observed spatial distribution of the selected copulas is analyzed through a random forest classifier (Breiman, 2001), in order to find evidence~~s~~ of reproducible patterns beyond simple chance.

As input features we consider a set of commonly available variables, such as: ground elevation, annual average temperature, annual total precipitation, precipitation seasonality (ratio between total precipitation in warm and cold months), annual average Normalized Difference Vegetation Index (NDVI), annual average soil moisture, and soil type. As hyperparameters for the random forest, we tuned the number of trees (ntree) and the number of features randomly sampled at each split (mtry) using the “randomForest” R package (Breiman, 2001).

3. Results

A preliminary analysis of the degree of correlation between the monthly ~~empirical frequencies of~~standardized 3-month precipitation and soil moisture (analogous to non-parametric SPI-3 and SMA) is tested on the full timeseries of each grid cell using the ~~non-parametric Kendall rank correlation coefficient (also known as~~ Kendall’s τ~~)~~, as depicted in Fig. 1 for the entire European domain.

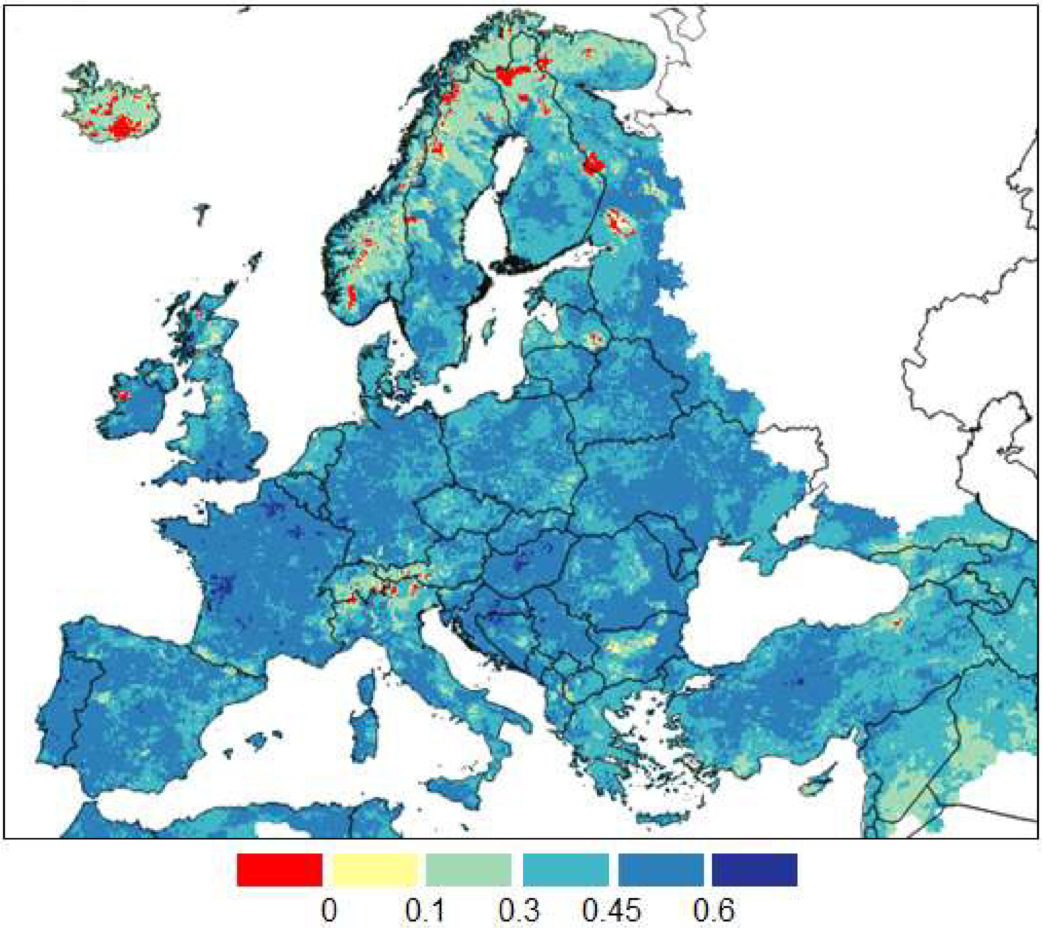

Fig. 1. Spatial distribution of the Kendall’s τ between monthly ~~empirical frequencies of~~standardized 3-month precipitation and soil moisture. Roughly, values lower than 0.1 are not statistically significant at $p = 0.05$ (two-tails).

The results reported in Fig. 1 confirms the expected direct relation between the two variables~~quantities~~, with a relatively homogeneous distribution of medium/high (between 0.3 and 0.5) ~~correlation~~ τ values ~~between 0.3 and 0.5~~ (τ = 0.42±0.1). Limited regions with low (and sometimes even slightly negative) τ values are sporadically observed, mostly ~~concentrated~~ over the Alps, Iceland and the coldest regions of the Scandinavia peninsula. Low ~~C~~correlations over these regions are likely ~~affected~~ related to~~by~~ the presence of snow coverage during extended periods of the year. Overall, the observed τ values cannot be considered statistically significant (at p = 0.05) only for less than 2% of the domain.

The analysis of the non-parametric tail-dependence values is summarized in the plot depicted in Fig. 2, where the cumulative frequency of the difference between the empirical λ_L and λ_U values is reported. The range of TD values in Fig. 2 for which it is possible to exclude significant asymmetry in the tail dependence coefficients ~~Symmetric behaviors in Fig. 2 can be~~is identified by setting a maximum value for $|\lambda_L - \lambda_U|$. To ~~identify~~ define this threshold, the non-parametric TD coefficients were re-computed on shuffled time series (to artificially reconstruct conditions of null ~~tail~~ dependence~~ies~~), and the $|\lambda_L - \lambda_U|$ value corresponding to a cumulative frequency of 90% of the grid cells after the shuffling was detected as threshold, corresponding to a value of 0.1. This value can be seen as a lower limit to identify symmetric dependence.

The plot in Fig. 2 highlights how the majority (about 50%) of the grid cells can be considered characterized by a symmetric behavior in the tail-dependence coefficients according to the above mentioned criterion ($|\lambda_L - \lambda_U| < 0.1$), whereas the rest of the grid cells are almost equally split between a predominance of the Upper Tail-Dependence (UTD, corresponding to negative differences) or a predominance of Lower Tail-Dependence (LTD, positive differences).

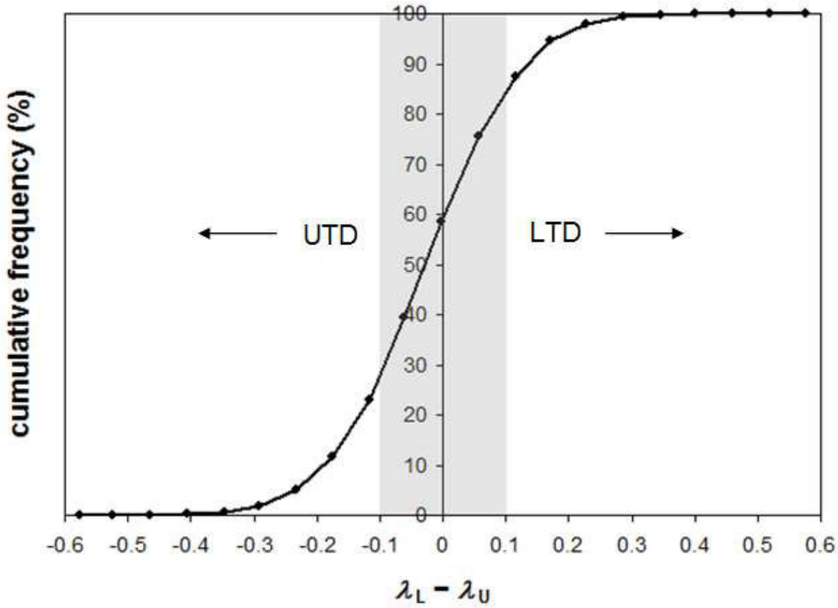

Fig. 2. Analysis of the frequency of the empirical tail-dependence coefficients. The plot shows the cumulative frequency distribution of the differences between the empirical λ_L and λ_U values computed according to Schmidt and Stadtmueller (2006). The domain with a roughly symmetric behavior ($|\lambda_L - \lambda_U| < 0.1$) is highlighted by the grey box area.

The results reported in Fig. 2 were used to divide the entire domain in three categories (symmetric, LTD and UTD) as depicted in Fig. 3. This map shows evidence of some coherent spatial patterns, such as the predominance of LTD in southern France, southern Italy, northern Germany and Denmark, and western Ukraine (among others), and a clustering of UTD in Poland, Czechia ~~Republic~~, southern Scandinavia, and Greece. The symmetric condition seems overall more spread across the entire domain, also thanks to the higher frequency, with a slightly predominance over northern Europe (i.e., northern Scandinavian peninsula and Iceland).

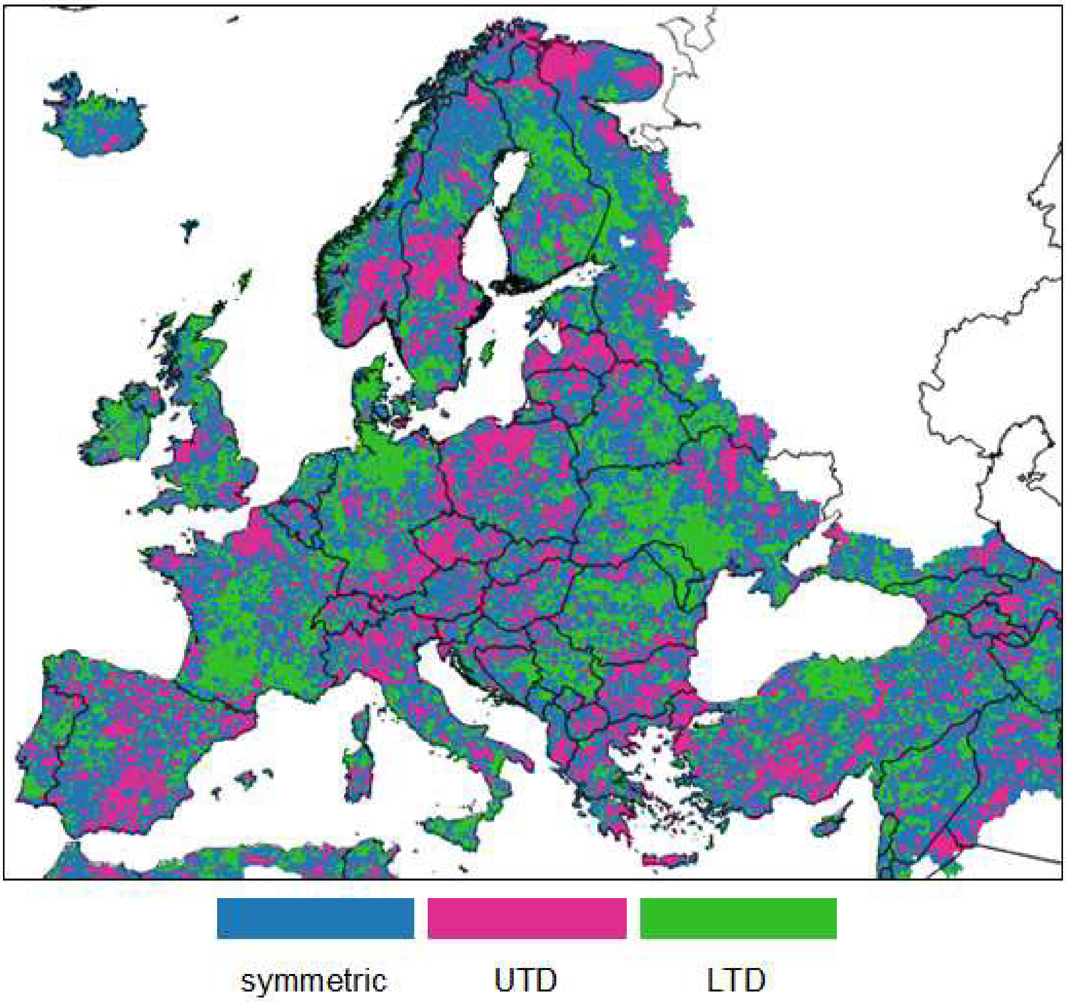

Fig. 3. Spatial distribution of the three categories derived from the differences in the empirical tail-dependenc~~i~~e coefficients.

Given the results ~~observed in terms~~ of the tail-dependence assessment, ~~is~~ it is useful to focus the copula parametric analysis on the capability to reproduce such patterns instead of finding the single copula that can perform reasonably well over the entire domain. Indeed, the search for the optimal copula based on the minimum AIC returns the BB7 as the optimal one~~solution~~ in about 80% of the domain (not shown). This result is ~~thanks~~ a consequence of ~~to~~ the BB7 flexibility ~~of its formulation~~ (being derived from~~as~~ a combination of two purely asymmetric functions), which allows reproducing both symmetric and asymmetric tail-dependence coefficients~~ies~~ according to~~depending on~~ the values assumed by the two parameters. However, the fact that a single flexible copula works well over a large range of conditions may hide the key spatial patterns observed in the TD analysis. These patterns may be better reproduced~~, which can be highlighted instead~~ by adopting a limited number of more specialized ~~common~~ copulas ~~specialized in reproducing specific behaviors~~.

By limiting the search to a subset of copula functions, comprising only purely symmetric or purely asymmetric tail behaviors, more interesting results are obtained, as summarized by the frequency plot in Fig. 4. The grid cells where symmetric tail behavior copulas are selected as optimal are about 55% of the domain (see Fig. 4b), with a predominance of Student-t copula but also with a non-negligible fraction of cells (23%) where the Gaussian (symmetric and without tail-dependence~~s~~) is chosen (see Fig. 4a). The remaining grid cells are almost equally split between upper and lower tail-dependence~~s~~, with Gumbel (and its rotated counterpart, Gumbel 180) as the most selected among the asymmetric options.

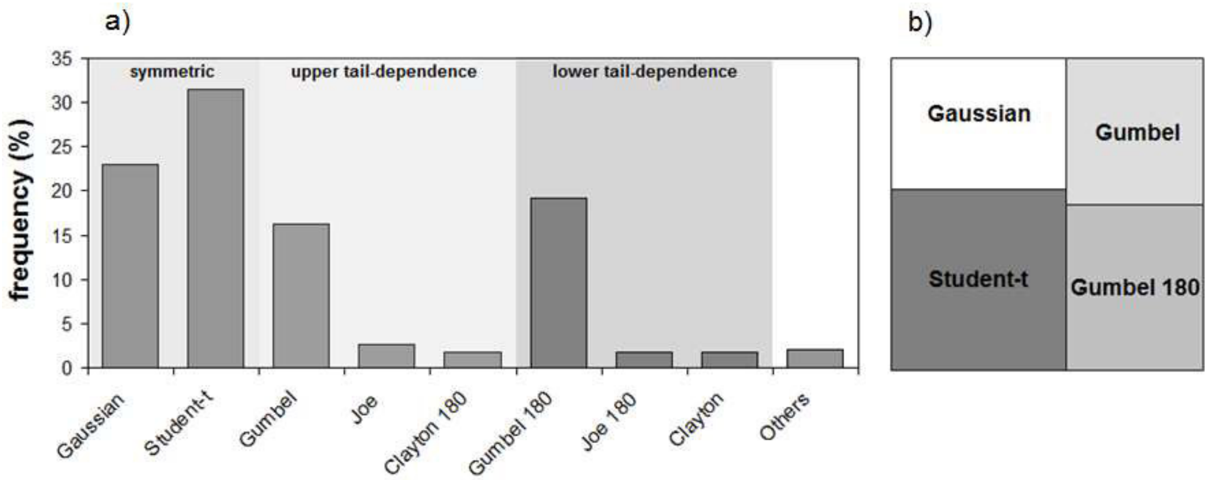

Fig. 4. Frequency of the optimal copulas based on the minimum AIC. The barplot in panel a) shows the frequency of~~for~~ each copula, while~~ereas~~ the box in panel b) reports a ~~synthetic~~ compact description of the subdivision of the entire domain among the 4 most frequent copulas.

The spatial distribution of these optimal copulas (Fig. 5) ~~confirms~~ mostly agree with ~~of~~ the patterns observed in Fig. 3, supporting the findings ~~as a converge of evidence~~ on the ~~TD~~ spatial distribution of TD coefficients~~ of the regions where either of the two TD is predominant~~. In addition, this result further confirms~~ation~~ that a rather limited range of simple copula functions is able to capture the overall dynamics of dependence between precipitation and soil moisture over the entire European domain. Despite the observed spatial clusters in the obtained optimal copulas, the overall patterns observed in Fig. 5 are still rather noisy and may be difficult to interpret. This erratic behavior can be partially explained by the fact that different copulas may perform quite similarly over some grid cells, hence the AIC of the optimal copula (AIC_{min}) may not differ significantly from the AIC of other functions.

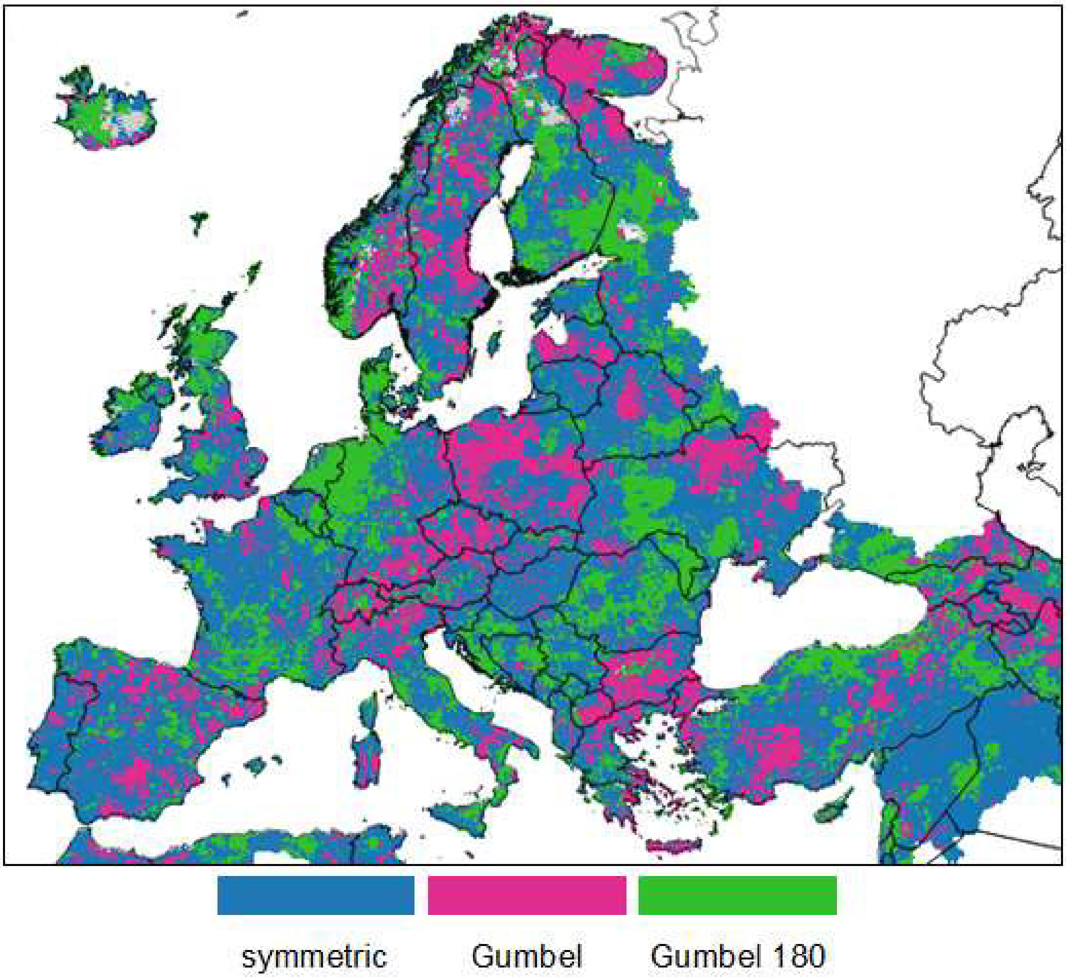

Fig. 5. Spatial distribution of the optimal copulas obtained by minimizing the AIC. The symmetric tail behavior class includes both Gaussian and Student-t copulas.

To further investigate this hypothesis, we evaluated the possibility to replace the optimal copulas with either a Student-t or a Gumbel (direct and rotated) over the entire domain. The Gaussian copula was excluded from this analysis under the assumption that the no tail-dependence of the Gaussian can be adequately reproduced by the Student-t with a small enough tail-dependence. The plots in Fig. 6 reports the relative likelihood for the Student-t (panel a) and Gumbel families (panel b) compared to the locally selected optimal copulas. Low values of this metric correspond to conditions where the optimal copula cannot be replaced by the alternative function (being either the Student-t or the Gumbel).

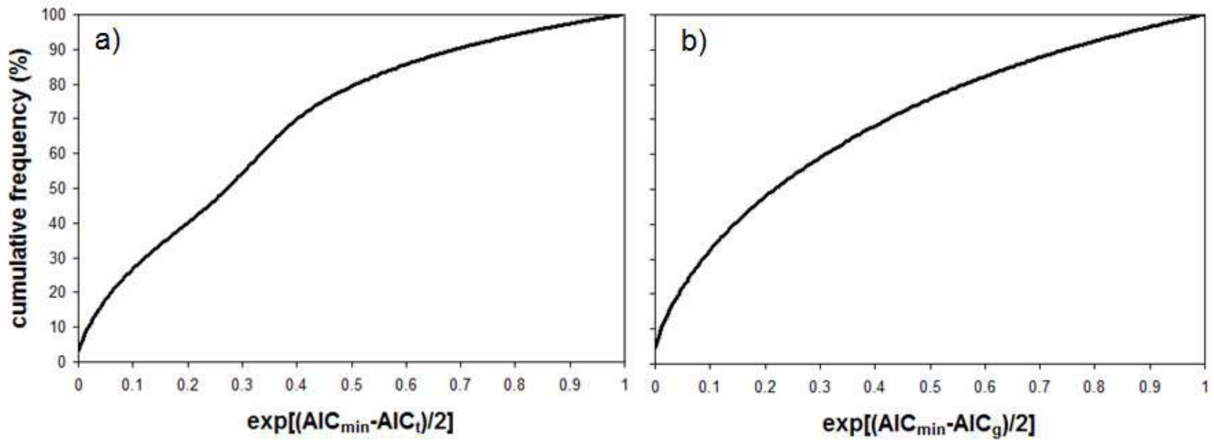

Fig. 6. Frequency analysis of the relative likelihood computed between the optimal AIC (AIC_{min}) and: a) Student-t (AIC_t), or b) Gumbel (AIC_g) families. The grid cells where either the Student-t or the Gumbel was already the optimal solution were excluded from the respective frequency analysis.

The results in Fig. 6 show that, if we assume a relative likelihood of 0.1 as a threshold to detect a statistically significant difference, the Student-t cannot reasonably replace the local optimal copula in about 18% of the entire domain (Fig. 6a), whereas this fraction is about 17% for the Gumbel family (Fig. 6b). ~~It is also possible to observed that how~~ It emerges that the Gumbel family is the optimal one~~copula~~ in almost the totality (about 99%) of the grid cells where the Student-t is not a suitable replacement of the local optimal, whereas almost only symmetric copulas (63% Student-t and 34% Gaussian) are the optimal functions where the Gumbel family is not a suitable replacement. Overall, these results suggest that the selection of the optimal copula is "univocal" (i.e., cannot be reasonably replaced by another function) in about 35% (18+17) of the domain, whereas either the Student-t or the Gumbel families can be adopted in the remain fraction of the domain with similar performances in terms of AIC (and no clear TD behavior). ~~It is worth mentioning how t~~This analysis also confirms the assumption that all the areas where the Gaussian was chosen as optimal copula can be satisfactory modeled ~~also~~ by using the Student-t (i.e. without a statistically significant increase in AIC).

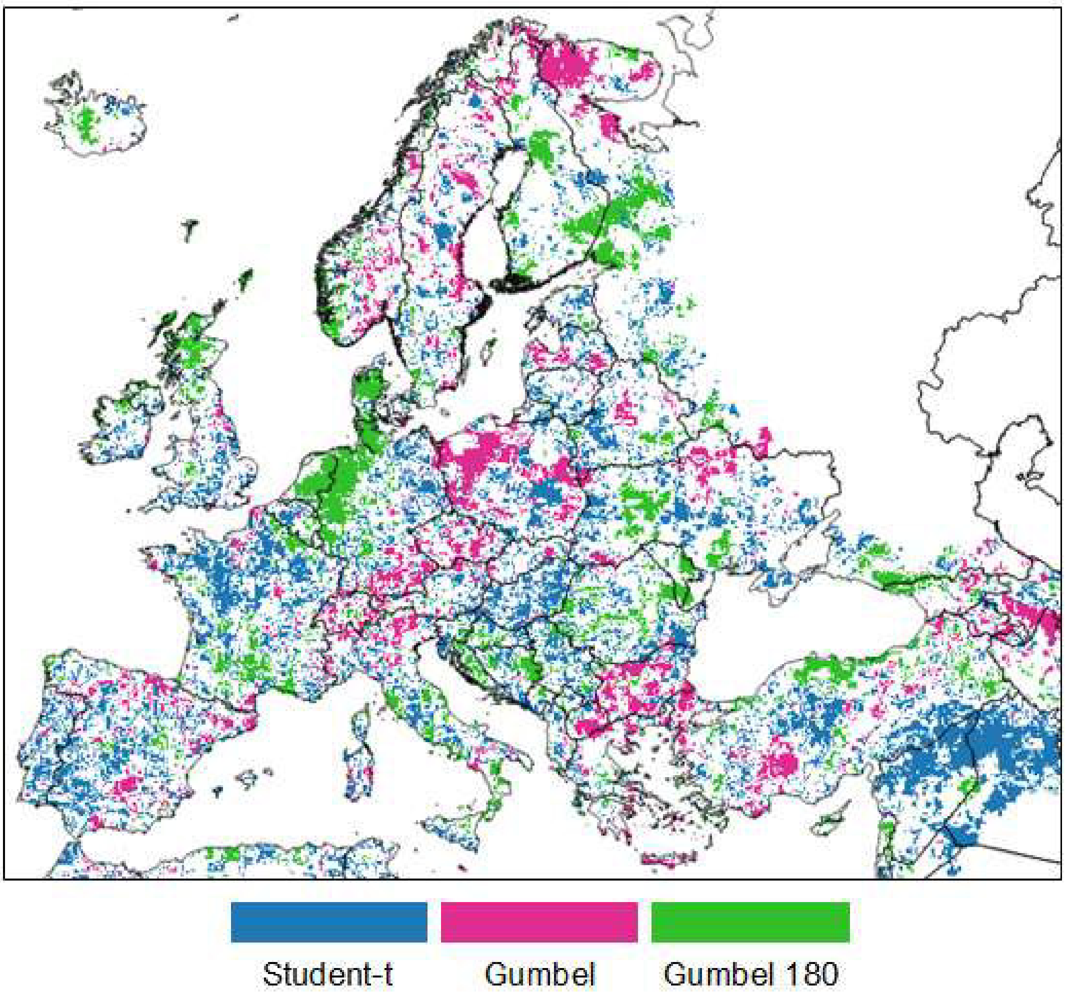

Fig. 7. Spatial distribution of the grid cells where the selection of the optimal copula is "univocal" according to the relative likelihood criterion.

The "univocal" areas derived from the previous analysis are mapped in Fig. 7, highlighting some of the more consistent spatial clusters already observed in both Figs. 3 and 5, as well as a large fraction of cells in northern Europe where a "univocal" optimal copula cannot be selected. These grid cells with "univocal" copula are used as a starting point for the random forest classification, given the robustness in their signal, and the agreement in the outcome of both parametric and non-parametric TD ~~s~~behaviors.

A sample corresponding to ~~of~~ 25% of the "univocal" grid cells (~~corresponding to~~ about 8% of the entire domain) was used to train the random forest, adopting a number of trees (ntree) of 80 and a single feature randomly sampled at each split (mtry = 1). The training size and the minimum values of hyperparameters were chosen to reduce the problem of overfitting. Among the possible features, three variables were selected by analyzing the variable importance plots, as well as the ease of access: annual average temperature, annual total precipitation, and precipitation seasonality. The trained classifier was then applied to the testing subset (the remaining 75% of the "univocal" grid cells) and the outcomes were analyzed by mean of a confusion matrix, which results are summarized in Table 2. Overall, the obtained classification has a very satisfactory matching with the test subset, with a general high accuracy (ACC = 0.86) and with all the metrics pointing toward a significant improving in the performance compared to the reference No-Information-Rate (NIR) (i.e., small p-values) and a high probability to have correct modeled values compared to simple chance (i.e., high Cohen's K).

Table 2. Summary of the confusion matrix analysis applied to the trained random forest on the testing subset.

Accuracy (ACC)	0.86
No-Information-Rate (NIR)	0.50
p-value (ACC > NIR)	$< 2.2 \times 10^{-16}$
McNemar's test p-value	3.44×10^{-5}
Cohen's kappa statistic (K)	0.78

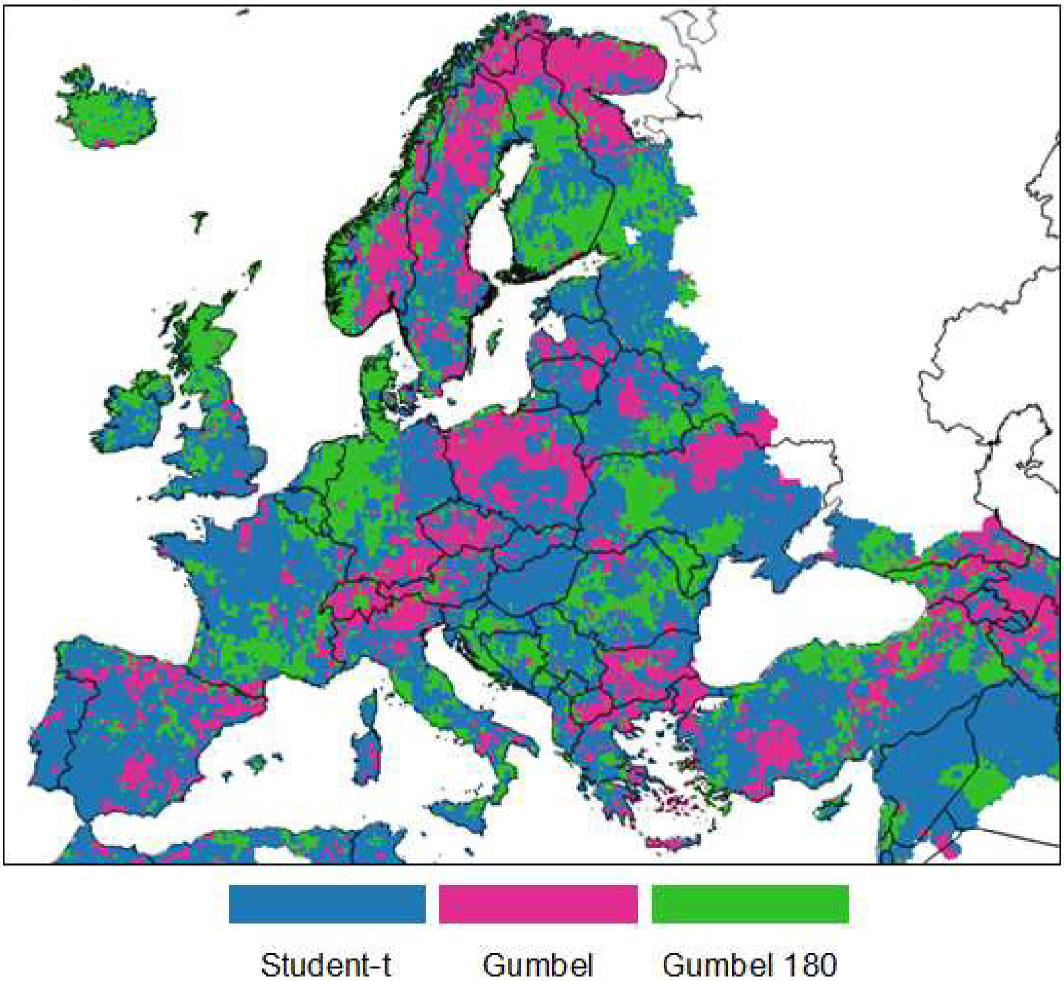

Fig. 8. Map of the optimal copula as modeled by the trained random forest classifier.

Finally, the trained classifier was ~~extended~~ applied to the entire ~~domain~~ dataset to obtain a classification of the ~~entire~~ European domain in term of the expected optimal copula and the corresponding TD behavior. This map, reported in Fig. 8, shows ~~bears~~ a strong resemblance to both the empirically-derived map in Fig. 3 and the optimal AIC fitting in Fig. 5. Beside this overall agreement, some notable discrepancies can be observed over northern Scandinavia and Iceland, two regions where low Kendall's τ and a small fraction of "univocal" selected copulas were already identified.

4. Discussion

The overarching goal of the study is to investigate the joint probability of two standardized variables aiming at capturing agricultural drought conditions, hence the overall agreement between these two quantities is a fundamental prerequisite. A ~~The expected~~ direct relationship between standardized 3-month cumulated precipitation and soil moisture is expected, ~~as~~ since both SPI-3 and SMA are similarly-used agricultural drought indices, and this can support the identification of the most suitable set of copula families (Salvadori et al., 2007; Genest et al., 2007). This ~~behavior~~ direct relationship is overall confirmed by the positive Kendall's τ values estimated~~observed~~ over most of the domain ($\tau = 0.42 \pm 0.1$). Moderately high correlation values of ~~were observed between~~ standardized precipitation and soil moisture were estimated also in other studies. Kwon et al. (2018) reported~~ing~~ Pearson's r values between 0.4 and 0.6 for 55 stations in South Korea, albeit with seasonal patterns; Gaona et al. (2022) ~~reported~~ found similar values over the Ebro basin with both land-surface modeled and satellite soil moisture, and Sepulcre-Cantó et al. (2012) obtained an average value of r of about 0.6 over nine stations across Europe.

Sehler et al. (2019) studied the correlation between remote sensing-based precipitation and soil moisture, finding moderate correlation over southern Europe, and a weak (often not significant) correlation in central Europe. However, central Europe is close to the upper limit of the analyzed remote sensing products, which can explain such low performance. Limited correlation even among different soil moisture products has been observed in northern Europe in other studies (Almenda-Martín et al., 2022), confirming the difficulty to model ~~the~~ soil moisture dynamics over this region.

The obtained values for the Kendall's τ fall in a somewhat optimal range for the analysis of the joint probability, since they ~~τ values~~ are statistically significant almost everywhere (i.e., ~~consistency in the produced outcomes of~~ the two indices are to a certain degree consistent) but not too high to make meaningless any joint use of the two datasets (i.e., the two indices are too similar and provide the same information ~~products~~).

~~Even more interesting is t~~The outcome of the tail-dependence analysis is even more interesting, given the role that such metric ~~quantity, (and in particular the low tail),~~ plays in the detection of ~~drought~~ extreme events (and in particular the low-tail for droughts). The TD investigation is sometimes overlooked in the development of multivariate drought ~~analyses~~indices, where previous studies often focused on optimizing the copula to the local data without analyzing the implicit assumption on the TD, the consistency with the ~~empirical~~ non-parametric TD, and the implications of the associated dependenci~~es~~ ~~forfor~~ ~~the modeling of drought conditions~~. ~~Indeed,~~ ~~p~~Previous ~~references in the scientific literature for~~ studies on the joint probability of precipitation and soil moisture are rather scarce, and TD is rarely the focus of ~~extensive~~ such analyses or, at least, limited to specific areas and/or conditions ~~not to, such an extend as the one results reported in this study for the entire Europe~~. ~~, and previous references in the scientific literature for precipitation and soil moisture are rather scarce.~~

As an example, Manning et al. (2018) performed a very detailed analysis over 11 FluxNet sites in Europe on the role of precipitation and evapotranspiration on soil moisture drought, based on pair copula constructions, but the authors did not provide any indication on which bivariate copula was the optimal one for each site. Kwon et al. (2018) reported that Frank copula was the most frequent optimal choice in their study over South Korea.~~,~~ However, ~~but~~ some clear spatial patterns ~~were also~~ observed in their outcomes were~~are~~ not discussed, with Frank being the selected copula mostly in the central area of the domain, but ~~also~~ with Gumbel and Student-t performing the best in the southern and eastern coasts, respectively.

Dash et al. (2019) found Frank (among the Archimedean copulas) working the best for 3-month precipitation and soil moisture over an Indian basin;~~,~~ while Hao and AghaKouchak (2013) highlighted the good performance of Frank and Gumbel in five regions of~~in~~ California, even if neither Gaussian nor Student-t were considered. In all these applications, no specific considerations on the ~~observed~~ TD behaviors were reported, even if a common trend seems to be the good performance of Frank copula.~~,~~ This ~~which~~ is in contrast with our results, where the Frank was very rarely selected as optimal (less than 1% of the domain). A possible explanation of these results may be our focus on empirical marginal frequencies rather than theoretical ones, given the well-documented increasing uncertainty in parametric fitting in the tails (Farahmand and AghaKouchak, 2015; Laimighofer and Laaha, 2022). As a possible confirmation of this hypothesis, a good performance of Gumbel and Gaussian has been observed over Iran by Bateni et al. (2018), similarly to our results, when a non-parametric form for SPI and SSI (Standardized Soil Moisture Index) was used.

The absence of a strict standard procedure to investigate tail-dependence may be another factor affecting the limited focus on the topic ~~of~~ in many studies on multivariate drought ~~studies~~indices. Non-parametric TD has the clear advantage to avoid any alteration of the data due to the fitting procedure, but the outcomes in this study also show a high degree of spatial noise likely due to the intrinsic nature of non-parametric analyses, the large uncertainty in non-parametric methods (Serinaldi et al., 2015), as well as the effects of ~~to~~ the limited sample size

~~which affects the estimates of TD.~~ (~~F~~for this last issue see also the illustration 3.18 in Salvadori et al., ~~(~~2007). The threshold used here to define a symmetric behavior, based on a random shuffling of the data, seems to successfully overcome the difficulty to define a self-consistent maximum difference in TD values, but it cannot be seen as a reliable approach to easily identify TD symmetry without the support of further evidence (e.g., by theoretical analyses ~~or other means~~).~~.~~

In this regard, ~~T~~the fitting of parametric copula functions returns ~~more consistent~~ spatial patterns in ~~our study~~TD coefficients similar to the ones obtained with the non-parametric approach. However, ~~, but evidence on~~ the absence of “univocal” fittings can be observed for large areas, as well as some contrasting results compared to the non-parametric TD especially over northern Europe (areas with low correlation). The grid cells where a given copula clearly outperforms the alternative options is limited to roughly 1/3 of the domain, further stressing the evidence that clear-cut outcomes are difficult to infer from a single methodology. Thus~~In this regard~~, it seems reasonable to state that ~~infer that only a~~ only a critical concerted analysis of both parametric and non-parametric TDs can return robust practical indications based on a converge of evidence.

A clear outcome of our study is the predominance of regions with symmetric tail-dependence~~s~~ coefficients, where the Student-t copula is suitable to reproduce the joint probability of standardized precipitation and soil moisture. An even split of the remaining domain between areas with either lower or upper tail-dependence ~~only~~ is also observed, where the Gumbel copula (in either is direct or 180 rotated forms) is proven to be a suitable option. These results are crucial in defining the role of standardized precipitation and soil moisture datasets in detecting drought events, and to which extent they can work in synergy in a drought monitoring system. ~~In fact,~~ ~~w~~While the correlation between the two datasets highlights the extent of their overall agreement, which in this study was somewhat uniform across most of the domain (τ ranging between 0.3 and 0.5), very different degrees of tail-consistency can be obtained for similar Kendall’s τ if the TDs differ substantially. Regions with higher LTD will have a higher agreement in the detection of drought extremes compared to ~~the~~ areas with a UTD predominance, hence a low number of false alarm and a higher signal-to-noise ratio may be~~is~~ expected.

To further explore this behavior~~stress on this issue~~, the time series of standardized variables were converted in binary vectors based on ~~a~~the commonly used standardized drought threshold of -1 (corresponding to an empirical frequency of 0.16). On these data, the pair-wise binary correlation coefficient, $\rho(-1)$, was computed separately for the grid cells with LTD and UTD.

Results are shown~~reported~~ in Fig. 9, for grid-cells with low (0.1 < τ ≤ 0.4, panel a) and high (τ > 0.4, panel b) overall correlation, respectively. They show a net increase in the pairwise binary correlation for the grid cells with LTD (of about 0.15 in both cases) compared to the cases with UTD, even if the overall correlation is comparable. This increase in ρ(-1) translates in a stronger agreement in the detection of extremes when a low tail-dependence is observed~~ (i.e., less uncertainty)~~, resulting in a more robust detection of the drought conditions thanks to the concurrency of extreme conditions in both drought indices (i.e., convergence of evidence).

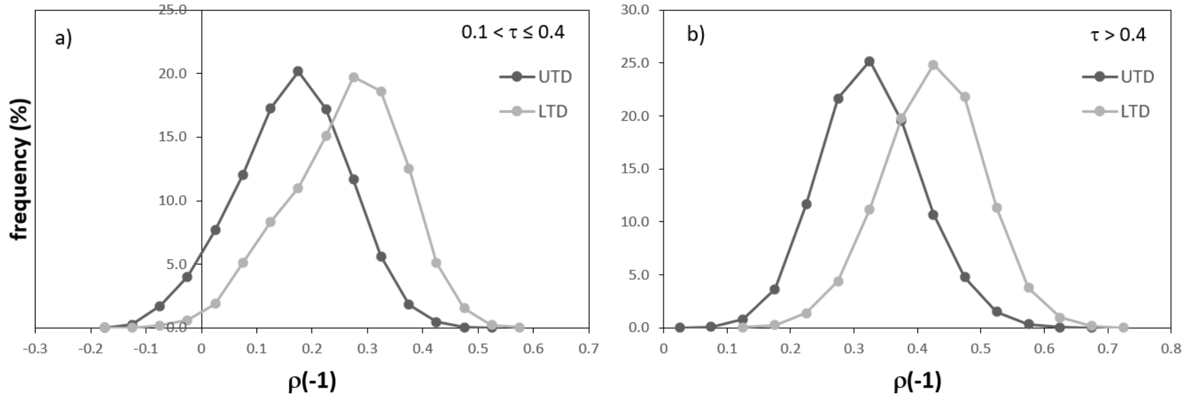

Fig. 9. Frequency distribution of the pairwise binary correlation between standardized precipitation and soil moisture lower than -1, computed separately for grid cells with UTD (dark grey lines) and LTD (light grey lines). Panel a) reports the results for the grid cells with low overall correlation (0.1 < τ ≤ 0.4), while~~ereas~~ panel b) reports the results for the grid cells with high correlation (τ > 0.4).

~~In this regard, r~~Regions such as southern France, northern UK, northern Germany and Denmark~~,~~ (where a strong LTD is observed, ~~(~~see Fig. 8)~~,~~ are appropriate candidates for a robust assessment of agricultural drought conditions based on a joint precipitation-soil moisture index, whereas some regions in central Europe (i.e., Poland, Czechia~~ Republic~~, Switzerland) may not equally benefit from the use of a joint index due to the lower importance of~~absence of~~ LTD.

Overall, ~~the fact that~~ the parametric copula fittings confirm most of the non-parametric TD patterns suggesting~~s~~ that a parametric approach is suitable for an operational implementation of a precipitation-soil moisture joint drought index over most of Europe. This implies that~~, and that~~ the proposed procedure, based on the combination of parametric and non-parametric analyses, can be considered ~~as well as~~ a reliable tool to provide meaningful insight on the potential application~~ity~~ of joint probability as detector of extreme droughts.

~~Even if, a~~At first glance, it may seem difficult to assign an ~~meaningful~~ explanation for the observed spatial patterns in LTD and UTD.~~,~~ However, the proven possibility to reasonably reconstruct these spatial patterns with a random forest classifier, starting from only a small sample of robust training data (less than 10% of the domain) and with commonly available driving features, suggests that the observed clusters are unlikely to be caused only by chance and that hidden structures may be present and may be further explored. This result is encouraging for an extension of the derived approach~~considerations~~ to other ~~spatial~~ regions of the world.

5. Summary and Conclusions

The use of combined indices based on copula seems a promising development in the field of drought detection and monitoring. In this study, we analyzed the joint probability of two variables~~quantities~~ commonly used ~~to~~ in agricultural ~~derive~~ drought ~~indices~~analyses:~~,~~ the empirical frequencies of 3-month cumulated precipitation and soil moisture.~~,~~ We~~with a special~~ focus on the probabilistic characteristics being ~~that are~~ key for ~~their usage in~~ agricultural drought ~~analyses~~studies.

The overall agreement in the marginal probability of the two standardized variables suggests that they are indeed valid candidates for the development of a joint drought index over the European domain. However, an in-depth analysis of the tail-dependence, derived with both non-parametric and parametric approaches, shows some clear spatial patterns, which have direct repercussion on the capability of such data to provide robust and coherent estimates of ~~the~~ drought extremes. In this regard, regions such as southern France, northern UK, northern Germany, and Denmark may benefit more from the joint use of the two standardized variables thanks to the observed strong low tail-dependence (i.e., increasing agreement on the left ~~ower~~ tail extremes). The joint dependence of standardized precipitation and soil moisture is well reproduced by using three common ~~nly~~ ~~used~~ copulas (Student-t, Gumbel and 180 rotated Gumbel), ~~which~~ with spatial patterns that were successfully reconstructed with a random forest classification, suggesting the presence of a structure in the outcomes not related to chance.

Code availability: The codes used for this analysis can be provided upon request via the corresponding author.

Data availability: All the data used in this study can be accessed and retrieved through the European Drought Observatory (EDO) web portal (https://edo.jrc.ec.europa.eu/gdo/php/index.php?id=2112).

Author contribution: CC designed the experiments, with inputs from AT and CDM. CC developed the codes and performed the analyses. CC prepared the manuscript, which was expanded and revised by all co-authors.

Competing interests: At least one of the (co-)authors is a member of the editorial board of Hydrology and Earth System Sciences.

References

Aas, K., Czado, C., Frigessi, A., Bakken, H.:~~, 2009.~~ Pair-copula constructions of multiple dependence,~~.~~ Ins.~~urance:~~ Math. Econ., 44(2), 182-198~~.~~, ~~doi:~~https://doi.org/https://doi.org/10.1016/j.insmatheco.2007.02.001, 2009.

Aghakouchak, A., Ciach, G., Habib, E.: Estimation of tail dependence coefficient in rainfall accumulation fields, Adv. Water Resour., 33, 1142-1149, https://doi.org/10.1016/j.advwatres.2010.07.003, 2010.

Almenda-Martín, L., Martínez-Fernández, J., Piles, M., González-Zamora, A., Benito-Verdugo, P., Gaona, J.:~~, 2022.~~ Influence of atmospheric patterns on soil moisture dynamics in Europe,~~.~~ Sci. Tot. Environ., 846, 157537~~.~~, ~~doi:~~https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.157537, 2022.

Anderson, M. C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J. R., Kustas, W. P.:~~, 2011.~~ Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States,~~.~~ J. Climate, 24(8), 2025-2044~~.~~, ~~doi:~~https://doi.org/https://doi.org/10.1175/2010JCLI3812.1, 2011.

Arnal, L., Asp, S. -S., Baugh, C., de Roo, A., Disperati, J., Dottori, F., Garcia, R., Garcia Padilla, M., Gelati, E., Gomes, G., Kalas, M., Krzeminski, B., Latini, M., Lorini, V., Mazzetti, C., Mikulickova, M., Muraro, D., Prudhomme, C., Rauthe-Schöch, A., Rehfeldt, K., Salamon, P., Schweim, C., Skoien, J. O., Smith, P., Sprokkereef, E., Thiemig, V., Wetterhall, F., Ziese, M.:~~, 2019~~. EFAS upgrade for the extended model domain – technical documentation, JRC Technical Reports, EUR 29323 EN, Publications Office of the European Union, Luxembourg, 58 pp~~.~~, ~~doi:~~https://doi.org/https://doi.org/10.2760/806324, 2019.

Bachmair, S., Tanguy, M., Hannaford, J., Stahl, K.~~, 2018.~~: How well do meteorological indicators represent agricultural and forest drought across Europe? Environ. Res. Lett., 13, 034042~~.~~, ~~doi:~~https://doi.org/10.1088/1748-9326/aaafda, 2018.~~.~~

Bateni M. M., Behmanesh J., De Michele C., Bazrafshan J., Rezaie H.~~, 2018.~~: Composite agrometeorological drought index accounting for seasonality and autocorrelation,~~.~~ J. Hydrol. Eng., 23(6), 04018020~~.~~, ~~doi:~~https://doi.org/10.1061/(ASCE)HE.1943-5584.0001654, 2018.

Box G. E. P., Jenkins G. M.: Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco, pp. 64-65, 1976.

Breiman, L.:~~, 2001.~~ Random forests,~~.~~ Machine Learn., 45, 5-32~~.~~, ~~doi:~~https://doi.org/10.1023/A:1010933404324, 2001.

Brown, J. F., Wardlow, B. D., Tadesse, T., Hayes, M. J., Reed, B. C.~~, 2008.~~: The Vegetation Drought Response Index (VegDRI): A~~a~~ new integrated approach for monitoring drought stress in vegetation,~~.~~ GISci. Remote Sens., 45(1), 16-46,~~.~~ ~~doi:~~https://doi.org/10.2747/1548-1603.45.1.16, 2008.

Burnham, K. P., Anderson, D. R.~~, 2002.~~: Model Selection and Multimodel Inference: A practical information-theoretic approach, Springer-Verlag, 488 pp,~~.~~ 2002.

Cammalleri, C., Micale, F., Vogt, J.~~, 2015.~~: On the value of combining different modelled soil moisture products for European drought monitoring,~~.~~ J. Hydrol., 525, 547-558~~.~~, ~~doi:~~https://doi.org/10.1016/j.jhydrol.2015.04.021, 2015.

Cammalleri, C., Micale, F., Vogt, J.:~~, 2016.~~ A novel soil moisture-based drought severity index (DSI) combining water deficit magnitude and frequency,~~.~~ Hydrol. Process., 30(2), 289-301~~.~~, ~~doi:~~https://doi.org/10.1002/hyp.10578, 2016.

Cammalleri, C., Vogt, J.V., Bisselink, B., de Roo, A.~~, 2017.~~: Comparing soil moisture anomalies from multiple independent sources over different regions across the globe,~~.~~ Hydrol. Earth Syst. Sci., 21, 6329-6343~~.~~, ~~doi:~~https://doi.org/10.5194/hess-21-6329-2017, 2017~~.~~.

Cammalleri, C., Arias-Muñoz, C., Marinho Ferreira Barbosa, P., De Jager, A., Magni, D., Masante, D., Mazzeschi, M., Mccormick, N., Naumann, G., Spinoni, J. and Vogt, J.~~, 2021a.~~: A revision of the Combined Drought Indicator (CDI) as part of the European Drought Observatory (EDO),~~.~~ Nat. Haz. Earth Syst. Sci., 21(2), 481-495,~~.~~ ~~doi:~~https://doi.org/10.5194/nhess-21-481-2021, 2021a.

Cammalleri, C., Spinoni J., Barbosa, P., Toreti, A., Vogt, J. V.:~~, 2021b.~~ The effects of non-stationarity on SPI for operational drought monitoring in Europe,~~.~~ Int. J. Climatol., 21, 1-13~~.~~, ~~doi:~~https://doi.org/10.1002/joc.7424, 2021b.

Carrão, H., Russo, S., Sepulcre-Canto, G., Barbosa, P.~~, 2016.~~: An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data~~.~~, Int. J. Appl. Earth Obs. Geoinf., 48, 74-84,~~.~~ ~~doi:~~https://doi.org/10.1016/j.jag.2015.06.011, 2016.

Chen, L., Guo, S.~~, 2019.~~: Copulas and Its Application in Hydrology and Water Resources. Springer Water, 290 pp, 2019.

Dash, S._S., Sahoo, B., Raghuwanshi, N._S.:~~, 2019.~~ A SWAT-Copula based approach for monitoring and assessment of drought propagation in an irrigation command,~~.~~ Ecol. Eng., 127, 417-430~~.~~, ~~doi:~~https://doi.org/ 10.1016/j.ecoleng.2018.11.021, 2019.

De Michele, C., Salvadori, G.:~~, 2003.~~ A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas~~.~~, J. Geophys. Res. Atmos., 108 (D2), 4067~~.~~, ~~doi:~~https://doi.org/10.1029/2002JD002534, 2003.

De Roo, A._P._J., Wesseling, C., van Deusen, W.:~~, 2000.~~ Physically based river basin modelling within a GIS: The LISFLOOD model,~~.~~ Hydrol. Process., 14, 1981-1992,~~.~~ ~~doi:~~https://doi.org/10.1002/1099-1085(20000815/30)14:11/12<1981::AID-HYP49>3.0.CO;2-F, 2000.

Dißman, J., Brechmann, E._C., Czado, C., Kurowicka, D.~~, 2013.~~: Selecting and estimating regular vine copulae and application to financial returns,~~.~~ Comput. Stat. Data Analysis, 59, 52-69~~.~~, ~~doi:~~https://doi.org/10.1016/j.csda.2012.08.010, 2013.

Dixit, S., Jayakumar, K._V.~~, 2021.~~: Spatio-temporal analysis of copula-based probabilistic multivariate drought index using CMIP6 model,~~.~~ Int. J. Climatol., 42(8), 4333-4350~~.~~, ~~doi:~~https://doi.org/10.1002/joc.7469, 2021.

Dutra, E., Viterbo, P., Miranda, P._M._A.~~, 2008.~~: ERA-40 reanalysis hydrological applications in the characterization of regional drought,~~.~~ Geophys. Res. Lett., 35(19), L19402~~.~~, ~~doi:~~https://doi.org/10.1029/2008GL035381, 2008.

Farahmand, A., AghaKouchak, A.:~~, 2015.~~ A generalized framework for deriving nonparametric standardized drought indicators,~~.~~ Adv. Water Resour., 76, 140-145~~.~~, ~~doi:~~https://doi.org/10.1016/j.advwatres.2014.11.012, 2015.

Frahm, G., Junker, M., Schmidt, R.: Estimating the tail-dependence coefficient: properties and pitfalls, Insur. Math. Econ., 37(1), 80-100, https://doi.org/10.1016/j.insmatheco.2005.05.008, 2005.

Gaona, J., Quintana-Seguí, P., Escorihuela, M._J., Boone, A., Llasat, M._C.~~, 2022.~~: Interactions between precipitation, evapotranspiration and soil-moisture-based indices to characterize drought with high-resolution remote sensing and land-surface model data,~~.~~ Nat. Hazard Earth Syst. Sci., 22, 3461-3485~~.~~, ~~doi:~~https://doi.org/10.5194/nhess-22-3461-2022, 2022.

Genest, C., Favre, A. C., Béliveau, J., Jacques, C.~~, 2007.~~ : Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data,~~.~~ Water Resour. Res., 43(9), 1-12~~.~~, ~~doi:~~https://doi.org/10.1029/2006WR005275, 2007.

Halwatura, D., McIntyre, N., Lechner, A. M., Arnold, S.~~, 2017.~~: Capability of meteorological drought indices for detecting soil moisture droughts,~~.~~ J. Hydrol. Reg. Studies, 12, 396-412~~.~~, ~~doi:~~https://doi.org/10.1016/j.ejrh.2017.06.001, 2017.

Hao, Z., AghaKouchak, A.~~, 2013.~~: Multivariate Standardized Drought Index: A parametric multi-index model,~~.~~ Adv. Water Resour., 57, 12-18, ~~doi:~~https://doi.org/10.1016/j.advwatres.2013.03.009, 2013.

Hao, Z., Singh, V. P.~~, 2015.~~: Drought characterization from a multivariate perspective: A review,~~.~~ J. Hydrol., 527, 668-678,~~.~~ ~~doi:~~https://doi.org/10.1016/j.jhydrol.2015.05.031, 2015.

Ji, L., Peters A. J.:~~, 2003.~~ Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices,~~.~~ Remote Sens. Environ., 87, 85-98~~.~~, ~~doi:~~https://doi.org/10.1016/S0034-4257(03)00174-3, 2003.

Joe, H.:~~, 2015.~~ Dependence Modeling with Copulas,~~.~~ CRC Press, Taylor and Francis, 480 pp, 2015.

Kanthavel, P., Saxena, C. K., Singh, R. K.~~, 2022.~~: Integrated drought index based on vine copula modelling, ~~.~~Int. J. Climatol., 42(16), 9510-9529~~.~~, ~~doi:~~https://doi.org/10.1002/joc.7840, 2022.

Kao, S. C., Govindaraju, R. S.~~, 2010.~~: A copula-based joint deficit index for droughts,~~.~~ J. Hydrol., 380, 121–134,~~.~~ ~~doi:~~https://doi.org/10.1016/j.jhydrol.2009.10.029, 2010.

Kwon, M., Kwon, H. -H., Han, D.~~, 2019.~~: Spatio-temporal drought patterns of multiple drought indices based on precipitation and soil moisture: A case study in South Korea~~.~~, Int. J. Climatol., 39(12), 4669-4687,~~.~~ ~~doi:~~https://doi.org/ 10.1002/joc.6094, 2019.

Laimighofer, J., Laaha, G.:~~, 2022.~~ How standard are standardized drought indices? Uncertainty components for the SPI & SPEI case~~.~~, J. Hydrol., 613(A), 128385~~.~~, ~~doi:~~https://doi.org/10.1016/j.jhydrol.2022.128385, 2022.

Manning, C., Widmann, M., Bevacqua, E., van Loon, A. F., Maraun, D., Vrac, M.:~~, 2018.~~ Soil moisture drought in Europe: A compound event of precipitation and potential evapotranspiration on multiple time scales,~~.~~ J. Hydrometeorol., 19(8), 1255-1271,~~.~~ ~~doi:~~https://doi.org/10.1175/JHM-D-18-0017.1, 2018.

Mishra, A. K., Singh, V. P.:~~, 2010.~~ A review of drought concepts,~~.~~ J. Hydrol., 391, 202-216~~.~~, ~~doi:~~https://doi.org/10.1016/j.rse.2016.02.064, 2010.

Mo, K. C., Lettenmaier, D. P.~~, 2013.~~: Objective drought classification using multiple land surface models,~~.~~ J. Hydrometeorol., 15, 990-1010~~.~~, ~~doi:~~https://doi.org/10.1175/JHM-D-13-071.1, 2013.

Mo, K. C., Lyon, B.~~, 2015.~~: Global meteorological drought prediction using the North American multi-model ensemble,~~.~~ ——J. Hydrometeorol., 16, 1409–1424~~.~~, ~~doi:~~https://doi.org/10.1175/JHM-D-14-0192.1, 2015.

Mohammed, S., Alsafadi, K., Enaruvbe, G. O., Bashir, B., Elbeltagi, A., Széles, A., Alsalman, A., Harsanyi, E.~~, 2022.~~: Assessing the impacts of agricultural drought (SPI/SPEI) on maize and wheat yields across Hungary,~~.~~ Sci. Rep. 12, 8838~~.~~, ~~doi:~~https://doi.org/ 10.1038/s41598-022-12799-w, 2022.

Nelsen, R. G.:~~, 2006.~~ An introduction to copulas,~~.~~ Springer Series in Statistics, Springer-Verlag, New York, 272 pp,~~.~~ ~~doi:~~https://doi.org/10.1007/0-387-28678-0, 2006.

Panu, U. S., Sharma, T. C.~~, 2002.~~: Challenges in drought research: Some perspectives and future directions~~.~~, Hydrol. Sci. J., 47, S19-S30~~.~~, ~~doi:~~https://doi.org/10.1080/02626660209493019, 2002.

Pieper, P., Düsterhus, A., Baehr, J.~~, 2020.~~: A universal Standardized Precipitation Index candidate distribution function for observations and simulations,~~.~~ Hydrol. Earth Syst. Sci., 24, 4541-4565,~~.~~ ~~doi:~~https://doi.org/10.5194/hess-24-4541-2020, 2020.

Poulin, A., Huard, D., Favre, A. C., Pugin, S.: Importance of tail dependence in bivariate frequency analysis, J. Hydrol. Eng., 12(4), 394-403, https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(394), 2007.

Quiring, S. M., Papakryiakou, T. N.:~~, 2003.~~ An evaluation of agricultural drought indices for the Canadian prairies,~~.~~ Agr. Forest Meteorol., 118(1-2), 49-62,~~.~~ ~~doi:~~https://doi.org/10.1016/S0168-1923(03)00072-8, 2003.

Ravelo, A. C., Decker, W. L.~~, 1979.~~: The probability distribution of a soil moisture index~~.~~, Agr. Meteorol., 20(4), 301-312,~~.~~ ~~doi:~~https://doi.org/10.1016/0002-1571(79)90004-9, 1979.

Rembold, F., Meroni, M., Urbano, F., Csak, G., Kerdiles, H., Perez-Hoyos, A., Lemoine, G., Leo, O., Negre, T.~~, 2019.~~: ASAP: A new global early warning system to detect anomaly hot spots of agricultural production for food security analysis,~~.~~ ~~—~~ Agr. Syst. 168, 247–257~~.~~, ~~doi:~~https://doi.org/10.1016/j.agsy.2018.07.002, 2019.

Ribeiro, A., Pires, C.~~, 2016.~~: Seasonal drought predictability in Portugal using statistical–dynamical techniques,~~.~~ Phys. Chem. Earth, 94, 155–166,~~.~~ ~~doi:~~https://doi.org/10.1016/j.pce.2015.04.003, 2016.

Sadri, S., Pan, M., Wada, Y., Vergopolan, N., Sheffield, J., Famiglietti, J. S., Kerr, Y., Wood, E. F.~~, 2020.~~: A global near-real-time soil moisture index monitor for food security using integrated SMOS and SMAP~~.~~, Remote Sens. Environ., 246, 111864~~.~~, ~~doi:~~https://doi.org/10.1016/j.rse.2020.111864, 2020.

Salvadori G., De Michele C.~~, 2004.~~: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events,~~.~~ Wat. Resour. Res., 40, W12511, ~~doi:~~https://doi.org/10.1029/2004WR003133, 2004.

Salvadori, G., De Michele, C., Kottegoda, N. T., Rosso, R.~~, 2007.~~: Extremes in Nature: An approach using Copulas,~~.~~ Water Science and Technology Library Series, vol. 56,~~.~~ Springer, Dordrecht, 292 pp.,~~.~~ ISBN: 978-1-4020-4415-1, 2007.

Schmidt, R., Stadtmueller, U.: Non-parametric estimation of tail dependence, Scandinav. J. Stat., 33(2), 307-335, https://doi.org/10.1111/j.1467-9469.2005.00483.x, 2006.

Sehler, R., Li, J., Reager, J. T., Ye, H.:~~, 2019.~~ Investigating relationship between soil moisture and precipitation globally using remote sensing observations~~.~~, J. Cont. Water Res. Edu., 168(1), 106-118,~~.~~ ~~doi:~~https://doi.org/10.1111/j.1936-704X.2019.03324.x, 2019.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., Teuling, A. J.~~, 2010.~~: Investigating soil moisture–climate interactions in a changing climate: ~~a~~A review,~~.~~ Earth-Sci. Rev., 99, 125-161,~~.~~ ~~doi:~~https://doi.org/10.1016/j.earscirev.2010.02.004, 2010.

Sepulcre-Cantó, G., Horion, S., Singleton, A., Carrão, H., Vogt, J.~~, 2012.~~: Development of a c~~C~~ombined d~~D~~rought i~~I~~ndicator to detect agricultural drought in Europe,~~.~~ Nat. Hazard Earth Syst. Sci., 12, 3519-3531,~~.~~ ~~doi:~~https://doi.org/10.5194/nhess-12-3519-2012, 2012.

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. Environ. Res. Risk Assess., 22(6), 671–688, https://doi.org/10.1007/s00477-007-0176-4, 2008.

Serinaldi, F., Bárdossy, A., Kilsby, C. G.: Upper tail dependence in rainfall extremes: would we know it if we saw it? Stoch. Environ. Res. Risk Assess., 29, 1211-1233, https://doi.org/10.1007/s00477-014-0946-8, 2015.

Sheffield, J., Goteti, G., Wen, F., Wood, E. F.~~, 2004.~~: A simulated soil moisture based drought analysis for the United States, J. Geophys. Res., 109, D24108,~~.~~ ~~doi:~~https://doi.org/10.1029/2004JD005182, 2004.

Sheffield, J., Wood, E. F.~~, 2007.~~: Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle~~.~~, J. Geophys. Res., ~~-~~112, D17115,~~.~~ ~~doi:~~https://doi.org/10.1029/2006JD008288, 2007.

Sims, A. P., Niyogi, D. S., Raman, S.~~, 2002.~~: Adopting drought indices for estimating soil moisture: A North Carolina case study,~~.~~ Geophys. Res. Lett., 29(8), 24-1-24-4~~.~~, ~~doi:~~https://doi.org/10.1029/2001GL013343, 2002.

Sivakumar, M. V. K., Motha, R. P., Wilhite, D. A., Wood, D. A.: ~~(Eds.). 2011.~~ Agricultural Drought Indices~~.~~, Proceedings of the WMO/UNISDR Expert Group Meeting on Agricultural Drought Indices, 2-4 June 2010, Murcia, Spain: Geneva, Switzerland: World Meteorological Organization~~.~~, AGM-11, WMO/TD No. 1572,~~;~~ WAOB-2011,~~.~~ 197 pp, 2011.

~~Schmidt, R., Stadtmueller, U., 2006.: Non-parametric estimation of tail dependence,. Scandinav. J. Stat., 33(2), 307-335,. doi:https://doi.org/10.1111/j.1467-9469.2005.00483.x, 2006.~~

Soľáková, T., De Michele, C., Vezzoli, R.: Comparison between parametric and nonparametric approaches for the calculation of two drought indices: SPI and SSI, J. Hydrol. Eng., 19(9), 04014010, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000942, 2014.

Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., van Loon, A. F., Stahl, K.~~, 2015.~~: Candidate distributions for climatological drought indices (SPI and SPEI),~~.~~ Int. J. Climatol., 35, 4027-4040,~~.~~ ~~doi:~~https://doi.org/10.1002/joc.4267, 2015.

Stoica, P., Selen, Y.~~, 2004.~~: Model-order selection: a review of information criterion rules,~~.~~ IEEE Signal Proces. Mag., 21(4), 36-47~~.~~, ~~doi:~~https://doi.org/10.1109/MSP.2004.1311138, 2004.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D.~~, 2002.~~: The drought monitor,~~.~~ Bull. Am. Meteorol. Soc., 83, 1181-1190~~.~~, ~~doi:~~https://doi.org/10.1175/1520-0477-83.8.1181, 2002.

Thielen, J., Bartholmes, J., Ramos, M. -H., De Roo A. P. J.:~~, 2009.~~ The European flood alert system – part 1: concept and development,~~.~~ Hydrol. Earth Syst. Sci., 13, 125-140~~.~~, ~~doi:~~https://doi.org/10.5194/hess-13-125-2009, 2009.

Thieming, V., Gomes, G. N., Skøien, J., Ziese, M., Rauthe-Schöch, A., Rustemeier, E., Rehfeldt, K., Walawender, J. P., Kolbe, C., Pichon, D., Schweim, C., Salamon, P.~~, 2022.~~: EMO-5: a high-resolution multi-variable gridded meteorological dataset for Europe~~.~~, Earth Syst. Sci. Data, 14, 3249-3272,~~.~~ ~~doi:~~https://doi.org/10.5194/essd-14-3249-2022, 2022.

Tian, L., Yuan, S., Quiring, S. M.:~~, 2018.~~ Evaluation of six indices for monitoring agricultural drought in the south-central United States,~~.~~ Agr. For. Meteorol., 249, 107-119~~.~~, ~~doi:~~https://doi.org/10.1016/j.agrformet.2017.11.024, 2018.

van der Wiel, K., Batelaan, T. J., Wanders, N.~~, 2022.~~: Large increases of multi-year droughts in north-western Europe in a warmer climate,~~.~~ Clim. Dynam., 60, 1781-1800~~.~~, ~~doi:~~https://doi.org/10.1007/s00382-022-06373-3, 2022.

Vicente-Serrano S. M., Beguería, S., López-Moreno, J. I.~~, 2010.~~: A Multi-scalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index ~~–~~ SPEI,~~.~~ J. Climate, 23, 1696-1718~~.~~, ~~doi:~~https://doi.org/10.1175/2009JCLI2909.1, 2010.

Wang, H., Rogers, J. C., Munroe, D. K.~~, 2015.~~: Commonly used drought indices as indicators of soil moisture in China,~~.~~ Hydrometeorol., 16(3), 1397-1408~~.~~, ~~doi:~~https://doi.org/10.1175/JHM-D-14-0076.1, 2015.

Wilhite, D. A., Glantz, M. H.~~, 1985.~~: Understanding the drought phenomenon: The role of definitions,~~.~~ Water International, 10(3), 111–120, 1985.

World Meteorological Organization (WMO)~~, 2012.~~: Standardized Precipitation Index User Guide (WMO n. 1090), Geneva, 24 pp, 2012.

World Meteorological Organization (WMO), Global Water Partnership (GWP)~~, 2016.~~: Handbook of Drought Indicators and Indices (M. Svoboda and B.A. Fuchs). Integrated Drought Management Programme (IDMP), Integrated Drought Management Tools and Guidelines Series 2,~~.~~ Geneva, 52 pp, 2016.

Xia, Y., Ek, M. B., Peters-Lidard, C. D., Mocko, D., Svoboda, M., Sheffield, J., Wood, E. F.~~, 2014.~~: Application of USDM statistics in NLDAS-2: optimal blended NLDAS drought index over the continental United States,~~.~~ J. Geophys. Res. Atmos., 119(6), 2947-2965~~.~~, ~~doi:~~https://doi.org/10.1002/2013JD020994, 2014.

Yuan, X., Wood, E. F.~~, 2013.~~: Multimodel seasonal forecasting of global drought onset~~.~~, Geophys. Res. Lett., 40, 4900–4905,~~.~~ ~~doi:~~https://doi.org/10.1002/grl.50949, 2013.

Zargar, A., Sadiq, R., Naser, B., Khan, F. I.~~, 2011.~~: A review of drought indices~~.~~, Environ. Rev., 19, 333-349,~~.~~ ~~doi:~~https://doi.org/ 10.1139/A11-013, 2011.