



# Evaluating skill and robustness of seasonal meteorological and hydrological drought forecasts at the catchment scale – Case Catalonia (Spain)

Theresa C. Van Hateren<sup>a,b,\*</sup>, Samuel J. Sutanto<sup>b</sup>, Henny A.J. Van Lanen<sup>b</sup>

<sup>a</sup> Remote Sensing and Natural Resources Modelling, Department of Environmental Research and Innovation, Luxembourg Institute of Science and Technology, 41 Rue du Brill, L-4422 Belvaux, Luxembourg

<sup>b</sup> Department of Hydrology and Quantitative Water Management, Environmental Sciences Group, Wageningen University & Research, Droevendaalsesteeg 3a, 6708 PB Wageningen, the Netherlands

## ARTICLE INFO

Handling Editor: Zhen (Jason) He

### Keywords:

Hydro-meteorological droughts  
Catchment scale  
Forecasting skill  
Sensitivity analysis

## ABSTRACT

Robust sub-seasonal and seasonal drought forecasts are essential for water managers and stakeholders coping with water shortage. Many studies have been conducted to evaluate the performance of hydrological forecasts, that is, streamflow. Nevertheless, only few studies evaluated the performance of hydrological drought forecasts. The objective of this study, therefore, is to analyse the skill and robustness of meteorological and hydrological drought forecasts on a catchment scale (the Ter and Llobregat rivers in Catalonia, Spain), rather than on a continental or global scale. Meteorological droughts were forecasted using downscaled (5 km) probabilistic weather reforecasts (ECMWF-SEAS4). These downscaled data were also used to produce hydrological drought forecasts, derived from time series of streamflow data simulated with a hydrological model (LISFLOOD). This resulted in seasonal hydro-meteorological reforecasts with a lead time up to 7 months, for the time period 2002–2010. These monthly reforecasts were compared to two datasets: (1) droughts derived from a proxy for observed data, including gridded precipitation data and discharge simulated by the LISFLOOD model, fed by these gridded climatological data; and (2) droughts derived from in situ observed precipitation and discharge. Results showed that the skill of hydrological drought forecasts is higher than the climatology, up to 3–4 months lead time. On the contrary, meteorological drought forecasts, analysed using the Standardized Precipitation Index (SPI), do not show added value for short accumulation times (SPI1 and SPI3). The robustness analysis showed that using either a less extreme or a more extreme threshold leads to a large change in forecasting skill, which points at a rather low robustness of the hydrological drought forecasts. Because the skill found in hydrological drought forecasts is higher than the meteorological ones in this case study, the use of hydrological drought forecasts in Catalonia is highly recommended for management of water resources.

## 1. Introduction

Drought events are one of the most costly weather-related natural hazards, because their effects can be widespread and long-lasting. For instance, the 2003, 2015, and 2018 droughts affected and triggered damages in large parts of Europe (EEA, 2010; Van Lanen et al., 2016; WMO, 2019). Even though the term *drought* might seem straightforward, it is not unambiguous, as many different definitions have been proposed. Here, we use the definition proposed by Tallaksen and Van Lanen (2004): *a sustained period of below-normal water availability*. The following three types of natural droughts can be determined (Wilhite, 2000; Tallaksen and Van Lanen, 2004): (1) meteorological drought: below-normal precipitation; (2) soil moisture drought: below-normal

soil moisture content; and (3) hydrological drought: below-normal (ground)water levels and discharge. According to Heinrich and Gobiet (2012); Orłowsky and Seneviratne (2013); Russo et al. (2013); Prudhomme et al. (2014); Pascual et al. (2015); Wanders and Van Lanen (2015); Wanders et al. (2015) and Van der Wiel et al. (2019), climate change will lead to drier conditions in many regions and river basins, causing drought events to occur more frequently, and increasing their impacts. Therefore, reliable and robust sub-seasonal and seasonal drought forecasts are essential for water managers and stakeholders.

In Europe, seasonal meteorological drought forecasts currently range up to seven months (e.g. Hwang and Carbone, 2009; Dutra et al., 2014; Mo and Lyon, 2015), and are usually based on the Standardized Precipitation Index (SPI, McKee et al., 1993) or the Palmer Drought

\* Corresponding author at: Remote Sensing and Natural Resources Modelling, Department of Environmental Research and Innovation, Luxembourg Institute of Science and Technology, 41 Rue du Brill, L-4422 Belvaux, Luxembourg.

E-mail address: [tessa.vanhateren@list.lu](mailto:tessa.vanhateren@list.lu) (T.C. Van Hateren).

<https://doi.org/10.1016/j.envint.2019.105206>

Received 2 April 2019; Received in revised form 23 August 2019; Accepted 20 September 2019

Available online 01 November 2019

0160-4120/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Severity Index (PDSI, Palmer, 1965). Pre-operational seasonal hydro-meteorological drought forecasts have recently also been made available for Europe and on a regional scale for Catalonia under the EU-funded project ANYWHERE ([www.anywhere-h2020.eu](http://www.anywhere-h2020.eu)) (Sutanto and Van Lanen, in preparation). Though hydro- and meteorological drought forecasts are both available, so far, they were mostly discussed separately. Therefore, in this paper, a comparison is made between hydrological and meteorological droughts.

In the early days of hydrological forecasting, deterministic weather forecasts were used as input for a hydrological model. The essence of deterministic forecasting is that, based on the current state of the system, there is only one way in which the system will develop, resulting in just one forecast without providing information on uncertainty. The first one who introduced the concept of uncertainties in hydrological forecasting was Krzysztofowicz (2001). Since then, a shift towards probabilistic hydrological modelling has been taking place (Schaake et al., 2007; Cloke and Pappenberger, 2009). Probabilistic forecasts indicate that a system can develop in many different ways, representing uncertainty in the forecasts. Thus, in probabilistic hydrological forecasting, multiple ensembles of numerical weather forecasts are used as input for a hydrological model, resulting in the same number of forecasted hydrological ensembles. The use of probabilistic hydrological forecasts differs from Ensemble Streamflow Prediction (ESP, introduced by Day, 1985). In probabilistic hydrological forecasts, a hydrological model is forced by probabilistic weather forecasts. On the other hand, ESP indicates that a hydrological model is initialized with current hydrological conditions, that subsequently is driven with different samples of historical meteorological data (Wood and Lettenmaier, 2008; Shukla and Lettenmaier, 2011; Staudinger and Seibert, 2014; Arnal et al., 2018; Monhart et al., 2019). In this study, we use probabilistic hydrological forecasts, which are, from now on, referred to as “hydrological drought forecasts”. Though all forecast ensembles together can indicate the *uncertainty* of the hydrological drought forecasts, each ensemble in itself differs from the reference data. The *skill* of the forecast is a measure for the magnitude of the difference between the forecast and the reference data (Bradley et al., 2008). In this paper, the skill of the forecast ensembles will be discussed.

In Europe, work on the improvement of hydrological forecasts, that is, prediction of river flow, is done in multiple projects, such as the Hydrological Ensemble Prediction EXperiment (HEPEX, Schaake et al., 2007), the European Flood Awareness System (EFAS, Burek et al., 2011), Improving PRedictions and management of hydrological EXtremes (IMPRES, Van den Hurk et al., 2016), and the End-to-end Demonstrator for improved decision-making in the water sector in Europe (EDgE, Samaniego et al., 2018; Samaniego et al., in press). The early projects were set up to decrease the impact of flooding, but later these projects also included hydrological forecasts and drought projections under different climate change scenarios at the pan-European scale (e.g. Marx et al., 2018; Thober et al., 2018; Wanders et al., 2019). However, none of them addresses hydrological drought forecasts. Drought forecasting requires an additional step using the forecasted time series of a hydrological variable, such as river flow. The hydrological drought community makes a clear difference between, for instance, low river flow (which is hydrology), which is described by the k-day lowest flows each year, and drought in river flow (river flow drought) that is characterized by, for instance, deficit volumes, drought duration and drought intensity (e.g. Hisdal et al., 2004). Thus, whereas low flows occur every year at some stage due to normal seasonal variations, hydrological droughts indicate an anomaly from the normal situation (i.e. climatology, hydrological regime) and, therefore, do not happen every year. Accurate prediction of hydrological droughts, including water deficits, which is not directly provided by low flow forecasting, is therefore of prime importance to water managers in drought-prone regions, who can take accurate measures to alleviate drought impacts. The difference between low streamflow and river flow drought

appeared to result in a substantially different skill between projections of low flow and drought (Alderlieste et al., 2014; Van Lanen et al., 2018). Hence, skills of hydrological forecasts, which are known (e.g. Wanders et al., 2019), and skills of hydrological drought forecasts are anticipated to be different.

Next to the projects which are merely restricted to hydrological forecasting, a few projects have developed operational hydrological drought forecasts at the pan-European scale, namely the European Drought Observatory (EDO, Vogt et al., 2011, <http://edo.jrc.ec.europa.eu/>) within the Copernicus programme, and the aforementioned H2020 project EnhANCing emergency management and response to extreme WeaTHER and climate Events (ANYWHERE, Sutanto et al., in preparation, <http://anywhere-h2020.eu/>). Both projects use simulated time series of hydrological variables with the LISFLOOD model (Burek et al., 2013). EDO forecasts droughts based on soil moisture anomalies with a lead time of 7 days and a meteorological drought forecast with a lead time of 3 months. The ANYWHERE drought early warning system produces probabilistic hydro-meteorological drought forecasts, e.g. droughts in precipitation, soil moisture, runoff, streamflow, and groundwater, with lead times up to 7 months. These indices have a spatial resolution of  $5 \times 5$  km, which is assumed to be meaningful for the water management scale, i.e. the catchment scale. The forecasts are updated at the beginning of each month and can be found online on <http://a4cat.hydsdev.net/login> for registered users.

The aim of this study was to investigate the skill and robustness of ANYWHERE hydrological drought forecasts at the catchment scale. Our study was performed on two small sub-catchments of the Ter and Llobregat river basins, located in the north of Catalonia, because the use of a delineated area allowed for a comparison with in situ observations. Both rivers are of prime importance regarding the water supply to, among others, the Barcelona metropolitan area, and can therefore benefit from improved drought forecasting. This study differs from the few existing hydrological drought forecast studies at the catchment or river basin scale (Fundel et al., 2013; Zappa et al., 2014; Trambauer et al., 2015) in the sense that first, it provides a more comprehensive analysis of the skill and robustness of drought forecasts. Second, in addition to obtaining the drought forecasting skills by comparing reforecasts against a proxy for reality, as done in the aforementioned studies, a comparison was made between reforecasts and in situ observations. This aids the interpretation of the results, as it gives a direct indication of the utility of reforecasts in actual situations. Finally, not only have hydrological drought forecasts been analysed using the threshold method (Yevjevich, 1967), but also meteorological drought forecasts have been included in the comparison, through the Standardized Precipitation Index (SPI, McKee et al., 1993). These together give a better insight on the skill and robustness of hydro-meteorological drought forecasts, addressing different types of drought occurring at the catchment scale.

## 2. Methods & data

### 2.1. Methods

#### 2.1.1. Data comparison

Gridded observed precipitation and discharge data spanning the period between 1990 and 2016, obtained from the LISFLOOD model (Simulation Forced with Observations, hereafter referred to as “SFO” for both precipitation and discharge, Section 2.2.2) were used as proxy for observations to calculate the skill of reforecasts of drought in precipitation (meteorological drought) and drought in discharge (hydrological drought), respectively, for the years 2002 to 2010. The main reason for this is that both the proxy for observations and the reforecasts were calculated using the same approach, resulting in time series of precipitation and discharge at a  $5 \times 5$  km grid. Therefore, the bias of the reforecasts, compared to the SFO, did not influence the results of the study (Wetterhall and Giuseppe, 2018). Moreover, the quality of the

SFO data is equal throughout the entire studied period, whereas in situ observed data contain missing data (Section 2.2.3). A comparison between the SFO and actual observations was performed, to make sure that the use of the SFO as a proxy for observations produces a valid result.

To quantify differences between the SFO and observations (precipitation, discharge), efficiency parameters were calculated separately for both studied catchments: the coefficient of determination ( $R^2$ , Wright, 1921), the Nash Sutcliffe Efficiency (NSE, Nash and Sutcliffe, 1970) and the Kling-Gupta Efficiency (KGE, Gupta et al., 2009).

### 2.1.2. SPI

Meteorological droughts, which are associated with a deviation from the normal precipitation, were analysed using the Standardized Precipitation Index (SPI, McKee et al., 1993). This widely used method was chosen, because it gives a good indication on the occurrence of meteorological droughts. Additional advantages of this method include its ability to quantify meteorological drought severity for different time scales and its low complexity compared to indices such as the PDSI, since it only needs time series of precipitation (e.g. McKee et al., 1993; Hayes et al., 1999; Hayes et al., 2011).

Fig. 1a indicates the steps taken for this part of the study. Every step in the figure is indicated in the text below by square brackets, e.g. [1a.1] or [1a.a]. The SFO precipitation data (Section 2.2.2) were analysed first [1a.1]. Precipitation data from the grid cells located in the Ripoll and Guardiola catchments were extracted for the entire time period (1990–2016). A spatial average of precipitation in each grid cell was taken to get the precipitation over the entire catchment [mm/day]. Then, daily averaged precipitation data were added up to monthly totals [mm/month].

The next step was to calculate the SPI [1a.2] using the SPEI package (Beguiría and Vicente-Serrano, 2017). In this package, monthly data is transformed into 12 gamma distributions, one for every month of the year. Every gamma distribution can be described by two parameters:  $\alpha$  (the shape parameter) and  $\beta$  (the inverse scale parameter). The gamma distribution was used as it has been proved to be well suited for calculating SPI across Europe for all accumulation periods (Stagge et al., 2015). Data resulting from the calculations thus consist of monthly values of SPI and 12  $\alpha$  and  $\beta$  values [1a.3] for each catchment. This procedure can be performed for different accumulation times ( $i$ , in months). An accumulation time, in this case, means that the precipitation is summed over the current month and the past  $i-1$  months. In this study, the SPI was calculated for accumulation times of 1, 3, 6, and 12 months, using SFO data as a proxy for observations.

For the next step, the reforecasted precipitation data (available for the period 2002–2010, Section 2.2.4) had to be organized. In the calculation of SPI3, SPI6, and SPI12, SFO precipitation data had to be added to the reforecast data, to be able to get SPI $i$  values in all reforecast months (Yuan et al., 2013a, 2013b; Dutra et al., 2014). This means that, for instance in the SPI6 reforecasted dataset, the SPI6 for a lead time of one month (LT = 1) consists of five months of SFO precipitation data and one month of reforecasted precipitation data. Similarly, for LT = 3 in the same dataset, the SPI6 consists of three

months of SFO precipitation data and three months of reforecasted precipitation data.

Once the data were ready for analysis, the SPI was calculated [1a.b] using the  $\alpha$  and  $\beta$  values generated from the SFO precipitation data [1a.3]. This resulted in monthly SPI values for the reforecasted data [1a.c]. The reforecasted data were fitted to the probability distributions of the SFO precipitation data, because the time period of the reforecasted data (9 years) is too small to create meaningful monthly probability distributions.

In steps [1a.4] and [1a.c], drought events were identified. A threshold had to be incorporated in the SPI analysis (Section 2.1.4), to be able to calculate the forecast skill and compare it to the results from the threshold analysis (Section 2.1.3). The last step in the approach was the calculation of the skill score [1a.5]. This calculation and the sensitivity analysis are discussed in Sections 2.1.4 and 2.1.5, respectively.

### 2.1.3. Threshold method

Hydrological droughts can be analysed using indices, such as the Standardized Runoff Index (Shukla and Wood, 2008), the Standardized Streamflow Index (Vicente-Serrano et al., 2012), the Standardized Groundwater Index (Bloomfield and Marchant, 2013), or the threshold method. In this study, the threshold method (introduced by Yevjevich (1967) and elaborated by Zelenhasić and Salvai (1987)) was selected to analyse droughts in discharge (Drought<sub>Q</sub>). The difference between the methods is that in the standardized indices, an anomaly in discharge is analysed, whereas in the threshold method, the focus is on the water deficit. Using the threshold method on discharge time series is particularly useful for the catchments in northern Catalonia: knowing when discharge will be below a certain threshold, helps water managers to take measures regarding water demand and water allocation. Moreover, data analysed using the threshold method has the potential for analysis of deficit volumes (Tallaksen and Van Lanen, 2004)

Drought determination using the threshold method can be performed using either a variable or a constant threshold (Hisdal et al., 2004). Heudorfer and Stahl (2017) studied the (dis)advantages of both types of thresholds. Based on their results, it was decided to use the variable threshold method in this study, to account for the seasonality in discharge in the study area. The threshold can be at any percentile, but for perennial streams, it is common to use a percentile ranging between 5% and 30% (Meigh et al., 2002; Hisdal et al., 2004; Fleig et al., 2006). The 20th percentile is used more often (Van Loon, 2015), and was used in this study as well. A threshold at the 20th percentile indicates that the discharge on that day is equalled or exceeded 80% of the time in other years on the same day, and is referred to as the Q80 threshold. Similarly, when speaking of a Q70 or Q90 threshold, the discharge is equalled or exceeded 70% or 90% of the time. Fig. 1b shows the flow chart for analysis of the threshold method. Every step of this analysis is described below. Numbers and letters between brackets (e.g. [1b.1] or [1b.a]) refer to the steps indicated in Fig. 1b.

From the SFO dataset, discharge was extracted for two locations [1b.1]. These locations are based on the coordinates of the Ripoll and Guardiola flow gauging stations. For every day of the year, the Q80 threshold was identified [1b.2], using the entire SFO dataset

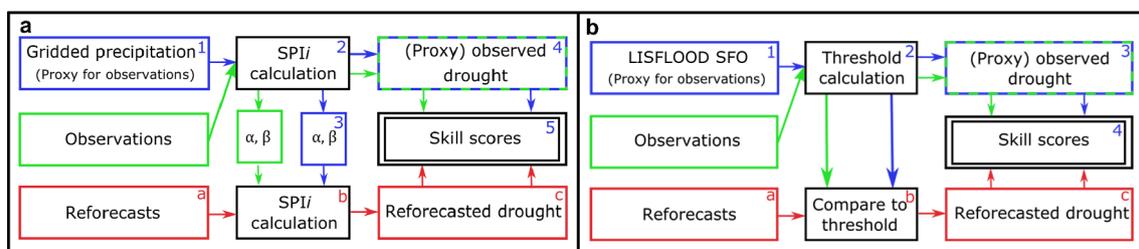


Fig. 1. Steps taken to convert different data to skill scores, using the SPI (a) and variable threshold (b) methods. The analysis of the observations (green) is part of the sensitivity analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(1990–2016). A two-sided 31-day moving average was then applied to create a smoothed threshold curve. Then, the time series of SFO discharge data were compared to the smoothed Q80 threshold curve, to identify droughts in this dataset [1b.3]. The next step was to extract the reforecasted discharge data for the same two locations from the seasonal reforecasts [1b.a]. The smoothed 31-day moving average, which was applied to SFO data, was also applied to the reforecast data. This enabled a fair comparison between hydrological drought forecasts and meteorological drought forecasts, which are accumulated over at least one month. Moreover, smoothing the forecasts avoided short wet spells to influence the results. Applying the two-sided moving average led to missing values for the first and last 15 days of every reforecast. To overcome this, SFO discharge data were added to the reforecast for the 15 days leading up to the first reforecast day, before applying a moving average. The moving average then gave values starting from the first day of the reforecasts. SFO discharge data were not added to the last 15 days of the reforecasts, as that would hugely inflate the skill of the reforecast in the last month. The reforecast data were then compared to the Q80 threshold, calculated from the SFO discharge data [1b.b], to identify droughts [1b.c]. The calculation of the skill score [1b.4] and the sensitivity analysis are discussed in Sections 2.1.4 and 2.1.5, respectively.

#### 2.1.4. Forecast skill analysis

As a last step in the SPI [1a.5] and threshold method [1b.4] the forecast skill was calculated. We chose to use the Brier Skill Score (BSS, Brier, 1950). It is a comprehensive skill score, as it gives an indication of the skill of the forecast, compared to the use of climatology as forecasts. The BSS is widely used in atmospheric science, because it is able to handle probabilistic forecasts.

The calculation of the BSS consists of two equations: Eq. (1), the calculation of the Brier Skill for both the climatology ( $BS_{\text{clim}}$ ) and the forecasts ( $BS_f$ ); and Eq. (2), the calculation of the Brier Skill Score using these two values.

$$BS_f, BS_{\text{clim}} = \frac{1}{N} \sum_1^N (p - o)^2 \quad (1)$$

$$BSS = 1 - \frac{BS_f}{BS_{\text{clim}}} \quad (2)$$

The  $BS_f$  was calculated using the probability of an event ( $p$ ), based on the number of ensemble members forecasting a drought, divided by the total number of ensemble members ( $N$ ), in this case 15 ( $p$  varies from 0 to 1). The parameter called Observed ( $o$ ) was assigned either 1, when a drought was observed, or 0, if no drought was observed. The  $BS_{\text{clim}}$  is based on the climatology, in this case the threshold. For the precipitation dataset, a threshold of  $SPI = -0.5$  was chosen, because this gives a good balance between capturing either too many minor droughts or too few droughts, the latter leading to a small sample size (Trambauer et al., 2015). The  $-0.5$  threshold value stems from a normal distribution, and therefore the value for  $p$  is 0.3085 for every month. For the discharge dataset, the Q80 threshold was used, which gives  $p = 0.2$ . If the BSS is close to 1, the second term in its equation is small, indicating that  $BS_f < BS_{\text{clim}}$  and thus that the forecast is more certain than the climatology. If the BSS is smaller than zero, the opposite is true. In that case, it would be better to use the climatology than to use the forecasts. Thus, the goal of functional forecasting is to get at least a BSS that is positive, but preferably higher.

The calculation of the BSS was done in a similar fashion for droughts in the SPI and discharge datasets. The only difference is that, after processing the data, discharge drought data consist of daily values, whereas SPI drought data consist of monthly values. To be able to compare droughts in discharge to droughts in SPI, the BSS values of discharge data were averaged into monthly values.

Every seasonal reforecast ensemble contains 7 months of data and results may vary considerably between these months. Results were

therefore compared for each of the 7 months (lead times, LT) separately. Furthermore, reforecasts initiated in different seasons were analysed separately and for the whole year. Skills of reforecasts initiated in December, January, and February (DJF, winter) are combined. Similarly, March, April, and May (MAM, spring), June, July, and August (JJA, summer) and September, October, and November (SON, autumn) are combined. As an example: a forecast done in February is part of the winter forecasts, even though the forecast also includes August (LT = 7 months).

#### 2.1.5. Robustness

A sensitivity analysis was performed to see how robust the forecasts are, if: (1) the analysis was performed using in situ observations (precipitation and discharge) instead of the SFO data; and (2) the analysis was performed using different thresholds:  $SPI = 0$  and  $SPI = -1$  for the standardized drought indices (meteorological droughts), and (3) Q70 and Q90 for the threshold method (hydrological droughts).

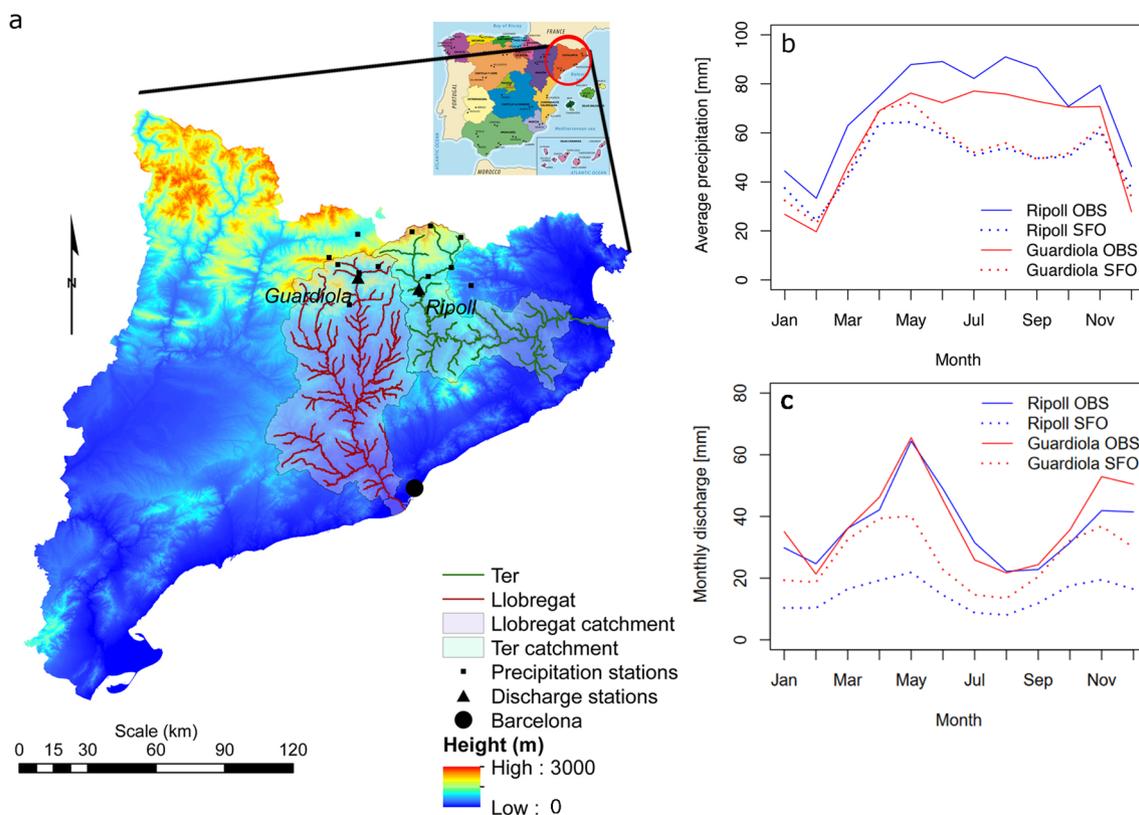
The results of the sensitivity analysis for the whole year are visualised in a colour-coded table. Four arbitrary categories were distinguished, based on what was considered to be an acceptable value. The BSS was considered to be good if the values are in between 0.50 and 1.00, acceptable if the values are in between 0.00 and 0.50, poor if the values are in between  $-0.50$  and 0.00 and bad if the values are in between  $-\infty$  and  $-0.5$ . This classification was chosen arbitrarily, based on the information that  $BSS = 1$  indicates perfect skill, and  $BSS \leq 0$  indicates no skill. Delgado et al. (2018) also used the BSS with an arbitrary classification, but opted not to differentiate between values below zero. We did want to make a distinction between very negative and slightly negative values in our study, and therefore decided to decide on the aforementioned classification.

## 2.2. Data

### 2.2.1. Study area

Catalonia is located in northeastern Spain and is characterised by a mountainous area in the north, and a lowland region in the south. Southeast Catalonia borders with the Mediterranean Sea, the north with the French Pyrenees and west Catalonia borders with inland Spain. Because of the highly varying topography, there are three main climates in Catalonia. The Pyrenees are characterised by a mountainous climate, leading to high precipitation amounts, cold winters, and mild summers. The coastline is influenced by the Mediterranean Sea. Temperatures in summer are high and winters are mild. Inland Catalonia has a more or less continental climate with warm and dry summers and cool winters. Heavy precipitation in the area is usually associated with convective events, which occur in spring, late summer, and autumn. During the winter season, precipitation is often classified as stratiform, with only a small percentage of convective events (Barrera Escoda and Llasat Botija, 2015). Droughts occur regularly in the region, and have a large impact on society, especially in agricultural areas, where economical effects are largest (Llasat et al., 2009). Water demand is already close to the available water resources (Agència Catalana de l'Aigua, 2008), and in the future, Catalonia will likely suffer from an increase in drought occurrence, due to climate change (Pascual et al., 2015). The most severe drought that hit Catalonia in recent years was the drought occurring in the years 2007–2008. To prevent Catalanian water taps from running dry during this drought, extreme measures were taken, such as importing drinking water from different regions in 2008, when reservoir levels dropped below the 20% capacity emergency level (Martin-Ortega et al., 2012).

In this study, we use two sub-catchments of the Llobregat and Ter river basins, located in the north of Catalonia. A map, showing the locations of the river basins and the corresponding river networks, is shown in Fig. 2a. The discharge in the mid- and downstream parts of both catchments is highly managed, to maintain a constant water supply for agriculture and inhabitants of the region. Therefore, we



**Fig. 2.** Catchment properties. **a** shows Catalonia, with the locations of the Guardiola and Ripoll sub-catchments. Elevation data was taken from the European Digital Elevation Model (EU-DEM v1.1) with a resolution of 25 m, obtained from the EEA Copernicus project. The map of Spain was obtained from [www.vectorstock.com](http://www.vectorstock.com). **b** and **c** show the average monthly precipitation (1996–2016) and discharge (1990–2016), respectively, per catchment and for the different datasets (OBS: observed, and SFO: proxy observed).

decided to analyse forecasted drought in precipitation and discharge from two relatively unmanaged sub-catchments, that is: Ripoll (737 km<sup>2</sup>, in the Ter river) and Guardiola (333 km<sup>2</sup>, in the Llobregat river). The sub-catchments are located upstream of the Sau and Baells dams, respectively. They thus reflect a rather natural discharge time series and therefore give a better indication of natural droughts than time series measured by discharge stations downstream of the dams.

Both the Ripoll and Guardiola catchments house perennial streams, indicating that at least part of the water ending up at the gauging stations, originates from subsurface water storage. Hydrogeological maps (ICGC and ACA, 2017) show that aquifers in the Guardiola catchment mainly consist of carbonated consolidated material with high porosity, due to karstification. The Ripoll catchment, on the other hand, generally has low permeability, with presence of aquifers only locally. It is therefore expected that storage plays a more prominent role in the hydrology of the Guardiola catchment than in the Ripoll catchment.

Fig. 2b shows the average monthly precipitation throughout the year for both sub-catchments. From April until November, the observed precipitation is rather constant, around 70 and 90 mm/month for Guardiola and Ripoll, respectively. A clear dip in precipitation is observed in the winter season, which corresponds with the results of Barrera Escoda and Llasat Botija (2015). There is no large difference in normalized precipitation rates between the two sub-catchments. Fig. 2c shows the average monthly discharge throughout a year for both discharge stations. What can be clearly distinguished is the peak in spring (May) in both catchments, caused by snowmelt in the Pyrenees. After that, discharge decreases, even though the precipitation remains constant. The reason for the decrease is the high evapotranspiration rates in summer. The next peak, at the end of autumn, is caused by a decrease in evapotranspiration, while precipitation remains constant. There is only a slight difference in specific discharge between the two stations.

### 2.2.2. The SFO dataset

The SFO dataset was provided by the European Flood Awareness System (EFAS, [www.efas.eu](http://www.efas.eu)), which is part of the Copernicus Emergency Management Services (CEMS). The dataset contains gridded time series of hydrological and meteorological variables. Meteorological variables were obtained by spatially interpolating in situ observations to a 5 × 5 km grid, using inverse distance weighting (Ntegeka et al., 2013; Smith et al., 2016). Hydrological variables, e.g. discharge, were simulated with the LISFLOOD model, using observed meteorological variables as model input and previously simulated hydrological variables as initial conditions. LISFLOOD is a physically-based, spatially distributed, rainfall-runoff model (Burek et al., 2013). As its main input, LISFLOOD uses soil properties (King et al., 1994), land cover (EEA, 2012), elevation data (Farr et al., 2007), and meteorological variables: precipitation, potential evapotranspiration and average daily temperature (Ntegeka et al., 2013). LISFLOOD was designed to simulate water balances for large European river basins. Originally it was set up to simulate floods, but it has already successfully been used to analyse streamflow droughts in other studies (e.g. Feyen and Dankers, 2009; Forzieri et al., 2014; Cammalleri et al., 2017). The model has also previously been validated, e.g. by Zajac et al. (2013) and Smith et al. (2016), who indicated that model efficiency was relatively low on the Iberian Peninsula. Nonetheless, a comparison of both meteorological and hydrological drought forecasts, resulting from both in situ observed and synthetic data, might lead to new insights on the performance of the model in drought forecasting. In this study, a daily SFO dataset was used, covering the pan-European region on a 5 × 5 km spatial resolution from January 1, 1990, to December 31, 2016. From this dataset, discharge (m<sup>3</sup>/s) and precipitation (mm/d) were selected to represent hydrological and meteorological conditions, respectively.

### 2.2.3. In situ observations

In situ observed precipitation and discharge were obtained from the Catalan Water Agency (Agència Catalana de l'Aigua, ACA). Daily precipitation data were available for 13 different stations in and around the Ripoll and Guardiola catchments (1996-present), whose locations are shown in Fig. 2a. Additionally, Fig. A.1a in the appendix shows a timeline with the availability of in situ precipitation data per station. Precipitation data were interpolated to catchment averages using Thiessen polygons, because of the simplicity and the short computation time of the method. For each time step in the dataset, polygons were computed based on the stations which had data available. An example for a time step for which all stations have data, is given in Fig. A.1b

Daily discharge observation data were supplied for two gauging stations: Ripoll and Guardiola (1916-present). Missing data were filled in using the percentile method, outlined in Tallaksen and Van Lanen (2004, pp 120).

### 2.2.4. Reforecasts

The final dataset used in this study consists of gridded seasonal hydro-meteorological reforecasts, i.e. precipitation and discharge, at a  $5 \times 5$  km scale. These reforecasts were calculated by the LISFLOOD model in a similar way as the SFO dataset (Section 2.2.2), but it used downscaled ECMWF-SEAS4 meteorological reforecasts as input, instead of gridded observed meteorological data. The grid cells in this dataset are therefore located on the same locations as in the SFO dataset. Reforecast data were provided for the years 2002 to 2010, including a severe drought event in the Iberian Peninsula (2007–2008, Stagge et al., 2013; Ionita et al., 2017; Gonzáles-Hidalgo et al., 2018). ECMWF-SEAS4 is a dataset of probabilistic seasonal weather reforecasts, produced by the ECMWF (Molteni et al., 2011). Seasonal in this case means that ensemble reforecasts with a lead time of 215 days (about 7 months) were available for each month. The seasonal weather reforecasts used in this study consist of 15 ensemble members. Each of these ensemble members provided a time series of reforecasted meteorological data, such as precipitation, temperature, and evapotranspiration, which were run through LISFLOOD, returning a daily time series of reforecasted discharge. In total, we used 15 daily time series of reforecasted precipitation and discharge (each 215 days long) for each of the 108 months in the period 2002–2010 for the evaluation of the skill and robustness.

## 3. Results

### 3.1. Data comparison

The efficiency parameters ( $R^2$ , KGE, and NSE) between the time series of daily SFO and in situ observed precipitation and discharge are summarised in Table 1. If the SFO and the in situ observed precipitation data were compared (upper part of Table 1) using  $R^2$ , analyses with and without moving averages show similar correlation ( $\pm 0.5$ ). This is not the case for the NSE. In the Guardiola catchment, both precipitation

**Table 1**

Efficiencies of the water balance variables compared to the in situ observations in the period 2002–2010. Efficiencies for both the precipitation (upper) and discharge (lower) are shown. “GU” and “RI” are abbreviations for Guardiola and Ripoll, respectively, “MA” indicates that a moving average was applied to both datasets before the efficiencies were calculated.

		$R^2$	NSE	KGE
Precipitation	RI	0.49	-0.04	0.25
	MA RI	0.60	-0.45	0.30
	GU	0.49	0.14	0.52
	MA GU	0.62	0.39	0.65
Discharge	RI	0.37	-6.25	-1.61
	MA RI	0.58	-0.25	0.03
	GU	0.28	-0.13	0.25
	MA GU	0.57	0.41	0.44

NSEs are positive and improve if a moving average is applied, because the peaks are smoothed. In the Ripoll catchment, on the other hand, the NSE for the moving average is smaller than the non-moving average, and both are negative. Both the NSE and the KGE are higher in the Guardiola than in the Ripoll catchment. Moreover, in both catchments, the KGE is higher when a moving average was applied.

The lower part of Table 1 shows the results for the discharge data. It indicates that the SFO precipitation data is closer to the in situ observed values than the SFO discharge: before applying a moving average, some of the discharge efficiency parameters are very low. LISFLOOD constantly underestimates the discharge in the Ripoll catchment. In the Guardiola catchment, the main trends are well modelled, the model only misses some peaks. For both stations,  $R^2$  is higher if a moving average was applied. The NSE is higher in the Guardiola than in the Ripoll catchment. Likewise, the KGE shows that the LISFLOOD model performs better in the Guardiola than in the Ripoll catchment. The difference between the two catchments can be explained by the difference in catchment behaviour: the Guardiola catchment is more storage-dominated than the Ripoll catchment (Section 2.2.1), which usually improves discharge forecasts, due to a more predictable time lag between precipitation and discharge peaks.

### 3.2. Forecast skill

Fig. 3a-d shows that meteorological drought forecasts (SPI1, SPI3), as compared to the SFO data, do generally not outperform the climatology for these short accumulation times (shown by  $BSS < 0$ ). Skill increases as accumulation time increases (SPI6, SPI12, Fig. 3e-h). The BSS is positive for short LT and long accumulation times. However, the positive BSS is mainly caused by the addition of SFO data to the reforecast data (Section 2.1.2), which was necessary to calculate the SPI with large accumulation times (Dutra et al., 2013; Yuan et al., 2013a, 2013b).

The seasonal results show that: (1) Winter (DJF, red) often has a higher BSS than the annual skill for almost all SPIs and LTs; (2) Spring (MAM, green) often shows lower values than the annual BSS; (3) Summer (JJA, cyan) generally shows similarity to the annual BSS, except for Guardiola at SPI6, where summer skill is lower than annual skill, and Ripoll at SPI12, where the summer BSS is higher than the annual values; and (4) Autumn (SON, purple) is comparable to the annual BSS, except for SPI6, where values are considerably lower for short LTs, and higher for long LTs.

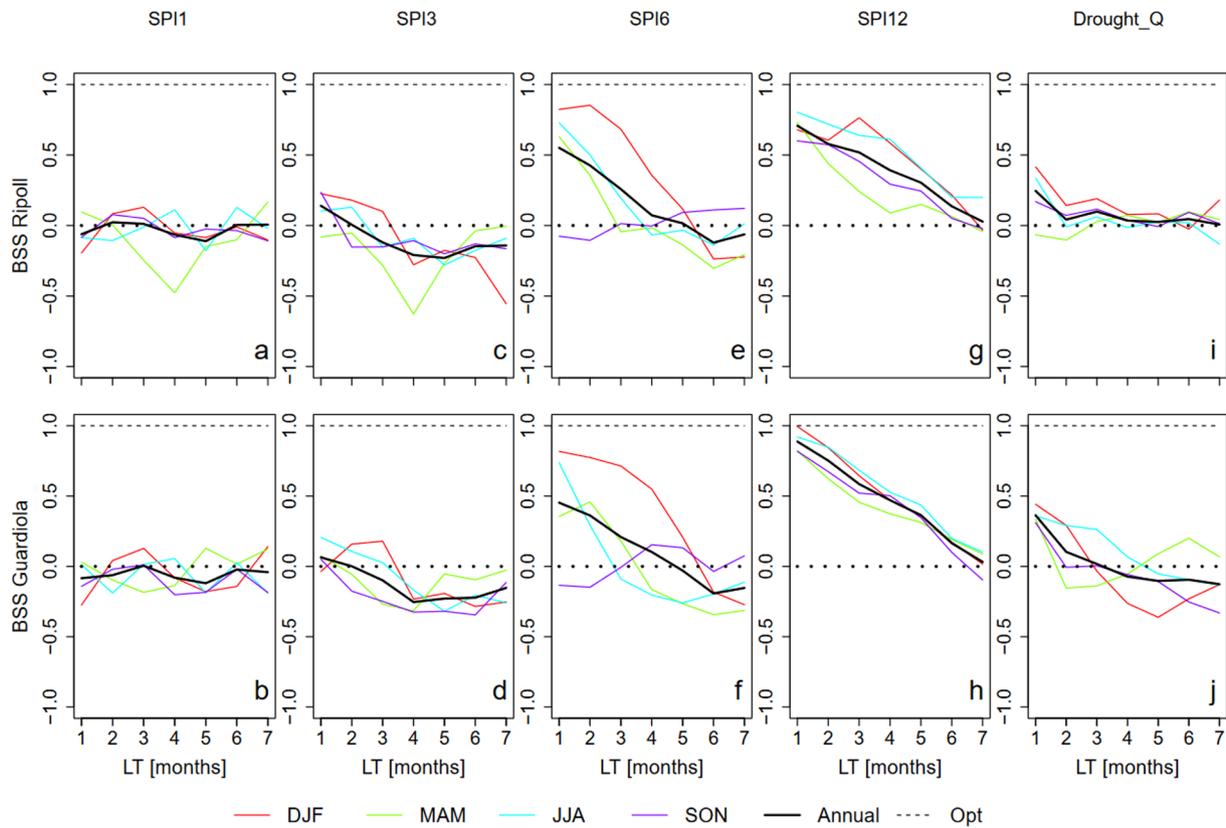
Hydrological drought reforecasts outperform the climatology up to 3–4 months LT (shown by  $BSS > 0$  in Fig. 3i,j). Seasonal hydrological drought forecasts are comparable to the annual values for both catchments. The higher performance of the hydrological drought forecasts than the meteorological drought forecasts is caused by memory of the catchments, that is, storage capacity, which does not play a role in meteorological forecasts.

### 3.3. Robustness

Table 2 shows the results of the sensitivity analysis to assess aspects of robustness. The BSS was calculated using different thresholds, and using in situ observed data instead of SFO data both for the meteorological drought (SPI) and for the hydrological drought (Drought\_Q). Changes in BSS are visible, which points at a relatively low robustness of the model. In general, if the skill is higher in the first three months (LT = 1–3), the skill is also higher in the second part of the reforecasts (LT = 4–7 months), and vice versa.

When the observed datasets were used (OBS in Table 2), instead of the SFO data, the meteorological drought forecasts show lower skill in most cases for both catchments, especially in the Ripoll catchment. This is in line with Trambauer et al. (2015), who hypothesised that forecasts using observed datasets compared to SFO data would show the highest possible skill.

The sensitivity analysis also shows that the use of a less extreme threshold value (SPI = 0, Q70 in Table 2) leads to decreased skill in the



**Fig. 3.** Brier Skill Scores (BSS) in SPI1 (a,b), SPI3 (c,d), SPI6 (e,f), SPI12 (g,h) and in discharge drought, Drought\_Q (i,j), for the Ripoll (a-i) and Guardiola (b-j) catchments. The coloured and black lines show the BSS per season and for the entire year, respectively. The dashed line shows the optimum of the BSS (1) and the dots at BSS = 0 support to recognize when the model skill outperforms the climatology.

hydrological drought forecasts and improved skill in the meteorological drought forecasts, for all accumulation times. Oppositely, the use of a more extreme threshold (SPI = -1, Q90 in Table 2) leads to a higher skill for the hydrological drought forecasts and a lower skill in the meteorological drought forecasts.

#### 4. Discussion

##### 4.1. Uncertainty in data

The performance of the EFAS datasets, precipitation and discharge (SFO), was assessed by comparing these data to in situ observed data, using several efficiency metrics. Both the discharge and precipitation

simulations show higher efficiencies in the Guardiola catchment than in the Ripoll catchment (Table 1). It was expected that the precipitation would have a higher efficiency coefficient than the discharge, as the gridded precipitation, which was used as input for LISFLOOD, was produced from observations (Section 2.2). Nonetheless, precipitation efficiency coefficients were lower than expected, likely caused by the different methods used for spatial interpolation in the two datasets and the number and distribution of stations. The in situ observations were interpolated using Thiessen polygons (TP), a quick and simple method, whereas the SFO precipitation data were interpolated using inverse distance weighting (IDW). Vicente-Serrano et al. (2003) studied different interpolation methods on precipitation data in the Ebro delta (Spain) and found that the use of IDW resulted in  $r^2 = 0.90$ , whereas

**Table 2**

Summary of the sensitivity analysis of the forecast skills of meteorological drought (SPI) and hydrological drought (Drought\_Q), for the Ripoll and Guardiola catchments. The BSS (entire year) is divided into two lead time periods: per catchment, the upper row shows the average BSS in the first 3 months lead time, whereas the lower row shows the average BSS in the last 4 months lead time. The BSS is considered to be good (green colour) if the values are in between 0.50 and 1.00, acceptable (yellow colour) if the values are in between 0.00 and 0.50, poor (orange colour) if the values are in between -0.50 and 0.00 and bad (red colour) if the values are in between  $-\infty$  and -0.5. OBS stands for in situ observations and SFO for simulation forced with observations.

LT	SFI	SPI=-0.5, SFO (reference)				SPI=-0.5, OBS				SPI=0, SFO				SPI=-1, SFO				Drought_Q			
		SPI1	SPI3	SPI6	SPI12	SPI1	SPI3	SPI6	SPI12	SPI1	SPI3	SPI6	SPI12	SPI1	SPI3	SPI6	SPI12	Q80, SFO (reference)	Q80, OBS	Q70, SFO	Q90, SFO
RI	1-3	-0.01	0.01	0.41	0.60	-0.39	-0.29	-0.23	0.62	0.17	0.28	0.52	0.80	-0.08	-0.06	0.36	0.50	0.13	-1.15	0.08	0.16
	4-7	-0.04	-0.18	-0.02	0.21	-0.52	-0.65	-0.63	0.16	0.11	0.06	0.23	0.50	-0.07	-0.23	-0.05	0.02	0.03	-0.99	0.06	0.01
GU	1-3	-0.05	-0.01	0.34	0.74	-0.10	-0.12	-0.24	0.64	0.19	0.35	0.53	0.80	-0.10	-0.02	0.26	0.56	0.16	0.10	0.05	0.22
	4-7	-0.07	-0.21	-0.07	0.26	-0.10	-0.26	-0.21	0.21	0.14	0.16	0.14	0.46	-0.14	-0.24	-0.17	0.02	-0.10	-0.04	-0.25	-0.02

the use of TP resulted in  $r^2 = 0.86$ . The different interpolation methods could thus have led to different values for precipitation in the two datasets, and therefore to lower efficiency coefficients.

#### 4.2. Drought forecasts skill and robustness

In this study, for the very first time, hydrological and meteorological forecasts from the LISFLOOD model were compared to both in situ observations and a proxy for observations on a basin scale. This new approach showed that the skill of meteorological drought (SPI) forecasts, calculated from gridded precipitation data, is lower than climatology for short accumulation periods (SPI1-3). For SPI data with longer accumulation periods, the model skill is higher, indicated by a positive BSS. This is to be expected, as longer accumulation periods have a higher proportion of SFO data, so the skill is artificially inflated. Hydrological drought (Drought<sub>Q</sub>) forecasts simulated by the LISFLOOD model show higher skill than the climatology, for lead times up to 3–4 months. The difference between skill in meteorological and hydrological drought forecasts can be explained by catchment memory, that is, prevailing catchment storage capacity (soil water, groundwater). Soil water and groundwater explain over 50% of the variance in discharge forecasts for short LTs (1–2 months) and remains about 40% up to LT = 3–4 months (Wanders et al., 2019).

The hydrological drought forecasts in Fundel et al. (2013), show high skill throughout the 1 month lead time forecast. This study shows that forecasting skill remains positive for lead times up to 3–4 months, and consequently, that it proves to be useful to look at longer time scales for hydrological drought forecasts. A comparison between the flood forecast results of Bartholmes et al. (2009) and the drought forecast results of this study, shows that drought forecasts have a higher skill than flood forecasts. The BSS at LT = 1 month in seasonal drought forecasts, as analysed here, is higher than the BSS at LT = 10 days in the 10-day flood forecasts (Bartholmes et al., 2009). An explanation is given by Hirschberg et al. (2011), who indicate that processes on a longer time scale usually have a lower uncertainty than processes occurring in a relatively short time scale. Floods are often related to hydrological processes and catchment state variables (e.g. surface runoff, topsoil wetness) that drive quick response of discharge on a precipitation event. Floods thus act on a short time scale, and consequently are expected to have high uncertainty. During a drought event, on the other hand, the low discharge is more linked to processes and state variables (e.g. base flow, groundwater table depth) that control slow response. Droughts, therefore, act on a longer time scale, and consequently are expected to have low uncertainty (Hirschberg et al., 2011).

The robustness analysis showed that in general, changes in BSS occur, which points at a rather low robustness of the model. Furthermore, in general, the use of a more extreme threshold led to increased skill in hydrological drought forecasts and decreased skill in meteorological drought forecasts. Trambauer et al. (2015) stated that forecast skills calculated using an SFO dataset should be treated as the upper skill limit. Our results show that this holds for droughts in the two case study catchments as well (Table 2).

#### 4.3. The use of LISFLOOD and ECMWF SEAS4

As described in Section 2.2, the hydrological forecasts and the SFO data, which were used as a proxy for observations, were simulated by the LISFLOOD model. Output from different continental-scale hydrological models could have been used, but literature shows that it is unlikely that this would have increased drought forecast performance (Kauffeldt et al., 2016; Trambauer et al., 2013). Kauffeldt et al. (2016) studied 24 large-scale hydrological models to be encapsulated in an operational forecasting tool, and concluded that other pan-European models than LISFLOOD do not result in a better framework. A similar conclusion was drawn by Trambauer et al. (2013), who performed a review of 16 continental-scale hydrological models on the African

continent for drought forecasting. Trambauer et al. (2015) used a fine resolution version of the PCR-GLOBWB model, and compared three seasonal drought forecasting methods for both streamflow and Standardised Runoff Index data: (1) ECMWF SEAS-4 (S4); (2) Ensemble Streamflow Prediction (ESP) approach using resampled historical data; and (3) an ESP approach conditional on the El Niño Southern Oscillation (ESP-cond). They found that S4 performs better than both ESPs at longer lead times. Contrarily, Arnal et al. (2018), who studied LISFLOOD seasonal streamflow forecasts over Europe with a spatial resolution of  $0.7^\circ$ , found that, on average, ESPs perform better than S4 for longer lead times, though this differs per season. It remains to be explored if this would also be the case for high resolution data in Catalonia.

LISFLOOD was calibrated for the EFAS by Zajac et al. (2013), whose results were also discussed by Smith et al. (2016). Results showed that NSE was very low on the Iberian Peninsula, compared to the rest of Europe, during the validation period. Moreover, several studies on LISFLOOD model validation in more than 231 river basins across Europe, show that LISFLOOD can produce reasonable river discharge compared to observations (Feyen et al., 2007, 2008; Feyen and Dankers, 2009; Van der Knijff et al., 2010). This is promising, as they indicate that forecast skill in catchments outside the Iberian Peninsula is likely to be higher than in the two studied pilot sites.

#### 4.4. Limitations of the study

Amongst others, Kumar (2009) and Shi et al. (2015) stated that skill scores, calculated based on small verification time series, can substantially deviate from their expected values because of sampling errors. The Brier Skill Score, which we used in this study to analyse forecast skill, is dependent on the number of droughts per period, as discussed by Stanski et al. (1989). If relatively few droughts occur, the BSS is often low due to a very small  $BS_{\text{clim}}$ , which is  $(0.2-0)^2 = 0.04$  if no drought occurs (Eq. (1)). Oppositely, if more droughts occur, the BSS automatically increases, as the  $BS_{\text{clim}}$  increases to  $(0.2-1)^2 = 0.64$  during a drought event. According to Stanski et al. (1989), this issue can only be overcome with large sample sizes. A small sample size can thus have a large impact on the generated results. In this study, a reforecast period of nine years was used, which is substantially smaller than the 30 years, which are recommended for a skill score unaffected by sampling errors (Kumar, 2009). It is important to realize that the small sample size could have affected the results presented in this paper, and that a larger reforecast dataset would likely have improved reliability of the results.

The relatively low correlation parameters in Table 1 show that some improvements can still be made in the forecasts. Bias correction is an interesting starting point for further research. Shi et al. (2008) studied the effect of bias correction on seasonal streamflow forecasts and found that the effect of bias correction on skill scores is almost equally large as the effect of model calibration. Furthermore, Wang et al. (2011) found that bias correction can be especially helpful if the NSE and KGE are much lower than the  $R^2$ , which is the case in this study (Table 1). This indicates that bias correction could decrease the bias and, consequently, will likely increase drought forecasting skill in Catalonia. The skill in the results shown here can presumably be further increased by e.g.: (1) improving the quality of the initial conditions (Koster et al., 2010; Singla et al., 2012; Fundel et al., 2013); (2) using a multi-model ensemble of both climate models and hydrological models (Thober et al., 2015; Samaniego et al., 2017; Thober et al., 2018; Wanders et al., 2019); (3) using improved meteorological reforecasts, i.e. ECMWF SEAS-5, which became operational in November 2017 (Yuan et al., 2011; Singla et al., 2012; Meißner et al., 2017); (4) using a small-scale hydrological model, as several physical processes occur on a smaller scale than the spatial scale of  $5 \times 5$  km, which was used in this study (Zappa et al., 2014); and (5) taking into account the coupling between for instance soil moisture and air temperature, so that meteorological

variables are influenced by hydrology, and not only the other way around.

## 5. Conclusions

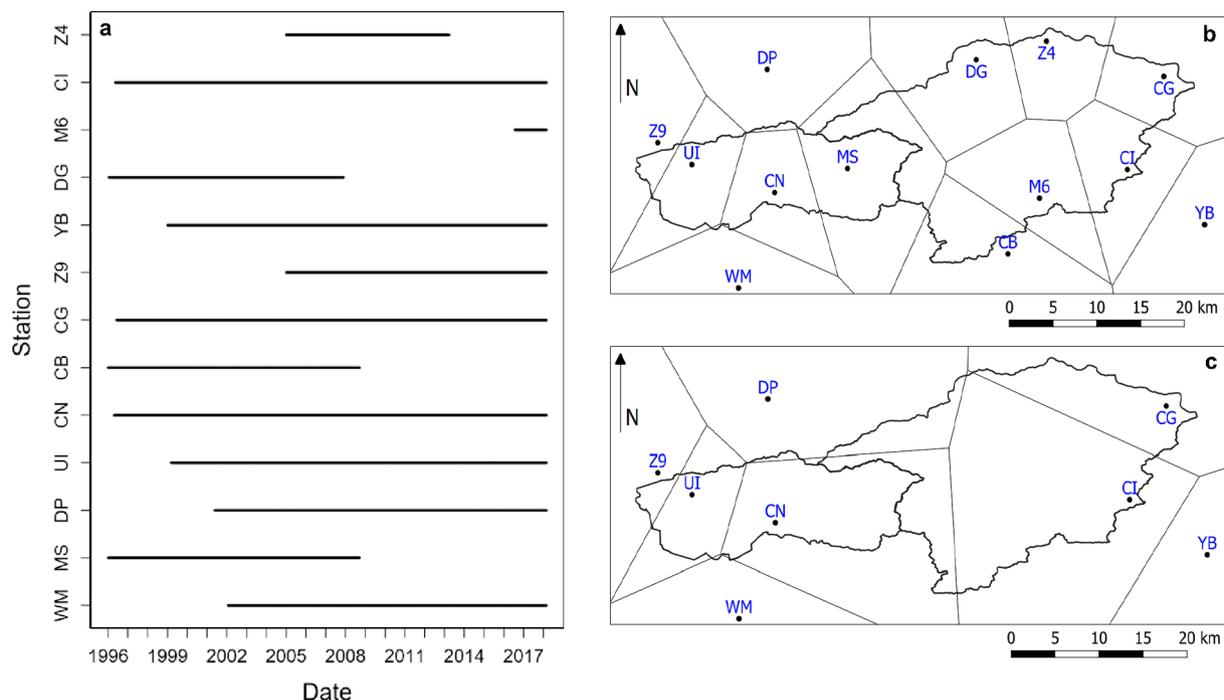
This study led to improved knowledge on drought forecasting on a catchment scale in Catalonia. First, results show that both seasonal meteorological and hydrological drought forecasts show highest skill when they are done in winter, whereas forecasts performed in spring show the lowest skill. Second, it can be concluded that meteorological drought forecasts have higher uncertainty than hydrological drought forecasts.

In this study, we analysed the skill and robustness of meteorological drought and hydrological drought forecasts in two catchments in Catalonia, using the LISFLOOD hydrological model driven by probabilistic ECMWF-SEAS (S4) seasonal weather forecasts. The seasonal hydro-meteorological forecasting with lead times up to 7 months is part of the Multi-Hazard Early Warning System (MH-EWS) developed in the framework of the ANYWHERE project (<http://anywhere-h2020.eu/>). It is the newest pre-operational MH-EWS currently available at the pan-European scale, and the first one that provides seasonal drought forecasts rather than only hydrological forecasts (soil moisture, discharge).

Forecasting both meteorological and hydrological drought enabled intercomparison, which shows that skills of hydrological drought forecasting in the case study catchments are higher than skills of meteorological forecasting, except for precipitation accumulation periods of 6 and 12 months. This is an important finding for Catalonian water resources management, which should rely more on hydrological drought forecasts. If these hydrological forecasts are unavailable, SPI forecasts with long accumulation times can be used as an alternative, but it remains a challenge to find the right SPI accumulation time,

## Appendix A

See Fig. A.1 and Table A.1.



**Fig. A.1.** The bars in Figure a show when data are available for each of the thirteen precipitation stations in the Guardiola (left in b and c) and the Ripoll catchment (right in b and c). The location of each precipitation station is shown in Figure b, including the corresponding Thiessen polygons, which were used for spatial interpolation of precipitation data. Figure c shows an example of the Thiessen polygons which were used if only the eight shown stations contained data. The abbreviations used in this figure belong to meteorological stations given in Table A.1.

which is unique for every river basin.

Skill and robustness of hydrological and meteorological drought forecasts may be increased if bias is further reduced, ECMWF-SEAS5 reforecasts are used, a new EFAS system with improved LISFLOOD model calibration and parameterization is included, and hydrological initial conditions are improved.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The research was partly supported by the ANYWHERE project (Grant Agreement No.: 700099), which is funded within EU's Horizon 2020 research and innovation program ([www.anywhere-h2020.eu](http://www.anywhere-h2020.eu)). The first author was additionally supported by the Fonds National de la Recherche Luxembourg (FNR) (PRIDE15/10623093 – HYDRO-CSI). The hydro-meteorological SFO and reforecasts originated from the EFAS Computational Centre, which is part of the Copernicus Emergency Management Service (EMS) and Early Warning Systems (EWS) funded by framework contract number 198702 of the European Commission. We thank Paul Smith (ECMWF, Reading, UK) for providing these SFO and reforecast data, and Arnau Cangros Alonso (ACA, Barcelona, Spain) for providing the in situ observed datasets. This research is part of the Wageningen Institute for Environment and Climate Research (WIMEK-SENSE) and it supports the work of UNESCO EURO FRIEND-Water and the IAHS Panta Rhei program.

**Table A.1**  
Station names corresponding to the abbreviations used in Fig. A.1.

Code	Station name
Z4	Ulldeter
CI	Sant Pau de Segúries
M6	Sant Joan de les Abadesses
DG	Núria
YB	Olot
Z9	Cadí Nord – Prat d’Aguiló
CG	Molló
CB	Montesquiu
CN	Guardiola de Berguedà
UI	Gisclareny
DP	Das - Aeròdrom
MS	Castellar de n’Hug – el Clot del Moro
WM	Santuari de Queralt

## References

- Agència Catalana de l’Aigua, 2008. Water in Catalonia: Diagnosis and proposed actions. Significant water management issues raised within the compilation of the river basin district management plan for Catalonia.
- Alderlieste, M.A.A., Van Lanen, H.A.J., Wanders, N., 2014. Future low flows and hydrological drought: how certain are these for Europe? In: Daniell, T.M., Van Lanen, H.A.J., Demuth, S., Laaha, G., Servat, E., Mahe, G., Boyer, J-F, Paturel, J-E, Dezetter, A., Ruelland, D. (Eds.), *Hydrology in a Changing World: Environmental and Human Dimensions*, IAHS Publ. No. 363, pp. 60–65.
- Arnal, L., Cloke, H.L., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., Krzeminski, B., Pappenberger, F., 2018. Skilful seasonal forecasts of streamflow over Europe? *Hydrolog. Earth Syst. Sci.* 22 (4).
- Barrera Escoda, A., Llasat Botija, M.D.C., 2015. Evolving flood patterns in a Mediterranean region (1301–2012) and climatic factors - the case of Catalonia. *Hydrology Earth Syst. Sci.* 19 (1), 465–483.
- Bartholmes, J.C., Thielen, J., Ramos, M.H., Gentilini, S., 2009. The European flood alert system EFAS – Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrolog. Earth Syst. Sci.* 13 (2), 141–153.
- Beguieria, S., Vicente-Serrano, S.M., 2017. SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index. R package version 1.7.1. <http://sac.csic.es/spei>.
- Bloomfield, J.P., Marchant, B.P., 2013. Analysis of groundwater drought building on the standardized precipitation index approach. *Hydrolog. Earth Syst. Sci.* 17, 4769–4787.
- Bradley, A.A., Schwartz, S.S., Hashino, T., 2008. Sampling uncertainty and confidence intervals for the brier score and brier skill score. *Weather Forecast.* 23 (5), 992–1006.
- Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* 78 (1), 1–3.
- Burek, P., Thielen-del Pozo, J., Thiemig, V., Salamon, P., De Roo, A., 2011. Das Europäische Hochwasser-Frühwarnsystem (EFAS). *Korrespondenz Wasserwirtschaft* 4 (11), 193–199.
- Burek, P., Van der Knijff, J., De Roo, A., 2013. LISFLOOD, distributed water balance and flood simulation model - Revised user manual. 10.2788/24719. Publications Office of the European Union. Directorate-General Joint Research Centre, Institute for Environment and Sustainability.
- Cammalleri, C., Vogt, J., Salamon, P., 2017. Development of an operational low-flow index for hydrological drought monitoring over Europe. *Hydrolog. Sci. J.* 62 (3), 346–358.
- Cloke, H.L., Pappenberger, F., 2009. Ensemble flood forecasting: a review. *J. Hydrolog.* 375 (3), 613–626.
- Day, G.N., 1985. Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.* 111 (2), 157–170.
- Delgado, J.M., Voss, S., Bürger, G., Vormoor, K., Murawski, A., Rodrigues Pereira, J.M., Martins, E., Vasconcelos Júnior, F., Francke, T., 2018. Seasonal drought prediction for semiarid northeastern Brazil: verification of six hydro-meteorological forecast products. *Hydrolog. Earth Syst. Sci.* 22 (9).
- Dutra, E., Di Giuseppe, F., Wetterhall, F., Pappenberger, F., 2013. Seasonal forecasts of drought in Africa basins using the Standardized Precipitation Index. *Hydrolog. Earth Syst. Sci.* 17, 2359–2373.
- Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., Barbosa, P., Vogt, J., Pappenberger, F., 2014. Global meteorological drought – Part 2: Seasonal forecasts. *Hydrolog. Earth Syst. Sci.* 18, 2669–2678.
- EEA, 2010. Mapping the impacts of natural hazards and technological accidents in Europe. An overview of the last decade. EEA Technical report No. 13/2010, Copenhagen, Denmark.
- EEA, 2012. Corine Land Cover Dataset (CLC 2012) (Copernicus). URL: <https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012>.
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The shuttle radar topography mission. *Rev. Geophys.* 45 (2).
- Feyen, L., Dankers, R., 2009. Impact of global warming on streamflow drought in Europe. *J. Geophys. Res.: Atmospheres* 114 (D17).
- Feyen, L., Kalas, M., Vrugt, J.A., 2008. Semi-distributed parameter optimization and uncertainty assessment for large-scale streamflow simulation using global optimization. *Hydrolog. Sci.–J.–des Sci. Hydrolog.* 53 (2), 293–308.
- Feyen, L., Vrugt, J.A., Nualláin, B.Ó., Van der Knijff, J., De Roo, A., 2007. Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *J. Hydrology* 332, 276–289.
- Fleig, A.K., Tallaksen, L.M., Hisdal, H., Demuth, S., 2006. A global evaluation of streamflow drought characteristics. *Hydrolog. Earth Syst. Sci.* 10, 535–552.
- Forzieri, G., Feyen, L., Rojas, R., Flörke, M., Wimmer, F., Bianchi, A., 2014. Ensemble projections of future streamflow droughts in Europe. *Hydrolog. Earth Syst. Sci.* 18 (1), 85.
- Fundel, F., Jörg-Hess, S., Zappa, M., 2013. Monthly hydrometeorological ensemble prediction of streamflow droughts and corresponding drought indices. *Hydrolog. Earth Syst. Sci.* 17 (1), 395–407.
- González-Hidalgo, J.C., Vicente-Serrano, S.M., Peña-Angulo, D., Salinas, C., Tomas-Burguera, M., Begueria, S., 2018. High-resolution spatio-temporal analyses of drought episodes in the western Mediterranean basin (Spanish mainland, Iberian Peninsula). *Acta Geophys.* 1–12.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrolog.* 377 (1–2), 80–91.
- Hayes, M.J., Svoboda, M.D., Wall, N., Widhalm, M., 2011. The Lincoln declaration on drought indices: Universal meteorological drought index recommended. *Bull. Am. Meteorol. Soc.* 92, 485–488.
- Hayes, M.J., Svoboda, M.D., Wilhite, D.A., Vanyarko, O.V., 1999. Monitoring the 1996 drought using the standardized precipitation index. *Bull. Am. Meteorol. Soc.* 80 (3), 429–438.
- Heinrich, G., Gobiet, A., 2012. The future of dry and wet spells in Europe: a comprehensive study based on the ENSEMBLES regional climate models. *Int. J. Climatol.* 32 (13), 1951–1970.
- Heudorfer, B., Stahl, K., 2017. Comparison of different threshold level methods for drought propagation analysis in Germany. *Hydrolog. Res.* 48 (5), 1311–1326.
- Hirschberg, P.A., Abrams, E., Bleistein, A., Bua, W., Monache, L.D., Dulong, T.W., Gaynor, J.E., Glahn, B., Hamill, T.M., Hansen, J.A., Hilderbrand, D.C., Hoffman, R.N., Hearn Morrow, B., Philips, B., Sokich, J., Stuart, N., 2011. A weather and climate enterprise strategic implementation plan for generating and communicating forecast uncertainty information. *Bull. Am. Meteorol. Soc.* 92 (12), 1651–1666.
- Hisdal, H., Tallaksen, L.M., Clausen, B., Peters, E., Gustard, A., 2004. Ch. 4 Hydrological drought characteristics. In: Tallaksen, L.M., Van Lanen, H.A.J. (Eds.), *Hydrological Drought. Processes and Estimation Methods for Streamflow and Groundwater*. Developments in Water Science, vol. 48. Elsevier Science B.V., pp. 139–198.
- Hwang, Y., Carbone, G.J., 2009. Ensemble forecasts of drought indices using a conditional residual resampling technique. *J. Appl. Meteorol. Climatol.* 48, 1289–1301.
- Institut Cartogràfic i Geològic de Catalunya (ICGC) & Agència Catalana de l’Aigua (ACA), 2017. Map of Hydrogeological Areas of Catalonia 1:250,000 (MAH250M v1.0, 2017). Available online via URL: <http://www.icgc.cat/en/Public-Administration-and-Enterprises/Downloads/Geological-and-geothematic-cartography/Hydrogeological-cartography/Map-of-Hydrogeological-Areas-of-Catalonia-1-250-000-MAH250M-v1.0-2017>.
- Ionita, M., Tallaksen, L.M., Kingston, D.G., Stagge, J.H., Laaha, G., Van Lanen, H.A.J., Scholz, P., Chelcea, S.M., Haslinger, K., 2017. The European 2015 drought from a climatological perspective. *Hydrolog. Earth Syst. Sci.* 2 (3), 1397–1419.
- Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., Thielen, J., 2016. Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environ. Modell. Software* 75, 68–76.
- King, D., Daroussin, J., Tavernier, R., 1994. Development of a soil geographic database from the soil map of the European Communities. *Catena* 21 (1), 37–56.
- Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Balsamo, G., Berg, A.A., Boisserie, M., Dirmeyer, P.A., Doblas-Reyes, F.J., Drewitt, G., Gordon, C.T., Guo, Z., Jeong, J.H., Lawrence, D.M., Lee, W.S., Li, Z., Luo, L., Malyshev, S., Merryfield, W.J., Seneviratne, S.I., Stanelle, T., van den Hurk, B.J.J.M., Vitart, F., Wood, E.F., 2010. Contribution of

- land surface initialization to hydrological forecast skill: First results from a multi-model experiment. *Geophys. Res. Lett.* 37 (L02402).
- Kumar, A., 2009. Finite samples and uncertainty estimates for skill measures for seasonal prediction. *Mon. Weather Rev.* 137, 2622–2631.
- Krzysztofowicz, R., 2001. The case for probabilistic forecasting in hydrology. *J. Hydrol.* 249 (1), 2–9.
- Llasat, M.C., Llasat-Botija, M., Barnolas, M., López, L., Altava-Ortiz, V., 2009. An analysis of the evolution of hydrometeorological extremes in newspapers: the case of Catalonia, 1982–2006. *Nat. Hazards Earth Syst. Sci.* 9, 1201–1212.
- Martin-Ortega, J., González-Eguino, M., Markandya, A., 2012. The costs of drought: the 2007/2008 case of Barcelona. *Water Policy* 14, 539–560.
- Marx, A., Kumar, R., Thober, S., Rakovec, O., Wanders, N., Zink, M., Wood, E.F., Pan, M., Sheffield, J., Samaniego, L., 2018. Climate change alters low flows in Europe under global warming 1.5, 2, and 3 °C. *Hydrol. Earth Syst. Sci.* 22, 1017–1032.
- McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales.
- Meigh, J., Tate, E., McCartney, M., 2002. Methods for identifying and monitoring river flow drought in southern Africa. In: van Lanen, H.A.J., Demuth, S. (Eds.), *FRIEND 2002 – Regional Hydrology: Bridging the Gap Between Research and Practice*, IAHS Publ. No. 274, pp. 181–188.
- Meißner, D., Klein, B., Ionita, M., 2017. Development of a monthly to seasonal forecast framework tailored to inland waterway transport in Central Europe. *Hydrol. Earth Syst. Sci.* 21 (12), 6401–6423.
- Mo, K.C., Lyon, B., 2015. Global meteorological drought prediction using the north American multi-model ensemble. *J. Hydrometeorol.* 16, 1409–1424.
- Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T., Vitart, F., 2011. The new ECMWF seasonal forecast system (System 4). *ECMWF Tech. Memorandum* 656, 1–49.
- Monhart, S., Zappa, M., Spirig, C., Schär, C., Bogner, K., 2019. Subseasonal hydro-meteorological ensemble predictions in small- and medium-sized mountainous catchments: benefits of the NWP approach. *Hydrol. Earth Syst. Sci.* 23, 493–513.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* 10 (3), 282–290.
- Ntegeka, V., Salamon, P., Gomes, G., Sint, H., Lorini, V., Zambrano-Bigiarini, M., Thielen, J., 2013. EFAS-Meteo: A European daily high-resolution gridded meteorological data set for 1990–2011. Report EUR, 26408 EN. Luxembourg: Publications Office of the European Union.
- Orlowsky, B., Seneviratne, S.I., 2013. Elusive drought: uncertainty in observed trends and short- and long-term CMIP5 projections. *Hydrol. Earth Syst. Sci.* 17, 1765–1781.
- Palmer, W.C., 1965. Meteorological drought. U.S. Weather Bureau, Research Paper No. 45.
- Pascual, D., Pla, E., Lopez-Bustins, J.A., Retana, J., Terradas, J., 2015. Impacts of climate change on water resources in the Mediterranean Basin: A case study in Catalonia, Spain. *Hydrol. Sci. J.* 60 (12), 2132–2147.
- Prudhomme, C., Giuntoli, I., Robinson, E.L., Clark, D.B., Arnell, N.W., Dankers, R., Fekete, B.A.Z.M., Franssen, W., Gerten, D., Gosling, S.N., Hagemann, S., Hannah, D.M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., Wisser, D., 2014. Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proc. Nat. Acad. Sci.* 111 (9), 3262–3267.
- Russo, S., Dosio, A., Sterl, A., Barbosa, P., Vogt, J., 2013. Projection of occurrence of extreme dry-wet years and seasons in Europe with stationary and nonstationary Standardized Precipitation Indices. *J. Geophys. Res.: Atmospheres* 118 (14), 7628–7639.
- Samaniego, L., Kumar, R., Breuer, L., Chamorro, A., Flörke, M., Pechlivanidis, I.G., Schäfer, D., Shah, H., Vetter, T., Wortmann, M., Zeng, X., 2017. Propagation of forcing and model uncertainties on to hydrological drought characteristics in a multi-model century-long experiment in large river basins. *Clim. Change* 141 (3), 435–449.
- Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E.F., Marx, A., 2018. Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Clim. Change* 8, 421–426.
- Samaniego, L., Thober, S., Wanders, N., Pan, M., Rakovec, O., Sheffield, J., Wood, E.F., Prudhomme, C., Rees, G., Houghton-Carr, H., Fry, M., Smith, K., Watts, G., Hisdal, H., Estrela, T., Buontempo, C., Marx, A., Kumar, R., 2019. Hydrological forecasts and projections for improved decision-making in the water sector in Europe. *Bull. Am. Meteorol. Soc.* (in press).
- Schaake, J.C., Hamill, T.M., Buizza, R., Clark, M., 2007. HEPEx: the hydrological ensemble prediction experiment. *Bull. Am. Meteorol. Soc.* 88 (10), 1541–1547.
- Shi, W., Schaller, N., MacLeod, D., Palmer, T.N., Weisheimer, A., 2015. Impact of hindcast length on estimates of seasonal climate predictability. *Geophys. Res. Lett.* 42 (5), 1554–1559.
- Shi, X., Wood, A.W., Lettenmaier, D.P., 2008. How Essential is Hydrologic Model Calibration to Seasonal Streamflow Forecasting? *J. Hydrometeorol.* 9 (6), 1350–1363.
- Shukla, S., Lettenmaier, D.P., 2011. Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill. *Hydrol. Earth Syst. Sci.* 15 (11), 3529–3538.
- Shukla, S., Wood, A.W., 2008. Use of a Standardized Runoff Index for characterizing hydrologic drought. *Geophys. Res. Lett.* 35 (L02405).
- Singla, S., Céron, J.P., Martin, E., Regimbeau, F., Déqué, M., Habets, F., Vidal, J.P., 2012. Predictability of soil moisture and river flows over France for the spring season. *Hydrol. Earth Syst. Sci.* 16 (1), 201–216.
- Smith, P.J., Pappenberger, F., Wetterhall, F., del Pozo, J.T., Krzeminski, B., Salamon, P., Muraro, D., Kalas, M., Baugh, C., 2016. On the operational implementation of the European Flood Awareness System (EFAS). *ECMWF Tech. Memorandum* 778, 1–34.
- Stagge, J.H., Tallaksen, L.M., Gudmundsson, L., Van Loon, A.F., Stahl, K., 2015. Candidate distributions for climatological drought indices (SPI and SPEI). *Int. J. Climatol.* 35, 4027–4040.
- Stagge, J.H., Tallaksen, L.M., Kohn, I., Stahl, K., van Loon, A.F., 2013. A European Drought Reference (EDR) Database: Design and Online Implementation. *DROUGHT-R&SPI Tech. Report No. 12*, pp. 1–47.
- Stanski, H.R., Burrows, W.R., Wilson, L.J., 1989. Survey of common verification methods in meteorology. WMO World Weather Watch Technical Report No.8, WMO/TD No. 358.
- Staudinger, M., Seibert, J., 2014. Predictability of low flow – An assessment with simulation experiments. *J. Hydrol.* 519, 1383–1393.
- Sutanto, S.J., Van Lanen, H.A.J., 2019. Hydrological drought forecasts outperform meteorological drought forecasts. *Sci. Rep.* (in preparation).
- Sutanto, S.J., Van Lanen, H.A.J., Wetterhall, F., Lior, X., 2019. Potential of pan-European seasonal hydro-meteorological drought forecasts obtained from a Multi-Hazard Early Warning System. *Bull. Am. Meteorol. Soc.* (in preparation).
- Tallaksen, L.M., Van Lanen, H.A.J., 2004. *Hydrological drought. Processes and Estimation Methods for Streamflow and Groundwater*. Elsevier, Amsterdam.
- Thober, S., Kumar, R., Wanders, N., Marx, A., Pan, M., Rakovec, O., Samaniego, L., Sheffield, J., Wood, E.F., Zink, M., 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.
- Thober, S., Kumar, R., Sheffield, J., Mai, J., Schäfer, D., Samaniego, L., 2015. Seasonal soil moisture drought prediction over Europe using the north American multi-model ensemble (NMME). *J. Hydrometeorol.* 16, 2329–2344.
- Trambauer, P., Maskey, S., Winsemius, H., Werner, M., Uhlenbrook, S., 2013. A review of continental scale hydrological models and their suitability for drought forecasting in (sub-Saharan) Africa. *Phys. Chem. Earth, Parts A/B/C* 66, 16–26.
- Trambauer, P., Werner, M., Winsemius, H.C., Maskey, S., Dutra, E., Uhlenbrook, S., 2015. Hydrological drought forecasting and skill assessment for the Limpopo River basin, southern Africa. *Hydrol. Earth Syst. Sci.* 19 (4), 1695–1711.
- Van den Hurk, B.J.J.M., Bouwer, L.M., Buontempo, C., Döschner, R., Erçin, E., Hananel, C., Hunink, J.E., Kjellström, E., Klein, B., Manez, M., Pappenberger, F., Pouget, L., Ramos, M.H., Ward, P.J., Weerts, A.H., Wijngaard, J.B., 2016. Improving predictions and management of hydrological extremes through climate services: www.impres.eu. *Clim. Serv.* 1, 6–11.
- Van der Knijff, J.M., Younis, J., De Roo, A.P.J., 2010. LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *Int. J. Geograph. Inform. Sci.* 24, 189–212.
- Van der Wiel, K., Wanders, N., Selten, F.M., Bierkens, M.F.P., 2019. Added value of large ensemble simulations for assessing extreme river discharge in a 2 °C warmer world. *Geophys. Res. Lett.* 46, 2093–2102.
- Van Lanen, H.A.J., Laaha, G., Kingston, D.G., Gauster, T., Ionita, M., Vidal, J.-P., Vlnas, R., Tallaksen, L.M., Stahl, K., Hannaford, J., Delus, C., Fendekova, M., Mediero, L., Prudhomme, C., Retz, E., Romanowicz, R.J., Gailliez, S., Wong, W.K., Adler, M.-J., Blauhut, V., Caillouet, L., Chelcea, S., Frolova, N., Gudmundsson, L., Hanel, M., Haslinger, K., Kireeva, M., Osuch, M., Sauquet, E., Stagge, J.H., Van Loon, A.F., 2016. Hydrology needed to manage droughts: the 2015 European case. *Hydrol. Process.* 30, 3097–3104.
- Van Lanen, H.A.J., Prudhomme, C., Wanders, W., Van Huijgevoort, M.H.J., 2018. Future of drought. In: Iglesias, A., Assimacopoulos, D., Van Lanen, H.A.J. (Eds.), *Drought: Science and Policy*. Wiley Blackwell, pp. 69–92.
- Van Loon, A.F., 2015. Hydrological drought explained. *Wiley Interdiscipl. Rev.: Water* 2 (4), 359–392.
- Vicente-Serrano, S.M., López-Moreno, J.I., Beguería, S., Lorenzo-Lacruz, J., Azorin-Molina, C., Morán-Tejeda, E., 2012. Accurate computation of a Streamflow Drought Index. *J. Hydrol. Eng.* 17 (2), 318–332.
- Vicente-Serrano, S.M., Saz-Sánchez, M.A., Cuadrat, J.M., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Climate Res.* 24 (2), 161–180.
- Vogt, J., Barbosa, P., Hofer, B., Singleton, A., 2011. Developing a European drought observatory for monitoring, assessing and forecasting droughts across the European continent. *Geophys. Res. Abstracts* 13 (EGU2011-6350).
- Wanders, N., Thober, S., Kumar, R., Pan, M., Sheffield, J., Samaniego, L., Wood, E., 2019. Development and evaluation of a Pan-European multi-model seasonal hydrological forecasting system. *J. Hydrometeorol.* 20, 99–115.
- Wanders, N., Van Lanen, H.A.J., 2015. Future discharge drought across climate regions around the world modelled with a synthetic hydrological modelling approach forced by three general circulation models. *Nat. Hazards Earth Syst. Sci.* 15 (3), 487–504.
- Wanders, N., Wada, Y., Van Lanen, H.A.J., 2015. Global hydrological droughts in the 21st century under a changing hydrological regime. *Earth Syst. Dyn.* 6, 1–15.
- Wang, E., Zhang, Y., Luo, J., Chiew, F.H.S., Wang, Q.J., 2011. Monthly and seasonal streamflow forecasts using rainfall-runoff modeling and historical weather data. *Water Resour. Res.* 47 (W05516).
- Wetterhall, F., Giuseppe, F.D., 2018. The benefit of seamless forecasts for hydrological predictions over Europe. *Hydrol. Earth Syst. Sci.* 22 (6), 3409–3420.
- Wilhite, D.A. (Ed.), 2000. Droughts as a natural hazard: concepts and definitions. In: *DROUGHT, A Global Assessment, Vol I and II*, Routledge Hazards and Disasters Series, Routledge, London.
- WMO, 2019. WMO Statement on the State of the Global Climate in 2018. WMO-No. 1233, World Meteorological Organization, Geneva.
- Wood, A.W., Lettenmaier, D.P., 2008. An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.* 35 (L14401).
- Wright, S., 1921. Correlation and causation. *J. Agric. Res.* 20 (7), 557–585.
- Yevjevich, V.M., 1967. An objective approach to definitions and investigations of continental hydrologic droughts. Hydrology papers (Colorado State University); No. 23.
- Yuan, X., Wood, E.F., Chaney, N.W., Sheffield, J., Kam, J., Liang, M., Guan, K., 2013a. Probabilistic seasonal forecasting of African drought by dynamical models. *J.*

- Hydrometeorol. 14, 1706–1720.
- Yuan, X., Wood, E.F., Luo, L., Pan, M., 2011. A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophys. Res. Lett.* 38 (13).
- Yuan, X., Wood, E.F., Roundy, J.K., Pan, M., 2013b. CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States. *J. Clim.* 26 (13), 4828–4847.
- Zajac, Z., Zambrano-Bigiarini, M., Salamon, P., Burek, P., Gentile, A., Bianchi, A., 2013. Calibration of the LISFLOOD hydrological model for Europe, Calibration Round 2013. JRC technical report, European Commission. Joint Research Centre, Ispra Italy.
- Zappa, M., Bernhard, L., Spirig, C., Pfander, M., Stahl, K., Kruse, S., Seidl, I., Stähli, M., 2014. A prototype platform for water resources monitoring and early recognition of critical droughts in Switzerland. In: *Evolving Water Resources Systems: Understanding, Predicting and Managing Water-Society Interactions*, Bologna, Italy, IAHS Publ. vol. 364, pp. 492–498.
- Zelenhasić, E., Salvai, A., 1987. A method of streamflow drought analysis. *Water Resour. Res.* 23 (1), 156–168.