Soil Moisture Drought in Europe: A Compound Event of Precipitation and Potential Evapotranspiration on Multiple Time Scales

COLIN MANNING AND MARTIN WIDMANN
School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom

EMANUELE BEVACQUA
Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

ANNE F. VAN LOON
Department of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom

DOUGLAS MARAUN
Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

MATHIEU VRAC
Laboratoire des Sciences du Climat et de l’Environnement (LSCE-IPSL), Centre d’Etudes de Saclay, Gif-sur-Yvette, France

(Manuscript received 27 January 2018, in final form 26 April 2018)

ABSTRACT

Compound events are extreme impacts that depend on multiple variables that need not be extreme themselves. In this study, we analyze soil moisture drought as a compound event of precipitation and potential evapotranspiration (PET) on multiple time scales related to both meteorological drought and heat waves in wet, transitional, and dry climates in Europe during summer. Drought indices that incorporate PET to account for the effect of temperature on drought conditions are sensitive to global warming. However, as evapotranspiration (ET) is moisture limited in dry climates, the use of such drought indices has often been criticized. We therefore assess the relevance of the contributions of both precipitation and PET to the estimation of soil moisture drought. Applying a statistical model based on pair copula constructions to data from FluxNet sites in Europe, we find at all sites that precipitation exerts the main control over soil moisture drought. At wet sites PET is additionally required to explain the onset, severity, and persistence of drought events over different time scales. At dry sites, where ET is moisture limited in summer, PET does not improve the estimation of soil moisture. In dry climates, increases in drought severity measured by indices incorporating PET may therefore not indicate further drying of soil but the increased availability of energy that can contribute to other environmental hazards such as heat waves and wildfires. We therefore highlight that drought indices including PET should be interpreted within the context of the climate and season in which they are applied in order to maximize their value.

1. Introduction

Soil moisture plays a critical role in agriculture and the variability of temperature (Seneviratne et al. 2006). As soil moisture observations are sparse, soil moisture drought must be monitored and quantified using indirect methods. These include land surface models that provide physically based estimates of soil moisture (Mitchell et al. 2004; Sheffield et al. 2014) and drought indices that are used as a proxy of soil moisture (Dai et al. 2004; Vicente-Serrano et al. 2012). The simplicity of drought indices is advantageous, but it also leaves their output open to interpretation, which we assess in this article.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Colin Manning, cjm317@bham.ac.uk

DOI: 10.1175/JHM-D-18-0017.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
Many studies have highlighted the multivariate nature of soil moisture drought as well as the importance of incorporating temperature in drought analysis (Seneviratne et al. 2012b; Teuling et al. 2013; AghaKouchak et al. 2014). Through the inclusion of temperature via potential evapotranspiration (PET) in drought indices such as the standardized precipitation evapotranspiration index (SPEI; Vicente-Serrano et al. 2010), the Palmer drought severity index (PDSI; Palmer 1965), and the reconnaissance drought index (RDI; Tsakiris and Vangelis 2005), studies have been able to analyze how drought conditions may change in a warming climate at regional and global scales (Dai et al. 2004; Dai 2011; Sheffield et al. 2012; Dai 2013; Trenberth et al. 2014; Törnros and Menzel 2014; Vicente-Serrano et al. 2014; Zarch et al. 2015; Stagge et al. 2017).

With increasing temperatures, drought events are expected to set in quicker (Trenberth et al. 2014) and become more severe based on indices incorporating PET (Törnros and Menzel 2014; Zarch et al. 2015; Vicente-Serrano et al. 2014). However, the meaning of this increase in severity for soil moisture according to these indices can be quite unclear due to the differing contribution of PET to soil moisture drought in moisture-limited and energy-limited climates (Seneviratne et al. 2010). Understanding these differences can help in the interpretation of future changes depicted by drought indices and the potential implications they bring at a regional level where impacts of drought are felt.

Soil moisture drought refers to moisture deficits in the upper layer of soil known as the root zone. Soil moisture in the root zone is primarily controlled by antecedent precipitation while excesses in evapotranspiration (ET), related to high temperatures, are required to explain the severity of a negative soil moisture anomaly (Teuling et al. 2013; Seneviratne et al. 2012b). The contribution of ET to soil moisture drought depends on the availability of moisture in the soil for ET to take place (Seneviratne et al. 2010). PET measures the evaporative demand of the atmosphere and indicates the amount of ET that would occur given an unlimited water supply. Under moisture-limited conditions, values of PET and ET can diverge where ET may verge to zero while PET can continue to rise with an increase in temperature (Seneviratne et al. 2010). In such dry conditions, PET and temperature can therefore have little contribution to the estimation of soil moisture (Luo et al. 2017) and lead to drying biases in terms of moisture levels in soil when incorporated into drought indices (Sheffield et al. 2012; Seneviratne 2012).

Describing soil moisture with drought indices requires one to account for antecedent meteorological conditions that soil moisture holds memory of. This is done using integrations of a climatic water balance (precipitation minus PET) varying in length from 1 to 24 months (e.g., SPEI), or through the use of recursive models (e.g., PDSI). The selection of this integration length for indices such as the SPEI is important; a length that is too short will not capture drought persistence while longer periods can include redundant information (Törnros and Menzel 2014). Studies using the SPEI or RDI to represent soil moisture generally use integration periods between 3 and 6 months (Hirschi et al. 2011; Törnros and Menzel 2014). The PDSI is calculated with monthly integrations and it can hold memory of the previous winter and spring in summer months (Dai et al. 2004).

The use of a climatic water balance implies that PET influences soil moisture over the same time scale as precipitation. However, drying of soil occurs on a daily time scale where excesses in ET can be driven by days of extreme temperature that are filtered out through the use of longer integration periods. Such a feature of long integrations of the climatic water balance can lead to an inability to capture both future changes in drying that may cause droughts to set in quicker in a warmer climate and the occurrence of flash droughts associated with short periods of warm temperature and rapidly decreasing soil moisture (Mo and Lettenmaier 2016).

High temperatures driving excesses in ET can be partly attributed to land–atmosphere interactions induced by deficits in precipitation. By leading to dry soil conditions, low antecedent precipitation is associated with an increased probability of hot days (Hirschi et al. 2011; Mueller and Seneviratne 2012; Whan et al. 2015; Ford and Quiring 2014), amplified extreme temperatures, and the persistence of heat waves (Miralles et al. 2014; Lorenz et al. 2010) that, in turn, can further deplete soil moisture where moisture is available.

The individual roles of precipitation and PET, and that of their dependence driven through land–atmosphere interactions, highlight the compound nature of soil moisture drought. Compound events are a class of events receiving an increased amount of attention in recent times. They encompass a broad range of impacts whose risk is influenced by the dependence between their drivers (Wahl et al. 2015; Hillier et al. 2015; Martius et al. 2016; Bevacqua et al. 2017). Understanding the dependence between hot and dry conditions and their impacts is of great importance. Overlooking nonlinear dependence between hot and dry conditions and crop yields leads to an underestimation of risk in reduced crop yields (Zscheischler et al. 2017), while the bivariate risk of hot and dry summers is underestimated when treating them independently (Zscheischler and Seneviratne 2017). Underlining this importance are findings of an increase in the concurrence of drought and heat wave events (Mazdiyasni and AghaKouchak 2015). Such an increase brings a potential rise in the risk of associated impacts, as the impact arising...
from the combination of drought and heat wave events can be greater than the sum of the impacts arising from individual events (Hegerl et al. 2011; Zscheischler et al. 2014).

In this study, we analyze soil moisture drought as a compound event of meteorological drought and heat waves in Europe. We use a conceptual framework developed in Bevacqua et al. (2017) and apply it to a system in which we describe soil moisture as a function of precipitation integrated over preceding months and PET integrated over recent days. This conceptual framework allows us to capture days of extreme temperature within the PET variable and its dependence on antecedent conditions. The framework is implemented via a multivariate statistical model based on pair copula constructions (PCCs; Aas et al. 2009). Copula-based methods, in particular PCCs, provide much flexibility in modeling multidimensional systems (Aas et al. 2009; Bevacqua et al. 2017), including the representation of nonlinear dependence associated with hot and dry conditions (Stagge et al. 2015; Zscheischler and Seneviratne 2017). Their use has therefore become quite prominent in the analysis of compound events (Serinaldi et al. 2009; AghaKouchak et al. 2014; Bevacqua et al. 2017; Zscheischler et al. 2017; Zscheischler and Seneviratne 2017). More details on copula, PCCs, and the advantages of their use can be found in the following sections.

We assess the compound nature of soil moisture drought in Europe during the summer months June–August (JJA) at locations in wet, transitional, and dry climates. We aim to demonstrate the individual contributions of precipitation and PET to the estimation of soil moisture drought and highlight where, when, and over what integration period lengths PET and its dependence with precipitation are important for the estimation of soil moisture in a statistical setting. In doing so, we aim to characterize the compound nature of soil moisture drought in differing climates during summer to provide information that may aid with the interpretation of drought indices incorporating PET and allow further insight to be gained from such indices.

The paper is organized as follows: the data employed in this study as well as the statistical methods involved are described in section 2, the main results are presented in section 3, and a summary and conclusions are provided in section 4.

2. Data and methods
a. Dataset

We employed the FluxNet dataset (Baldocchi et al. 2001) for this study using 11 stations situated across Europe. The selection of sites was based both on an initial review of data quality and length across many sites as well as the recommendations of Rebel et al. (2012). Table 1 provides a summary of the site characteristics. To aid the interpretation of the results, we classify the sites as wet or dry based on values of soil moisture. Locations are provided in Fig. 1. At each site, soil moisture measurements from the top 30 cm of soil are provided along with precipitation data as well as the variables required for the calculation of PET via the reference crop Penman–Monteith equation, as described in Zotarelli et al. (2010). These variables include incoming solar radiation, temperature, wind speed, and relative humidity. Among the selected sites, two general land cover types are available: grassland and forest. The data used here are at a daily resolution. We use soil moisture values for the summer months of JJA. For the contributing meteorological variables, we used observations that extend back into previous months in order to calculate integration periods prior to a given soil moisture observation.

b. Conceptual model

We design a conceptual model, based on a framework developed by Bevacqua et al. (2017), in which we describe soil moisture $h$ as an impact of contributing

<table>
<thead>
<tr>
<th>Site</th>
<th>Site name</th>
<th>Lat</th>
<th>Lon</th>
<th>Site type</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Dripsey, Ireland</td>
<td>51.99°N</td>
<td>8.75°W</td>
<td>Grassland</td>
</tr>
<tr>
<td>b</td>
<td>Hainich, Germany</td>
<td>51.08°N</td>
<td>10.45°E</td>
<td>Forest</td>
</tr>
<tr>
<td>c</td>
<td>Klingenberg, Germany</td>
<td>50.89°N</td>
<td>13.52°E</td>
<td>Grassland</td>
</tr>
<tr>
<td>d</td>
<td>Oensingen, Switzerland</td>
<td>47.28°N</td>
<td>7.73°E</td>
<td>Grassland</td>
</tr>
<tr>
<td>e</td>
<td>Pang/Lambourne, United Kingdom</td>
<td>51.45°N</td>
<td>1.27°W</td>
<td>Forest</td>
</tr>
<tr>
<td>f</td>
<td>Le Bray, France</td>
<td>44.72°N</td>
<td>0.77°W</td>
<td>Forest</td>
</tr>
<tr>
<td>g</td>
<td>Amplero, Italy</td>
<td>41.9°N</td>
<td>13.6°W</td>
<td>Grassland</td>
</tr>
<tr>
<td>h</td>
<td>Las Majadas del Tietar, Spain</td>
<td>39.94°N</td>
<td>5.77°W</td>
<td>Savanna/Grassland</td>
</tr>
<tr>
<td>i</td>
<td>Bugacpuszta, Hungary</td>
<td>46.69°N</td>
<td>19.6°E</td>
<td>Grassland</td>
</tr>
<tr>
<td>j</td>
<td>Mitra IV Tojal, Portugal</td>
<td>38.48°N</td>
<td>8.02°W</td>
<td>Grassland</td>
</tr>
<tr>
<td>k</td>
<td>Vall d’Alinya, Spain</td>
<td>42.15°N</td>
<td>1.45°E</td>
<td>Grassland</td>
</tr>
</tbody>
</table>
meteorological variables $Y$. The contributing meteorological variables include a short-term precipitation variable ($Y_1^{PS}$), a long-term precipitation variable ($Y_2^{PL}$), and a PET variable ($Y_3^{PET}$) that are integrated over periods $L_1$, $L_2$, and $L_3$, respectively. A schematic representation of the variables modeled is given in Fig. 2.

The quantities $Y_1^{PS}$ and $Y_2^{PL}$ respectively represent the most recent and antecedent precipitation that influence the short- and long-term variability of soil moisture. Their respective integration periods $L_1$ and $L_2$ are nonoverlapping. Two precipitation variables are required to better capture the temporal distribution of precipitation that would otherwise be lost using one long-term integration only.

The quantity $Y_3^{PET}$ represents PET integrated over the period $L_3$. PET is often employed as an estimate of ET in drought indices given the lack of ET data. We calculate PET using the reference crop Penman–Monteith equation as defined in Zotarelli et al. (2010), where it is derived from incoming solar radiation, temperature, wind, and the actual and saturation vapor pressures. The quantity $Y_3^{PET}$ includes temperature within its calculation and so can capture heat waves that influence the drying of soil moisture. Depending on the question at hand, the integration length $L_3$ is varied; more details of this are given in section 2e.

c. Copula

A copula is a multivariate distribution function that describes the dependence structure between random variables independent of their marginal behavior. The selection of structure of dependence, defined by the given copula family, is hence not constrained by the choice of the marginal distribution functions. This feature provides much flexibility in modeling multivariate distributions as it allows for the application of complex marginal distributions (Salvadori et al. 2007). According to Sklar’s theorem (Sklar 1959), the joint cumulative distribution function (CDF) $F$ of an $n$-dimensional random vector $Y = (Y_1, \ldots, Y_n)$ with marginal CDFs $F_1, \ldots, F_n$ can be written as

$$F(y_1, \ldots, y_n) = C(u_1, \ldots, u_n),$$

where $C$ is an $n$-dimensional copula and $u_i = F_i(y_i)$ are uniformly distributed variables in the domain $[0, 1]$. Provided the marginal distributions $F_i$ are continuous, the multivariate probability density function (PDF) may be decomposed as

$$f(y_1, \ldots, y_n) = f_1(y_1) \times \cdots \times f_n(y_n) \times c[F_1(y_1), \ldots, F_n(y_n)],$$

where $c$ is the copula density and $f_i$ are marginal PDFs.

There exists a large number of bivariate copula families that each provide an explicit formulation for a given structure of dependence. However, the number of copula families applicable to a dimension of three or higher is quite limited (Aas et al. 2009) and in contrast to reality, where heterogeneous dependence structures often exist, each copula will usually assume the same structure of dependence between all marginals (Aas et al. 2009; Acar et al. 2012; Noh et al. 2013; Bevacqua et al. 2017). We therefore employ PCCs that provide higher flexibility than multivariate copulas and more simplicity in terms of the selection of dependence structure (Aas et al. 2009; Noh et al. 2013).

d. Pair copula constructions

PCCs, initially proposed by Joe (1997), allow us to mathematically decompose an $n$-dimensional copula density into a product of $n(n-1)/2$ bivariate copulas, of which some are conditional. They allow much flexibility in modeling multidimensional distributions (Aas et al. 2009; Bevacqua et al. 2017) and provide a means to easily calculate quantiles of the multivariate conditional distribution of an impact $h$ given values of $Y$ (Noh et al. 2013; Bernard and Czado 2015; Kraus and Czado 2017; Fischer et al. 2017).
For a high-dimensional distribution, there exists a significant number of decompositions of a multivariate PDF into a PCC that are each mathematically equivalent to one another (Aas et al. 2009). Two special types of decompositions called vines exist for PCCs, the canonical vine (C-vine) and the D-vine (Kurowicka and Cooke 2005). Throughout this study we employ a D-vine decomposition. For the four possible D-vine special types of decompositions called vines exist for PDF into a PCC that are each mathematically equivalent to one another (Aas et al. 2009). Two bivariate copulas $c_{32}$, $c_{21}$, and $c_{1h}$ are fitted to capture the respective pairwise dependencies of the variables $u_3$, $u_2$, $u_1$, and $u_h$. Second, the conditional bivariate copulas $c_{31|2}$ and $c_{21|1}$ are then fitted to the respective conditional probabilities $u_{3|2}$, $u_{1|2}$, $u_{2|1}$, and $u_{h|1}$. These variables are obtained from the conditional distributions given by the partial differentiation of the respective unconditional bivariate copula with respect to the conditioning variable:

$$u_{3|2} = F_{3|2}(u_3|u_2) = \frac{\partial}{\partial u_2} C_{32}(u_3, u_2),$$

$$u_{1|2} = F_{1|2}(u_1|u_2) = \frac{\partial}{\partial u_2} C_{21}(u_2, u_1),$$

$$u_{2|1} = F_{2|1}(u_2|u_1) = \frac{\partial}{\partial u_1} C_{21}(u_2, u_1),$$

and

$$u_{h|1} = F_{h|1}(u_h|u_1) = \frac{\partial}{\partial u_1} C_{1h}(u_1, u_h).$$

(4)

In the final step of the estimation procedure, a copula $c_{3h|21}$ is fitted to the conditional probabilities $u_{3|21}$ and $u_{h|21}$. These conditional probabilities are obtained from the conditional distributions given from the partial differentiation of the respective conditional bivariate copula with respect to the conditioning variable:

$$u_{3|21} = F_{3|21}(u_3|u_2, u_1) = \frac{\partial C_{3|21}[F_{3|2}(u_3|u_2), F_{2|1}(u_1|u_2)]}{\partial F_{1|2}(u_1|u_2)},$$

and

$$u_{h|21} = F_{h|21}(u_h|u_2, u_1) = \frac{\partial C_{h|21}[F_{2|1}(u_2|u_1), F_{h|1}(u_h|u_1)]}{\partial F_{2|1}(u_2|u_1)}.$$  

(5)

From the conditional copula $c_{3h|21}$, the conditional CDF $F_{h|321}$ can be obtained through partial differentiation of $c_{3h|21}$ with respect to $F_{3|21}$:

$$F_{h|321}(u_h|u_3, u_2, u_1) = \frac{\partial C_{3|21}[F_{3|2}(u_3|u_2), F_{2|1}(u_1|u_2)]}{\partial F_{3|21}(u_3|u_2, u_1)}. $$

(6)
As will be shown, all unconditional and conditional bivariate CDFs described above are required when sampling from the PCC.

2) SAMPLING FROM PCC

Sampling variables \( u_3, u_2, u_1, \) and \( u_h \) from the four-dimensional D-vine PCC repeatedly results in four uniformly distributed variables that exhibit a dependence structure specified by the given PCC. Algorithms proposed by Aas et al. (2009) provide a convenient means of sampling variables \( u_3, u_2, u_1, \) and \( u_h \). Within these algorithms, variables \( w_3, w_2, w_1, \) and \( w_h \) are first drawn independently from a random uniform distribution on \([0, 1]\). Then, \( u_3, u_2, u_1, \) and \( u_h \) are determined as

\[
\begin{align*}
  u_3 &= w_3, \\
  u_2 &= F_2^{-1}(w_2 | u_3), \\
  u_1 &= F_1^{-1}[F_1^{-1}(w_1 | u_2, u_3)], \quad \text{and} \\
  u_h &= F_h^{-1}[F_h^{-1}[F_h^{-1}(w_h | u_3, u_2, u_1)]].
\end{align*}
\]

(7)

Given specified values of \( Y \), the model may be used to sample \( h = F_h^{-1}(u_h) \) from a conditional distribution defined by the given \( Y \) values. In this case, the variables \( u_3, u_2, u_1, \) and \( u_h \) are obtained as

\[
\begin{align*}
  u_3 &= F_3(y_3), \\
  u_2 &= F_2(y_2), \\
  u_1 &= F_1(y_1), \quad \text{and} \\
  u_h &= F_h^{-1}[F_h^{-1}[F_h^{-1}(w_h | u_3, u_2, u_1)]].
\end{align*}
\]

(8)

Throughout this study, we use an algorithm proposed by Bevacqua et al. (2017) to sample from Eq. (8) and carry out all simulations using the CDVineCopulaConditional R package (Bevacqua 2017).

When sampling from Eq. (8) given an observed \( Y \), we produce a stochastic time series of \( h \). Repeated simulations conditioning on the observed \( Y \) will produce multiple time series with varying statistics and agreement with observed \( h \) values (Pham et al. 2016). Throughout this study, given an observed time series of \( Y \), we produce an ensemble consisting of 1000 members of \( h \) time series and obtain a probabilistic forecast of \( h \) at each time step.

e. Model construction

In this section, we lay out the procedure taken for selecting integration period lengths \( L_i \) for the contributing meteorological variables \( Y_i \) (Fig. 2). We also provide details of the selection procedure for the D-vine decomposition of the PCC and the selection of copula families within the PCC.

1) METEOROLOGICAL PREDICTOR SELECTION

We describe soil moisture \( h \) as a function of two precipitation variables, \( Y_{1ps} \) and \( Y_{2pl} \), integrated over periods \( L_1 \) and \( L_2 \), and a PET variable, \( Y_{3PET} \) integrated over the period \( L_3 \). By developing a statistical model with these variables and soil moisture, we look to answer the following three questions:

1) What are the individual contributions of the meteorological variables \( Y_i \) to the estimation of soil moisture \( h \) on time scales related to meteorological drought and heat waves?
2) What relevance does the dependence between antecedent precipitation \( (Y_{2pl}) \) and recent PET \( (Y_{3PET}) \) have for the estimation of low soil moisture values?
3) What relevance does PET have for the estimation of soil moisture over varying integration lengths \( L_3 \)?

To answer these questions, we propose two sets of \( Y \) variables, S1 and S2. Questions 1 and 2 are then approached using variable set S1 while Question 3 is approached using variable set S2. The difference between S1 and S2 is the integration \( L_3 \) chosen at each site. A short integration period is considered for PET in S1, while a long integration period is considered for PET in S2. For each value of \( L_i \) used, the contributing meteorological variable \( Y_i \) may be defined as

\[
\begin{align*}
  Y_{1ps}(t) &= \sum_{r=-L_1}^{I} p(t), \\
  Y_{2pl}(t) &= \sum_{r=-(I-L_2)-1}^{I-L_2} p(t), \quad \text{and} \\
  Y_{3PET}(t) &= \sum_{r=-L_3}^{I-L_3} \text{pet}(t),
\end{align*}
\]

(9)

where \( p(t) \) and \( \text{pet}(t) \) are daily precipitation and PET, respectively.

We address the first two questions with variable set S1. The selected \( L_i \) for S1 must result in \( Y \) variables that provide satisfactory estimates of soil moisture \( h \), hold physically meaningful dependencies, and capture time scales relevant for both meteorological drought and heat waves. Physically meaningful dependencies are obtained by constraining \( L_i \) such that \( L_1 = L_3 \) and through ensuring that there is no overlap between \( L_2 \) and the short-term integrations.

Based on the analysis described below, we find a difference between grassland sites and forest sites. Forest sites require a longer integration \( L_1 \). This is possibly explained by the deeper root systems at forest sites, which filter the influence of short-term variability in rainfall on the integrated soil column. We therefore
choose two sets of $L$: LG and LF, for grassland and forest sites, respectively. At all grassland (forest) sites, the same LG (LF) are used.

We choose integrations of $LG_1 = LG_3 = 7$ and $LG_2 = 63$ for grassland sites. For forest sites, we choose integrations of $LF_1 = LF_3 = 30$ and $LF_2 = 60$. We thus use information of precipitation over the previous 70 and 90 days for each daily soil moisture observation at grassland and forest sites, respectively.

To select $LG_i$ ($LF_i$) in S1, we first calculate the Spearman correlation between $Y_i(t)$ and $h(t)$ for multiple integrations within a window of 120 days prior to day $t$. We then choose the integration length that maximizes the Spearman correlation for each $Y_i$. Integration periods are then constrained such that $LG_1 = LG_3 = 7$ and $LF_1 = LF_3 = 30$. This ensures physically meaningful dependencies and avoids arbitrary dependencies that would otherwise arise between differing $LG_1$ ($LF_1$) and $LG_3$ ($LF_3$).

The sensitivity of the conditional model’s performance, in representing $h$ conditioning on $Y$, to changes in LF (LG) is tested by varying the short-term LG (LF) by ±4 days while the long-term integration LG (LF) is varied by ±10 days. Changes in performance are found to be minimal (not shown). Assuming the same LG (LF) at all grassland (forest sites) and constraining the integration periods is therefore expected to have little weight in the outcome of this analysis.

We acknowledge in S1 that the influence of most recent daily temperature extremes on soil moisture is potentially filtered out at forest sites by setting $LF_3 = 30$. This is addressed in variable set S2 where we assess the relevance of the selection of $L_3$ to the estimation of $h$ (question 3). In S2, two models are constructed using a short- and long-term integration of $L_3$. The same LG, $LF_1$, $LF_2$, and $LF_3$ as S1 are used while LG and LF are set to 7 and 70 days and 7 and 90 days, respectively.

As the variables are all calculated on a daily resolution, from day $t$ to day $t + 1$, there will be an overlap of LG, $LF_i = 1$ or $LF_i = 1$ mutual days used in the integration periods associated with two consecutive days. We thus violate the assumption that data are independent and identically distributed, which the statistical methods used here are based upon. It should therefore be noted that the performance of the model as well as any estimated dependence between variables may be overestimated.

2) STATISTICAL INFERENCE OF THE MULTIVARIATE PDF

The parameters of each bivariate copula in Eq. (3) are estimated based on the marginal variables $u_i$ drawn from the marginal CDFs $F_i$. We use a kernel density estimate for all marginal distributions. All marginal densities are estimated using the ks R package (Duong 2017), which employs the bandwidth selector of Wand and Jones (1994).

The estimation of copula parameters requires that no equal ranks are present in $u_i$. We follow the approach used in Pham et al. (2016) to remove ties from the data. In this approach, a small random noise is drawn from a uniform distribution on $[-0.001, 0.001]$ and added to $Y_{1pt}$ and $Y_{2pt}$, values greater than zero. For values equal to zero, we add a random noise drawn from the uniform distribution on $[0, 0.001]$.

The use of kernel density estimates provides a convenient way of estimating the marginal distribution of $h$. Soil moisture has natural upper and lower bounds, according to its wilting and saturation points, respectively, and can also exhibit a bimodal distribution (Porporato and D’Odorico 2004; D’Andrea et al. 2006).

The selection of the D-vine decomposition in Eq. (3) is based on an initial test in which we assess the performance of each of the six possible decompositions in their ability to represent $h$ when conditioning on the observed $Y$. At all sites we fit a PCC for each of the six decompositions and use the Akaike information criterion (AIC) when selecting the type of copulas to be used. The selection of copula families and the estimation of their parameters is carried out at each site separately. Each copula is chosen from a range of copulas provided by the VineCopula R package (Schepsmeier et al. 2017). To assess each of the six possible decompositions, a probabilistic forecast of $h$ consisting of 1000 members is produced at all sites. These are compared with observed soil moisture using the root-mean-square error. We then select the decomposition that generally shows the highest explanatory power of $h$ at all sites.

After selecting the decomposition to apply, the goodness of fit (GoF) of the selected copulas is tested. Copulas initially selected according to the AIC did not always provide a satisfactory fit. For this reason we use two criteria in the selection of a copula for each pair in the PCC. This procedure is carried out sequentially as outlined in section 2d(1), where unconditional copulas are first selected followed by the conditional copulas. For each pair, we first select the top three copulas according to the AIC and second test the GoF of each using K plots (Genest and Favre 2007; Bevacqua et al. 2017). We then select the highest ranked copula according to the AIC that shows satisfactory compliance in the K plots.

A K plot is a plot of the Kendall function $K(w) = P(C_{ij}(u_i, u_j) \leq w)$ obtained from the fitted copula against $K(w)$ computed with the empirical copula obtained.
using the observed data. Similarly to a Q–Q plot for
univariate distributions, the K plot indicates good
quality of fit when points follow the diagonal. These
plots provide uncertainties around the empirical copula
as well as a qualitative idea of the quality of fit of each
copula (Bevacqua et al. 2017). Most selected copulas
show good agreement according to the K plots (not
shown) where parametric $K(w)$ values generally follow
the mean of the empirical values and mostly remain
within the uncertainty intervals calculated from 1000
simulations. Some small problems are found with the
copulas at sites e and f, which may limit the strength of
conclusions drawn from these sites.

f. Model evaluation metrics

The model simulations are evaluated overall and in
their ability to represent low values of soil moisture $h$.
Using the Brier score (BS), we evaluate the accuracy of
probabilistic predictions of low $h$ values defined as those
below the 15th percentile of observed soil moisture. The
closer BS is to zero, the better the predictions. The BS is
defined as:

$$BS = \frac{1}{N} \sum_{t=1}^{N} (p_t - o_t)^2,$$

where $p_t$ is the probability of getting a simulated value of
$h$ below the observed 15th percentile from the model at
time $t$, while $o_t$ is 1 if observed soil moisture $h_{obs}(t)$ is
below the 15th percentile and 0 otherwise. Along with BS
we calculate the associated Brier skill score (BSS) that
evaluates the model relative to a reference model $BS_{ref}$:

$$BSS = 1 - \frac{BS}{BS_{ref}},$$

We consider the climatology as the reference model in
which the probability of a value occurring below the
15th percentile is always 0.15.

The model is also evaluated in its ability to capture the
presence of drought conditions by comparing the
autocorrelation function (ACF) and using an empirical
order 1 persistence probability (PP). Both are derived from
the observed values and the mean of the simulated
values. We choose an order 1 persistence after assessing
partial autocorrelation function (PACF) at each site,
which only showed significant correlations for order 1.
The quantity PP is defined as

$$PP = \Pr[h_{t+1} < F_h^{-1}(0.15) \mid h_t < F_h^{-1}(0.15)].$$

The PP may be interpreted as the probability that $h_{t+1}$
will be below the 15th percentile given that $h_t$ is below
the 15th percentile.

3. Results

The set of variables $S_1$, described in section 2e(1), are
employed to evaluate the contributions of the individual
$Y$ variables and that of their dependence structure to
soil moisture. To achieve this we perform a number of
sensitivity simulations and compare them with a control
simulation (CTRL). All simulations carried out are
done through a K-fold cross validation to avoid over-
fitting. Parameter $K$ here is the number of summers in a
time series at a given site. In each simulation, we thus
remove one summer at a time when fitting the copula
rather than the entire multivariate statistical model. For
each simulation, we then produce a probabilistic fore-
cast of $h$ consisting of 1000 members through condition-
on on specified values of $Y$.

a. Model performance

The CTRL simulation is performed through sampling
$h$ conditioned on observed values of $Y$ [Eq. (8)]. The
performance of CTRL may be qualitatively gauged from
Fig. 3. Plots shown in Figs. 3a–e are results from
wet sites while those from Figs. 3f–k are results from dry
sites. The mean value of $h$ from CTRL at each time step
can be seen to follow the temporal evolution of ob-
served soil moisture ($h_{obs}$) quite well, while $h_{obs}$ is
generally found within the 95% confidence interval of
CTRL. Also shown within each panel in Fig. 3 are the
order 1 persistence probabilities of low $h$ for observed
($PP_{obs}$) and mean simulated ($PP_{sim}$) levels.

We consider the climatology as the reference model in
which the probability of a value occurring below the
15th percentile is always 0.15.

The model is also evaluated in its ability to capture the
presence of drought conditions by comparing the
autocorrelation function (ACF) and using an empirical
order 1 persistence probability (PP). Both are derived from
the observed values and the mean of the simulated
values. We choose an order 1 persistence after assessing
partial autocorrelation function (PACF) at each site,
which only showed significant correlations for order 1.
The quantity PP is defined as

$$PP = \Pr[h_{t+1} < F_h^{-1}(0.15) \mid h_t < F_h^{-1}(0.15)].$$

The PP may be interpreted as the probability that $h_{t+1}$
will be below the 15th percentile given that $h_t$ is below
the 15th percentile.
indicating that the proposed model and variables included do not predict soil moisture correctly at all sites. However, with satisfactory results generally obtained at most sites, we employ the model for use in sensitivity analysis in a number of tests presented below.

**b. Assessment of contributing variables to soil moisture**

We test the contribution of $Y_{1PS}$ (short-term precipitation), $Y_{2PL}$ (long-term precipitation), and $Y_{3PET}$ (PET) to the estimation of $h$ in three sensitivity simulations

Fig. 3. Observed time series (red) alongside the cross-validation time series of the CTRL mean (black) and the 95% prediction interval (gray), obtained from 1000 simulations, at (a)–(e) wet sites and (f)–(k) dry sites. Also provided within each panel are the order 1 persistence probabilities calculated from the observed (PP$_{obs}$) and CTRL mean (PP$_{sim}$) time series.
SENS-Y$_{1PS}$, SENS-Y$_{2PL}$, and SENS-Y$_{3PET}$, respectively. For each sensitivity simulation, $h$ is sampled conditioning on the median value of the respective variable to be tested and the observed values of the other two $Y$ variables. To assess the contributions of all variables, we compare the mean of each simulation with the CTRL mean. We also compare the probabilistic forecasts from SENS-Y$_1$ with CTRL using the BS, BSS, and the mean ensemble bias computed for values of $h$ below the observed 15th percentile (Table 2).

At wet sites, precipitation is generally seen to have the most influence on low soil moisture values, while PET can act to amplify the low soil moisture anomaly during drought periods. Comparing the means of the three sensitivity simulations with the mean of CTRL (Fig. 4), larger overestimations of low $h$ values with respect to CTRL are generally seen in either of the simulations assessing the influence of a precipitation variable, SENS-Y$_{1PS}$ or SENS-Y$_{2PL}$, than is, seen in SENS-Y$_{3PET}$.

Underlining this are larger changes in positive bias of low soil moisture values seen from SENS-Y$_{1PS}$ or SENS-Y$_{2PL}$ than from SENS-Y$_{3PET}$ (Table 2). A comparison of BSS for each simulation in Table 2 also shows a larger reduction in skill of forecasting values below 15th percentile in either SENS-Y$_{1PS}$ or SENS-Y$_{2PL}$ than in SENS-Y$_{3PET}$. Focusing on drought events at wet sites a, b, and d in 2003 and 2006, years in which heat waves have also occurred (Ciais et al. 2005; Rebetez et al. 2009), we see from the mean of the simulations (Fig. 5) that removing the influence of precipitation can lead to the misspecification of a drought event with the green line largely above the black line (CTRL). On the other hand, removing the influence of PET can result in the underestimation of the severity of the event, with the blue line only just higher than the black during a drought event.

At dry sites, we see that precipitation again holds the main influence over soil moisture while PET generally offers little added benefit to the estimation of soil moisture. The main differences of CTRL with SENS-Y$_{1PS}$ and SENS-Y$_{2PL}$ are found for high values of soil moisture (Fig. 4). Low values in these sensitivity simulations are generally equivalent with CTRL, as the medians of $Y_{1PS}$ and $Y_{2PL}$ are associated with relatively low values due to the positively skewed nature of the variables’ distributions. Little or no difference is seen between SENS-Y$_{3PET}$ and CTRL simulations for low values of soil moisture. Large percentage changes in bias for low soil moisture values are seen at sites f and i, though the actual changes in soil moisture are relatively low (Table 2). This would be expected at dry sites during summer where soil moisture normally reaches low levels such that ET is moisture-limited and will diverge from PET. Extremes of PET driven by extreme temperatures would then have little added effect to the severity of soil moisture drought in dry locations.

c. Assessing the relevance of $Y$ dependence structure

The contribution of the dependence between $Y_{2PL}$ and $Y_{3PET}$ to the estimation of low $h$ values is assessed using the sensitivity simulation IND-Y$_{2PL}$, which is used to highlight where interactions between drought and heat wave conditions, arising through land–atmosphere interactions, act to amplify drought conditions. To illustrate the dependence between $Y_{2PL}$ and $Y_{3PET}$, we calculate Spearman’s $\rho$ and a measure of tail dependence $\lambda_{\rho}$, calculated as

<table>
<thead>
<tr>
<th>Site</th>
<th>Score</th>
<th>CTRL</th>
<th>SENS-Y$_{1PS}$</th>
<th>SENS-Y$_{2PL}$</th>
<th>SENS-Y$_{3PET}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>BS</td>
<td>0.09</td>
<td>0.10</td>
<td>0.12</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.25</td>
<td>0.21</td>
<td>0.01</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>3.89</td>
<td>+5%</td>
<td>+113%</td>
<td>+45%</td>
</tr>
<tr>
<td>b</td>
<td>BS</td>
<td>0.11</td>
<td>0.13</td>
<td>0.10</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.15</td>
<td>-0.01</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>4.33</td>
<td>+107%</td>
<td>-1%</td>
<td>+33%</td>
</tr>
<tr>
<td>c</td>
<td>BS</td>
<td>0.12</td>
<td>0.14</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.04</td>
<td>-0.1</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>10.66</td>
<td>+51%</td>
<td>-24%</td>
<td>+2%</td>
</tr>
<tr>
<td>d</td>
<td>BS</td>
<td>0.06</td>
<td>0.06</td>
<td>0.09</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.51</td>
<td>0.49</td>
<td>0.26</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>3.52</td>
<td>+45%</td>
<td>+149%</td>
<td>+83%</td>
</tr>
<tr>
<td>e</td>
<td>BS</td>
<td>0.09</td>
<td>0.10</td>
<td>0.09</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.28</td>
<td>0.25</td>
<td>0.27</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>3.62</td>
<td>-15%</td>
<td>+78%</td>
<td>+81%</td>
</tr>
<tr>
<td>f</td>
<td>BS</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.36</td>
<td>0.43</td>
<td>0.53</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.37</td>
<td>-215%</td>
<td>+207%</td>
<td>+720%</td>
</tr>
<tr>
<td>g</td>
<td>BS</td>
<td>0.09</td>
<td>0.09</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.31</td>
<td>0.24</td>
<td>0.15</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>1.48</td>
<td>+43%</td>
<td>+86%</td>
<td>+6%</td>
</tr>
<tr>
<td>h</td>
<td>BS</td>
<td>0.12</td>
<td>0.13</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.04</td>
<td>0.00</td>
<td>0.05</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>3.28</td>
<td>+9%</td>
<td>-3%</td>
<td>+4%</td>
</tr>
<tr>
<td>i</td>
<td>BS</td>
<td>0.12</td>
<td>0.13</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.07</td>
<td>0.002</td>
<td>0.04</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>1.24</td>
<td>-9%</td>
<td>+5%</td>
<td>+52%</td>
</tr>
<tr>
<td>j</td>
<td>BS</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.09</td>
<td>0.10</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>2.8</td>
<td>-8%</td>
<td>+10%</td>
<td>+16%</td>
</tr>
<tr>
<td>k</td>
<td>BS</td>
<td>0.08</td>
<td>0.13</td>
<td>0.12</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>BSS</td>
<td>0.36</td>
<td>0.01</td>
<td>0.09</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>1.05</td>
<td>+205%</td>
<td>+146%</td>
<td>+10%</td>
</tr>
</tbody>
</table>
\[ \lambda_q = \Pr \left[ Y_{3\text{PET}} > F_3^{-1}(q) \mid Y_{2\text{PL}} < F_2^{-1}(1 - q) \right], \]  \hspace{1cm} (13) \]

where \( q = 0.9 \) in this case. The quantity \( \lambda_{0.9} \) can be interpreted as the fraction of days when \( Y_{3\text{PET}} \) was greater than its observed 90th percentile when \( Y_{2\text{PL}} \) was less than its 10th percentile. For two independent variables, the expected value of \( \lambda_q \) is \( 1 - q \). Values of \( \rho \) and \( \lambda_q \) for each site are given in Fig. 6. At many sites we observe a negative dependence between \( Y_{2\text{PL}} \) and \( Y_{3\text{PET}} \), as measured by \( \rho \), and an increased probability of extreme PET (\( Y_{3\text{PET}} \)) when antecedent precipitation (\( Y_{2\text{PL}} \)) had been extremely low.

To test the relevance of such dependence in IND-\( Y_{2\text{PL}} \), we break the dependence between \( Y_{2\text{PL}} \) and the short-term variables \( Y_{1\text{PS}} \) and \( Y_{3\text{PET}} \). This is achieved by shuffling \( Y_{2\text{PL}} \) such that it is randomly associated with them. A probabilistic forecast of \( h \), consisting of 1000 members, is then produced, sampling from the multivariate distribution where we condition on the observed values of \( Y_{1\text{PS}} \) and \( Y_{3\text{PET}} \) and the shuffled \( Y_{2\text{PL}} \). To account for sampling variability of the shuffling process, we produce 1000 IND-\( Y_{2\text{PL}} \) probabilistic forecasts.

We obtain a kernel density estimate of the PDF produced from each of the 1000 IND-\( Y_{2\text{PL}} \) simulations. The mean density and the 95% confidence interval of IND-\( Y_{2\text{PL}} \) PDFs are calculated and presented alongside the PDFs of CTRL and \( h^{\text{obs}} \) (Fig. 6). The statistical significance of the difference between the CDFs of CTRL and IND-\( Y_{2\text{PL}} \) is assessed at the 5th, 10th, and 15th percentiles of observed soil moisture. CTRL is considered significantly different for a given percentile if the associated soil moisture value of CTRL is less than the lower bound of 95% confidence interval of that percentile from IND-\( Y_{2\text{PL}} \). This would signify that the probability
of values below that percentile are underestimated when the dependence between $Y_{2PL}$ and $Y_{3PET}$ is broken.

Statistically significant differences are found between all three percentiles at site d, where we also see a noticeable difference between PDFs (Fig. 6d). A negative dependence as well as a significant dependence in the tails is also observed here. Site d lies in a transitional region where land–atmosphere interactions can lead to the mutual reinforcement of drought and heat wave events (Seneviratne et al. 2010). This result highlights the importance of the interplay between drought and heat wave conditions, driven by land–atmosphere interactions, to the reinforcement of drought conditions in such locations.

Statistically significant differences between the percentiles tested are also found at wet sites a, b, and e and dry sites g and j, though relatively little difference is observed between CTRL and IND-$Y_{2PL}$ PDFs at these sites for values below the tested percentiles (Fig. 6). We observe negative dependencies ($\rho$) and tail dependencies ($\lambda_{\alpha}$) at these sites, which highlights that the concurrence of such conditions may be important for the estimation of low values of soil moisture. These dependencies are also observed at other dry sites, but no significant differences between assessed percentiles are found. Such dependencies at these sites are perhaps of little relevance for soil moisture during summer, as extremes of PET may be energy limited in wet climates while soil in dry climates may have little available moisture for ET. In dry conditions then, extremes of PET in combination with extremely low antecedent precipitation will have little effect on moisture levels in soil.

d. Relevance of PET over short and long integration periods

The variable set S2, as described in section 2e(1), is used to demonstrate the relevance of PET, integrated over various durations $LG_3$ and $LF_3$, to the estimation of soil moisture $h$. We fit two models at wet sites a, b, and d, where we see contributions of PET to the estimation of soil moisture drought in variable set S1 (Fig. 5). The integration periods used for precipitation variables $Y_{1PS}$ and $Y_{2PL}$ in S1 remain the same.

Based on the mean of the simulations (Fig. 7), better representation of drought onset can be seen at sites a and d in the PET-INT$_S$ simulations where the black line generally follows red (observed) at the beginning of an event when initial drying is taking place. On the other hand, drought persistence is generally captured better.
by the PET-INT$_L$ simulation where the blue line remains low with the red line in comparison to the black. Better BSSs are found for simulations using a long-term integration of PET at sites a and b. Increases of BSS, from PET-INT$_S$ to PET-INT$_L$, from 0.24 to 0.36 and from 0.18 to 0.25 are found at each site, respectively, while little difference is seen between simulations at site d with BSS equal to 0.51 and 0.52.

Although these results are somewhat qualitative, they highlight that both short- and long-term integrations of PET are important for the estimation of drought events in this framework. Longer integrations are generally better in capturing the persistence of drought conditions as they can account for the memory soil moisture holds of drying during the event. Short-term integrations, however, are better in capturing drought onset as they are able to account for short intense periods of drying that can accelerate the propagation of meteorological drought to soil moisture drought. With drought events expected to set in quicker in a warming climate (Trenberth et al. 2014), it will be important to detect such changes in the intensity of drying over short periods in spring and summer that are filtered out in longer integrations of PET. This may be of particular relevance in Europe, where early onset of drought conditions can have large implications for extreme temperatures in summer (Vautard et al. 2007).

4. Summary and conclusions

Compound events are multivariate extreme events in which the contributing variables need not be extreme themselves, but their joint dependent concurrence produces an extreme impact (Leonard et al. 2014; Bevacqua et al. 2017). We have analyzed soil moisture drought over Europe as a compound event of variables employed in common drought indices, namely, precipitation and PET, and assessed the individual roles of these variables and that of their dependence structure to the estimation of soil drought.
moisture. The overall aim was to explore the compound nature of soil moisture drought and the differences that exist between wet and dry climates.

To achieve our aim, we developed a statistical model based on pair copula constructions. Within the model we considered precipitation and PET over time scales related to meteorological drought and heat waves, respectively. These time scales were considered to assess the influence of heat wave conditions on soil moisture, as well as dependencies driven by land–atmosphere interactions that can cause a mutual reinforcement between drought and heat wave events in Europe. We applied the model to data from 11 FluxNet sites situated in wet, transitional, and dry climates in Europe and generally found satisfactory performance of the model. We thus employed it in a number of sensitivity experiments to assess the relevance of contributing variables and their dependence structure to the estimation of soil moisture drought.

Results obtained from sensitivity experiments were in line with previous studies. Precipitation was found to hold the main control over soil moisture drought. PET was required only when it departs from normal conditions (Vicente-Serrano et al. 2010) to partly explain the severity of drought conditions in wet climates (Seneviratne et al. 2012b; Teuling et al. 2013), while little or no contribution was found in dry climates (Luo et al. 2017) during summer. The concurrence of extremely low antecedent precipitation with extremely high PET was found to be most relevant at a site situated in a transitional climate region between wet and dry climates where land–atmosphere interactions are most relevant for the development of soil moisture drought (Seneviratne et al. 2006, 2012a). The concurrence of these conditions was also seen at many dry sites, though they were found to have little relevance for soil moisture. This lack of relevance at dry sites is presumably related to the limited availability of moisture in soil for actual ET to occur, such that PET and extremes of PET could have little influence to a low soil moisture anomaly.

The aforementioned contribution of PET is based on a short-term integration period that was used to capture the influence of heat waves on soil moisture. At wet sites, this short integration period is found to be effective in describing the onset of drought events as it can capture initial drying that occurs on a daily basis. It can, however, be ineffective in capturing the persistence of drought conditions, which longer integrations can better account for, as it neglects the memory soil moisture may hold of PET and the intense drying that may have occurred throughout a drought event. The differences found between short and long integrations of PET may become relevant in the analysis of changes in the

Fig. 7. Mean cross-validation time series of simulations from models PET-INT$_S$, in which PET is considered over a short integration period (black), and PET-INT$_L$, in which PET is considered over a long integration period (blue), along with the observed time series (red) for the (top) 2003 and (bottom) 2006 drought events at wet sites a, b, and d.
onset of drought events using drought indices. A warmer climate may cause droughts to set in quicker (Trenberth et al. 2014) and lead to flash droughts (Mo and Lettenmaier 2016). Such drying may be hidden through the use of longer integration periods of PET in an index such as the SPEI or through a recursive model used for the PDSI that retains memory of PDSI values from previous time steps.

Advantages of using drought indices include the simplicity they offer and the widespread availability of meteorological datasets compared to those of soil moisture. Although they are not specifically designed to represent soil moisture (Seneviratne et al. 2010), indices such as the SPEI, PDSI, and RDI provide a convenient means of combining precipitation and PET into a kind of impact function that may be implicitly linked to soil moisture.

However, soil moisture drought is not a simple phenomenon to characterize with drought indices due to differing contributions and relevant integration periods of meteorological variables in wet and dry climates. The use of a climatic water balance (precipitation – PET) in the SPEI and PDSI assumes oversimplified relationships between precipitation, PET, and soil moisture (Seneviratne 2012) and implies that the statistical relevance of precipitation and PET to the estimation of soil moisture are the same over a given integration period. With such simplifications comes a loss of information, such as short intense periods of drying that may be filtered out through the inclusion of redundant information when using a long integration period for PET.

Through the inclusion of PET, these indices are expected to provide a better picture of changes in drought conditions in a warming climate than indices that use precipitation alone such as the standardized precipitation index (SPI). Ubiquitously applying indices that incorporate PET across different climates can provide a general overview of the response of drought conditions to global warming. It is, however, important to note that severe drought, as depicted by these indices, will have a different meaning for soil moisture drought in wet and dry climates. ET is limited by moisture availability and so will diverge from PET in dry conditions, leading to an overestimation of the actual drying taking place with respect to soil. In contrast, land surface models account for this moisture limitation by capturing the physical relationship between PET and soil moisture; they can therefore provide a more reliable estimate. Their use within coupled climate models to study changes in soil moisture drought is particularly advocated for by Berg et al. (2017), who also demonstrate the added complexity of diverging changes to soil moisture at different soil depths that cannot be disentangled using drought indices.

Despite discrepancies between PET and ET in dry conditions, extremes of PET will still be indicative of the drying potential of the atmosphere, provided it is calculated using a reliable physically based method such as the Penman–Monteith equation. Such atmospheric drying potential may possibly have adverse effects on crop yields and contribute to other environmental hazards such as wildfires that are mediated by the availability of moisture in vegetation (Gudmundsson et al. 2014; Ruffault et al. 2016).

Much information of soil moisture and other drought impacts may be deduced from drought indices and their response to a warming climate. To do so requires careful interpretation and detailed knowledge of the involved variables’ influence on soil moisture in a given climate. It is therefore important that drought indices incorporating PET are interpreted within the context of the climate in which they are applied, while also keeping in mind the applications they are designed for.

In our impact focused approach, we have made use of the little soil moisture data that are available across different locations and climate types in Europe to demonstrate the compound nature of soil moisture drought during summer. These results provide further insight into the relationship between soil moisture and drought indices that incorporate PET. It is hoped that this insight will aid with the interpretation of drought indices in a given climate and season so that as much information as possible may be gained from their application.

Acknowledgments. Colin Manning received funding from the Volkswagen Foundation’s CE: LLO project (Az.: 88469), which also supported project meetings. The authors thank Prof. Arnoldo Frigessi and Prof. Ingrid Hobaek Haff for fruitful discussions and the hosting of a project meeting at the Norwegian Computing Center. The authors would also like to sincerely thank the anonymous reviewers for their constructive comments that have enhanced the quality of the manuscript.

REFERENCES


