BAMS Article

Potential of Pan-European Seasonal Hydrometeorological Drought Forecasts Obtained from a Multihazard Early Warning System

Samuel Jonson Sutanto, Henny A. J. Van Lanen, Fredrik Wetterhall, and Xavier Llort

ABSTRACT: Drought early warning systems (DEWS) have been developed in several countries in response to high socioeconomic losses caused by droughts. In Europe, the European Drought Observatory (EDO) monitors the ongoing drought and forecasts soil moisture anomalies up to 7 days ahead and meteorological drought up to 3 months ahead. However, end users managing water resources often require hydrological drought warning several months in advance. To answer this challenge, a seasonal pan-European DEWS has been developed and has been running in a preoperational mode since mid-2018 under the EU-funded Enhancing Emergency Management and Response to Extreme Weather and Climate Events (ANYWHERE) project. The ANYWHERE DEWS (AD-EWS) is different than other operational DEWS in the sense that the AD-EWS provides a wide range of seasonal hydrometeorological drought forecasting products in addition to meteorological drought, that is, a broad suite of drought indices that covers all water cycle components (drought in precipitation, soil moisture, runoff, discharge, and groundwater). The ability of the AD-EWS to provide seasonal drought predictions in high spatial resolution (5 km × 5 km) and its diverse products mark the AD-EWS as a preoperational drought forecasting system that can serve a broad range of different users' needs in Europe. This paper introduces the AD-EWS and shows some examples of different drought forecasting products, the drought forecast score, and some examples of a user-driven assessment of forecast trust levels.

https://doi.org/10.1175/BAMS-D-18-0196.1
Corresponding author: S. J. Sutanto, samuel.sutanto@wur.nl
Supplemental material: https://doi.org/10.1175/BAMS-D-18-0196.2

In final form 6 November 2019

AFFILIATIONS: Sutanto and Van Lanen—Hydrology and Quantitative Water Management Group, Wageningen University and Research, Wageningen, Netherlands; **Wetterhall**—European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom; **Llort**—Hydrometeorological Innovative Solutions, Barcelona, Spain

The losses caused by natural hazards have continued to rise worldwide over the past three decades. According to the data compiled by FAO (2017), the impact of natural disasters in the past 10 years (2003–13) estimates a total economic damage of $1.53 trillion (U.S. dollars) and in total 1.16 million human casualties. Due to the high losses triggered by disasters, implementation of the United Nation Sendai Framework for Disaster Risk Reduction 2015–30 is encouraged (Poljanšek et al. 2017). This framework has the overall objective of reducing disaster risks and losses, and highlights the needs of multihazard early warning systems (MH-EWS) as one of the important Sendai's global targets. The weather-induced EWS that have been developed, however, are mostly designed to predict high-flow extremes, such as floods and flash floods (Alfieri et al. 2011; Hurford et al. 2012; Laiolo et al. 2014; Smith et al. 2016), landslides (Osanai et al. 2010; Baum and Godt 2010), or storm surge (de Vries 2009; Bajo and Umgiesser 2010). While EWS for floods and landslides require forecasts at shorter lead times, such as hours and days, a drought EWS requires longer forecasts lead times ranging from monthly up to seasonal because of its creeping nature (Tallaksen and Van Lanen 2004). The historical lack of skill to produce robust monthly precipitation forecasting products (Vitart 2004; Buizza and Leutbecher 2015) is the main reason why seasonal drought EWS is less developed than many other short-time-scale hazard warning systems.

Recently, the use of probabilistic seasonal weather forecasting systems based on numerical weather prediction models has become more common, since the skill of these models have greatly improved (Barnston et al. 2012; Yuan et al. 2013a). The improvement of the probabilistic seasonal weather forecasts has led to the development of operational seasonal drought EWS in some regions, such as the United States and Africa. Those EWS, however, have been developed mainly using precipitation, temperature, and soil moisture forecast data (Pozzi et al. 2013; Sheffield et al. 2014). In the United States, real-time drought monitoring has been operational and provides information on the current drought conditions as well as the seasonal drought forecasts using the standardized precipitation index (SPI)-6, and the anomalies of precipitation, temperature, soil moisture, and streamflow (Svoboda et al. 2002; Ek et al. 2010). In Africa, there are two drought monitoring and EWS: 1) the African flood and drought monitor developed by Princeton University (Sheffield et al. 2014) and 2) the African Drought Observatory (ADO) developed by the EU Joint Research Centre (JRC) (Barbosa et al. 2013). These two drought EWS in principle produce a warning signal based on the forecasted SPI, soil moisture anomaly, and streamflow anomaly up to 6 months (Pozzi et al. 2013; Sheffield et al. 2014). In China, Yuan et al. (2016, 2017) developed an experimental seasonal hydrological forecasting system over the Yellow River basin. The system produces seasonal hydrological drought forecasts for streamflow based on the North American Multi-Model Ensemble (NMME) climate forecast models and the Variable Infiltration Capacity (VIC) hydrological model.

In Europe, the need for a seasonal drought EWS have risen with the increase in the numbers of drought occurrences over the past decades (Stahl et al. 2010; Spinoni et al. 2017). From several major drought events that affected Europe (Feyen and Dankers 2009; Mishra and Singh 2010), one of the costliest drought events were in 2003, where the economic cost of the damage was estimated at least EUR 8.7 billion. In the past 30 years, the economic loss due to droughts was estimated to be over EUR 100 billion (EU 2007). The losses will likely worsen in the future because of climate change and increase of water use (Lehner et al. 2006). Southern

and southeastern Europe will be more prone to a rise in drought frequencies (Prudhomme et al. 2014; Wanders and Van Lanen 2015; Wanders et al. 2015).

Because of the big economic losses caused by drought in Europe, several countries in Europe have developed a drought monitor and warning system, such as Switzerland and Germany (Zappa et al. 2014; Zink et al. 2016). The German drought monitoring system provides daily up-to-date agricultural drought information based on a soil moisture index (SMI; Zink et al. 2016), and the Swiss water resources monitoring and early recognition of critical droughts provide a hydrological drought monitoring and warning system based on hydrological variables, such as soil moisture, runoff, groundwater level, and snow water equivalent on a weekly basis up to 3 months in advance (Zappa et al. 2014). At the pan-European level, JRC has developed the European Drought Observatory (EDO) in the framework of the 2007 EU Water Scarcity and Drought Communication to provide European-wide up-to-date information on the occurrence and severity of droughts (Vogt et al. 2013). Beside its function to monitor the ongoing drought, the EDO also forecasts the soil moisture anomaly up to 7 days ahead at 5-km spatial resolution with 2 days of delay, and each month the meteorological drought (SPI-1) with a 3-month lead time. Thus, EDO provides forecasts of soil water drought events with a temporal resolution of one week, but no hydrological drought forecasts (e.g., streamflow and groundwater). Drought events in Europe may sustain from months up to seasons and cover large areas. A new drought EWS, which can provide the forecasted drought at longer time scales from months to seasons and including more hydrometeorological variables than soil moisture is potentially a useful tool to reduce the impact of drought events in Europe.

The multihazard EWS

To answer the challenge in providing operational drought forecasts for Europe with a temporal resolution of more than a week, including all water cycle components, a seasonal drought EWS has been developed under the EU-funded Enhancing Emergency Management and Response to Extreme Weather and Climate Events (ANYWHERE) innovation project (www.anywhere-h2020.eu). The drought EWS has been running in a preoperational mode since mid-2018. The objective of ANYWHERE is to employ cutting-edge innovative technologies to build a pan-European MH-EWS for faster analysis and anticipation of the risk before the event occurrence, improved coordination of the emergency actions, and assist to raise the self-preparedness. This platform will address all major weather-related natural hazards in Europe, such as floods, flash floods, debris flows, landslides, storm surges, heat waves and air quality, wildfires, droughts, convective storms, severe winds, and heavy snowfall (see ANYWHERE website for all products: http://anywhere-h2020.eu/catalogue/).

The drought EWS (DEWS) developed under the ANYWHERE project [ANYWHERE DEWS (AD-EWS)] is a part of the MH-EWS and it is complementary to the current pan-European operational system (EDO) by producing probabilistic hydrometeorological drought forecasts with lead times up to 7 months. The drought monitoring products of the EDO are also integrated into the MH-EWS. The current AD-EWS is driven by the latest ECMWF seasonal forecast system [SEAS 5 (S5)] and a state-of-the-art hydrological model LISFLOOD (Bartholmes et al. 2008; Van Der Knijff et al. 2010; Burek et al. 2013) to predict seasonal hydrometeorological ensemble drought forecasts at pan-European scale. The seasonal hydrological forecasts are provided through the European Flood Awareness System (EFAS; www.efas.eu), which is part of the Copernicus Emergency Management Services (CEMS). The AD-EWS generates a comprehensive list of drought indices (standardized drought indices, threshold drought indices, and areal drought indices). For the standardized drought indices, the SPI-*x* (McKee et al. 1993), the standardized precipitation evaporation index (SPEI-*x*; Vicente-Serrano et al. 2010), the standardized runoff index (SRI-*x*; Shukla and Wood 2008), and the standardized groundwater index approaches (SGI-1; Bloomfield and Marchant 2013) have been implemented with accumulation periods *x*

The EDgE project

The End-to-End Demonstrator for Improved Decision-Making in the Water Sector in Europe (EDgE) project funded by the Copernicus Climate Change Service program provides knowledge that is relevant for further development of drought forecasting. The project has published knowledge on drought projections and seasonal hydrological reforecasts (using hindcast) at the pan-European scale (http://edge.climate.copernicus.eu/Tools/). They use four GCMs (HadGEM2-ES, IPSL-CM5A-LR, MIROC-ESM-CHEM, GFDL-ESM2M, and NorESM1-M) as forcing for three hydrological models [mesoscale hydrologic model (mHM), Noah-MP, and PCRaster Global Water Balance (PCR-GLOBWB)] at the daily time scale under three representative concentration pathways (RCPs 2.6, 6.0, and 8.5) to simulate the future droughts under global warming (Marx et al. 2018; Thober et al. 2018; Samaniego et al. 2018, 2019). For the seasonal hydrological reforecasting study, they use four GCMs to produce seasonal meteorological reforecasts (CanCM4, GFDL-FLOR, ECMWF-S4, and LFPW) fed into four hydrological models (mHM, Noah-MP, PCR-GLOBWB, and VIC) (Wanders et al. 2019). Beside the drought projection under climate change scenarios, the EDgE project also provides low streamflow-related indices, as well low levels of other hydrological variables (http://edge.climate.copernicus.eu/Publications/EDgE_D2.1_Stage_2_SCIIs_Final.pdf, accessed on 29 April 2019). The EDgE output could be used as a basis to identify both standardized and threshold-based drought indices. The use of multimodel seasonal hydrological forecast systems is encouraged to better understand model structure uncertainty (Wanders et al. 2019). However, it requires intensive data processing and computational power, and therefore likely it is not fit for operational services yet. For an operational system, a balance always has to be found between computer constraints and data collation, running a climate and hydrological model (single model chain)/climate and hydrological models (multi-model chain), calculation of drought indices for each ensemble member. With the development of computing capacity in the future, the use of a multimodel framework for a drought EWS is foreseen (Bauer et al. 2015). Like the EDgE project (e.g., Marx et al. 2018; Thober et al. 2018; Samaniego et al. 2018; Wanders et al. 2019; Samaniego et al. 2019), several studies already produced seasonal forecasts of hydrological variables for the past (hindcasts) and drought projections (e.g., Prudhomme et al. 2014; Bell et al. 2017; Arnal et al. 2018), which could serve as a basis for further development of hydrological drought forecasting. One should realize that forecasted drought in hydrological variables, for example, streamflow drought forecasts deviate from streamflow forecasts (or generally hydrological forecasts). Drought forecasting requires an additional step using the forecasted time series of a hydrological variable; that is, forecasted drought is a derived product from hydrological forecasts. The hydrological drought community makes a clear difference between, for instance, low streamflow (which is hydrology) and drought in streamflow. Low streamflow is not necessarily categorized as a drought; however, drought in streamflow can be caused by a long period of below-normal/low streamflow (e.g., www.epa.gov/ceam/definition-and-characteristics-low-flows#low). This difference appears to result, for instance, in substantial differences in projections of low streamflow and drought in streamflow (Alderlieste et al. 2014; Van Lanen et al. 2018). Hence, hydrological drought forecasts and hydrological forecasts needed to be distinguished.

= 1, 3, 6, and 12 months. The threshold approach (Yevjevich 1967; Hisdal et al. 2004) uses the variable threshold method (VTM) to identify drought events and quantify its characteristics, such as duration, onset, termination, and deficit volume or average cumulative deviation from threshold for state variables (e.g., soil moisture and groundwater). The areal drought indices provide a summary of areas affected by drought and their characteristics, such as average duration in the affected area.

The AD-EWS is different compared to other operational drought EWS in the sense that 1) the AD-EWS provides a wide range of seasonal hydrometeorological ensemble drought forecasting products that covers most of the water cycle components (drought in precipitation, soil moisture, runoff, groundwater, and discharge), and 2) the AD-EWS has a nominal spatial resolution of 5 km × 5 km, the same resolution to EDO. Only catchment-scale drought EWS have a higher spatial resolution (see Zappa et al. 2014). The up-to-date drought forecasting produces each month a comprehensive suite of forecasting products and drought identification approaches at a high resolution. The current generation of operational drought EWS only uses, in general, the SPI, soil moisture anomaly, runoff anomaly, and streamflow anomaly as drought indices with a rather coarse spatial resolution of 30 km (Ek et al. 2010; Pozzi et al. 2013; Sheffield et al. 2014).

The AD-EWS platform

The MH-EWS, as the main platform for the AD-EWS, is defined as a software as a service (SaaS) cloud-based platform that essentially integrates the models developed within ANYWHERE

and existing pan-European platforms [e.g., EDO, EFAS, the European Forest Fire Information System (EFFIS)]. The MH-EWS computes the hazard products and serves the generated information to its users. The AD-EWS drought identification algorithms are encapsulated and executed by the MH-EWS, whereas the ECMWF, EDO, and EFAS products are incorporated in the MH-EWS (Fig. 1). The drought outputs of the MH-EWS are categorized into gridded maps and summary of areal droughts (time series and bar diagrams). The general schematization of the AD-EWS components, as part of the MH-EWS, is presented in Fig. 1.

The LISFLOOD model and ECMWF forecasting system

LISFLOOD, the hydrological model used within EFAS (Thielen et al. 2009), is a spatially distributed raster-based rainfall-runoff model, which was developed by the JRC (Bartholmes et al. 2008; Van Der Knijff et al. 2010; Burek et al. 2013). It is driven by meteorological forcing data (precipitation, temperature, potential evapotranspiration, and evaporation rates for open water and bare soil surfaces) interpolated to a similar grid as the model. LISFLOOD calculates a complete daily water balance for every grid cell within the model domain. Land surface processes, such as snowmelt, soil freezing, surface runoff, infiltration into the soil, preferential flow, redistribution of soil moisture within the soil profile, drainage of water to the groundwater system, groundwater storage, and groundwater base flow, are simulated for each grid cell. The groundwater storage in the LISFLOOD model is modeled based on two parallel linear reservoirs. The upper zone controls the quick runoff component, which includes fast groundwater and subsurface flow, for example, through macropores in the soil. Runoff from the upper zone ceases during a dry period. The lower zone drives the slow groundwater component that generates the base flow dependent on the stored water in this zone. Runoff generated for every grid cell is routed through the river network using a kinematic wave approach, with considering water retention in lakes and reservoirs.

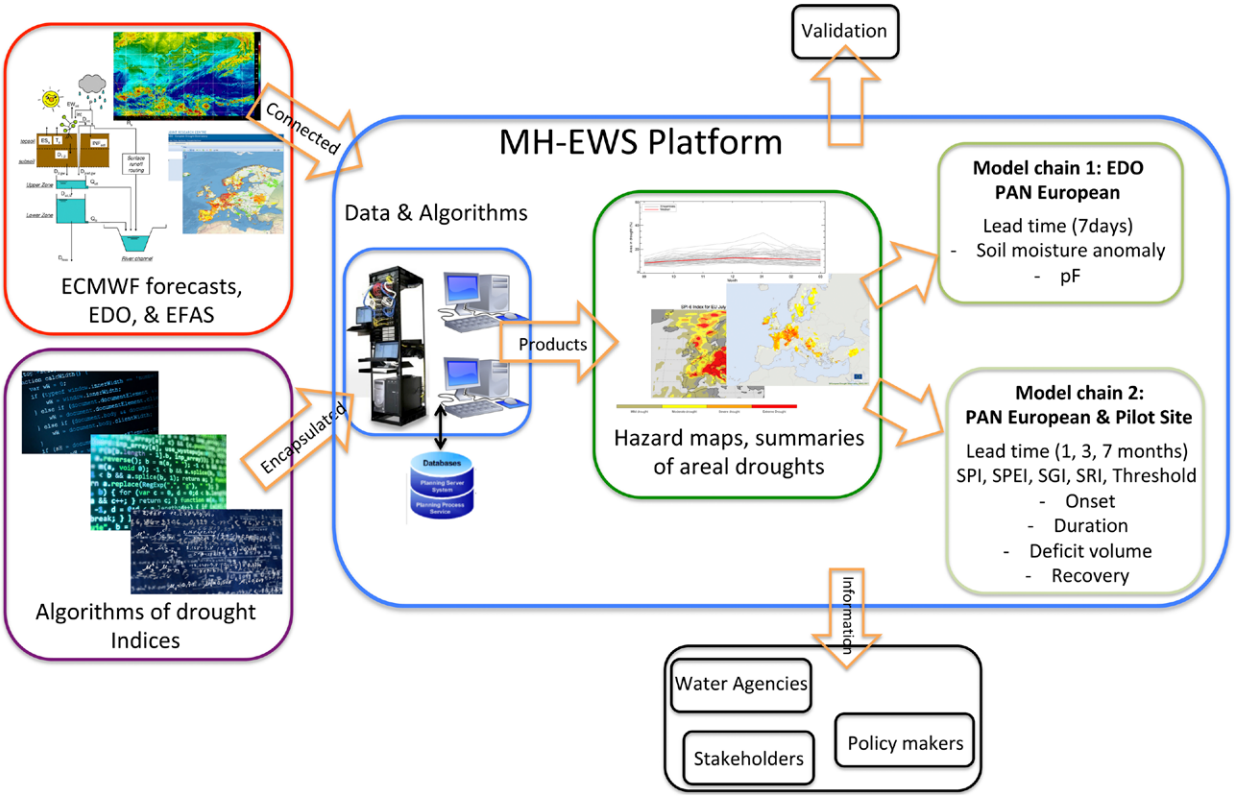

Fig. 1. General schematization of the AD-EWS.

The LISFLOOD model setup for Europe uses spatially distributed thematic data on soil, vegetation, and land use obtained from European datasets (Bartholmes et al. 2008; Van Der Knijff et al. 2010). Time series of river discharges in 258 catchments were used for calibration by tuning a number of parameters that control snowmelt, overland flow, river flow, infiltration, and residence times in the soil and subsurface (Feyen and Dankers 2009; Forzieri et al. 2014). LISFLOOD obtained a median Nash-Sutcliffe coefficient of efficiency (NSE) of 0.57 over the validation period. Basins with large discrepancies between the observed and simulated flow statistics were situated mainly in the south of Europe (Iberian Peninsula) and in the Baltic coast region (Zajac et al. 2013; Smith et al. 2016). The LISFLOOD model was selected for operational flood and drought forecasting after considering many factors in the model's selection, including, for example, model's performance, cost of implementation, and feasibility of technical implementation (Kauffeldt et al. 2016). It has been tested in Europe to produce robust high-flow forecasts and showed a promising result (Pappenberger et al. 2011; Arnal et al. 2018). Although originally the LISFLOOD model was developed for flood forecasting, the model is also capable to realistically produce low flows in Europe (Sepulcre-Canto et al. 2012; Forzieri et al. 2014; Wetterhall and Di Giuseppe 2018), as well as in Africa (Trambauer et al. 2013).

The hydrological seasonal forecasts in EFAS use the ECMWF System 5 (SYS5) seasonal forecast as forcing (Stockdale et al. 2018). The forecasting system covers the global domain and consists of several models: an atmospheric general circulation model, an ocean general circulation model, an ocean wave model, and a land surface model. It also contains perturbation models for the data assimilation and generation of forecast ensembles (ENS; 51 ensemble members). Postprocessing to remove model bias in the forecast system was carried out using two slightly different methods, with a posteriori correction based on the assumption of a quasi-linear behavior of the atmosphere and ocean anomalies. The mean bias of the model relative to observations estimated as a function of lead time and calendar month for the difference between model's reforecast values and verifying analyses is applied to Niño indices. For all other predicted variables, biases are only considered for model anomalies concerning to the model mean state. Detailed information on the ECMWF SEAS5 and postprocessing can be downloaded online (www.ecmwf.int/sites/default/files/medialibrary/2017-10/System5_guide.pdf).

SEAS5 are available on the tenth day of each month and include forecasted weather variables from day 1 to day 215 (~7-month lead time) for 51 ensemble members. For reforecasts, only 15 ensemble members are provided. A SPHEREMAP interpolation technique was applied in the current version of EFAS since 2019, replacing the inverse distance weighting scheme (Arnal et al. 2018; Smith et al. 2016), to all meteorological variables before these can be used as input for the hydrological model LISFLOOD. This method generates grids more reliably and with lower overall uncertainty, in particular, compared with ordinary kriging (Arnal et al. 2019). However, this method has not been tested against the external drift kriging (EDK) method yet, which proves to be an effective method for reducing orographic effects in the spatial distribution of precipitation and temperature (Zink et al. 2017). The temperature data are also corrected using the surface elevation. The downscaling technique must be carried out since the ECMWF SEAS5 products have a coarser spatial resolution (~35 km) than the LISFLOOD model (5 km), which hence is suited to better analyze hazard impacts at a finer resolution. The use of finer resolution in the hydrological model is anticipated to increase the forecast skill caused by this model (Wanders et al. 2019). Compared to ensemble streamflow prediction (ESP; Day 1985), the ECMWF S4 is, in general, more skillful, particularly for certain seasons and lead times (Arnal et al. 2018; Trambauer et al. 2015; Wanders et al. 2019; Samaniego et al. 2019).

A hydrological climatology, referred to as the simulation forced with observation (SFO), was used as a reference. It was created by forcing LISFLOOD with gridded meteorological

forcing data to provide a continuously updated simulation of current conditions with the model, as done by Yuan et al. (2016) for a Chinese region. The gridded meteorological observation data are collected from ground observations (>5,000 synoptic stations), obtained from various sources, such as the Global Telecommunication System of the WMO, the JRC meteorological database, and high-resolution data received from the National member States institutions (Pappenberger et al. 2011; Wetterhall and Di Giuseppe 2018). SFO is used to 1) provide hydrological initial conditions (HIC) for the forecasts and 2) monitor the current state of the European hydrological system. Both SFO and gridded meteorological observations are referred as a proxy data for hydrological and meteorological observed data (from 1990 to the present).

Examples of AD-EWS drought forecasting products

The AD-EWS provides a wide range of hydrometeorological drought forecasting products to fulfill the needs of different users and to cover the multifaceted drought phenomenon. In the AD-EWS, we include meteorological, soil moisture and hydrological drought [see Tallaksen and Van Lanen (2004) and Mishra and Singh (2010) for drought types]. With the wide range of drought products provided in the AD-EWS, different drought impacts can be investigated. Every impact requires its drought index or own set of drought indices since not every drought impact can be derived from each drought index (Wanders et al. 2017). Thus, it is very important to deliver suitable indices for studying the targeted impact. For the agricultural sector, information of water availability for plants to grow (precipitation and soil moisture drought indices in the upcoming months) is more relevant than, for instance, information on drought indices addressing runoff. This information can be obtained from, for example, the SPI-*x*, SPEI-*x*, precipitation and soil moisture threshold products. For water resources managers, on the other hand, the discharge, runoff, and groundwater drought forecasting products (e.g., discharge and runoff threshold indices, SRI-*x*, SGI-1) are far more important than products based on precipitation forecasts (e.g., SPI-*x*, precipitation threshold) (see appendix A for drought indices). One should note that the proper accumulation periods x and delay time depend on catchment characteristics and the impacted sector (Bloomfield and Marchant 2013; Kumar et al. 2016). With a wide range of different indices and accumulation periods provided by the AD-EWS, the users can decide by themselves which indices are best suited for their needs. However, suggestions are given by Kruse and Seidl (2013), who specify information about drought variables needed from the perspective of user groups.

In the AD-EWS, drought events in each grid cell in the European Union with lead times up to 7 months can be obtained from different standardized products in different percentiles of the 51 ensemble members (median and 10th, 25th, 75th, and 90th percentiles). All standardized products are provided for different accumulation periods (x = 1, 3, 6, and 12 months), except for the groundwater index (SGI) since groundwater at a certain time has a long-term memory of previous events (Bloomfield and Marchant 2013). We applied different accumulation periods to runoff using the SRI index because runoff can react differently to precipitation, for example, a fast to slow response to previous precipitation events (Van Loon and Van Lanen 2012). Figure 2 shows some examples of the forecasted severity of European drought events calculated from the standardized drought products (SPI-6, SPEI-6, SRI-6, and SGI-1) based on the median of the 51 ensemble members for May through July 2018. The standardized indices were calculated using hydrometeorological data forecasted at the beginning of May 2018.

All drought forecasting products using different standardized indices produce similar drought pattern in Europe. Northern, central, and northeast European countries and the United Kingdom were forecasted to experience mild to extreme drought events based on SPI-6, SPEI-6, and SRI-6 (Figs. 2a–i). The forecasted groundwater index (SGI; Figs. 2j–l), however,

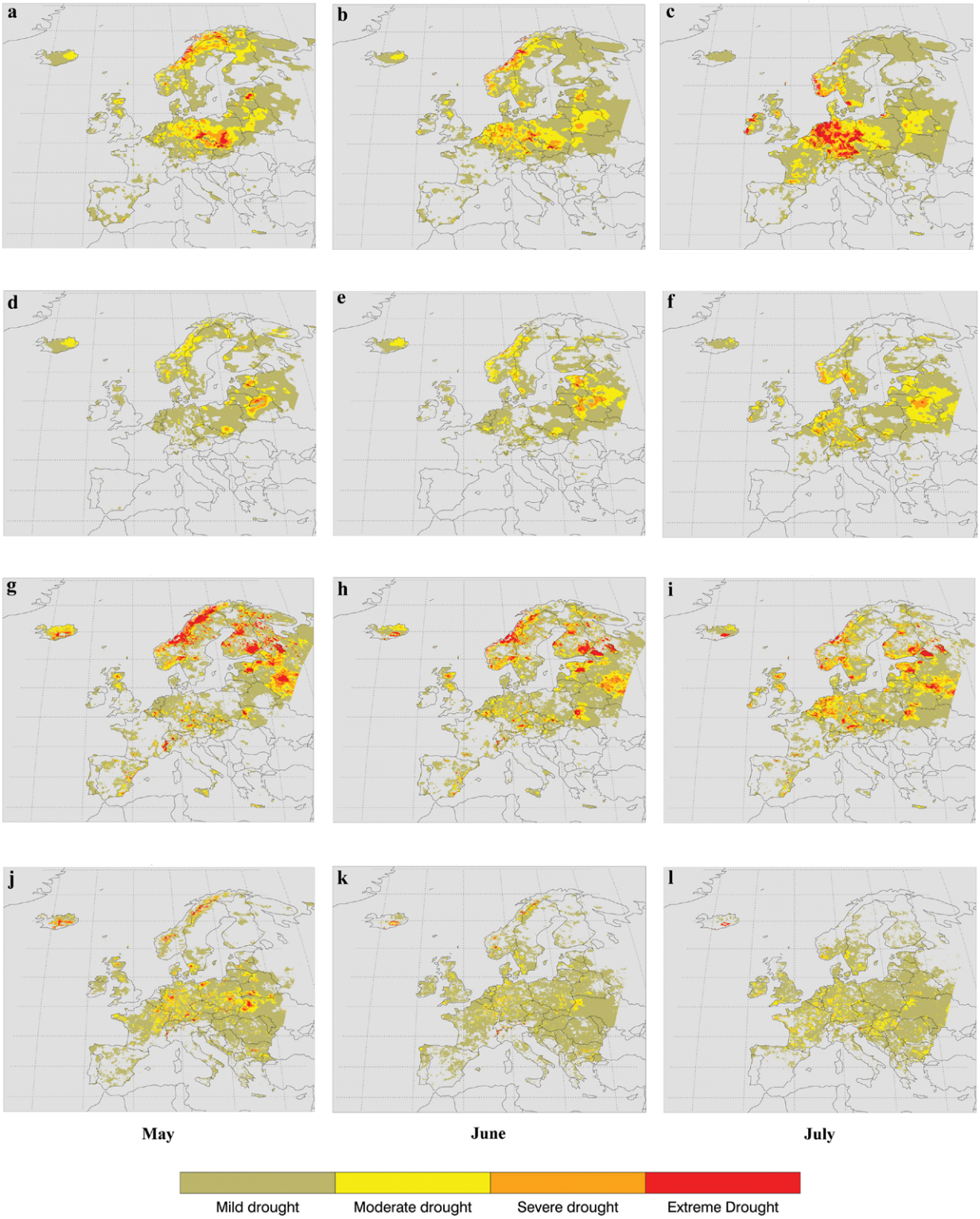

Fig. 2. Forecasted drought severity class (median of the 51 ensemble members) in several hydrometeorological variables using (a)–(c) SPI-6, (d)–(f) SPEI-6, (g)–(i) SRI-6, and (j)–(l) SGI-1 (left) May, (center) June, and (right) July 2018 obtained from the forecast done on 2 May 2018.

only shows mild to moderate drought for the same drought regions as for the other standardized indices. The drought severity was forecasted to increase in the second and third months (June and July 2018) in all products. The SPI produces higher drought severity compared to the others, indicated by more reddish colors on the maps. The higher drought severity shown by meteorological drought indices (i.e., SPI) than hydrological drought indices (e.g., SRI and SGI) indicates drought propagation from meteorological to hydrological drought (Peters et al.

2003; Van Loon et al. 2012). Figure 2 also demonstrates that the extreme drought in northern and central Europe in July 2018 has been forecasted 3 months ahead. For the standardized indices, one should note that we define mild drought for indices lower than 0 up to –1, following McKee et al. (1993). In the SPI user guide published by WMO (2012), however, SPI values from –0.99 to 0.99 are categorized as near-normal condition.

In addition to the standardized drought indices products, which give monthly results, the threshold drought method allows us to enrich the drought forecasting products by providing daily drought information. The onset and termination of the longest drought (day number) within the forecast period (7 months), the total drought duration (in days rather than in months as for the standardized indices), the deficit volume (for fluxes), and average cumulative deviation from threshold (for state variables) on daily basis for 7 months are forecasted for precipitation, soil moisture, runoff, and groundwater (see Fig. 3 for examples). For drought in discharge, there are three products available, showing the number of ensemble members in drought (*n* out of 51), drought duration, and drought deficit volume in the longitudinal profile of the major European river (see Figs. 3e,f). The forecasted drought onset and termination will give information to the users on when the longest drought will start in a certain region and when it will end. For example, in the south of Norway and Sweden, the soil moisture drought is forecasted to start on day 90 and will end on day 200 (Figs. 3a,b). With this information, users can improve water resources management in their regions to cope with the foreseen drought. The forecasted total drought duration will give additional information on how long all drought events in the 7-month period will strike. Users need to note that the drought duration gives information on the total drought duration, including minor drought events, while drought onset and termination give information only for the longest drought event within the 7 months. The forecasted drought deficit volume/average cumulative deviation from threshold is potentially useful information especially for water resources managers. It provides insight on how much water is predicted to be missing for the coming months. Thus, they can decide to release less water from the reservoirs, for instance, to maintain ecological minimum flow or to put a restriction on groundwater abstraction for irrigation. The standardized drought indices cannot determine the drought deficit volume.

In addition to pan-European maps with forecasted drought occurrence and characteristic (Figs. 2 and 3), the AD-EWS also provides areal drought indices, which are time series of the average percentage area in drought for each month over the 7-month forecast period (only for moderate to extreme drought forecasted by the standardized indices; Fig. 4). Moreover, bar diagrams are presented that denote for each of the 51 ensemble members the probability of the total drought duration of the area in drought and the total deficit volume/average cumulative deviation from the threshold of the area in drought. The time series of the area in drought and total drought duration of the area in drought are forecasted using both the standardized indices and threshold drought indices. Threshold drought indices are only used to forecast the total deficit volume/average cumulative deviation from the threshold of the area in drought. Examples of the areal drought forecast products are given in Fig. 4.

Figure 4a shows that around 5%–13% of the pan-European region is predicted to experience drought in the runoff in May 2018 and the situation is expected to slightly improve in June. However, there is a probability that the drought area will slightly increase again from 6% in June 2018 to 12% by the end of the forecast period, based on the median of the ensemble. Figure 4b provides the forecasted 7-month drought deficit volume in precipitation for the area in drought of each ensemble member. This forecast provides information on how severe the precipitation drought across the European continent is. For example, the median of ensemble member 28 says that 7 mm of rain is missing in the coming 7 months, and 28 mm of rain is missing according to ensemble member 4. To that end, the areal drought information provided by the AD-EWS summarizes the foreseen droughts properties for the

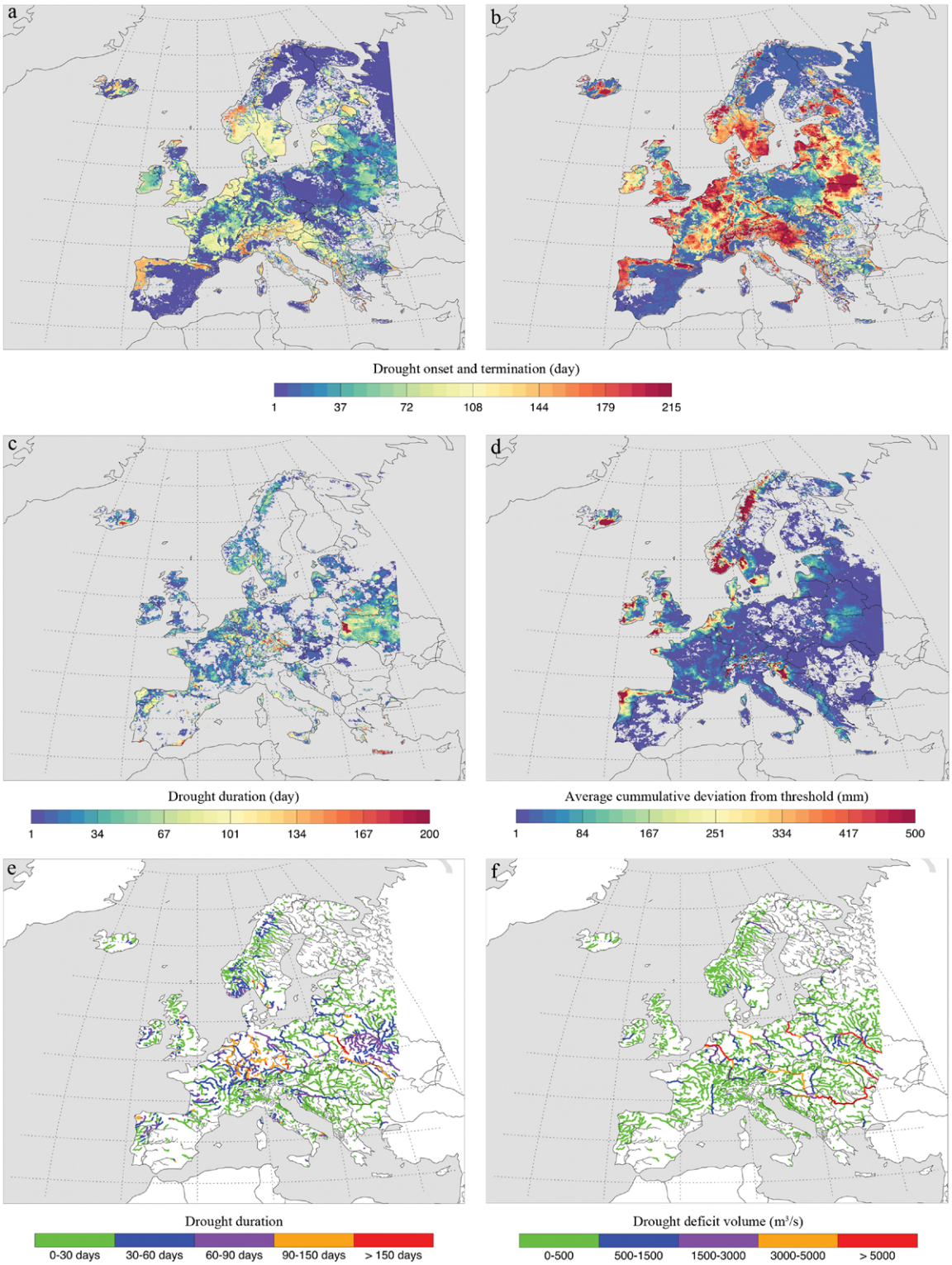

Fig. 3. Examples of drought forecast products for different hydrological variables. (a) Drought onset for the longest soil moisture drought event, (b) drought termination for the longest soil moisture drought event, (c) drought duration in the runoff, (d) average cumulative deviation from the threshold in groundwater, (e) drought duration in discharge, and (f) cumulative drought deficit volume in discharge. All data hold for the median of the 51 ensemble members within the forecast period of 7 months and obtained from the forecast on 2 May 2018.

whole Europe, and it might be useful for regional users or EU services and statistics. Table 1 summarizes all the drought forecast products provided by the AD-EWS for a wide range of stakeholders and users.

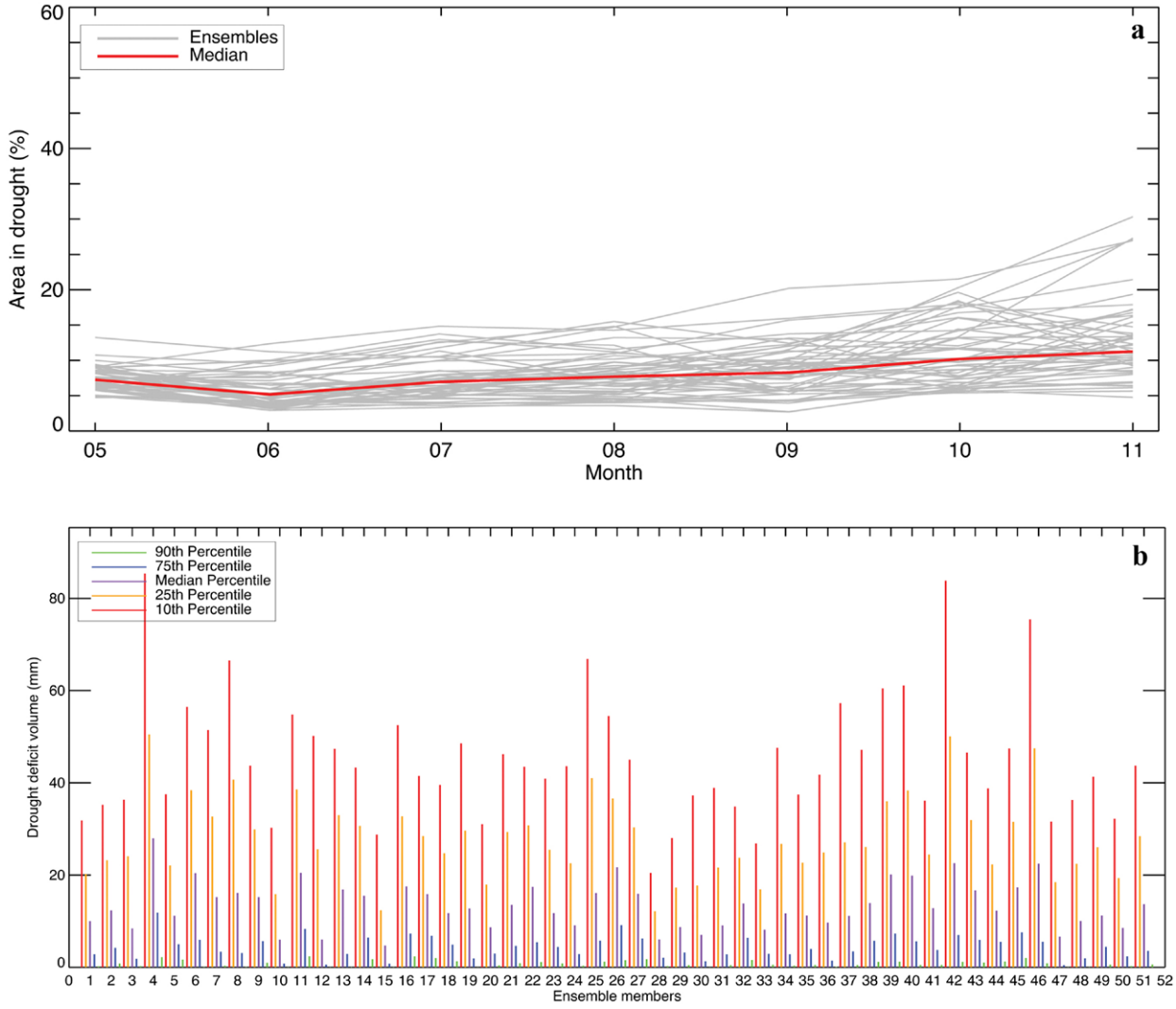

Fig. 4. Examples of areal drought forecast products. (a) Forecasted monthly area in drought (percentage of Europe) for moderate to extreme runoff drought and (b) forecasted 7-month drought deficit volume in precipitation (median and 10th, 25th, 75th, and 90th percentiles) for each ensemble member. All data are obtained from the forecast done on 2 May 2018.

The pan-European 2003 drought

The 2003 drought that covered a large area in Europe (e.g., Schär et al. 2004; Fink et al. 2006; EEA 2010; Ionita et al. 2017; Laaha et al. 2017) was used to validate AD-EWS. The 2003 drought that particularly was severe in central Europe caused economic losses estimated at least EUR 8.7 billion and heat-related dead toll exceeded 70,000 casualties (EU 2007; Robine et al. 2008). Thus, we selected two indices, SPI and SRI, from a wide range of drought indices available in the AD-EWS to represent the meteorological drought (SPI) and hydrological drought (SRI) that was hit Europe in 2003. Figure 5 shows the hydrometeorological drought class for the SPI-3 (Figs. 5a,b) and SRI-3 (Figs. 5d,e), derived from the median ensemble forecast and proxy observations. Forecasted SPI-3 and SRI-3 show severe to extreme droughts in the west and central Europe, such as France, Germany, and Benelux. The droughts derived from proxy observed data show wider affected areas by extreme drought, especially in SPI-3 (Fig. 5b). To investigate the skill of the forecasts, we present a new approach based on a categorical drought classification (see appendix A for the method). This simple approach has been introduced to evaluate drought forecasting score that is well understood by end users compared to common skill metrics, such as Brier skill score (BSS; Brier. 1950), relative operating characteristic (ROC) curve (Mason 1982), and the equitable threat score (Rogers et al. 1995). In several international

Table 1. A summary of drought products.

No.	Method	Resolution	Nature	Description
1	*Standardized indices*			
	SPI-$x_{j,k}$	Monthly	Raster maps	Calculated for multiple time scales except SGI (x = 1, 3, 6, and 12 months), lead times (j = 1, 2, 3, …, 7), and the indices will be provided for k = the median and the 10th, 25th, 75th, and 90th percentiles obtained from 51 ensemble members.
	SPEI-$x_{j,k}$	Monthly	Raster maps	
	SRI-$x_{j,k}$	Monthly	Raster maps	
	SGI-$1_{j,k}$	Monthly	Raster maps	
2	*Threshold indices*			
	Drought deficit volume	Daily	Raster maps	Calculated for precipitation, soil moisture, groundwater, and runoff, and provided for k = the median and the 10th, 25th, 75th, and 90th percentiles obtained from 51 ensemble members. Maps show total drought deficit volume (mm) during the 7-month forecast period.
	Drought duration	Daily	Raster maps	Calculated for precipitation, soil moisture, groundwater, and runoff, and provided for k = the median and the 10th, 25th, 75th, and 90th percentiles obtained from 51 ensemble members. Maps show total drought duration (days) during the 7-month forecast period.
	Drought onset and termination	Daily	Raster maps	Calculated for precipitation, soil moisture, groundwater, and runoff, and provided for k = the median and the 10th, 25th, 75th, and 90th percentiles obtained from 51 ensemble members. Maps show the day when the maximum drought started and terminated (day number) during the 7-month forecast.
	Drought in discharge	Daily	Raster maps	Drought characteristic calculated from discharge in major and runoff, rivers and provided for k = the median and the 10th, 25th, 75th, and 90th percentiles obtained from 51 ensemble members. Maps show for each river node: 1) number of ensemble members in drought by the end of 30, 90, and 215 days; 2) total drought duration (days) during the 7-month forecast period; and 3) total drought deficit volume during the 7-month forecast period.
3	*Areal indices*			
	Area in drought	Monthly	Time series	Forecasted of temporal (monthly) evolution of area in drought (standardized drought indices, multiple time scales of x = 1, 3, 6, and 12 months; threshold drought indices) during the 7-month forecast period (median and 51 ensemble members).
	Drought duration of area in drought	—	Bar diagram	Forecasted total drought duration of area in drought (standardized drought indices, multiple time scales of x = 1, 3, 6, and 12 months; threshold drought indices) during the 7-month forecast period in probabilistic terms (percentiles derived from cells in drought for each ensemble member).
	Drought deficit volume of area in drought	—	Bar diagram	Forecasted of total drought deficit volume of area in drought (standardized drought indices, multiple time scales of x = 1, 3, 6, and 12 months; threshold drought indices) during the 7-month forecast period in probabilistic terms (percentiles derived from cells in drought for each ensemble member).

EU-funded projects [e.g., Exercise to Assess Research Needs and Policy Choices in Areas of Drought (XEROCHORE), Fostering European Drought Research and Science-Policy Interfacing (DROUGHT-R&SPI), ANYWHERE], end users said that the abovementioned skill metrics, which are common in scientific literature, are difficult to understand and to transfer to their local conditions.

Figure 5 also shows the forecast scores for meteorological drought (Fig. 5c) and hydrological drought (Fig. 5f) (see appendix B for the method). In 58% of the European area, the forecast score of the meteorological drought is 0, indicating the forecasted drought class equal to the observed class, and in 19% (20%) the score is +1 (–1), meaning one class difference with a bias to overestimation (underestimation) of the observed drought. Figure 5c also shows that in some European regions (southeast France), AD-EWS is unable to produce the same drought class as

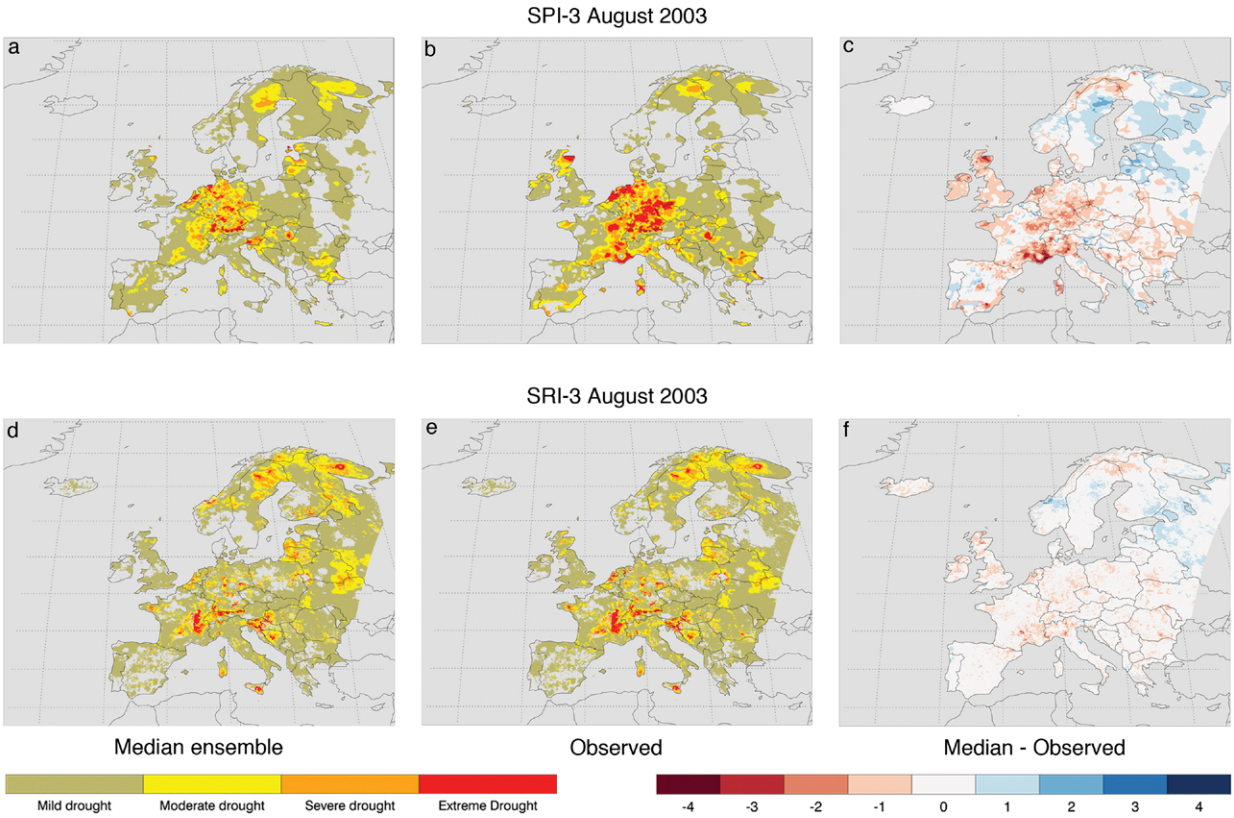

Fig. 5. Hydrometeorological drought in 2003 expressed as drought severity classes using (top) the standardized precipitation index and (bottom) the standardized runoff index accumulated over 3 months (SPI-3 and SRI-3, respectively): (a) forecasted SPI-3 done in early August 2003 (median of 15 ensembles) for a lead time of 1 month, (b) SPI-3 obtained from observations for August 2003 (SFO), (c) drought forecasts score expressed as the difference in drought class between the forecast and the observed. (d)–(f) As in (a)–(c), but for SRI-3. Reddish colors in (c) and (f) indicate that the forecast underestimates the drought class and vice versa for bluish colors.

observed (difference up to 4 classes). The forecasting scores for the hydrological drought are better (Fig. 5f). In 85% of the European area the score is 0, whereas in 10% (5%) of the area the score is +1 (–1). A detailed analysis of the forecast class scores of the seasonal hydrometeorological drought forecasts (Figs. 5c,f), denoted by SPI and SRI is presented in Fig. 6 and Fig. ES1 in the online supplemental material (https://doi.org/10.1175/BAMS-D-18-0196.2), respectively.

Figure 6 shows the forecast score of the meteorological drought (SPI with accumulation periods of 1, 3, 6, and 12 months) for the pan-European domain and for lead times of 1–5 months. In 40%–50% of the area, perfect forecasts (score = 0) for the SPI-1 are found for all lead times (except SON). For SPI-3 with a lead time of 1 month, the area increases to 50%–70%. As expected, the area decreases for longer lead times. For instance, for a lead time of 3 months, the area is between 30% and 50%. When assessing the forecast scores, we need to consider that some observational data are included in the SPI-3 with lead times of 1 and 2 months (Fig. A2). SPI-12, which has the longest accumulation period, produces the highest forecasting score with perfect forecasts for >60% of the area up to 3 months ahead in all seasons. As mentioned above, hydrological drought, denoted by SRI (Fig. ES1), shows better scores than meteorological drought, denoted by SPI (Fig. 6). The score class of SRI-1 is comparable with SPI-3 to SPI-6 and SRI-3 is comparable with SPI-6 and SPI-12, as expressed by similar values of perfect forecasts. Interestingly, meteorological drought forecasts tend to underestimate the drought (reddish color dominant for negative class differences) and hydrological drought forecasts tend to overestimate drought (reddish color dominant for positive class differences).

Season	Class difference	SPI-1 with lead times of (month)					SPI-3 with lead times of (month)					SPI-6 with lead times of (month)					SPI-12 with lead times of (month)				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Winter (DJF)																					
	none	47.7	37.4	34.6	46.1	43.9	63.9	42.2	34.2	39.0	40.7	75.7	65.6	57.1	53.1	40.6	85.0	72.7	66.5	67.5	64.6
	+1	18.7	8.7	11.9	6.1	5.2	11.9	15.2	14.3	7.8	3.4	9.4	11.7	12.4	6.7	6.6	6.6	7.9	6.7	4.8	5.5
	+2	1.3	0.3	0.5	0.0	0.1	1.9	1.4	0.8	0.2	0.1	0.7	1.4	0.6	0.2	0.4	0.3	0.6	0.5	0.4	0.4
	+3	0.3	0.1	0.0	0.0	0.0	0.5	0.2	0.2	0.1	0.0	0.0	0.2	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	+4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	-1	19.1	29.6	35.1	36.0	29.4	15.9	25.8	27.4	29.8	37.9	10.0	18.1	23.9	25.9	33.8	7.9	17.6	21.8	22.6	25.1
	-2	5.8	11.1	10.1	7.3	11.1	4.8	9.0	13.3	12.0	10.2	0.3	1.7	6.8	10.1	10.2	0.1	1.1	4.1	4.6	3.9
	-3	4.1	7.1	4.6	3.2	6.8	1.0	4.6	7.0	7.8	5.2	0.0	0.1	2.0	3.9	4.4	0.0	0.0	0.3	0.3	0.4
	-4	2.8	5.6	3.5	1.2	3.5	0.0	1.5	2.6	3.2	2.4	0.0	0.0	0.1	0.7	1.5	0.0	0.0	0.0	0.0	0.0
Spring (MAM)																					
	none	42.8	44.3	43.2	44.5	41.2	51.0	38.2	39.6	40.9	42.5	71.7	53.3	43.5	39.3	39.1	85.3	74.0	65.2	59.5	55.1
	+1	8.8	4.7	10.7	11.2	17.2	8.4	6.9	5.2	6.4	10.8	6.4	6.2	7.3	6.4	5.8	3.8	4.4	5.7	5.9	6.4
	+2	0.5	0.1	0.4	0.7	1.3	0.9	0.3	0.5	0.6	2.1	0.5	0.4	0.7	0.7	0.6	0.0	0.2	0.4	0.4	0.4
	+3	0.1	0.0	0.0	0.2	0.2	0.3	0.1	0.1	0.1	0.4	0.0	0.0	0.1	0.1	0.2	0.0	0.0	0.0	0.0	0.0
	+4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	-1	30.5	33.4	29.4	28.2	26.8	30.4	31.3	34.7	32.9	28.4	19.9	29.0	31.8	33.0	33.1	10.8	20.4	24.8	28.2	31.1
	-2	9.3	10.1	9.2	8.7	7.7	7.6	14.3	11.1	10.9	8.8	1.3	9.1	11.3	12.2	11.0	0.0	1.0	3.7	5.3	5.8
	-3	5.0	5.0	4.8	4.3	3.7	1.4	7.0	6.1	5.9	4.7	0.0	1.7	4.4	6.5	6.5	0.0	0.0	0.2	0.6	1.0
	-4	3.1	2.2	2.1	2.1	1.7	0.1	1.9	2.6	2.3	1.8	0.0	0.1	0.8	1.6	3.4	0.0	0.0	0.0	0.0	0.0
Summer (JJA)																					
	none	42.1	42.6	43.9	43.9	41.0	57.1	49.7	44.5	42.8	41.2	61.6	55.0	51.5	45.9	45.9	75.8	68.9	62.2	56.4	50.6
	+1	13.1	15.3	25.1	18.9	28.9	11.5	12.2	16.5	21.4	27.3	10.4	11.3	15.4	17.9	20.2	8.3	7.9	13.3	14.9	18.6
	+2	1.1	1.3	0.9	0.9	1.5	1.3	2.4	3.1	3.4	3.9	1.1	1.7	3.2	3.9	3.4	0.5	0.4	1.1	2.3	3.1
	+3	0.2	0.2	0.1	0.1	0.2	0.2	0.4	0.6	0.6	1.6	0.2	0.3	0.9	1.3	1.8	0.0	0.0	0.0	0.2	0.4
	+4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.3	0.0	0.0	0.3	0.3	0.6	0.0	0.0	0.0	0.0	0.1
	-1	29.6	27.4	18.4	25.7	22.1	24.5	24.4	22.3	21.9	19.9	23.3	22.9	20.0	21.4	18.8	14.6	20.0	20.7	21.5	20.3
	-2	8.1	7.2	6.2	6.9	4.7	4.5	7.1	7.2	5.5	3.4	3.0	6.7	5.8	6.0	5.8	0.8	2.5	2.4	3.9	5.0
	-3	3.8	4.0	3.7	2.6	1.3	0.8	3.2	4.0	2.6	1.4	0.3	1.9	2.5	2.6	2.7	0.0	0.3	0.2	0.7	1.7
	-4	1.7	1.9	1.5	0.8	0.2	0.0	0.6	1.7	1.6	0.9	0.0	0.2	0.2	0.5	0.8	0.0	0.0	0.0	0.0	0.1
Autumn (SON)																					
	none	50.3	46.7	44.7	55.4	54.8	69.1	60.1	50.3	55.3	60.3	73.9	66.0	63.9	63.0	64.2	81.6	73.2	68.4	65.6	64.6
	+1	18.1	19.4	22.7	17.4	15.9	13.9	16.9	22.4	23.6	20.2	12.9	15.0	15.6	18.4	18.1	8.8	11.1	13.9	15.9	15.7
	+2	1.5	0.5	0.6	0.4	0.6	1.3	1.1	1.7	1.1	0.8	0.8	1.4	1.6	1.1	1.0	0.1	0.7	1.2	1.3	1.3
	+3	0.1	0.1	0.1	0.0	0.0	0.1	0.2	0.4	0.3	0.2	0.1	0.3	0.2	0.3	0.3	0.0	0.0	0.1	0.1	0.2
	+4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	-1	20.9	26.9	26.4	21.5	22.3	14.4	18.5	22.5	18.5	17.7	11.9	14.4	15.4	14.9	15.3	9.3	13.8	13.9	14.8	15.9
	-2	6.6	4.9	4.5	4.0	4.8	1.1	2.2	2.1	1.0	0.6	0.4	2.5	2.5	1.7	0.8	0.1	1.0	2.2	2.0	2.0
	-3	2.1	1.2	1.0	0.7	1.0	0.1	0.7	0.4	0.1	0.0	0.0	0.3	0.6	0.4	0.1	0.0	0.0	0.3	0.2	0.3
	-4	0.3	0.3	0.1	0.4	0.5	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0

Fig. 6. Meteorological drought (SPI) forecasting scores for the pan-European 2003 drought. Scores are expressed as the difference between the drought class derived from the median of 15 ensemble forecasts and the class obtained from the observed data. Scores, i.e., percentage of cells that agree (none; no class difference) and disagree (–4 to +4 class differences) are provided for SPI with accumulation periods of 1, 3, 6, and 12 months for the four seasons as the starting forecast months, with different lead times (1–5 months). The light green color indicates high forecasting scores, the light brown color indicates medium forecasting scores, the light red color indicates low forecasting scores, and the white color indicates 0% of the area (see appendix B for the detailed explanation of the color coding). Please note that we present Fig. 6 as a figure (not as a table) because we provide different colors indicating different forecasting scores.

An example of the evolution of the ensemble hydrological drought forecasts in 2003 for a randomly selected grid cell in central France is given in Fig. 7. The drought is represented by SRI-3 with a lead time of 1–7 months for 15 ensemble members initialized in the middle of each session [January for winter (DJF), April for spring (MAM), and so on]. In general, the observed hydrological drought event in 2003 is found within the ensemble spread. The median forecast initiated in January 2003 produces a higher drought index than the observed, and it is closer to the lower ensemble members than to the median. The median forecast indicates good skill for forecasts done in April and July 2003. The forecast done in October 2003 provides information that the drought in 2003 would be alleviated in 2004,

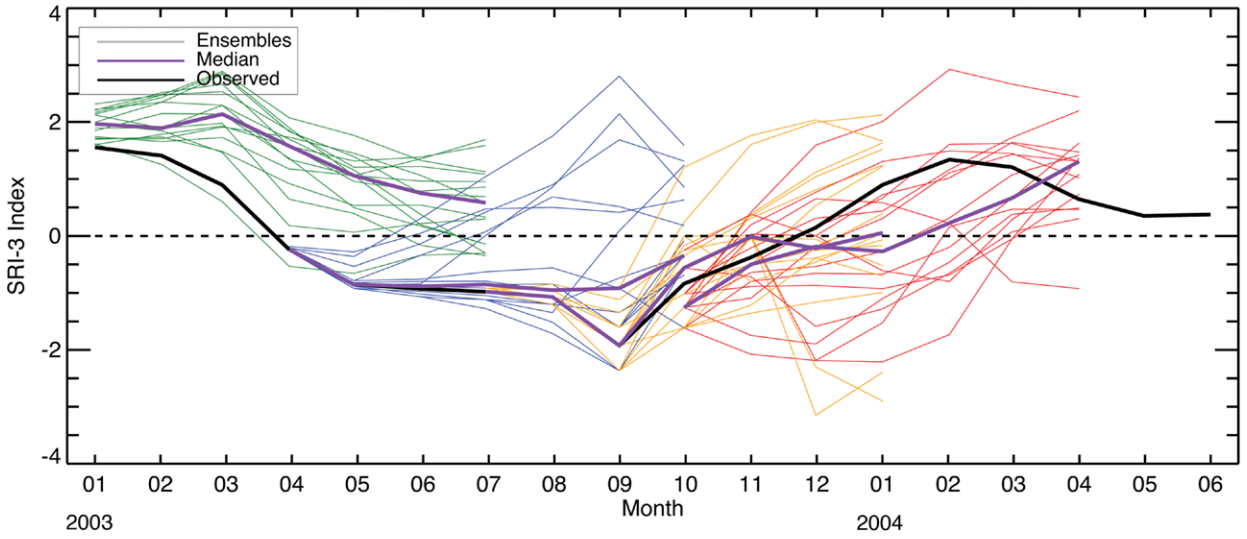

Fig. 7. Seven-month hydrological (SRI-3) forecasts (median in purple color, and ensemble members in gray color) and SRI-3 derived from proxy observed runoff of the European 2003 drought for one grid cell (5 km × 5 km) in central France. The 7-month forecasted is presented at four issue times: January (green), April (blue), July (orange), and October (red) 2003.

although the forecasted drought is still more severe than the observed drought. The difference between the forecasted drought class derived from the 25th percentile of 15 ensemble members and the class from the observed is presented in Fig. ES2. A comparison between the median (Fig. ES1) and the 25th percentile of 15 ensemble members (Fig. ES2) indicates that the median ensemble produces higher no class different (perfect forecast) than the 25th percentile, except during summer. This indicates that the median ensemble is better in representing the drought severity than the 25th percentile.

Forecasts user-defined trust level

The scores of hydrological drought forecasts are higher than for the meteorological drought forecasts. The higher forecast scores come from the hydrological initial condition and catchment storage. Even though the AD-EWS can provide different seasonal drought forecasting products with better skill than seasonal weather forecasts, end users are often questioning forecast developers with a simple, but tough question: How confident are you with the results? This question has no unique answer since it varies for every user. To date, we can relate the user-defined trust level of our forecasts with the forecast skill itself. Arnal et al. (2018) showed that the hydrological forecasts have limited skill beyond 8 weeks for some regions and seasons in Europe. However, some studies in the United States and Africa (e.g., Trambauer et al. 2015; Yuan et al. 2013a) showed that the skill of drought forecasts initiated during winter is higher than summer and encouraging for the first 2–3 months. Hydrological drought forecasts seem to have slightly higher skill than meteorological forecasts (an additional 1–2 months), as illustrated above through the forecast scores. Drought products should be used with this limitation in mind and be carefully tested before being put in use for decision-making.

Another factor beside the forecasting skill that may affect the trust level, that is, the feeling or belief that users can have faith in the forecast products is the personal judgment of the users. The selection made by the users on which components are more important than others can define the trust level. Here we give some examples of how the user-defined trust level can be analyzed based on the forecasts skill and user-driven decisions on what is important for them. We calculated the user-defined trust level using the forecast score (Fig. 6) in which

we determined the credibility of the forecasts to produce the drought classes (mild, moderate, severe, and extreme). The percentage of the area that is in agreement (i.e., the number of grid cells with perfect forecasts) or level of disagreement (mismatch for positive or negative signs) is combined with the weighting values defined by the users. A perfect forecasting score is achieved if there is no difference between drought classes derived from forecasts and observed. A positive (negative) forecast score means that the drought forecast is pointing at a more (less) severe drought than the one based on the observed data (overforecasting or underforecasting) (see appendix B). Using these two main components, the user-defined trust level of the forecasts can be calculated as follows:

$$y(t) = a_1(t)x_1(t) - \left[\sum_{k=2}^{5} a_k(t)x_k(t) + \sum_{l=6}^{9} a_l(t)x_l(t)\right], \quad (1)$$

where $y(t)$ is user-defined trust level in month/season t for a particular variable (hydrometeorological forecast product), $a_1(t)$ is the weighting factor for perfect forecasts in month/season t, $a_k(t)$ is the weighting factor for 1 to 4 class differences (overforecasting) in month/season t, $a_l(t)$ is the weighting factor for –1 to –4 class differences (underforecasting) in month/season t, $x_1(t)$ is percentage of area with perfect forecasts in month/season t for a particular variable, $x_k(t)$ is percentage of area for 1 to 4 class differences in month/season t for a particular variable, and $x_l(t)$ is percentage of area for –1 to –4 class differences in month/season t for a particular variable. The weighting factors [$a_1(t)$, $a_k(t)$, $a_l(t)$] (Fig. 8, left) are defined by the user, and reflect how much importance is given to, for instance, perfect forecasts or a certain level of mismatch (e.g., class differences of 4 or –4). The percentage of the areas with perfect/mismatch forecasts [$x_1(t)$, $x_k(t)$, $x_l(t)$] are taken from the forecast score that is presented in the previous section (Fig. 6). Therefore, the analysis of user-defined trust level is user dependent (how the user defines the weighting factors) and it is time dependent (the forecast has higher skill in winter than in other seasons). In the examples in this study, we used, for simplicity, the same weighting factor for plus and minus class differences ($a_k = a_l$), which implies that the end user gives equal weight to overforecasting and underforecasting. Total of weighting factors a for perfect forecasts and class differences is equal to 100%.

In this study we give three examples of user-defined trust levels for meteorological forecasts using SPI-x and cases that the users give more weight to 1) a perfect forecast, 2) a mismatch forecast, or 3) higher lead times. A higher weighting factor for perfect forecasts can be chosen when the users want to give more credit to the ability of forecasts to produce the same class as observed (Fig. 8a). In this decision, the highest weight is given for perfect forecasts and the lowest for the highest class difference (CD4). Figure 8c shows the weights if the users focus more on the mismatch between forecasts and observed. The weights for perfect forecasts are now lower and the weights for CD1 to CD4 are now higher than in Fig. 8a. Figure 8e is reversed from Fig. 8a when users want to give more credit to have perfect forecasts for higher lead times. This decision can be made if users want to have high user-defined trust level for longer lead times since drought event may last more than a season. However, users need to realize that giving more weight to longer lead times likely leading to lower user-defined trust levels due to low forecast skills for longer lead times.

Figures 8b, 8d, and 8f show that SPI with longer accumulation periods (i.e., SPI-12) has higher user-defined trust level than SPI with shorter accumulation periods (i.e., SPI-1), which comes from the fact that SPI-12 consists of mixing SFO data with forecasts. The high user-defined trust level in SPI-12 reflects the highest forecasting skill for SPI with longer accumulation periods. Giving more weight to perfect forecasts will result in high user-defined trust level from 35% up to 70% for a lead time of 1 month and for SPI-1 to SPI-12, respectively (Fig. 8b). More weight to class differences between forecasted and observed drought classes and less weight to perfect forecasts significantly reduces the user-defined trust level to around

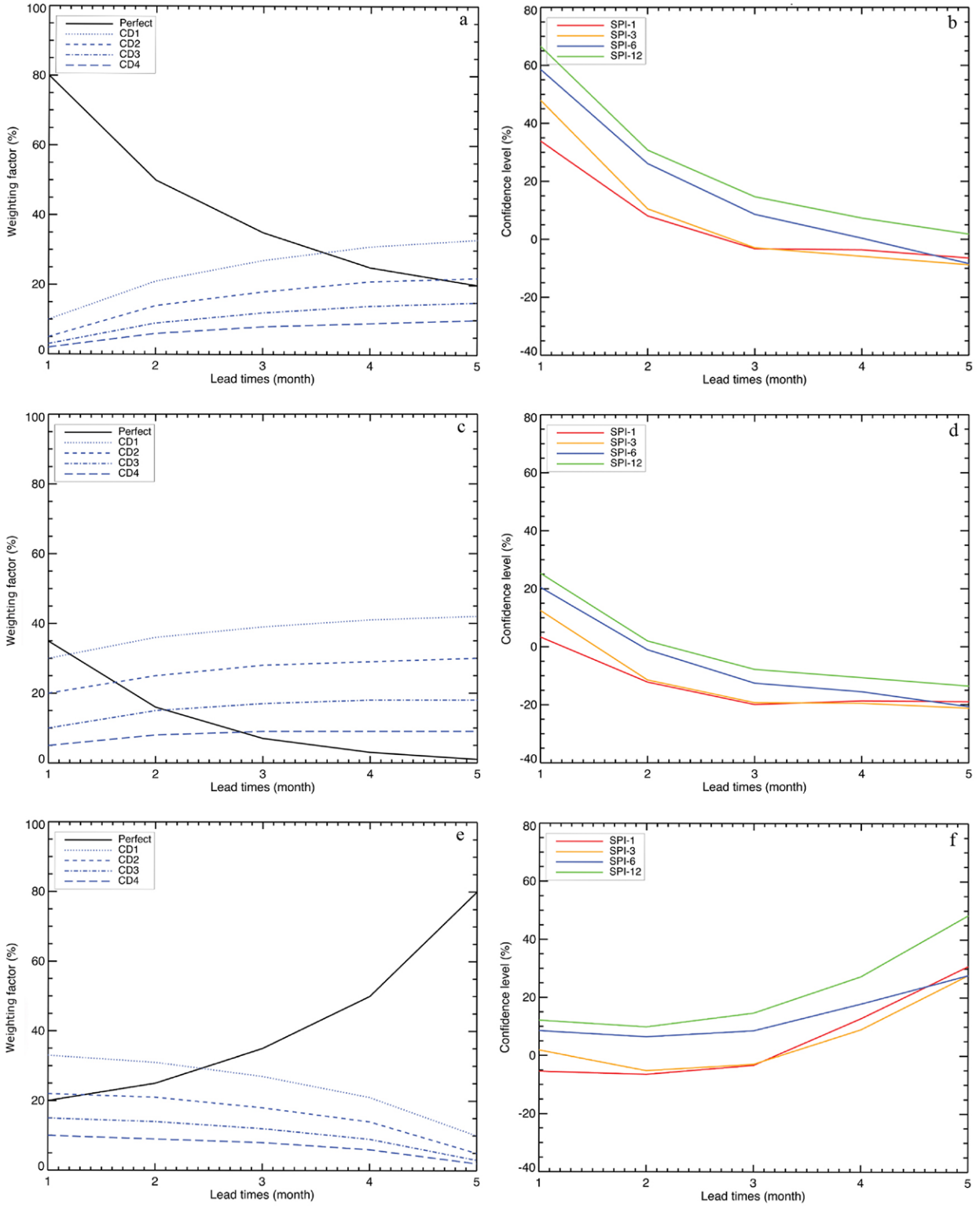

Fig. 8. Example of (left) predefined weights and (right) computed user-defined trust levels using Eq. (1) for SPI-*x* (*x* = 1, 3, 6, and 12 months) for lead times from 1 to 5 months: (a),(b) higher weight to perfect forecasts; (c),(d) higher weight to class differences (mismatch forecasts with observed); and (e),(f) higher weight to perfect forecasts for longer lead times. CD stands for class difference; for instance, CD4 means that the forecasted drought class differs four classes from the observed drought class.

10%–20% for a lead time of 1 month (Fig. 8d; SPI-1 even lower). This is plausible since class differences become more dominant (in total 65% of weight for a lead time of 1 month) compared to perfect forecasts (35% of weight). Figure 8f shows the highest user-defined trust level for the longest lead times. Although we gave a weight of 80% to perfect forecasts for

5-month lead times, the user-defined trust level for SPI-12 is not higher than 50%, which is smaller than for the SPI-12 for 1-month lead time in Fig. 8b (70%) with the same weight. Clearly, the lower skill of forecasts for higher lead times reduces the user-defined trust level at higher lead times.

Conclusions

A probabilistic seasonal pan-European hydrometeorological drought forecasting system (AD-EWS) has been developed with the aim to reduce drought impacts in Europe and to answer the challenge of the Sendai Framework for Disaster Risk Reduction 2015–30. This system has been in preoperational mode since mid-2018 as part of the ANYWHERE multihazard platform (www.aw-a4demos.eu and http://a4cat.hydsdev.net; see Fig. ES3 for an example). The AD-EWS provides users with comprehensive drought forecasting products, including a suite of different drought indices across Europe from a month up to seasons. This system provides hydrometeorological drought predictions of standardized indices (e.g., drought in precipitation, runoff, and groundwater) and threshold-based indices, such as drought onset, termination, deficit volume/average cumulative deviation from the threshold, and duration. The drought forecasting algorithms have been developed by utilizing the outputs of the LISFLOOD hydrological model fed by the ECMWF seasonal forecasting system SEAS5. These algorithms are encapsulated to the MH-EWS platform and run each month when the new seasonal weather forecast data become available. The developed AD-DEWS, which generates hydrological drought forecasts (e.g., drought deficit, or drought intensity in streamflow), is innovative by complementing existing hydrological forecasts (e.g., soil moisture, low streamflow).

In this paper, we illustrate the high potential of the AD-EWS with some examples from the wide range of drought forecasting products and its validation. The drought products represented by different indices are supporting each other in the sense that drought in the different components of the water cycle becomes plausible. The validation shows that AD-EWS can reproduce observed drought classes, especially for short lead times and for hydrological drought. The various drought forecast products provided by the AD-EWS cover the need of different users (e.g., decision-makers, water agencies, environment agencies), as identified in the EU project DROUGHT-R&SPI (Andreu et al. 2015) to anticipate possible drought conditions that may appear in upcoming months. Moreover, we present some examples of a user-driven assessment of forecasts trust levels by utilizing weights defined by the user and the forecasting skill. With this information, the users can take preliminary contingency actions, such as the implementation of water priority ranking, planning crop irrigation, energy supply (hydropower reduction, restricted release of cooling water), cargo planning for waterborne transport, water conservation for ecosystems, and more targeted monitoring of hydrometeorological variables, with a certain trust level. The ability of the AD-EWS to provide operational seasonal drought predictions in high spatial resolution and its diverse products mark the AD-EWS as the drought ensemble forecasting system that can serve a broad range of different users' needs in Europe.

Acknowledgments. The research is supported by the ANYWHERE project (Grant 700099), which is funded within the European Union's Horizon 2020 research and innovation program (www.anywhere-h2020.eu). The hydrometeorological output came from the EFAS computational center, which is part of the Copernicus Emergency Management Service (EMS) and early warning systems (EWS) funded by Framework Contract 198702 of the European Commission. This research is part of the Wageningen Institute for Environment and Climate Research (WIMEK-SENSE) and it supports the work of UNESCO EURO-FRIEND and the IAHS Panta Rhei program.

Appendix A: The drought algorithms

Standardized drought indices. Standardized indices are used to indicate anomalies (drought) in different domains of the water cycle (e.g., drought in precipitation, groundwater, and runoff). As mentioned before, the SPI-*x*, SPEI-*x*, SRI-*x*, and SGI-1 are used in the AD-EWS, where *x* indicates time scales or accumulation periods of 1, 3, 6, and 12 months (temporal resolution). The SPI allows determination of the probability of an ongoing drought at a given time scale (temporal resolution) for any site (e.g., grid cell) based on historical precipitation data (WMO 2012). However, it can also be applied in the forecasting mode (Yuan et al. 2013b; Dutra et al. 2014b), if precipitation forecasts are available. The SPI calculation for any grid cell is based on a long monthly precipitation record that is fitted to a probability distribution, which is then transformed into a normal distribution so that the median SPI for the site and desired period is zero (Edwards and McKee 1997). During the algorithms development period, 29 years of monthly precipitation data were used (from 1990 to 2018). Negative SPI values indicate that the precipitation on a certain month is lower than median precipitation and vice versa for positive SPI values. In the AD-EWS a drought event is assumed to occur when the SPI is less than 0 (i.e., that the event is dryer than the median for that grid cell) and ends when the SPI becomes greater or equal than 0. The drought event is categorized into four classes, which are (i) mild drought: $0 > \text{SPI} \geq -1$; (ii) moderate drought: $-1 > \text{SPI} \geq -1.5$; (iii) severe drought: $-1.5 > \text{SPI} \geq -2$; and (iv) extreme drought: $\text{SPI} \leq -2$. Each drought event has a duration defined by its beginning and end, and severity (i.e., magnitude of negative SPI) for each month that the event continues.

The SPI is solely based on precipitation, which can be interpreted as a weakness or strength (WMO 2012). As strength due to only precipitation data are required and as weakness due to no evapotranspiration and snow components are considered. Vicente-Serrano et al. (2010) have included the potential evapotranspiration and Staudinger et al. (2014) have included the snowmelt component, which is called the SPEI and SMRI, respectively. One should note that we do not include the SMRI index in our products. Standardized indices have also been developed for groundwater (SGI) and runoff (SRI) by Bloomfield and Marchant (2013) and Shukla and Wood (2008), respectively. The SGI and SRI include the snow component because LISFLOOD considers snow accumulation and melting. Differences in anomalies in precipitation, groundwater, and runoff reflect propagation of drought through the subsurface part of the hydrological cycle, which is relevant for drought impacts (Peters et al. 2003; Van Loon and Van Lanen 2012; Van Lanen et al. 2016). All standardized indices are calculated for the median and 10th, 25th, 75th, and 90th percentiles obtained from the 51 ensemble members.

Threshold-based drought indices. Similar to standardized indices, the threshold indices are also frequently used in particular to quantify water deficits (drought) in different domains of the water cycle (precipitation, soil moisture, runoff, discharge, and groundwater; Van Loon 2015). Compared to standardized indices, the threshold approach can quantify missing water volumes that are needed to recover from the drought. It is based on defining a threshold, below which the precipitation or other hydrological variables is considered in a drought (Yevjevich 1967; Hisdal et al. 2004). A drought event starts when the data falls below the defined threshold value and ends when the data rises above the defined threshold value. The following drought characteristics are estimated: (i) onset and termination (day), (ii) duration (days), and (iii) deficit volume (mm), and average cumulative deviation from the threshold (mm) for state variables, such as soil moisture and groundwater.

The threshold values are usually derived from the 80th percentile of the duration curve Q_{80}, which means the values that are equivalent or exceeded in 80% of the time (Hisdal et al. 2004; Van Loon and Van Lanen 2012). During a prolonged dry period, it is often observed that the data exceed the threshold level for a short period of time and thereby a large drought is divided into a number of minor droughts that are mutually dependent, which usually leads

to pooling droughts (Hisdal et al. 2004; Fleig et al. 2006). However, in the AD-EWS, we do not apply pooling to define an independent sequence of droughts. We count all drought events (daily resolution) in the 7-month forecast period in all our analyses, except for the drought onset and termination. Drought onset and termination time are determined only for a major drought event (i.e., event with the longest drought duration) during the 7-month forecast period, which also could occur at the start (i.e., already ongoing drought) or at the end of the 7-month record (i.e., a drought that continues after the 7-month forecast period). In this case, we apply the first day of the 7-month period as the onset date for the already ongoing drought and last day as the termination date to the continuing drought. Thus, users should take into account when either the onset is on the first day or the termination is on the last day that the major drought will last longer. We also apply a centered 30-day smoothing to all forecasted hydrometeorological variables to cope with frequent zero-precipitation days (Van Loon and Van Lanen 2012) and to reduce the number of minor droughts due to flashy behavior. All threshold indices are calculated for the median and 10th, 25th, 75th, and 90th percentiles obtained from the 51 ensemble members.

Areal drought indices. Gridded time series from forecasted spatially distributed drought indices are used as input to forecast the spatial characteristics of drought events in different variables. For the areal drought indices, the forecasted percentage of area in drought, the forecasted drought duration, and the forecasted total drought deficit volume are calculated as a summary for the whole Europe. The temporal evolution of area in drought in Europe over the 7-month forecast period is given for the four standardized drought indices (SPI-*x*, SPEI-*x*, SRI-*x*, SGI-1). The temporal evolution of the area in drought using the threshold indices (drought in precipitation, soil moisture, runoff, and groundwater) is obtained from the median of the 51 ensemble members. For the threshold indices, first, we averaged daily drought area into monthly. The drought duration of the area in drought in Europe is presented as a bar diagram that gives for each of the 51 ensemble members the probability of the drought duration of the area in drought for four standardized drought indices and four threshold drought indices during the 7-month forecast period. The drought deficit volume of area in drought provides a similar bar diagram, giving the probability of the deficit volume of the area in drought for each ensemble member. All areal indices are calculated for the median and 10th, 25th, 75th, and 90th percentiles obtained from the 51 ensemble members.

Development of drought algorithms. In the AD-EWS, the drought calculation is carried out in two steps (Fig. A1). The step (left side of Fig. A1) is developing algorithms to calculate the probability distributions for each of the standardized indices for different months *i* and aggregation levels *x*, and algorithms to calculate the monthly threshold values for each of the hydrometeorological variables. This step was carried out outside the MH-EWS using the LISFLOOD data forced with observational data [referred to as LISFLOOD simulation forced with observations (SFO)]. We use the LISFLOOD SFO as a proxy for the observed because a long time series of observed data for each grid cell does not exist. The probability distributions and threshold values for each month and variable were calculated using 29 years of daily SFO data. In the second step (right side of Fig. A1), we have developed drought algorithms to process the forecasted hydrometeorological data into drought products using all parameters that have been calculated in the first step. These algorithms are encapsulated in the MH-EWS and will be run every month when new forecast data are available.

Due to enormous datasets that need to be stored in the MH-EWS (including all other hazards), currently, we only keep the probability distributions, the threshold values, the monthly proxy observed data (gridded meteorological data and LISFLOOD SFO) from the last 11 months, and daily proxy observed data from the last 1 month. Each month these data

will be overwritten when new proxy data become available. Monthly proxy data from the last 11 months and daily data from the last 1 month are needed to calculate the droughts using standardized indices with different aggregation levels (i.e., $x > 1$) and to calculate the droughts using threshold method, respectively. For example, for SPI-12, 11 months of proxy data are blended with the first month of forecasted data to predict a possible drought event in the first month ($i = 1$). To predict a possible drought event in the second month ($i = 2$) using the same aggregation level ($x = 12$), 10 months of proxy data and 2 months of forecasted data are blended, and so on for other months (see Fig. A2). Therefore, the calculation of drought forecasts using standardized indices for $x > 1$ contains both observed data (or a proxy for observed, as in the AD-EWS) and forecast data. The same concept is applied for drought in runoff (SRI-x). The ratio of observed and forecast data depends on the selected temporal aggregation level and lead time. Higher aggregation level and shorter lead times require more observed data than lower aggregation levels and longer lead times (Yuan et al. 2013b; Dutra et al. 2014a). The merging of these products (proxy observed and SEAS5), which have the same climatological properties within proxy observed, reforecast, and with operational forecasts, is basically a fusing of the proxy observed with seasonal forecast of precipitation/hydrological variables.

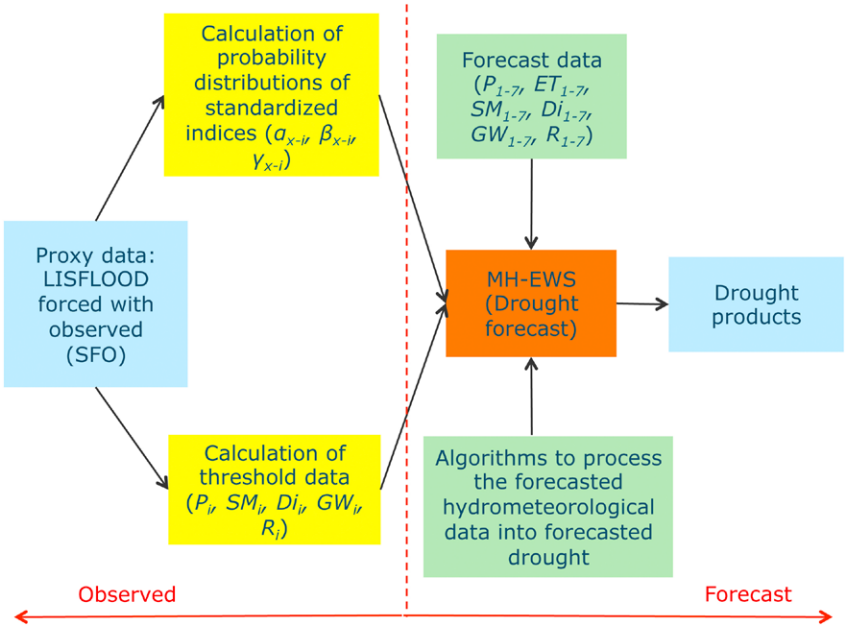

Fig. A1. Flowchart of drought algorithms, where α_{x-i}, β_{x-i}, and γ_{x-i} stand for standardized indices parameters alpha, beta, and gamma, i stands for months (1–12 months), and x stands for temporal aggregation levels (1, 3, 6, and 12 months); P is precipitation; SM is soil moisture; D is discharge; T is temperature; and R is runoff.

Although we store 1-month daily proxy data in the MH-EWS database for drought forecasting calculation using the threshold method, only the latest 15 days data are essential and used in the prediction. A centered moving average of 30 days method requires 15 days of observed data and 15 days of forecast data to predict a possible drought event on the first day. For example, to forecast a possible drought event on 1 January 2018, data from 17 to 31 December 2017 are moving averaged together with the forecasted data from 1 to 15 January 2018, and so on. This implies that for the threshold approach, no observed data are required after the middle of the first forecasted month.

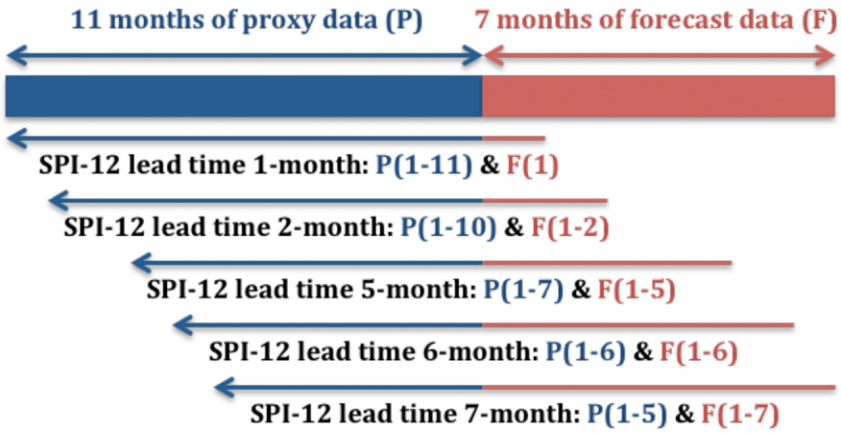

Fig. A2. Schematic presentation of the SPI-12 calculation using an accumulation of proxy data and forecast data.

Appendix B: Drought classes and forecasting score

In our paper, we present a new approach, that is, a categorical drought classification, to analyze the score of the forecasts by comparing the forecasts with the reference observed (here we call it proxy observed) and calculating the scores for forecasted hydrometeorological drought using different standardized indices for all water cycle components. We calculated meteorological droughts (e.g., SPI) using the gridded observed weather data prepared for the LISFLOOD model simulation. The outputs of the model were then used to calculate hydrological droughts (e.g., SRI). The use of LISFLOOD's SFO (proxy observations) as a reference to verify the forecasts has been done for Europe by Arnal et al. (2018), similarly to Yuan et al. (2017) for a Chinese river basin. This approach allows us to equivalently compare hydrometeorological drought forecasts score for all indices. This new approach uses drought class scores, which appeared to be better response to the requirements of end users. In our study, we used drought classes to describe the severity of drought in the various hydrometeorological variables and to determine the forecasting score. Drought classes are easier to understand by end users than an index (i.e., an SPI value of –1.5). For instance, the SPI drought severity classes are as follows: mild drought for $0 > \text{SPI} \geq -1$, moderate drought for $-1 > \text{SPI} \geq -1.5$, severe drought for $-1.5 > \text{SPI} \geq -2$, and extreme drought for $\text{SPI} \leq -2$ (McKee et al. 1993).

For each time step (month), a drought severity class expressed as a number was assigned as follows: 1, no drought; 2, mild drought; 3, moderate drought; 4, severe drought; and 5, extreme drought. Based on this scoring method, we identified the score of the forecast by comparing the drought class derived from the observations with the forecasts. The drought class difference [$\text{Diff}_k(x, y, t)$] was determined from the drought class number derived from the median of the ensemble of forecasted data [$D_{k,i}(x, y, t)$, with i = median ensemble] and the number obtained from the observed data [$D_{ok}(x, y, t)$]. A perfect forecasting score is achieved if there is no difference between both drought classes. A positive forecast score (bluish colors in the maps in Fig. 5) means that the drought forecast is pointing at a more severe drought than the one based on the observed data (overforecasting), and vice versa for the negative sign, that is, underforecasting (reddish colors in the maps in Fig. 5). The total percentage area for each drought class difference [$\text{TDiff}_k(t, c)$] and for a certain time step is calculated by summing up all land cells (x, y), which have same class difference {sum[$\text{Diff}_k(x,y,t)$] in x and y for each class difference $c = -4$ to $+4$} divided by total number of land grid cells, multiplied by 100%. Based on this analysis, including all grid cells (x, y), a high percentage area of no class difference ($c = 0$) and a low percentage of large class differences ($c = -4$ to -1 and $+1$ to $+4$) express a high forecasting score for a certain variable, time step, and lead time.

For seasonal analysis of SPI forecasts (Fig. 6), we averaged the monthly percentage class differences forecasted in winter (DJF), spring (MAM), summer (JJA), and autumn (SON). For instance, the winter skill with lead time of 1 month is obtained from the values of forecasted drought done in December, January, and February [TDiff_k(December–February), c], and of 5 months from the values of April (forecasted in December), May (forecasted in January), and June [forecasted in February; TDiff_k(April–June), c].

Color coding was applied to indicate scores using three different colors (Fig. 6). Light red, light brown, and light green colors indicate the percentage values of perfect forecast below the 25th percentile, at the 50th percentile, and above the 75th percentile derived from all seasons and all SPIs (4 seasons × 4 SPI indices × 5 lead times = 80 total values). We grouped all SPIs with different accumulation periods in the percentile calculation since SPIs were derived from one variable, which is precipitation. This allows us to compare the score of drought forecasts using different variables for both meteorology (precipitation) and hydrology (runoff). The colors for the $c = 0$ (top row for each season) reflect the score to produce the "perfect" forecast.

If the color is light green, it means that the percentage of the area (indices SPI-*x* or SRI-*x*), belongs to the 25% highest for all lead times and seasons. The percentage values of no class difference ($c = 0$) lying between the 25th and 50th percentiles are graded from light red to light brown. For percentage values of $c = 0$ lying between the 50th and 75th percentiles, the colors are graded from light brown to light green. The colors are inversed for $c = -4$ to $c = +4$, with light green color representing score below the 25th percentile and light red color representing score above the 75th percentile. The percentiles were calculated from the percentages of area ($n = 80$ for SPI and SRI), which have the same c for both plus and minus (e.g., $c = -1$ is grouped with $c = +1$, $c = -2$ is grouped with $c = +2$, and onward). The colors for c negative or positive (last eight rows for each season) reflect the score to produce the "imperfect" forecast. If the color is green, it means that the percentage of the area (indices SPI-*x* and SRI-*x*) belongs to the 25% lowest in these classes for all lead times and seasons. We did not apply any color (white) if the percentage of area is zero, indicating that class difference between observed and forecasted does not occur anywhere in Europe.

References

Alderlieste, M. A. A., H. A. J. V. Lanen, and N. Wanders, 2014: Future low flows and hydrological drought: How certain are these for Europe? IAHS Publ., 363, 60–65.

Alfieri, L., D. Velasco, and J. Thielen, 2011: Flash flood detection through a multi-stage probabilistic warning system for heavy precipitation events. *Adv. Geosci.*, **29**, 69–75, https://doi.org/10.5194/adgeo-29-69-2011.

Andreu, J., A. Solera, J. Paredes-Arquiola, D. Haro-Monteagudo, and H. A. J. Van Lanen, 2015: *Drought: Research and Science-Policy Interfacing.* CRC/Balkema Publishers, 514 pp.

Arnal, L., H. L. Cloke, E. Stephens, F. Wetterhall, C. Prudhomme, J. Neumann, B. Krzeminski, and F. Pappenberger, 2018: Skilful seasonal forecasts of streamflow over Europe? *Hydrol. Earth Syst. Sci.*, **22**, 2057–2072, https://doi.org/10.5194/hess-22-2057-2018.

——, and Coauthors, 2019: EFAS upgrade for the extended model domain. EU Joint Research Center Tech. Doc. EUR 29323 EN, 58 pp., https://doi.org/10.2760/806324.

Bajo, M., and G. Umgiesser, 2010: Storm surge forecast through a combination of dynamic and neural networks models. *Ocean Modell.*, **33**, 1–9, https://doi.org/10.1016/j.ocemod.2009.12.007.

Barbosa, P., G. Naumann, L. Valentini, J. Vogt, E. Dutra, D. Magni, and A. de Jager, 2013: A pan-African map viewer for drought monitoring and forecast. 14th Waternet Symp., Dar es Salaam, Tanzania, Waternet.

Barnston, A. G., M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. D. Witt., 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.,* **93,** 631–651, https://doi.org/10.1175/BAMS-D-11-00111.1.

Bartholmes, J., J. Thielen, and M. Kalas, 2008: Forecasting medium-range flood hazard on European scale. *Georisk*, **2**, 181–186, https://doi.org/10.1080/17499510802369132.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, https://doi.org/10.1038/nature14956.

Baum, R. L., and J. W. Godt, 2010: Early warning of rainfall-induced shallow landslides and debris flows in the USA. *Landslides*, **7**, 259–272, https://doi.org/10.1007/s10346-009-0177-0.

Bell, V. A., H. N. Davies, A. L. Kay, A. Brookshaw, and A. A. Scaife, 2017: A national-scale seasonal hydrological forecast system: Development and evaluation over Britain. *Hydrol. Earth Syst. Sci.*, **21**, 4681–4691, https://doi.org/10.5194/hess-21-4681-2017.

Bloomfield, J. P., and B. P. Marchant, 2013: Analysis of groundwater drought building on the standardised precipitation index approach. *Hydrol. Earth Syst. Sci.*, **17**, 4769–4787, https://doi.org/10.5194/hess-17-4769-2013.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3, https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

Buizza, R., and M. Leutbecher, 2015: The forecast skill horizon. *Quart. J. Roy. Meteor. Soc.*, **141**, 3366–3382, https://doi.org/10.1002/qj.2619.

Burek, P., J. V. D. Knijff, and A. D. Roo, 2013: LISFLOOD distributed water balance and flood simulation model. EU Joint Research Center Tech. Rep., 150 pp., https://doi.org/102788/24719.

Day, G. N., 1985: Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.*, **111**, 157–170, https://doi.org/10.1061/(ASCE)0733-9496(1985)111:2(157).

de Vries, H., 2009: Probability forecasts for water levels at the coast of the Netherlands. *Mar. Geod.*, **32**, 100–107, https://doi.org/10.1080/01490410902869185.

Dutra, E., F. Wetterhall, F. D. Giuseppe, G. Naumann, P. Barbosa, J. Vogt, W. Pozzi, and F. Pappenberger, 2014a: Global meteorological drought—Part 1: Probabilistic monitoring. *Hydrol. Earth Syst. Sci.*, **18**, 2657–2667, https://doi.org/10.5194/hess-18-2657-2014.

——, and Coauthors, 2014b: Global meteorological drought—Part 2: Seasonal forecasts. *Hydrol. Earth Syst. Sci.*, **18**, 2669–2678, https://doi.org/10.5194/hess-18-2669-2014.

Edwards, D. C., and T. B. McKee, 1997: Characteristics of 20th century drought in the United States at multiple time scales. Colorado State University Climatology Rep., 172 pp.

EEA, 2010: Mapping the impacts of natural hazards and technological accidents in Europe: An overview of the last decade. European Environment Agency Tech. Rep. 13/2010, 146 pp.

Ek, M., and Coauthors, 2010: NCEP/EMC NLDAS support for drought monitoring and seasonal prediction. NOAA Rep., 2 pp.

EU, 2007: Addressing the challenge of water scarcity and droughts in the European Union. Commission of the European Communities Rep., 14 pp.

FAO, 2017: The impact of disasters on agriculture and food security. Food and Agriculture Organization of the United Nations Rep., 54 pp.

Feyen, L., and R. Dankers, 2009: Impact of global warming on streamflow drought in Europe. *J. Geophys. Res.,* **114,** D17116, https://doi.org/10.1029/2008JD011438.

Fink, A. H., T. Brücher, A. Krüger, G. C. Leckebusch, J. G. Pinto, and U. Ulbrich, 2006: The 2003 European summer heatwaves and drought—Synoptic diagnosis and impacts. *Weather*, **59**, 209–216, https://doi.org/10.1256/wea.73.04.

Fleig, A. K., L. M. Tallaksen, H. Hisdal, and S. Demuth, 2006: A global evaluation of streamflow drought characteristics. *Hydrol. Earth Syst. Sci.*, **10**, 535–552, https://doi.org/10.5194/hess-10-535-2006.

Forzieri, G., L. Feyen, R. Rojas, M. Flörke, F. Wimmer, and A. Bianchi, 2014: Ensemble projections of future streamflow droughts in Europe. *Hydrol. Earth Syst. Sci.*, **18**, 85–108, https://doi.org/10.5194/hess-18-85-2014.

Hisdal, H., L. M. Tallaksen, B. Clausen, E. Peters, and A. Gustard, 2004: Hydrological drought characteristics. *Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater*, L. M. Tallaksen and H. A. J. Van Lanen, Eds., Development in Water Science, Vol. 48, Elsevier Science, 139–198, http://europeandroughtcentre.com/hydrological-drought-1st-edition-file-15/.

Hurford, A. P., S. J. Priest, D. J. Parker, and D. M. Lumbroso, 2012: The effectiveness of extreme rainfall alerts in predicting surface water flooding in England and Wales. *Int. J. Climatol.*, **32**, 1768–1774, https://doi.org/10.1002/joc.2391.

Ionita, M., and Coauthors, 2017: The European 2015 drought from a climatological perspective. *Hydrol. Earth Syst. Sci.*, **21**, 1397–1419, https://doi.org/10.5194/hess-21-1397-2017.

Kauffeldt, A., F. Wetterhall, F. Pappenberger, P. Salamon, and J. Thielen, 2016: Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environ. Modell. Software*, **75**, 68–76, https://doi.org/10.1016/j.envsoft.2015.09.009.

Kruse, S., and I. Seidl, 2013: Social capacities for drought risk management in Switzerland. *Nat. Hazards Earth Syst. Sci.*, **13**, 3429–3441, https://doi.org/10.5194/nhess-13-3429-2013.

Kumar, R., and Coauthors, 2016: Multiscale evaluation of the standardized precipitation index as a groundwater drought indicator. *Hydrol. Earth Syst. Sci.*, **20**, 1117–1131, https://doi.org/10.5194/hess-20-1117-2016.

Laaha, G., and Coauthors, 2017: The European 2015 drought from a hydrological perspective. *Hydrol. Earth Syst. Sci.*, **21**, 3001–3024, https://doi.org/10.5194/hess-21-3001-2017.

Laiolo, P., S. Gabellani, N. Rebora, R. Ruari, L. Ferraris, S. Ratto, H. Stevenin, and M. Caudoro, 2014: Validation of the flood-proofs probabilistic forecasting system. *Hydrol. Processes*, **28**, 3466–3481, https://doi.org/10.1002/hyp.9888.

Lehner, B., P. Doll, J. Alcamo, T. Henrichs, and F. Kaspar, 2006: Estimating the impact of global change on flood and drought risks in Europe: A continental, integrated analysis. *Climatic Change*, **75**, 273–299, https://doi.org/10.1007/s10584-006-6338-4.

Marx, A., and Coauthors, 2018: Climate change alters low flows in Europe under global warming of 1.5, 2, and 3 °c. *Hydrol. Earth Syst. Sci.*, **22**, 1017–1032, https://doi.org/10.5194/hess-22-1017-2018.

Mason, I. B., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291–303.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scale. *Proc. Eighth Conf. on Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–184.

Mishra, A. K., and V. P. Singh, 2010: A review of drought concepts. *J. Hydrol.*, **391**, 202–216, https://doi.org/10.1016/j.jhydrol.2010.07.012.

Osanai, N., T. Shimizu, K. Kuramoto, S. Kojima, and T. Noro, 2010: Japanese early-warning for debris flows and slope failures using rainfall indices with radial basis function network. *Landslides*, **7**, 325–338, https://doi.org/10.1007/s10346-010-0229-5.

Pappenberger, F., J. Thielen, and M. del Medico, 2011: The impact of weather forecast improvements on large scale hydrology: Analysing a decade of forecasts of the European flood alert system. *Hydrol. Processes*, **25**, 1091–1113, https://doi.org/10.1002/hyp.7772.

Peters, E., P. J. J. F. Torfs, H. A. J. Van Lanen, and G. Bier, 2003: Propagation of drought through groundwater—A new approach using linear reservoir theory. *Hydrol. Processes*, **17**, 3023–3040, https://doi.org/10.1002/hyp.1274.

Poljanšek, K., and Coauthors, 2017: Science for disaster risk management 2017: Knowing better and losing less. EU Joint Research Centre Rep., 60 pp., https://doi.org/10.2760/451402.

Pozzi, W., and Coauthors, 2013: Toward global drought early warning capability. *Bull. Amer. Meteor. Soc.*, **94**, 776–785, https://doi.org/10.1175/BAMS-D-11-00176.1.

Prudhomme, C., and Coauthors, 2014: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proc. Natl. Acad. Sci. USA*, **111**, 3262–3267, https://doi.org/10.1073/pnas.1222473110.

Robine, J.-M., S. L. K. Cheung, S. L. Roy, H. van Oyen, C. Griffiths, J.-P. Michel, and F. R. Herrmann, 2008: Death toll exceed 70,000 in Europe during the summer of 2003. *C. R. Biol.*, **331**, 171–178, https://doi.org/10.1016/j.crvi.2007.12.001.

Rogers, E., D. G. Deaven, and G. J. DiMego, 1995: The regional analysis system for the operational "early" eta model: Original 80-km configuration and recent changes. *Wea. Forecasting*, **10**, 810–825, https://doi.org/10.1175/1520-0434(1995)010<0810:TRASFT>2.0.CO;2.

Samaniego, L., and Coauthors, 2018: Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Climate Change*, **8**, 421–426, https://doi.org/10.1038/s41558-018-0138-5.

——, and Coauthors, 2019: Hydrological forecasts and projections for improved decision-making in the water sector in Europe. *Bull. Amer. Meteor. Soc.*, **100,** 2451–2472, https://doi.org/10.1175/BAMS-D-17-0274.1.

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature*, **427**, 332–336, https://doi.org/10.1038/nature02300.

Sepulcre-Canto, G., S. Horion, A. Singleton, H. Carrao, and J. Vogt, 2012: Development of a combined drought indicator to detect agricultural drought in Europe. *Nat. Hazards Earth Syst. Sci.*, **12**, 3519–3531, https://doi.org/10.5194/nhess-12-3519-2012.

Sheffield, J., and Coauthors, 2014: A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bull. Amer. Meteor. Soc.*, **95**, 861–882, https://doi.org/10.1175/BAMS-D-12-00124.1.

Shukla, S., and A. W. Wood, 2008: Use of a standardized runoff index for characterizing hydrologic drought. *Geophys. Res. Lett.,* **35,** L02405, https://doi.org/10.1029/2007gl032487.

Smith, P., and Coauthors, 2016: On the operational implementation of the European Flood Awareness System (EFAS). ECMWF Tech. Memo., 34 pp.

Spinoni, J., G. Naumann, and J. Vogt, 2017: Pan-European seasonal trends and recent changes of drought frequency and severity. *Global Planet. Change*, **148**, 113–130, https://doi.org/10.1016/j.gloplacha.2016.11.013.

Stahl, K., and Coauthors, 2010: Streamflow trends in Europe: Evidence from a dataset of near-natural catchments. *Hydrol. Earth Syst. Sci.*, **14**, 2367–2382, https://doi.org/10.5194/hess-14-2367-2010.

Staudinger, M., K. Stahl, and J. Seibert, 2014: A drought index accounting for snow. *Water Resour. Res.*, **50**, 7861–7872, https://doi.org/10.1002/2013WR015143.

Stockdale, T., S. Johnson, L. Ferranti, M. Balmaseda, and S. Briceag, 2018: ECMWF's new long range forecasting system SEAS5. *ECMWF Newsletter,* Vol. 154, ECMWF, Reading, United Kingdom, 15–20, https://doi.org/10.21957/tsb6n1.

Svoboda, M., and Coauthors, 2002: The Drought Monitor. *Bull. Amer. Meteor. Soc.*, **83**, 1181–1190, https://doi.org/10.1175/1520-0477-83.8.1181.

Tallaksen, L. M., and H. A. J. Van Lanen, 2004: *Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater,* L. M. Tallaksen and H. A. J. Van Lanen, Eds., Development in Water Science, Vol. 48, Elsevier Science, 579 pp.

Thielen, J., J. Bartholmes, M.-H. Ramos, and A. D. Roo, 2009: The European flood alert system—Part 1: Concept and development. *Hydrol. Earth Syst. Sci.*, **13**, 125–140, https://doi.org/10.5194/hess-13-125-2009.

Thober, S., and Coauthors, 2018: Multi-model ensemble projections of European river floods and high flows at 1.5, 2, and 3 degrees global warming. *Environ. Res. Lett.*, **13**, 014003, https://doi.org/10.1088/1748-9326/aa9e35.

Trambauer, P., S. Maskey, H. Winsemius, M. Werner, and S. Uhlenbrook, 2013: A review of continental scale hydrological models and their suitability for drought forecasting in (sub-Saharan) Africa. *Phys. Chem. Earth*, **66**, 16–26, https://doi.org/10.1016/j.pce.2013.07.003.

——, M. Werner, H. C. Winsemius, S. Maskey, E. Dutra, and S. Uhlenbrook, 2015: Hydrological drought forecasting and skill assessment for the Limpopo River basin, southern Africa. *Hydrol. Earth Syst. Sci.*, **19**, 1695–1711, https://doi.org/10.5194/hess-19-1695-2015.

Van Der Knijff, J. M., J. Younis, and A. P. J. D. Roo, 2010: LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *Int. J. Geogr. Inf. Sci.*, **24**, 189–212, https://doi.org/10.1080/13658810802549154.

Van Lanen, H. A. J., and Coauthors, 2016: Hydrology needed to manage droughts: The 2015 European case. *Hydrol. Processes*, **30**, 3097–3104, https://doi.org/10.1002/hyp.10838.

——, C. Prudhomme, W. Wanders, and M. H. J. Van Huijgevoort, 2018: Future of drought. *Drought: Science and Policy,* A. Iglesias, D. Assimacopoulos, and H. A. J. Van Lanen, Eds., Wiley Blackwell, 69–92.

Van Loon, A. F., 2015: Hydrological drought explained. *Wiley Interdiscip. Rev.: Water*, **2**, 359–392, https://doi.org/10.1002/wat2.1085.

——, and H. A. J. Van Lanen, 2012: A process-based typology of hydrological drought. *Hydrol. Earth Syst. Sci.*, **16**, 1915–1946, https://doi.org/10.5194/hess-16-1915-2012.

——, M. H. J. Van Huijgevoort, and H. A. J. Van Lanen, 2012: Evaluation of drought propagation in an ensemble mean of large-scale hydrological models. *Hydrol. Earth Syst. Sci.*, **16**, 4057–4078, https://doi.org/10.5194/hess-16-4057-2012.

Vicente-Serrano, S. M., S. Begueria, and J. I. López-Moreno, 2010: A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Climate*, **23**, 1696–1718, https://doi.org/10.1175/2009JCLI2909.1.

Vitart, F., 2004: Monthly forecasting at ECMWF. *Mon. Wea. Rev.*, **132**, 2761–2779, https://doi.org/10.1175/MWR2826.1.

Vogt, J., G. Sepulcre, D. Magni, L. Valentini, A. Singleton, F. Micale, and P. Barbosa, 2013: The European Drought Observatory (EDO): Current state and future directions. *EGU General Assembly*, Vienna, Austria, EGU, EGU2013-7374.

Wanders, N., and H. A. J. Van Lanen, 2015: Future discharge drought across climate regions around the world modeled with a synthetic hydrological modeling approach forced by three general circulation models. *Nat. Hazards Earth Syst. Sci.*, **15**, 487–504, https://doi.org/10.5194/nhess-15-487-2015.

——, Y. Wada, and H. A. J. Van Lanen, 2015: Global hydrological droughts in the 21st century under a changing hydrological regime. *Earth Syst. Dyn.*, **6**, 1–15, https://doi.org/10.5194/esd-6-1-2015.

——, A. F. Van Loon, and H. A. J. Van Lanen, 2017: Frequently used drought indices reflect different drought conditions on global scale. *Hydrol. Earth Syst. Sci. Discuss.*, https://doi.org/10.5194/hess-2017-512.

——, S. Thober, R. Kumar, M. Pan, J. Sheffield, L. Samaniego, and E. F. Wood, 2019: Development and evaluation of a pan-European multimodel seasonal hydrological forecasting system. *J. Hydrometeor.*, **20**, 99–115, https://doi.org/10.1175/JHM-D-18-0040.1.

Wetterhall, F., and F. Di Giuseppe, 2018: The benefit of seamless forecasts for hydrological prediction over Europe. *Hydrol. Earth Syst. Sci.*, **22**, 3409–3420, https://doi.org/10.5194/hess-22-3409-2018.

WMO, 2012: Standardized precipitation index user guide. WMO Rep. 1090, 24 pp.
Yevjevich, V., 1967: An objective approach to definition and investigations of continental hydrologic droughts. Colorado State University Hydrology Paper 23, 19 pp.
Yuan, X., E. F. Wood, J. K. Roundy, and M. Pan, 2013a: CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States. *J. Climate*, **26**, 4828–4847, https://doi.org/10.1175/JCLI-D-12-00683.1.
——, ——, N. W. Chaney, J. Sheffield, J. Kam, M. Liang, and K. Guan, 2013b: Probabilistic seasonal forecasting of African drought by dynamical models. *J. Hydrometeor.*, **14**, 1706–1720, https://doi.org/10.1175/JHM-D-13-054.1.
——, F. Ma, L. Wang, Z. Zheng, Z. Ma, A. Ye, and S. Peng, 2016: An experimental seasonal hydrological forecasting system over the Yellow River basin—Part 1: Understanding the role of initial hydrological conditions. *Hydrol. Earth Syst. Sci.*, **20**, 2437–2451, https://doi.org/10.5194/hess-20-2437-2016.
——, M. Zhang, L. Wang, and T. Zhou, 2017: Understanding and seasonal forecasting of hydrological drought in the Anthropocene. *Hydrol. Earth Syst. Sci.*, **21,** 5477–5492, doi:10.5194/hess-21-5477-2017.
Zajac, Z., M. Zambrano-Bigiarini, P. Salamon, P. Burek, A. Gentile, and A. Bianchi, 2013: Calibration of the LISFLOOD model for Europe: Calibration round 2013. EU Joint Research Centre Rep., 47 pp., http://ef8.commpla.com/sites/default/files/Manuals/JRC87717_efas_calibration_report%20_final_9_jan_2014.pdf
Zappa, M., L. Bernhard, C. Spirig, M. Pfaundler, K. Stahl, S. Kruse, I. Seidl, and M. Stähli, 2014: A prototype platform for water resources monitoring and early recognition of critical droughts in Switzerland. *Proc. Int. Assoc. Hydrol. Sci.*, **364**, 492–498, https://doi.org/10.5194/piahs-364-492-2014.
Zink, M., L. Samaniego, R. Kumar, S. Thober, J. Mai, D. Schäfer, and A. Marx, 2016: The German drought monitor. *Environ. Res. Lett.*, **11**, 074002, https://doi.org/10.1088/1748-9326/11/7/074002.
——, R. Kumar, M. Cuntz, and L. Samaniego, 2017: A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty. *Hydrol. Earth Syst. Sci.*, **21**, 1769–1790, https://doi.org/10.5194/hess-21-1769-2017.