



OPEN

Catchment memory explains hydrological drought forecast performance

Samuel Jonson Sutanto^{1,2,3,4}✉ & Henny A. J. Van Lanen^{1,4}

Hydrological drought forecasts outperform meteorological ones, which is anticipated coming from catchment memory. Yet, the importance of catchment memory in explaining hydrological drought forecast performance has not been studied. Here, we use the Baseflow Index (BFI) and the groundwater Recession Coefficient (gRC), which through the streamflow, give information on the catchment memory. Performance of streamflow drought forecasts was evaluated using the Brier Score (BS) for rivers across Europe. We found that BS is negatively correlated with BFI, meaning that rivers with high BFI (large memory) yield better drought prediction (low BS). A significant positive correlation between gRC and BS demonstrates that catchments slowly releasing groundwater to streams (low gRC), i.e. large memory, generates higher drought forecast performance. The higher performance of hydrological drought forecasts in catchments with relatively large memory (high BFI and low gRC) implies that Drought Early Warning Systems have more potential to be implemented there and will appear to be more useful.

Skillful Drought Early Warning Systems (DEWSs) to predict drought a few months in advance are of utmost importance to reduce the impacts of the drought hazard^{1,2}. Previous studies on meteorological drought forecasts using the Standardized Precipitation Index (SPI)³ show that (1) drought can be sufficiently predicted up to 1–3 months ahead, depending on the accumulation period x (e.g., SPI- x , with $x = 1, 3, 6,$ or 12 months) and (2) the highest drought forecast skill is obtained for SPI with longer accumulation periods^{4–7}. The monthly temporal resolution used in the seasonal drought forecasts, e.g. SPI that measures drought by the deviation of monthly precipitation values from long-term median precipitation for each month^{8,9}, is the main cause that escalates the weather forecast skill¹⁰, which commonly (in terms of below normal precipitation) has skill only for a few days up to a few weeks^{11,12}. Drought forecasts, however, do not require detailed day-to-day evolution, as demanded in conventional weather forecasts. Instead, drought predictions, such as SPI, need estimates of monthly total precipitation, that provide information on how likely the coming months will be wetter or drier than the median¹³. For example, over- or underestimation of precipitation by 10 mm on a daily basis is crucial for weather forecasting in general and flood forecasting in particular. For drought forecasting, missing or having more precipitation of 10 mm on a certain day in a month is less significant as long as the predicted monthly median of precipitation derived from the ensembles is fairly reliable, i.e. it can be counteracted by more or less precipitation on other days in the month. In addition to the monthly accumulation period, the blending of preceding monthly-observed data with forecasts to calculate drought indices with higher accumulation periods (e.g. SPI- x , $x > 1$) for short lead times (LT, e.g. LT = 1-month) mainly explains meteorological drought forecast skill^{4,7,14–16}.

The skill of hydrological drought forecasts, on the other hands, is even 2–3 months higher than the meteorological ones^{7,16}. Sutanto et al.⁷ show that hydrological drought forecasts with 1-month accumulation period identified using the Standardized Runoff Index (SRI-1)¹⁷ and Standardized Groundwater Index (SGI-1)¹⁸ and LT = 1 have perfect forecasts up to 71.5% and 73.2% of the pan-European area, respectively, which is higher than meteorological drought scores (53.7% for SPI). Another study by Van Hateren et al.¹⁶ supports this finding. They show that streamflow drought forecasts outperform climatology up to LT = 4-month while the skill of meteorological forecasts (SPI-1 and SPI-3) is higher than climatology only for LT = 1-month. In addition to

¹Hydrology and Quantitative Water Management Group, Environmental Sciences Department, Wageningen University and Research, Wageningen, The Netherlands. ²Water Systems and Global Change Group, Environmental Sciences Department, Wageningen University and Research, Droevendaalsesteeg 3, 6708 PB Wageningen, The Netherlands. ³Institute for Marine and Atmospheric research Utrecht, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands. ⁴These authors contributed equally: Samuel Jonson Sutanto and Henny A. J. Van Lanen. ✉email: samuel.sutanto@wur.nl

the accumulation period and the blending of observed and forecasted data (see above), the high skill in hydrological drought forecasts is anticipated coming from the land surface water storage/memory (e.g. lakes, soils, groundwater) that pools, attenuates, and lengthens the effect of the driving forces (i.e. precipitation)¹⁹. Recent studies by Pechlivanidis et al.²⁰ and Girons et al.²¹ on seasonal streamflow forecasts (not streamflow drought) also conclude that the predictability is higher in the slowly responding basin, associated with large catchment memory. The effect of catchment memory on hydrological drought forecasting performance, however, has never been revealed yet.

Here, we have conducted a pioneering study that investigates the importance of catchment memory on the forecast performance of streamflow drought across Europe. Please note that 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 flow time series need to be converted into a time series of drought events by applying drought identification approaches, e.g. the threshold approach or the standardized approach. In this study, we identify streamflow drought using the Standardized Streamflow Index²² for different subregions in Europe²³. The observed and forecasted streamflow droughts at major European rivers were derived from the streamflow data obtained from the European Flood Alert System (EFAS) driven by observed and forecasted weather data^{24,25}. Catchment memory was derived from: (1) the Baseflow Index (BFI)²⁶ that explains the capacity of a catchment to store and release water from catchment sources of water, and (2) the groundwater Recession Coefficient (gRC) that describes storage and releases of water from subsurface water. The forecast performance was evaluated using the Brier Score (BS) approach²⁷ and correlated with BFI and gRC (“Methods”). Our results show that the highest forecast performance is found in rivers with higher BFI and lower gRC values, associated with higher memory, which explain the importance of considering catchment memory in improving the performance of hydrological drought forecasts.

Results

The Baseflow Index (BFI) for major European rivers. The BFI is commonly used to indicate the portion of the flow, i.e. baseflow, that comes from groundwater storage or other delayed sources, which is derived from recessions in streamflow^{26,28,29} (“Methods”). Rivers with a flashy flow regime, associated with small catchment memory due to, e.g. less permeable soil layers, no or limited aquifer storage, force rainfall to flow quickly to the stream, have typically low BFI values³⁰. Vice versa, rivers in catchments with large memory yield high BFI values. The rivers located in mountain ranges or snow-dominated regions generate high BFI (Supplementary Fig. 1) that may come from water stored and released from snow and ice and the presence of lakes and wetlands^{20,31}. The BFI of rivers in other European regions ranges from 0.1 to 0.8. In these regions, BFI is predominantly controlled by other catchment properties. For example, a big difference in BFI values is seen for some rivers flowing in France and the UK both located in Western Europe. Rivers flowing in France have BFI values between 0.3 in the upstream and 0.7 in the downstream, while rivers in the UK have lower BFI values range from 0.1 to 0.6 (Supplementary Fig. 1). The aquifer types in these countries explain the different BFI values. North and west UK have little catchment storage compared to France, where major aquifer system occur in some areas^{32,33}.

Supplementary Figure 2b–g show examples of BFI varying from 0.78 to 0.19 for six rivers located in different European subregions²³, with at least one river represented in each region. The Danube and the Rhine Rivers have the highest BFI values (0.78 and 0.74, respectively). These rivers are identified as baseflow dominated rivers, which have long recessions and small response to precipitation²⁰. The annual climatic water balance (precipitation minus evapotranspiration) in these rivers is around 200 mm with a runoff coefficient of 0.45 (Supplementary Table 1). The Guadiana and the Tanaro Rivers located in the southern European region yield lower baseflow than the Danube and Rhine located in central Europe with BFI values of 0.66 and 0.51, respectively. These rivers, especially Guadiana, have lower precipitation and higher evapotranspiration than the Rhine and Danube Rivers. Although the Tanaro River has similar average baseflow to the Guadiana, it is flashier in the high flow range, resulting in a lower BFI (Supplementary Fig. 2d,e). The Tweed and Kolbacksan Rivers experience the lowest BFI values due to low baseflow and quick response to precipitation. These rivers situate in regions that are characterized by limited aquifer storage^{34,35} and hence have low catchment memory and fast catchment response, with a runoff coefficient up to 0.62 for the Tweed River (Supplementary Table 1).

Drought forecast performance for major European rivers and connection with BFI. Using the re-forecast and proxy observed data (SFO, see “Methods”), the drought forecast performance for major European rivers was calculated using the Standardized Streamflow Index with accumulation periods of 1 month (SSI-1)²² and of 3 months (SSI-3) with lead times (LTs) of 1 and 3-month (Fig. 1). The performance of drought forecasts was determined by using the Brier Score (BS), which indicates a higher performance when BS is low²⁷ (“Methods”). The highest performance is obtained in some rivers located in Finland, Hungary, Poland, north Germany, and east Romania (Fig. 1a,b,d). Rivers, situated in the mountainous regions, such as in the Alps and east Norway, have the lowest performance, although these rivers have fairly high BFI (Supplementary Fig. 1). The low performance of drought forecasts in these regions relates to the timing of snow occurrences. The mismatch in the prediction of too early or late snow accumulation or melt is caused by the mismatch in the prediction of temperature and precipitation, resulting in a bias in the temperature-dependent simulation of snowfall and snowmelt^{7,36,37}. For example, the prediction of too high temperature leads to early snowmelt. On the other hand, the prediction of too low temperature will result in more snow accumulation, which otherwise would have resulted in rainfall infiltrating into the soil. Moreover, the meteorological forecast data used as forcing in the EFAS are not bias corrected (ECMWF, personal communication). SSI-3 with 1-month LT attains the high-

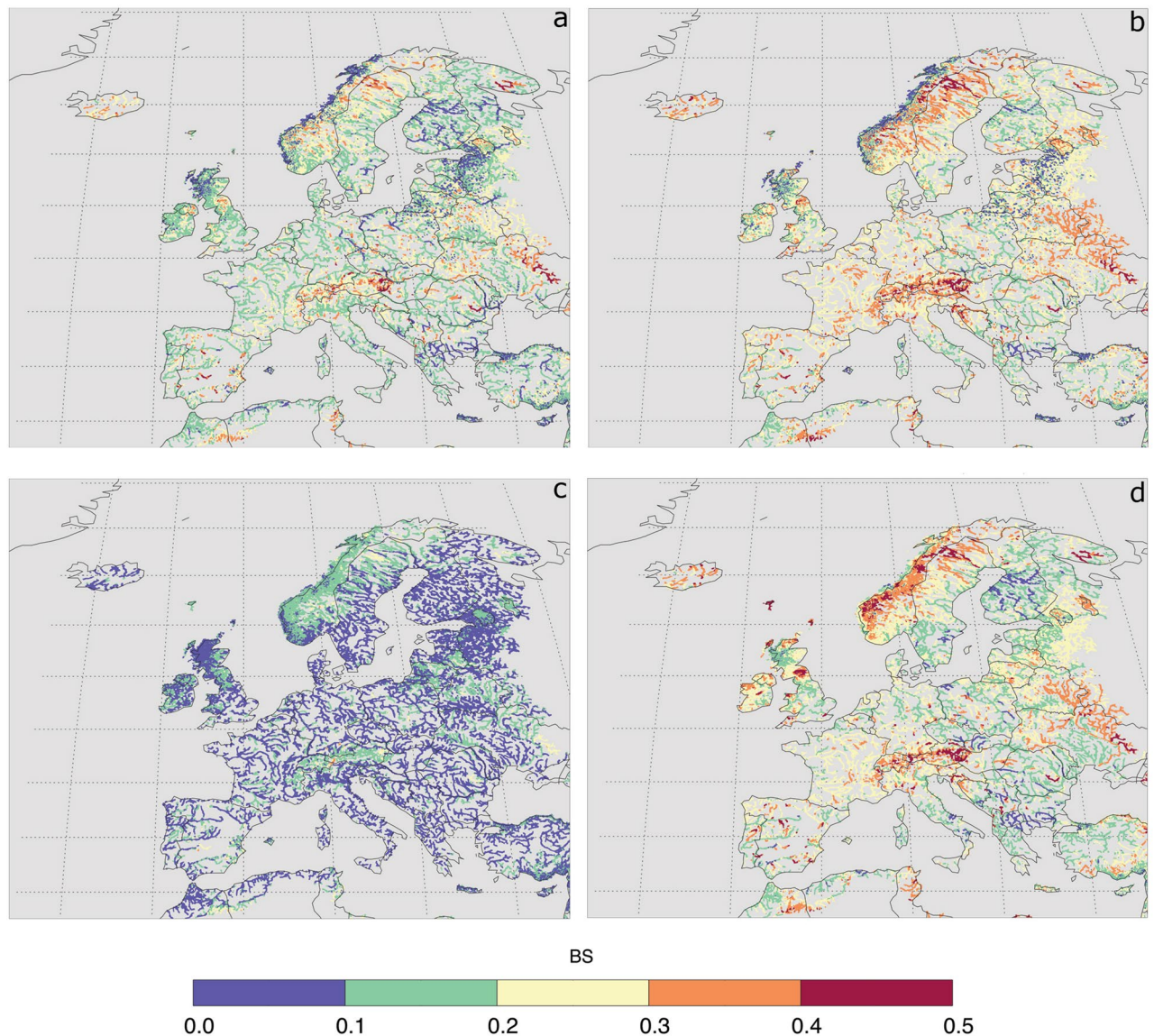


Figure 1. Drought forecast performance denoted by the Brier Score (BS) for major European rivers: (a) BS for SSI-1 with Lead Time (LT) = 1-month, (b) BS for SSI-1 with LT = 3-month, (c) BS for SSI-3 with LT = 1-month, and (d) BS for SSI-3 with LT = 3-month. Note all figures were plotted using the Interactive Data Language (IDL) software version 8.0 (<https://www.l3harrisgeospatial.com/Software-Technology/IDL>).

est performance due to the blending of 2 months observed (SFO) data with one month forecast (Fig. 1c). The performance deteriorates with an increased of LTs (Fig. 1a vs. b and c vs. d).

A summary of SSI-1 and SSI-3 drought forecast performance for major European rivers and for all LTs is presented in Fig. 2. A sharp decrease in forecast performance (increase of BS) is clearly seen from LT = 1 to LT = 3. For LTs beyond 3-month, the forecast performance does not significantly change, meaning that these drought forecasts do not provide added value, as it was also found in previous studies^{6,7,16}. Both SSI-1 and SSI-3 with LT > 3 show similar forecast performance (BS = 0.24–0.26 for ensemble median). As expected, the highest performance is obtained for SSI-3 due to the blending of observed data with forecasts (Fig. 2b). Furthermore, the performance of drought forecasts varies over seasons, depending on the accumulation period and LT. The highest performance derived from the median ensemble is obtained for SSI-3 with LT = 1 in summer, followed by spring, autumn, and winter seasons (BS < 0.07, Supplementary Fig. 3). For SSI-1 with LT = 1, the highest performance is obtained in summer followed by autumn, winter, and spring seasons (BS < 0.15). For the longer lead time (LT = 3), the highest/lowest performance is obtained in winter/autumn and spring/autumn for SSI-1 and SSI-3, respectively. In general, the drought forecasts perform well in summer for LT = 1 and in winter and spring for LT = 3.

The hydrological persistence of a river with high BFI causes high streamflow forecast performance^{20,21,38}. A high BFI means that there is a rather high probability that streamflow in the next days will not be too much different from today's streamflow. Likely, this also applies to the dry anomalies in streamflow (i.e. drought), that is, if the river is below normal today, then it is also expected to be below normal in the days after. However, the

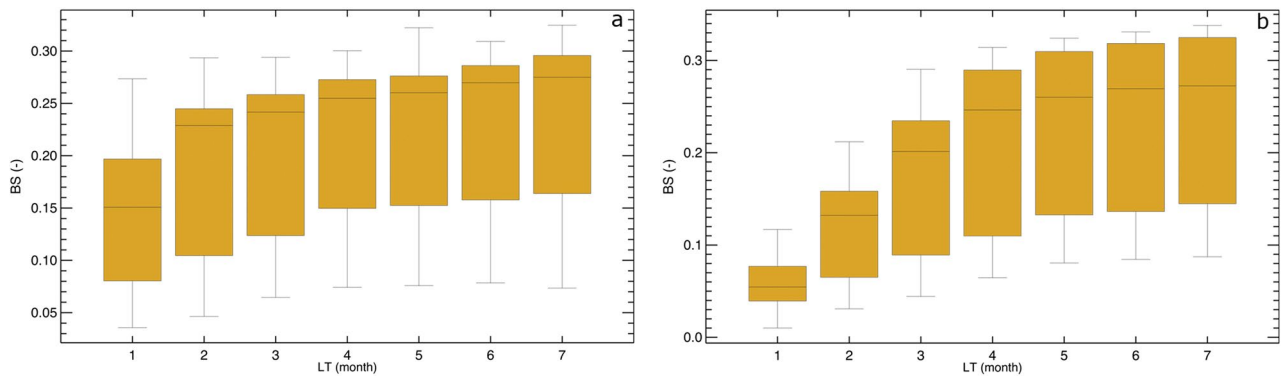


Figure 2. Box and whisker plot showing drought forecast performance denoted by the Brier Score (BS) for major European rivers: **(a)** BS for SSI-1 with LTs from 1- to 7-month and **(b)** BS for SSI-3 with LTs from 1- to 7-month. Lower box shows 25 percentile, middle line shows median, and upper box shows 75 percentile. The whiskers show the 10 and 90 percentiles.

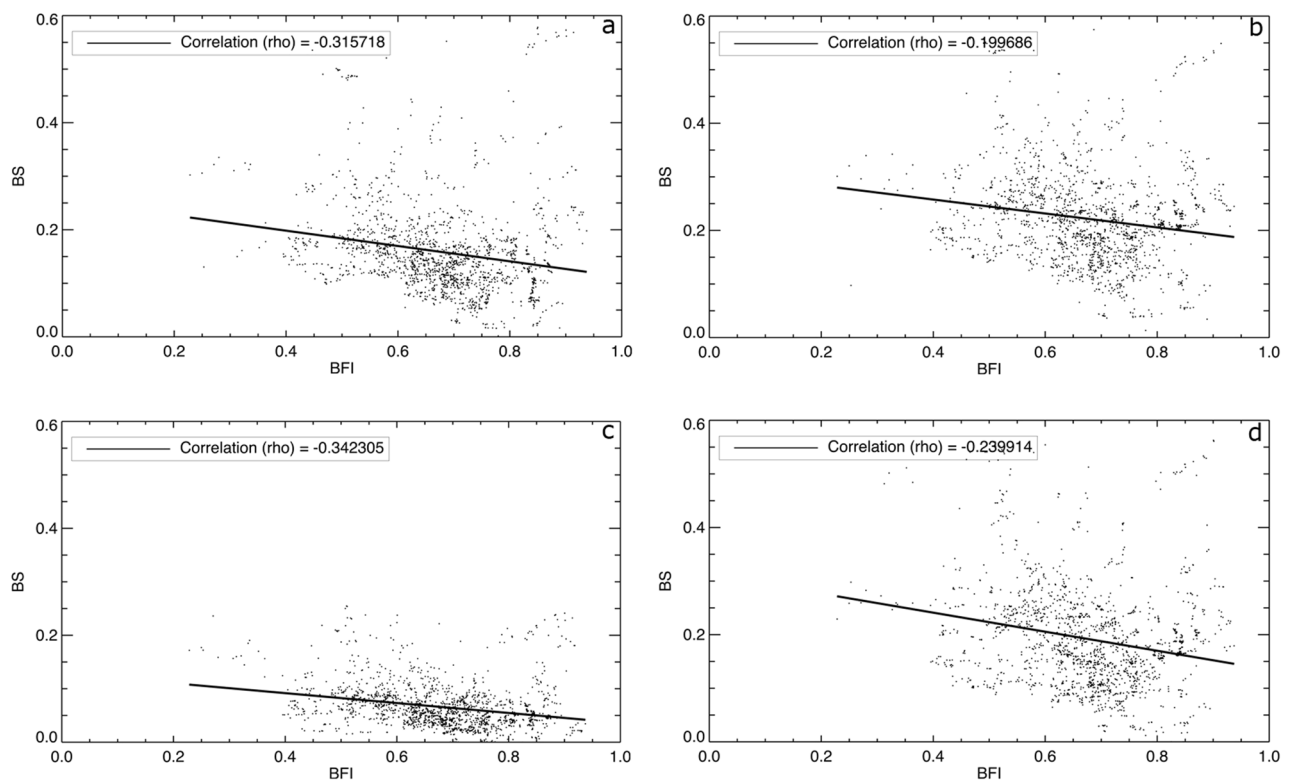


Figure 3. Correlation between BFI and BS for southern European river grid cells: **(a)** for SSI-1 LT = 1, **(b)** for SSI-1 LT = 3, **(c)** for SSI-3 LT = 1, and **(d)** for SSI-3 LT = 3. See Supplementary Table 2 for average BFI and BS values in each European subregion and Europe as a whole.

persistence in streamflow drought is observed³⁹, but not well understood yet. To prove our hypothesis that the higher performance of drought forecasts is obtained for rivers with higher BFI, we plotted the correlation between BFI and BS for all major river grid cells ($n = 10,106$) in southern Europe (Fig. 3 and Supplementary Table 2 for other European regions). Overall, significant negative correlations ($p < 0.05$) are obtained with higher Spearman's correlation values (Rho) for LT = 1 than LT = 3, as expected (Fig. 3a vs. b and c vs. d). Lower forecast performance for LT = 3 is illustrated by more spread in BS and lower Rho, particularly for SSI-1. Compared to SSI-1 LT = 1, the BS values for SSI-3 LT = 1 are less scattered, associated with higher forecast performance for SSI-3 than SSI-1. The blending of observed data for drought standardized indices (e.g., SPI, SSI, and SRI) with an accumulation period of more than one month enhances the forecast performance, in particular for meteorological drought forecasts^{4-7,15,16}. On average, the highest BFI values (0.71) and lowest BS (0.16) are found in rivers located in Central and South Europe, respectively (Supplementary Table 2), which indicate a potential for the development of drought early warning systems for rivers located in these regions. Conversely, rivers in West Europe have the

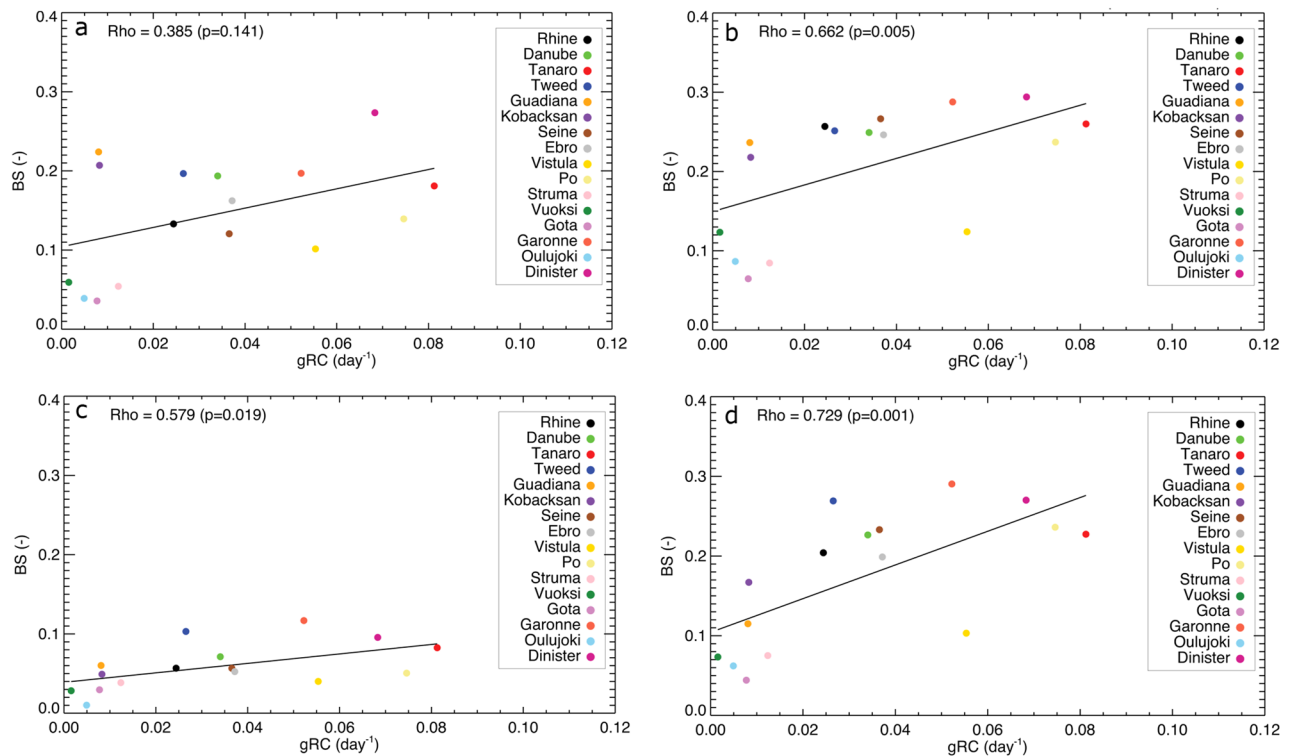


Figure 4. Correlation between BS and gRC for the 16 selected river catchments: (a) for SSI-1 LT = 1, (b) for SSI-1 LT = 3, (c) for SSI-3 LT = 1, and (d) for SSI-3 LT = 3.

lowest BFI (0.61), but medium forecast performance (BS = 0.19). The lowest forecast performance is found in rivers located in North Europe (BS = 0.21).

Some past studies correlated the streamflow forecast skill with BFI (e.g. Harrigan et al.³⁸ and Girons et al.²¹). They found that the forecast skill is high in catchments with high BFI. Our study on streamflow drought forecast performance supports their findings but with lower correlation coefficients. The somewhat lower correlation obtained in our study might be caused by that, in streamflow forecasting, the skill assessment is performed for all flow conditions, which include normal, high, and low flows. In streamflow drought forecasting, however, we evaluate the forecast performance only for extreme below normal events, i.e. anomalies called drought. The focus on only dry extremes is hypothesized to reduce forecast performance. Moreover, there are more indicators that can describe the catchment memory, such as hydrogeology, the relative area of lakes/wetland/bogs, topography, land use, and soil types that might affect correlation strength. Although the effects of these indicators are already to some extent embedded in the BFI, the use of more indicators than BFI may improve the correlation coefficient.

The impact of groundwater memory. In previous sections, the BFI is used as the first metric to investigate the impact of catchment memory. In this section, we use the groundwater Recession Coefficient (gRC, Supplementary Fig. 4a) that more specifically signifies groundwater, as a second metric to further study the impact of catchment memory on drought forecast performance⁴⁰. The performance was investigated by evaluating the relationship between groundwater gRC and BS (“Methods”). Here we consider 16 river catchments that are located in different European subregions (Supplementary Fig. 2a). These rivers have a broad range of BS and gRC values to derive the correlation. Unlike BFI that can be derived for all individual river cells, the gRC is investigated for river grid cells at the outlet, which is obtained by spatially averaging gRCs of all land cells upstream of the outlet (“Methods”) that contribute to groundwater flow (catchment, Supplementary Fig. 4b). The time-consuming spatial averaging is the main reason that estimating the correlation between gRC and BS and between gRC and BFI is not feasible for all river grid cells across Europe (“Methods”). Hence, we selected several catchments that reflect a broad range of physio-geographic conditions across Europe. We infer that slowly responding catchments (large groundwater memory) denoted by low gRC will generate high drought forecast performance as it was proved for streamflow forecast.

Figure 4 shows that catchment with lower gRC associated with a slow release of groundwater from the system produces higher drought forecast performance. Strong positive correlations between gRC and BS are seen for both SSI-1 and SSI-3 and for all LTs (Rho > 0.5, except for SSI-1 LT = 1). The correlation, however, becomes more robust (increase up to > 0.7) for SSI-3 (Fig. 4c,d) than SSI-1 (Fig. 4a,b), because two months preceded observation data are included. Rivers located in northern Europe, such as Vuoksi, Göta, and Oulujoki, have low gRC and high forecast performance. This might be due to the presence of lakes in their catchments that attenuate streamflow peaks owing to storage and evaporation⁴¹. Rivers located in the mid latitudes (Danube, Vistula, Rhine, Tweed, Seine, Dinister, and Garonne) in general have similar forecast performance and gRC, with higher performance relatively found in the rivers in east Europe. Mediterranean rivers (Struma, Guadiana, Ebro, Po, and Tanaro) have

a wide range of gRCs from 0.01 to 0.08 and BS from 0.06 to 0.22 (Fig. 4a). The Guadiana River in Spain and the Struma river in Greece that have slow catchment response (gRC = 0.008 and 0.012, respectively) yield different prediction performance (BS = 0.22 and 0.05, respectively, Fig. 4a). Anthropogenic activities to cope with water scarcity in the Mediterranean region, such as the rapid development of reservoirs in the past decades, modified the flow regime⁴². A clear example is seen in the Guadiana River, which has 39 man-made reservoirs located along the catchment⁴³. It seems that the reservoir regulation and management highly influence the flow regime and associated hydrological drought⁴⁴ in this river and thus reduce the forecast performance.

We also tested our hypothesis that catchment with lower gRC generates higher drought forecast performance using a dichotomous forecast method for comparison. The forecast performance is identified based on forecast accuracy and success ratio (SR), which is 1 for perfect forecast (see Supplementary Note). Supplementary Figure 5 shows negative correlations between gRC and forecast accuracy, as well as forecast SR ($Rho < -0.5$), which support our hypothesis using BS. Note that the correlation sign of accuracy and SR is opposite to the correlation using BS because higher forecast performance is reflected by higher accuracy and SR, unlike BS that shows lower performance for higher BS. On average, catchments that have gRC values lower than 0.04 (high memory) have higher prediction performance, with more than 80% of all forecasts that were correct (accuracy > 0.8, Supplementary Fig. 5a) and 70% of the forecasted drought events were actually observed (SR > 0.7, Supplementary Fig. 5c).

The relation between gRC and BFI is also provided in Supplementary Fig. 6. The correlation between gRC and BFI is negative. This is not surprising because catchments with high gRC are associated with a quick release of groundwater and fast responding streamflow (low BFI). The correlation ($Rho = -0.376$), however, is rather weak and not significant at $p < 0.05$. The weak correlation found in our study might be caused by the small number of river catchments (16 catchments) to confirm the dependency. Using 314 catchments in the UK, Harrigan et al.³⁸ found a strong correlation between BFI and catchment storages (soil moisture and groundwater), with a correlation coefficient up to 0.87. Although the BFI in many applications is used as a measure of streamflow that originates from the subsurface storage (e.g., groundwater), it also measures catchment characteristics, such as lake area, wetland area, land cover, catchment slope, and catchment shape and area, which result in different streamflow responses to climate i.e., precipitation and evaporation^{45,46}. Our study proved that groundwater memory (fast to slow release of groundwater), denoted by gRC, has a strong impact on drought forecast performance, in addition to other performance provided by other sources of water in the catchment, which are incorporated in the BFI.

Discussion and conclusions

Our study shows that catchment memory is a key component in the predictability of hydrological drought, since it strongly influences hydrological drought characteristics^{19,40,45}. Catchment memory pools, attenuates, and lengthens anomalies in driving forces i.e. precipitation.

The use of higher accumulation periods in hydrometeorological drought forecasts improves forecast performance. This is also shown in our study, where the SSI with 3 months accumulation period and lead time of 1-month (SSI-3, LT = 1) yields higher forecast performance than the SSI-1 for the same lead time, as shown in our study. However, multiple studies have indicated that hydrological drought forecasts that are not blended with observed data (e.g. SSI-1 and SRI-1) outperform meteorological drought forecasts (e.g. SPI-1) up to a few months ahead due to the non-linear physical processes in the catchment system that pool, attenuate, and lengthen the effect of the driving forces i.e. precipitation^{7,16}. The performance is even higher for catchment that has large memory, as it is shown in our paper. The catchment memory as the key component increases the predictability of the hydrological drought.

Our study on streamflow drought forecasting across Europe confirms that hydrological drought forecast performance is higher in a river draining in a slowly responding catchment, associated with larger memory or storage in groundwater (high BFI, low gRC). We also highlight the importance of other water bodies that store water, such as lakes that reduce temporal streamflow variability in describing the high performance of streamflow drought forecasts. The significant negative correlation between BFI and BS proves that high forecast performance can be obtained for rivers with high BFIs. This conclusion is further supported by a correlation analysis of gRC and BS. A positive correlation between gRC and BS for 16 river catchments across Europe provides clear evidence ($Rho > 0.4$) that the rate of release of groundwater storage to the river is an important factor in the predictive performance. We further explored the effect of catchment size on forecast performance by dividing the catchments into small and medium-sized catchments (7 catchments) and large catchments (9 catchments). This investigation confirms the conclusion that high streamflow drought forecast performance (low BS) comes hand in hand with rivers having high catchment memory (high BFI and low gRC), irrespective of catchment size (Supplementary Fig. 7).

The blending of observed data with forecasts, e.g. SSI-3 for short lead times (LT < 2 months), further improves the correlation between gRC and BS ($Rho > 0.6$). One should note that the use of longer accumulation periods ($x > 1$ month) in hydrological drought forecasting should be done with caution. Hydrological variables, e.g. streamflow already comprise some catchment memory aspects (delayed flow from groundwater). Hence some studies recommend to use only 1 month accumulation period for hydrological drought analysis^{18,47}. On the other hand, we need to realize that anomalies in the accumulated flow over a longer period (e.g. SSI-3) have relevance for some purposes, such as the management of surface water reservoirs.

Among different seasons, streamflow drought forecasts perform well in summer for LT = 1 and in winter/spring for LT = 3. Our finding that streamflow drought forecast has high performance in summer supports the study by Sutanto et al.⁷ on hydrological drought forecasts that the skill is higher in summer. However, some studies on streamflow forecast (note: not on streamflow drought) contradict with our findings (e.g., Crochemore et al.⁴⁸; Bell et al.⁴⁹; Arnal et al.⁵⁰). They concluded that the skill of streamflow prediction is higher in winter than

in summer. We argue that the results obtained from streamflow drought forecasts may differ from streamflow forecasts. In streamflow drought forecasts, the skill of forecast is determined by the ability to predict low streamflow (here SSI < -0.5), whereas the skill of streamflow forecasts is derived from the full flow spectrum from low to high flow. The moderate skill of winter hydrological drought forecasts might be linked to large-scale teleconnection patterns that affect European climate in winter. Some studies show that the North Atlantic Oscillation (NAO), the occurrence of blocking, and the North Atlantic jet stream are conjectured as strong predictors in winter meteorological forecasts in Europe^{33,51,52}. The high skill in prediction of winter precipitation and temperature is then translated to better prediction of streamflow and subsequently streamflow drought. However, the skill of meteorological forecasts is limited for long-term predictions (> 15 days) due to the chaotic nature of the atmosphere^{53,54}.

In this study, we used the large-scale hydrological model, LISFLOOD, with a spatial resolution of 5 km by 5 km. This implies that results were derived purely based on the large-scale model simulation. A model-based assessment may not fully replicate the observational-based assessment, especially for the extreme cases, such as drought, because of uncertainties of different modeling components; in particular, the sub-surface characteristics that are not specifically captured in large-scale continental-based modeling assessment²¹. The errors will be higher in the catchments that had no streamflow observations for model calibration and validation. Nevertheless, the LISFLOOD model was calibrated using 717 stations across Europe⁵⁵, which means that a reasonable number of catchments was covered. Although the uncertainties in modeling and forecasting streamflow drought at a continental scale, i.e. Europe, are expected to be higher than at catchment scale, the use of a large-scale model provides indicative results of drought forecasts for all major European rivers. This will fill the gap in providing seasonal drought forecasts in regions that have no DEWS, especially in regions with slowly responding hydrological systems that have high catchment memory.

Given drought hazard and its impacts are projected to increase under a future warmer climate in multiple regions across the globe^{56–58}, the development of DEWSs that also provides hydrological drought forecasts is highly recommended because they have higher skill than meteorological drought forecasts and they better serve different users managing water resources⁵⁹. Our results suggest that the performance of the hydrological drought forecast module of a DEWS would be higher for rivers draining catchments in East Europe, followed by rivers in South, Central, and North Europe (Supplementary Table 2). On the contrary, rivers situated in West Europe have the lowest BFI and highest BS due to the relatively fast response from the groundwater storage (high gRC, Supplementary Table 3) that explains the low predictability of hydrological drought.

Methods

Data. Proxy of daily observed streamflow data from 1990 to 2018 was simulated using the LISFLOOD hydrological model fed by gridded meteorological observation data (EFAS, 5 by 5 km) collected from > 5000 ground observations (hereafter called Simulation Forced by Observations, SFO or proxy observed). The same model was run with the re-forecast (known as hindcast) meteorological data obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) SEAS5 to simulate the re-forecast streamflow data from 2002 to 2016 (180 months). This period covers major drought events in Europe, e.g. west and central Europe in 2003, south Europe from 2006–2008, east Europe and Russia in 2011, and central and east Europe in 2015. The seasonal re-forecasts of streamflow have a lead-time of 215 days (circa 7 months) and consist of 25 ensemble members. The streamflow data, both SFO and re-forecast, were provided by the ECMWF as part of the European Flood Awareness System (EFAS system version 3)^{24,25,60}. In LISFLOOD, the potential evapotranspiration was calculated using the Penman–Monteith equation through the offline LISVAP pre-processor⁶¹. The soil moisture flux out of the subsoil and between the upper and lower soil layer are defined based on the Darcy's law. The groundwater storage and flow are modeled using two parallel linear reservoirs following to the approached used in the HBV-96 model. The groundwater upper zone represents a quick runoff component, including fast groundwater and subsurface flow through macro-pores in the soil. Whereas, the lower zone represents the slow groundwater component that generates the baseflow. Surface runoff and flow from the two groundwater zones (5 × 5 km grid cells) are routed to the river network. A kinematic wave approach was used for routing the water movement in the river network (~ 10,106 river grid cells). The LISFLOOD model used in our study incorporates water abstraction for irrigation, water demands for the livestock, energy production and cooling, and manufacturing industry, and 1454 reservoirs across Europe³⁵. The LISFLOOD model was calibrated by Arnal et al.⁵⁵ using over 700 streamflow observation data across Europe, which results in a Kling–Gupta Efficiency (KGE) of > 0.75 for more than 42% stations, KGE between 0.5 and 0.75 for 33% stations, and only 25% of stations show KGE below 0.54. LISFLOOD proved to be able to adequately simulate streamflow and hence the modelled results can mirror observed spatial and temporal streamflow patterns across Europe. Detailed information on SFO and re-forecasts data taken from EFAS are provided by Sutanto et al.¹ and Arnal et al.^{50,55}. All data and figures were processed using the Interactive Data Language (IDL) software version 8.0.

Standardized Streamflow Index (SSI). The Standardized Streamflow Index (SSI) identifies the severity of the drought in the river^{22,62}. The SSI indicates the degree of dryness by providing the streamflow deviation (anomaly) from the long-term median. The gamma distribution was applied to monthly streamflow data taken from the 1990–2018 records, which then is transformed into 12 normal distributions for each accumulation period and every month of the year. In this study, SSI-x was calculated for the accumulation periods of 1 and 3 months (SSI-1 and SSI-3, respectively). Then, these distribution parameters were used as a basis to identify drought events in the observed (1990–2018) and re-forecast time series (2002–2016). The gamma distribution was applied to all major European rivers since it is applicable for streamflow forecasting as shown in a previous study⁶³. We would like to note that the gamma distribution is typically used to model highly skewed time series

and it is not fully able to fit data at the lower tail (low streamflow in this study), leading to underestimation of drought⁶⁴. Furthermore, the use of the gamma distribution in rivers that encounter zero flow conditions, e.g. in semi-arid and arid regions (not in this study), also needs attention because the gamma distribution cannot model zero streamflow. Other distributions for streamflow, such as log normal, Generalized Extreme Value (GEV), Tweedie, and nonparametric method can be considered^{47,64}. However, studies done by Vicente-Serrano et al.⁶² and Tjrdeman et al.⁴⁷ show that none of the single distributions fits all streamflow time series across Europe and will fit well with all monthly streamflow data in all river grid cells, e.g., sample properties of streamflow in January likely will differ from those in August. Thus, we simply used the widely selected gamma distribution in our analysis.

The re-forecasted SSI-3 with a lead-time (LT) of k-month (k from 1 to 7-month) combines for $LT \leq 2$ month re-forecast data and observation data from preceding months (SFO streamflow), and uses for $LT > 2$ month only forecast data. For example, the re-forecasted SSI-3 for January 2003 with 1-month LT (forecast issued on 2nd January 2003) was estimated by combining the observed data from November and December 2002 with January 2003 re-forecast data ($LT = 1$). To re-forecast the SSI-3 for February 2003 with $LT = 2$ (forecast issued on 2nd January 2003), we combined the observed data of December 2002 with re-forecast data for January and February 2003. Thus, the re-forecasted SSI-3 with $LT = 3$ consists of only re-forecast data. Using SSI, a drought event was classified into four classes, which are: (i) mild drought: $0 > SSI \geq -1$, (ii) moderate drought: $-1 > SSI \geq -1.5$, (iii) severe drought: $-1.5 > SSI \geq -2$, and (iv) extreme drought: $SSI < -2$.

Baseflow Index (BFI). The BFI is considered a measure of streamflow that originates from stored sources of water in a catchment e.g. groundwater, lakes, bogs, glacier, and snow. It can also be used for indexing the effect of geology, land use, and catchment area on streamflow, which results in different streamflow responses^{45,46}. Thus, the effects of geology, land use, and other catchment characteristics are implicitly considered in the BFI. The BFI has been used in many studies to indicate catchment memory, i.e. storage and release properties of catchments, incl. groundwater^{30,31,38,45,46}. These studies show that catchments with high BFI values have large (groundwater) storage and slow release of water from other delayed sources (e.g. water stored in lakes, bogs, glaciers, and snow). A slowly responding catchment is associated with a high memory catchment. The opposite holds for a catchment with low BFI, which is more responsive to rainfall. The BFI is obtained by applying hydrograph separation and is calculated as:

$$BFI = \frac{V_{base}}{V_{total}}, \quad (1)$$

where V_{base} is the volume of water beneath the baseflow part of the hydrograph between the first and last date of interest. V_{total} is the volume of water beneath the recorded hydrograph. Hisdal et al.³⁰ provides a detailed procedure for BFI calculation. Examples of streamflow hydrograph and its baseflow for several selected rivers are presented in Supplementary Fig. 2b-g.

Groundwater Recession Coefficient (gRC). To classify the responsiveness of river catchments (rate of release of groundwater storage to streams⁶⁵, e.g. slowly or quickly declining groundwater levels), we used the groundwater Recession Coefficient (gRC), which was simulated by the PCR-GLOBWB hydrology and water resources model coupled with a MODFLOW groundwater module (Supplementary Fig. 4a)⁶⁶. We used the gRC from PCR-GLOBWB because LISFLOOD only simulates transient groundwater storage and not groundwater levels. Hence, we could not derive the gRC, i.e. the decline rate of the groundwater levels in dry periods. Note that we used the gridded groundwater Recession Coefficient (gRC) that was obtained from groundwater hydrographs rather than the more frequently used Recession Coefficient (RC), which is derived from recession limbs of the streamflow hydrograph³⁰. Low gRC indicates slow release from the groundwater storage to the river, meaning that the catchment has a high memory (large storage) and is categorized as a slowly responding catchment. The gRC values for the 16 selected river catchments were roughly estimated by averaging the gRC values from all grid cells upstream of the river grid cell that represents the outlet (see Supplementary Fig. 4b). The PCR-GLOBWB model used the aquifer properties taken from the International Hydrogeological Map of Europe (IHME)⁶⁶.

We would like to note that the gRC in a river grid cell represents only the groundwater response in that particular river grid cell. There is no influence of gRC values from the upstream cells in a catchment, which is required for this study to characterize memory in the whole catchment. The gRC in a grid cell is only determined by neighboring cells as inflow and outflow to and from the grid cell, respectively. Hence, to obtain the overall gRC of a catchment that will be assigned to the most-downstream river grid cell (catchment outlet), spatially averaging is required for the gRC of all grid cells (land and river cells) in the upstream catchment (note: not only river grid cells). Therefore, analyzing gRC values for all river grid cells ($\sim 10,106$) would require: (1) catchment delineation for each of these river grid cells, (2) determination of the gRC of all grid cells in the delineated catchment area, and (3) calculation of the catchment-averaged gRC. This was beyond the scope of this study, instead we obtained the gRC of 16 catchments across Europe to correlate these with the forecast performance (Fig. 4).

BFI and RC are commonly used to represent catchment response, whether it is fast or slow, which is associated with low or high catchment memory. Although BFI and RC are both derived from river flow hydrograph, they represent much more than just river system properties. In this study, we do not use the commonly applied RC, but the groundwater Recession Coefficient (gRC, see above), which even better represents catchment memory. We also note that indicators, such as hydrogeology (permeability and storage properties of aquifers and soils), the relative area of lakes/wetland/bogs, topography, land use that also describe catchment memory, however, are already implicitly embedded in the BFI and gRC.

Brier Score (BS). The performance of the drought forecasts was assessed using a commonly used metric called the Brier Score (BS)²⁷. The BS has been used among many studies dealing with probabilistic forecasts^{6,67,68}. The BS values for each LT is calculated as follow:

$$BS = \frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2, \quad (2)$$

where N is the number of members in the forecasting ensemble, in this case is 25, P is the forecast probability of an event (P varies from 0 to 1), and O is the observed probability. The observed probability (O) consists of binary numbers either 0 for no drought event and 1 for drought event. For BS analysis, drought occurs if the SSI-x value falls below -0.5 (mild drought)^{6,16}. The BS value ranges from 0 to 1, with 0 being the perfect forecast and 1 being the worst forecast. In this paper, we average the BS values from every forecast initiation (on the second of each month), from January 2002 to December 2016, into BS values for each LT (LT = 1 to 7-month). An exception was made for seasonal analysis, where we averaged the BS values from forecast initiations in December, January, and February (DJF) for winter, March, April, and May (MAM) for spring, June, July, and August (JJA) for summer, and September, October, and November (SON) for autumn. Note that we used the Brier Score (BS) in this study that describes forecast performance, and not the Brier Skill Score (BSS), which labels forecast skill.

Data availability

The EFAS streamflow data are accessible under a COPERNICUS open data license (<https://doi.org/10.24381/cds.e3458969>). In this study, we used EFAS version 3. The SSI-1 and SSI-3 datasets are available in the 4TU data center with doi: <https://doi.org/10.4121/17259449.v1>. Other data and codes generated and/or analyzed during this study are available from the corresponding author on request.

Received: 18 June 2021; Accepted: 1 February 2022

Published online: 17 February 2022

References

- Sutanto, S. J., Van Lanen, H. A. J., Wetterhall, F. & Lloret, X. Potential of pan-European seasonal hydro-meteorological drought forecasts obtained from a multi-hazard early warning system. *Bull. Am. Meteorol. Soc.* **101**, 368–393 (2020).
- Sheffield, J. *et al.* A drought monitoring and forecasting system for sub-Saharan African water resources and food security. *Bull. Am. Meteorol. Soc.* **95**, 861–882 (2014).
- McKee, T. B., Doesken, N. J. & Kleist, J. The relationship of drought frequency and duration to time scale. In *Proc. of 8th Conf. on Applied Climatology*, 179–184 (Amer. Meteor. Soc., Anaheim, California, 1993).
- Dutra, E., di Giuseppe, F., Wetterhall, F. & Pappenberger, F. Seasonal forecasts of droughts in African basins using the standardized precipitation index. *Hydrol. Earth Syst. Sci.* **17**, 2359–2373 (2013).
- Yuan, X. *et al.* Probabilistic seasonal forecasting of African drought by dynamical models. *J. Hydrometeorol.* **14**, 1706–1720 (2013).
- Trambauer, P. *et al.* Hydrological drought forecasting and skill assessment for the Limpopo river basin, Southern Africa. *Hydrol. Earth Syst. Sci.* **19**, 1695–1711 (2015).
- Sutanto, S. J., Wetterhall, F. & Van Lanen, H. A. J. Hydrological drought forecasts outperform meteorological drought forecasts. *Environ. Res. Lett.* **15**, 084010 (2020).
- Tallaksen, L. M. & Van Lanen, H. A. J. *Hydrological Drought. Processes and Estimation Methods for Streamflow and Groundwater Development in Water Science* (Elsevier Science B.V., 2004).
- Mishra, A. K. & Singh, V. P. A review of drought concepts. *J. Hydrol.* **391**, 202–216 (2010).
- Sutanto, S. J. & Van Lanen, H. A. J. Streamflow drought: Implication of drought definitions and its application for drought forecasting. *Hydrol. Earth Syst. Sci.* **25**, 3991–4023 (2021).
- Cui, B., Toth, Z., Zhu, Y. & Hou, D. Bias correction for global ensemble forecast. *Weather Forecast.* **27**, 396–410 (2012).
- Vitart, F. & Robertson, A. W. The sub-seasonal to seasonal prediction project (s2s) and the prediction of extreme events. *NPJ Clim. Atmos. Sci.* **1**, 3 (2018).
- Weisheimer, A. & Palmer, T. N. On the reliability of seasonal climate forecasts. *J. R. Soc. Interface* **11**, 20131162 (2014).
- Yuan, X. & Wood, E. F. Multimodel seasonal forecasting of global drought onset. *Geophys. Res. Lett.* **40**, 4900–4905 (2013).
- Dutra, E. *et al.* Global meteorological drought—Part 2: Seasonal forecasts. *Hydrol. Earth Syst. Sci.* **18**, 2669–2678 (2014).
- Van Hateren, T., Sutanto, S. J. & Van Lanen, H. A. J. Evaluating uncertainty and robustness of seasonal meteorological and hydrological drought forecasts at the catchment scale—case Catalonia (Spain). *Environ. Int.* **133**, 105206 (2019).
- Shukla, S. & Wood, A. W. Use of a standardized runoff index for characterizing hydrologic drought. *Geophys. Res. Lett.* **35**, L02405 (2008).
- Bloomfield, J. P. & Marchant, B. P. Analysis of groundwater drought building on the standardized precipitation index approach. *Hydrol. Earth Syst. Sci.* **17**, 4769–4787 (2013).
- Van Loon, A. F., Van Huijgevoort, M. H. J. & Van Lanen, H. A. J. Evaluation of drought propagation in an ensemble mean of large-scale hydrological models. *Hydrol. Earth Syst. Sci.* **16**, 4057–4078 (2012).
- Pechlivanidis, I. G., Crochemore, I., Rosberg, J. & Bosshard, T. What are the key drivers controlling the quality of seasonal streamflow forecasts?. *Water Resour. Res.* **56**, e2019WR026987 (2020).
- Crochemore, M. G. L. & Pechlivanidis, I. G. Benchmarking an operational hydrological model for providing seasonal forecasts in Sweden. *Hydrol. Earth Syst. Sci.* **25**, 1189–1209 (2021).
- Nalbantis, I. & Tsakiris, G. Assessment of hydrological drought revisited. *Water Resour. Manag.* **23**, 881–897 (2009).
- UNSD. Methodology: Standard country or area codes for statistical use (m49). Tech. Rep., United Nations Statistics Division (1999). <https://unstats.un.org/unsd/methodology/m49/> (accessed 17 Nov 2020).
- Pappenberger, F., Thielen, J. & del Medico, M. The impact of weather forecast improvements on large scale hydrology: Analysing a decade of forecasts of the European flood alert system. *Hydrol. Process.* **25**, 1091–1113 (2011).
- Cloke, H., Pappenberger, F., Thielen, J. & Thiemeig, V. In *Operational European Flood Forecasting. Environmental Modelling: Finding Simplicity in Complexity*, 2nd ed. (eds. Wainwright, J. & Mulligan, M.) (Wiley, 2013).
- IOH. Low flow studies report. Tech. Rep., Institute of Hydrology, Wallingford, UK (1980). https://nora.nerc.ac.uk/id/eprint/90931/Low_Flow_01.pdf (accessed 6 Aug 2020).
- Brier, G. W. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **78**, 1–3 (1950).
- Hall, F. R. Base-flow recession—A review. *Water Resour. Res.* **4**(5), 973–983 (1968).

29. Salinas, J. L. *et al.* Comparative assessment of predictions in ungauged basins—part 2: Flood and low flow studies. *Hydrol. Earth Syst. Sci.* **17**, 2637–2652 (2013).
30. Hisdal, H., Tallaksen, L. M., Clausen, B., Peters, E. & Gustard, A. Hydrological drought characteristics. Hydrological drought. Processes and estimation methods for streamflow and groundwater. In *Development in Water Science*, vol. 48 (eds. Tallaksen, L. M. & van Lanen, H. A. J.) (Elsevier Science B.V., 2004).
31. Beck, H. E. *et al.* Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resour. Res.* **49**, 7843–7863 (2013).
32. Barker, L. J., Hannaford, J., Chiverton, A. & Svensson, C. From meteorological to hydrological drought using standardized indicators. *Hydrol. Earth Syst. Sci.* **20**, 2483–2505 (2016).
33. Svensson, C. *et al.* Long-range forecasts of UK winter hydrology. *Environ. Res. Lett.* **10**, 064006 (2015).
34. Schneider, M. K., Brunner, F., Hollis, J. M. & Stamm, C. Towards a hydrological classification of European soils: Preliminary test of its predictive power for the base flow index using river discharge data. *Hydrol. Earth Syst. Sci.* **11**, 1501–1513 (2007).
35. Sprenger, C. *et al.* Inventory of managed aquifer recharge sites in Europe: Historical development, current situation and perspectives. *Hydrogeol. J.* **25**, 1909–1922 (2017).
36. Clark, M. P. & Hay, L. E. Global meteorological drought-part 1: Probabilistic monitoring. *J. Hydrometeorol.* **5**, 15–32 (2004).
37. Wetzol, M. *et al.* Mesoscale snowfall prediction and verification in mountainous terrain. *Weather Forecast.* **19**, 806–828 (2004).
38. Harrigan, S., Prudhomme, C., Parry, S., Smith, K. & Tanguy, M. Benchmarking ensemble streamflow prediction skill in the UK. *Hydrol. Earth Syst. Sci.* **22**(3), 2023–2039 (2018).
39. Tallaksen, L. M. & Stahl, K. Spatial and temporal patterns of large-scale droughts in Europe: Model dispersion and performance. *Geophys. Res. Lett.* **41**, 429–434 (2014).
40. Van Lanen, H. A. J., Wanders, N., Tallaksen, L. M. & Van Loon, A. F. Hydrological drought across the world: Impact of climate and physical catchment structure. *Hydrol. Earth Syst. Sci.* **17**, 1715–1732 (2013).
41. Zajac, Z. *et al.* The impact of lake and reservoir parameterization on global streamflow simulation. *J. Hydrol.* **548**, 552–568 (2017).
42. Lorenzo-Lacruz, J., Vicente-Serrano, S. M., González-Hindalco, J. C., López-Moreno, J. I. & Cortesi, N. Hydrological drought response to meteorological drought in the Iberian Peninsula. *Clim. Res.* **58**, 117–131 (2013).
43. Canuto, N. *et al.* Influence of reservoir management on Guadiana streamflow regime. *J. Hydrol. Reg. Stud.* **25**, 100628 (2019).
44. Van Loon, A. F. & Van Lanen, H. A. J. Making the distinction between water scarcity and drought using an observation-modeling framework. *Water Resour. Res.* **49**, 1483–1502 (2013).
45. Van Loon, A. F. & Laaha, G. Hydrological drought severity explained by climate and catchment characteristics. *J. Hydrol.* **526**, 3–14 (2015).
46. Bloomfield, J. P., Allen, D. J. & Griffiths, K. J. Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames basin, UK. *J. Hydrol.* **373**, 164–176 (2009).
47. Tjeldeman, E., Stahl, K. & Tallaksen, L. M. Drought characteristics derived based on the standardized streamflow index: A large sample comparison for parametric and nonparametric methods. *Water Resour. Res.* **56**, e2019WR026315 (2020).
48. Crochemore, L., Ramos, M.-H. & Pappenberger, F. Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts. *Hydrol. Earth Syst. Sci.* **20**, 3601–3618 (2016).
49. Bell, V. A., Davies, H. N., Kay, A. L., Brookshaw, A. & Scaife, A. A. A national-scale seasonal hydrological forecast system: Development and evaluation over Britain. *Hydrol. Earth Syst. Sci.* **21**, 4681–4691 (2017).
50. Arnal, L. *et al.* Skillful seasonal forecasts of streamflow over Europe?. *Hydrol. Earth Syst. Sci.* **22**, 2057–2072 (2018).
51. Scaife, A. A. *et al.* Skillful long-range prediction of European and north American winters. *Geophys. Res. Lett.* **41**, 2514–2519 (2014).
52. Madonna, E., Battisti, D. S., Li, C. & White, R. H. Reconstructing winter climate anomalies in the Euro-Atlantic sector using circulation patterns. *Weather Clim. Dyn. Discuss.* **2**, 777–94 (2021).
53. Vitart, F. Evolution of ECMWF sub-seasonal forecast skill scores. *Q. J. R. Meteorol. Soc.* **140**, 1889–1899 (2014).
54. Richardson, D., Fowler, H. J., Kilsby, C. G., Neal, R. & Dankers, R. Improving sub-seasonal forecast skill of meteorological drought: A weather pattern approach. *Nat. Hazards Earth Syst. Sci.* **20**, 107–124 (2020).
55. Arnal, L. *et al.* Efas upgrade for the extended model domain. JRC Tech. Rep. EUR29323EN, JRC, Ispra, Italy (2019). [58pp, ISBN978-92-79-92881-9].
56. Forzieri, G. *et al.* Ensemble projections of future streamflow droughts in Europe. *Hydrol. Earth Syst. Sci.* **18**, 85–108 (2014).
57. Prudhomme, C. *et al.* Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *PNAS* **111**(9), 3262–3267 (2014).
58. Wanders, N., Wada, Y. & Van Lanen, H. A. J. Global hydrological droughts in the 21st century under a changing hydrological regime. *Earth Syst. Dyn.* **6**, 1–15 (2015).
59. Van Lanen, H. A. J. *et al.* Hydrology needed to manage droughts: The 2015 European case. *Hydrol. Process.* **30**, 3097–3104 (2016).
60. Thielen, J., Bartholmes, J., Ramos, M.-H. & de Roo, A. The European flood alert system-part 1: Concept and development. *Hydrol. Earth Syst. Sci.* **13**, 125–140 (2009).
61. Burek, P., van der Knijff, J. & Ntegeka, V. Lisvap evaporation pre-processor for the lisflood water balance and flood simulation model. JRC Tech. Rep. EUR26167EN, JRC, Ispra, Italy (2013). [42pp, ISBN978-92-79-33199-2].
62. Vicente-Serrano, S. M. *et al.* Accurate computation of a streamflow drought index. *J. Hydrol. Eng.* **17**, 318–332 (2012).
63. Slater, L. J. & Villarini, G. Enhancing the predictability of seasonal streamflow with a statistical-dynamical approach. *Geophys. Res. Lett.* **45**(13), 6504–6513 (2018).
64. Svensson, C., Hannaford, J. & Prosdociimi, I. Statistical distributions for monthly aggregations of precipitation and streamflow in drought indicator applications. *Water Resour. Res.* **53**, 999–1018 (2017).
65. Tallaksen, L. M. A review of baseflow recession analysis. *J. Hydrol.* **165**, 349–370 (1995).
66. Sutanudjaja, E. H., Van Beek, L. P. H., de Jong, S. M., Van Geer, F. C. & Bierkens, M. F. P. Large-scale groundwater modeling using global datasets: A test case for the rhine-meuse basin. *Hydrol. Earth Syst. Sci.* **15**, 2913–2935 (2011).
67. Yossef, N. C., Van Beek, R., Weerts, A., Winsemius, H. & Bierkens, M. F. P. Skill of a global forecasting system in seasonal ensemble streamflow prediction. *Hydrol. Earth Syst. Sci.* **21**, 4103–4114 (2017).
68. Sutanto, S. J., van der Weert, M., Blauhut, V. & Van Lanen, H. A. J. Skill of large-scale seasonal drought impact forecasts. *Nat. Hazards Earth Syst. Sci.* **20**, 1595–1608 (2020).

Acknowledgements

The research is supported by the ANYWHERE project (Grant Agreement No.: 700099), which is funded within the EU's Horizon 2020 research and innovation program (www.anywhere-h2020.eu). 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. We thank Edwin Sutanudjaja (Utrecht University) for providing the PCR-GLOBWB groundwater Recession Coefficient (gRC) data.

Author contributions

Both authors conceived and implemented the research equally. Data analyses, model output analyses, and figures have been performed by S.J.S., S.J.S. and H.A.J.V.L. wrote the initial version of the paper, interpreting the results, discussion, and improving the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-06553-5>.

Correspondence and requests for materials should be addressed to S.J.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022