Seasonal forecasting of local-scale soil moisture droughts with Global BROOK90: A case study of the European drought of 2018.

Ivan Vorobevskii, Thi Thanh Luong, Rico Kronenberg

Faculty of Environmental Sciences, Department of Hydrosciences, Institute of Hydrology and Meteorology, Chair of Meteorology, Technische Universität Dresden, Tharandt, 01737, Germany

Correspondence to: Ivan Vorobevskii (ivan.vorobevskii@tu-dresden.de)

Abstract. Prolonged deficit of soil moisture can result in significant ecosystem and economical losses. General slowdown of vegetation growth and development, withering of foliage cover, reduction of carbon, nutrients and water cycling, increase of fire and insect outbreaks are just a few examples of soil moisture drought impacts. Thus, an early and timely warning via monitoring and forecast could help to prepare for the drought and manage its consequences.

In this study, a new version of Global BROOK90, an automated framework to simulate water balance at any location is presented. The new framework integrates seasonal meteorological forecasts (SEAS5 forecasting system) from European Centre for Medium-Range Weather Forecasts (ECMWF). Here we studied how well the framework can predict the soil moisture drought on a local scale. Twelve small European catchments (from 7 to 115 km²) characterised by various geographical conditions were chosen to reconstruct the 2018-2019 period, when a large-scale prolonged drought was observed in Europe. Setting the ERA5-forced soil moisture simulations as a reference, we analysed how the lead time of the SEAS5 hindcasts influences the quality of the soil moisture predictions under drought and non-drought conditions.

It was found that the hindcasted soil moisture fits well with the reference model runs only within the first (in some cases until second and third) month of lead time. Afterwards significant deviations up to 50% of soil water volume were found. Furthermore, within the drought period the SEAS5 hindcast forcing resulted in overestimation of the soil moisture for most of the catchment, indicating an earlier end of a drought period. Finally, it was shown that application of the probabilistic forecast using the ensembles’ quantiles to account for the uncertainty of the meteorological input is reasonable only for lead time up to three months.

Introduction and motivation

Drought is a complex, multifactorial phenomenon that includes climate, water resources, and socioeconomic factors and impacts on a community in short term as well as in long term (Crausby et al., 2017; Grillakis, 2019; Mueller and Zhang, 2016; Sheffield et al., 2012; Wanders et al., 2014). In two past decades Europe experienced a series of dry summers with significant impacts: in 2003 (Fischer et al., 2007; Schär and Jendritzky, 2004), 2010 (Barriopedro et al., 2011), 2015 (Moravec et al., 2021; Van Lanen et al., 2016), and 2018-2020 (Moravec et al., 2021; Peters et al., 2020; Rakovec et al., 2022). The European Commission reported 9 billion euro annual monetary losses across Europe due to drought in the current situation, which will increase up to 65 billion by the end of the century for the worst climate change scenario (Naumann et al., 2021). Among commonly accepted drought types, the soil drought typically causes most of the damages for agriculture, forestry and ecosystems in general (Mishra and Singh, 2010; Sutanto et al., 2019; Zink et al., 2016). Although significant efforts are being
made to develop drought monitoring and forecasting systems, the ability to forecast droughts is limited due to the inherent uncertainties of long-term weather forecasts (Sutanto and Van Lanen, 2022; Wanders and Van Lanen, 2015). Therefore, often multiple multi-year droughts are rarely mentioned in seasonal forecasts and only reported ‘post factum’ in observations, reports and reconstructions (Boeing et al., 2022; Boergens et al., 2020; Rakovec et al., 2022). Hence, accurate monitoring and seasonal forecasting of drought is beneficial for the development of early prevention, mitigation, and management strategies. Recently, with the improvement of computing infrastructure and capacity, the use of probability-based seasonal weather forecasts driven by numerical weather prediction models has become more popular and advanced (Samaniego et al., 2019). This has led to the development of drought warning systems at various spatial and temporal scales (Wanders et al., 2019). Several operational monitoring and forecasting systems exist on continental, national and regional scales. These systems are principally based on rainfall, temperature and hydrological gridded observed and modelled data (Otkin et al., 2018; Sheffield et al., 2012), although new approaches such as DroughtCast (Brust et al., 2021) implementing a machine learning algorithms have also been attempted. For the United States, a real-time drought monitoring system provides information on current, short- (up to 8 weeks) and long-term (3.5 months) predicted drought conditions in 0.12° spatial resolution (Lorenz et al., 2017; Svoboda et al., 2002). It uses a combination of precipitation anomalies, evaporative stress index as well as soil moisture tendencies on three levels and the input of regional and local experts. The African Flood and Drought Monitor (0.25° resolution with daily updates) was developed for the monitoring purposes and provides a set of drought indexes such as SPI (Standardised Precipitation Index), soil various moisture and vegetation indices, and streamflow percentiles and deficit (Sheffield et al., 2014). Swiss monitoring and forecasting system shows canton-based current precipitation and soil moisture deficit levels as well as gives a 5- and 30-day forecast (Zappa et al., 2014). The German drought monitor (Zink et al., 2016) provides daily drought information for topsoil and full soil column based on soil moisture anomalies on a 5 km grid scale. European Drought Observatory provides up-to-date information on the occurrence and severity of droughts across Europe with 5 km resolution based on a combination of SPI, soil moisture anomaly index and vegetation greenness (Cammalleri et al., 2021), as well as basic forecasts based on SPI (Wanders et al., 2019). Finally, the European Flood Awareness System’s (EFAS) forecasts produced with the LISTFLOOD model and based on SEAS5 forcing (Thielen et al., 2009) provides soil moisture forecasts for 3 standard soil layers on a 5 km grid and for 24 h time step since 2020 (Joint Research Centre (European Commission) et al., 2019). However, the major drawbacks of the most advanced existing frameworks is their inability to reach local-scale for the current and predicted conditions of soil moisture. Tackling the problem of achieving high resolution in monitoring and forecasting of water balance components in general and soil moisture in particular is an ongoing process. The local scale plays a special role (Figure 1), since this is often the scale, where the final decisions are made, measures are implemented and management is taking place (European Commission, 2015; Suárez-Almiñana et al., 2017; Wagner et al., 2009). Although there are several global and national datasets, which could be used for monitoring purposes, data quality and resolution often do not correspond to the local requirements. So far, the grid size for the state of the art up-to-date global reanalysis varies in a range of 10-50 km (Gelaro et al., 2017; Ebita et al., 2011; Martens et al., 2020) and regional models can reach 1-5 km (Zink et al., 2017). Therefore, the local scale is not and in the nearest future probably will not be considered for global or regional models (Beven and Cloke, 2012; Sood and Smakhtin, 2015; Wood et al., 2011). Finally, even in the presence of a dense network of long-term measurements, it is highly improbable
that the data of all observed variables will be available or transferable and thus representative for the desired location.

---

In 2020, a first version of a fully automated framework to simulate water balance in general, and soil moisture in particular on a local scale at any desired location with historical meteorological data called Global BROOK90 was released (Vorobevskii et al., 2020). The framework has thereafter been validated using runoff and evaporation components (Vorobevskii et al., 2021, 2022). Here we want to present an updated version of the framework, which integrates SEAS5 seasonal forecasts from the ECMWF (European Centre for Medium-Range Weather Forecasts) and thus allows long-term forecasts of the local water balance. Taking the European drought of 2018-2019 as a case study, we attempted to hindcast soil moisture conditions during this period on a local scale for several catchments located within different geographical conditions. We compared simulations forced with reanalysis and forecast meteorological datasets to answer the following research questions:

- Can a local-scaled soil moisture forecast be achieved from a global setup based on open sourced data and a 1D water balance model?
- How skilful are the local seasonal soil moisture estimates derived using long-term meteorological forecasts under drought conditions of the years 2018 and 2019 for European catchments?
- How does the lead time of the meteorological forecast affect soil moisture predictions on a local scale in normal state and under drought conditions, as well as for different landscapes?
- Does the usage of ensemble quantiles advance drought prediction compared to the ensemble mean?

1. Models and Data

1.1 Global BROOK90 v 2.0

The first version of Global BROOK90 framework was introduced in 2020 (Vorobevskii et al., 2020) and the new updated version was released in 2023, which is applied here to simulate soil moisture. The framework uses open-source global datasets to parameterise and force the water balance model in a fully automatic mode based on the input location and time-interval. As these input datasets are covering almost the entire terrestrial earth, the framework is applicable all over the globe including ungauged locations, which was one of the main goals of its development. The framework was tested for a variety of different geographical conditions (Vorobevskii et al., 2021).

The following datasets are incorporated to describe the characteristics of the location. For the canopy, identification and parameterization Copernicus Global land Cover 100 m (Buchhorn et al., 2020), MOD15A2H
MODIS Leaf Area Index/FPAR 8-Day L4 Global (Myneni et al., 2021) and Global Forest Canopy Height (Potapov et al., 2021) are used. The SoilGrids250 dataset (Hengl et al., 2017) provides global information on soil properties for seven standard layers (texture, depth to the bedrock, stone content). Digital elevation model is downloaded from the Mapzen Terrain Tiles (Larrick et al., 2020).

For meteorological forcing, reanalysis and forecast datasets are implemented. Historical runs could be performed with ERA5 (Copernicus Climate Change Service Information, 2018a) and MERRA-2 (Gelaro et al., 2017). ERA5 provides hourly temporal and 0.25° spatial resolution data and covers the time period from 1959, while MERRA-2 has a 0.5°×0.625° grid with 6-hour time-slices and starts from 1980. For seasonal forecasting the SEAS5 data from ECMWF (Copernicus Climate Change Service Information, 2018b) is integrated. It implements a 51-member ensemble meteorological forecast for 215 days on a 1° grid with daily temporal resolution and is released on the fifth day of each month. The dataset is also available in a hindcast mode starting from 1993. For the bias-correction of the forecast, empirical quantile mapping on monthly basis (Boé et al., 2007) is applied. The bias is calculated between each hindcast ensemble mean and reanalysis data on a monthly scale, and then averaged for each calendar month and lead time. Thereafter, the resulting bias is used to correct the forecast. As the temporal resolutions of the datasets are not the same, which might be of importance when sub-daily precipitation dynamics play a significant role in water redistribution in the system, an additional remark is needed. BROOK90 itself runs with a variable number of iterations per day, which is automatically determined and dictated by the equilibrium of the water balance equation inside the model system. The minimum iteration number is two, or in case of precipitation event, it equals the resolution of the meteorological data input. However, under some conditions (e.g. heavy rain, drought, complex soil profile) this number could be as high as 1000, independent from the input data resolution. Nevertheless, the model output is always of daily resolution. However, the BROOK90 allows to account for the problem of sub-daily precipitation absence (SEAS5 case) to create hourly data using daily precipitation input and monthly values of mean event duration in hours (DURATN parameter). It is then calculated from the available ERA5 hourly data and applied to the forecast forcing to improve the results. Finally, Global BROOK90 allows combining historical and forecast simulations if the continuous timeline is preserved.

This framework is based on the BROOK90 model (Federer et al., 2003) which is a one-dimensional physically-based model for the vertical water fluxes simulations in soil-plant-atmosphere systems. At first, the precipitation input goes through the canopy, where it could be either intercepted (and then evaporated) or passed through to the ground surface. Then, the portion reaching the ground level, could be infiltrated, frozen, evaporated, converted to surface flow, percolated or stored as soil moisture. The infiltrated volume is distributed between soil layers to macropore bypass and matrix flow using ‘top-down’ approach. Soil water movement in the model is based on the approximations of the Richard's equation (Richards, 1931), where the functional relationships between main soil parameters (soil-water content, matric potential and hydraulic conductivity) are estimated using Clapp and Hornberger parameterisation (Clapp and Hornberger, 1978). The soil column has groundwater, seepage and downslope outflow. Finally, soil water storage is used for evaporation from the top layers and root uptake for transpiration.

The scheme of the framework is presented on Figure 2. It applies a regular 50×50 m grid to identify hydro response units (HRUs) based on the downloaded characteristics of the input catchment. The model is then applied separately to each HRU, and then an area-weighted mean for each variable is calculated from HRU output data. A more detailed description of the framework is presented in Vorobevskii et al., 2020.
By means of an automatic workflow of all processes, the framework serves both professional and non-professional users. Further, as it is designed for a local scale, it does not require substantial resources regarding computational power, time and memory. For example, a 7 month water balance forecast for a small catchment is required. For a test case we can use a 4.6 km² catchment with 24 HRUs (Wernersbach Creek near Tharandt in Saxony, Germany) provided with the Global BROOK90 framework on GitHub. It takes around 5-10 min to download elevation, land cover and soil data. Depending on requested length and type of meteorological data, the download time can vary from half an hour to a few days due to system build-up of meteorological data providers. For instance, download of one year ERA5 data for the model warm-up and 7 months of SEAS5 forecasts without hindcasts for bias correction lasted 7 hours (6 hours from which were needed for the forecast request). Finally, computational time including data processing, modelling and saving the results on 3.4 GHz 16 GB RAM PC lasted approximately 30 min, from which the time needed to forecast one HRU with the abovementioned time interval yields 1.5 min.

### 2. Methodology

#### 2.1. Spatio-temporal extent of the soil moisture simulations

For the assessment of the soil moisture forecast with Global BROOK90 the time period of 2017-2020 was chosen. The choice of the time period was determined by the extent of the drought event including pre- and post-conditions. Thus, two years with ‘normal’ (2017, 2020) and two with ‘drought’ (2018-2019) conditions were considered to investigate the difference in forecasting skills for two hydrological conditions. According to ECMWF technical documentation, the used version of the forecast system (SEAS5) was operationally launched in 2017. The hindcast data (available from 1991) is advised to be used for the forecast bias correction purposes only since the ensemble number is significantly reduced in comparison to the operational dataset. SEAS5 hindcasts with 51 ensemble members on daily scale and 7 months lead time were applied for each month starting from July 2016 up to December 2020, so that for each month in a period of 2017-2020 all possible lead times (1-7 month) will appear. All forecasts were bias-corrected using hourly ERA5 data and the whole available length of hindcasts. Additionally, for all model runs one warm-up year was included. In case of a simulation with hindcast, corresponding ERA5 data was attached. Model input files and row simulation results with standard daily resolution output are available as Supplementary (Vorobevskii, 2023).
Simulations were conducted in twelve European catchments in total (Figure 3, Table 1). The focus of the study was intentionally put on predicting soil moisture in dry conditions since they are typically hard to forecast. The pilot catchments were chosen in such a way that they are distributed over Europe, ideally covered different geographical conditions and showed good discharge validation skill-score. Thus potentially giving reliable estimations of soil moisture with regard to site-specific parameterization. Chosen catchments range from 7 to 115 km² (average size 52 km²). They are covered with three different forest types (opened and closed, deciduous and evergreen, needle- and broadleaf), two short canopies types (grassland, crops) and have various soil textures. Available open-source satellite images and maps do not show significant signs of urbanisation (maximum values of 5% were identified for a few catchments) or hydraulic structures (except artificial channels in cultivated areas) which could noticeably influence a natural flow regime. Although numerous reports and research data of the 2018-2019 drought are available, evaluations of the drought spatial extension in 2018-19 over Europe differ significantly (Boergens et al., 2020; Buras et al., 2020; Hari et al., 2020; JRC European Drought Observatory, 2018, 2019). However, the majority of the selected locations appear within the commonly identified territories that were affected by the drought. The Kling-Gupta-Efficiency values for a daily scale discharge validation for the selected catchments (Vorobevskii et al., 2021) varies between 0.43 and 0.77 (with a mean of 0.57) for the evaluation time-period of 30-42 years (with a mean of 38).

![Figure 3. Overview on the selected locations](image)

**Table 1. Summary on the selected catchments**

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Country</th>
<th>Size [km²]</th>
<th>Dominating land cover types</th>
<th>Dominating soil texture</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ribeira de Sampaio - Cabriz</td>
<td>Portugal</td>
<td>10.8</td>
<td>Open forest (unknown), grassland</td>
<td>Sandy clay loam</td>
</tr>
<tr>
<td>2</td>
<td>Le Langelin - Briec</td>
<td>France</td>
<td>7.04</td>
<td>Agriculture, grassland, open forest (unknown)</td>
<td>Loam</td>
</tr>
<tr>
<td>3</td>
<td>La Glaeyle - Gluiras</td>
<td>France</td>
<td>70.9</td>
<td>Closed forest, deciduous broadleaf</td>
<td>Loam</td>
</tr>
<tr>
<td>4</td>
<td>Warleggan - Trengoffe</td>
<td>UK</td>
<td>25.3</td>
<td>Grassland, agriculture</td>
<td>Loam</td>
</tr>
<tr>
<td>5</td>
<td>Jiterka - Dolni Stepanice</td>
<td>Czech Republic</td>
<td>44.1</td>
<td>Closed forest (evergreen, needle leaf)</td>
<td>Loam</td>
</tr>
<tr>
<td>6</td>
<td>Ucja - Zaga</td>
<td>Slovenia</td>
<td>50.2</td>
<td>Closed forest (deciduous, broadleaf)</td>
<td>Loam</td>
</tr>
<tr>
<td>7</td>
<td>Grosse Ohe - Taferlruck</td>
<td>Germany</td>
<td>19.1</td>
<td>Closed forest (evergreen and deciduous, needle and broadleaf)</td>
<td>Sandy loam</td>
</tr>
<tr>
<td>8</td>
<td>Wartach - Wartach</td>
<td>Germany</td>
<td>34.5</td>
<td>Closed forest (evergreen, needle leaf)</td>
<td>Loam</td>
</tr>
<tr>
<td>9</td>
<td>Alto – Taglio-Isolaccio</td>
<td>France</td>
<td>114</td>
<td>Closed forest (deciduous, broadleaf)</td>
<td>Clay loam</td>
</tr>
<tr>
<td>10</td>
<td>Lenne - Oelkassen</td>
<td>Germany</td>
<td>65.6</td>
<td>Agriculture, closed forest (deciduous, broadleaf)</td>
<td>Loam</td>
</tr>
<tr>
<td>11</td>
<td>La Dragne - Vandenesse</td>
<td>France</td>
<td>115</td>
<td>Grassland, closed forest (deciduous, broadleaf)</td>
<td>Loam</td>
</tr>
<tr>
<td>12</td>
<td>Natzschung - Rothenthal</td>
<td>Germany</td>
<td>76.1</td>
<td>Closed forest (evergreen, needle leaf)</td>
<td>Sandy loam</td>
</tr>
</tbody>
</table>
2.2. Validation of the simulated soil moisture

2.2.1 Validation of ERA5-forced simulations against SMAP pseudo-observations

With the absence of available data of on-site soil moisture measurements, satellite based products were chosen for the validation. To benchmark the soil moisture simulations from Global BROOK90 the SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data V007 (Reichle et al., 2021a) was used. The product uses L-band brightness temperature data from satellites assimilated into a land surface model to estimate soil moisture on a 9 km spatial and 3 hours temporal scale globally from April 2015. The soil moisture is provided for the topsoil (0-5 cm), root zone (0-100 cm) and the full soil column (up to the model bedrock - e.g. 1.3-3.8 m for the study catchments depending on specific location with a mean of 1.9 m) according to product description (Reichle et al., 2021b).

For the validation, the catchment-weighted-mean of soil moisture simulated with ERA5 meteorological forcing was taken. As the thickness of standard soil layers in Global BROOK90 (dictated by SoilGrids250 dataset) does not provide a full match with the SMAP layout, it was decided to use the closest values of 0-2.5 cm for topsoil, 0-80 cm for root zone and 2 m for the whole soil profile. Since the area of one SMAP grid (81 km²) corresponds well with the sizes of the chosen catchments, the closest to the catchment centre grid was selected. Kling-Gupta-Efficiency (KGE) (Gupta et al., 2009) was chosen to show the agreement of volumetric water content from SMAP and Global BROOK90 on a monthly scale.

2.2.2 Validation of SEAS5-forced against ERA5-forced simulations

As Global BROOK90 initially provides soil moisture estimates applying ERA5 reanalysis as a forcing, it is reasonable to show the added value of forecasted (with SEAS5 forcing) soil moisture with regard to ERA5, since both datasets come from the same assimilation model. Daily and monthly catchment-weighted means as well as HRU-based absolute values (mm per layer) were used to compare soil moisture simulations using reanalysis and forecast forcings. For the calculation of differences between two forcings, monthly means from the 51 forecast ensemble runs were considered. Relative errors were calculated as absolute error divided by mean soil moisture of the specific catchment to account for differences in site-specific soil moisture conditions between the catchments. Furthermore, the results for topsoil (0-45 cm) and full soil column (up to 200 cm) were analysed separately. Drought periods were identified based on the Root Extractable Water coefficient (Eq. 1).

\[
REW = (\theta_C - \theta_{WP})/(\theta_{FC} - \theta_{WP})
\] (1)

where \(\theta\) is volumetric soil moisture at different states: C is current value at present conditions, WP is wilting point (-1500 kPa), and FC is field capacity. This coefficient is calculated along with soil moisture values on a daily scale in Global BROOK90 and is presented in the output. Various researchers state that the values of REW below 0.2-0.4 indicate the beginning of a water stress for vegetation (Bréda et al., 2006; Granier et al., 1999; Schmidt-Walter et al., 2019; Vilhar, 2016). Here, we have chosen a threshold of 0.3 to mark drought conditions.

3. Results and discussion

3.1 Comparison of SMAP and ERA5-forced soil moisture

Summarised performance of Global BROOK90 with regard to SMAP data for the study catchments is shown on Figure 4. The performance of the soil moisture simulations in the upper zone of 5 cm (which in SMAP is directly
assimilated from satellite data) was found better, than for the root zone and for the full profile (which in SMAP are already land surface model derivatives). The mean KGE value for the topsoil was 0.53, the lowest value 0.27 was found for the catchment 8 and the highest value of 0.82 for the catchment 11. All catchments showed high correlation coefficients with a mean KGE of 0.88. However, both Global BROOK90 underestimated the mean (mean bias 0.82) as well as the variance (mean variance ratio 0.66) in comparison to SMAP, except for two catchments. This could be partly explained by different framework setups, namely differences in the models themselves as well as underlying datasets used to derive soil properties. With increase of the soil depth, the agreement between two datasets decreased, leading to mean KGE values of 0.48 for the root zone and 0.34 for the full column. This is mainly due to decrease of correlation and variance ratio.

Overall, the best performance was achieved for the short canopies (cultivated and herbaceous land covers), where the satellite signal could penetrate deeper through the vegetation into soil (Babaeian et al., 2019). Thus, it is not evident that Global BROOK90 simulations in tall canopies have worse performance, rather that the uncertainty of SMAP data is much higher for these land covers.

Figure 4. Kling-Gupta-Efficiency and its components between SMAP and Global BROOK90P monthly soil moisture for twelve catchments

3.2 Drought monitoring and forecasting on a local scale

Several snapshots of spatial soil water deficit for 2018 in Natzschung catchment (12 on Figure 1) are shown in Figure 5 to present the advantages of drought monitoring with Global BROOK90 on a local scale. In January, the soil is close to saturation (REW values 0.7-0.8 for topsoil and 0.9-1 for full column). Exceptions are the few HRUs with urban areas, where the highly-sealed surfaces lead to blockage of moisture renewal. Six months later in June, when meteorological and hydrological droughts were already clearly noticeable, amounts of the soil water were reduced by around 40%, but remaining on a plant-demand level. Cultivated territories in the catchment are mainly planted with wheat, barley, oil fruits, silage maize and rye. As these cultures have a shallow effective root penetration, topsoil soil moisture (where REW values were found between 0.4 and 0.6) plays a higher role, than deeper horizons. The predominant forest species in the catchment is Norway Spruce also quite often has shallow rooting system and the majority of the root mass is concentrated in the upper soil layers (Puhe, 2003). Thus, it could have already experienced some water stress by June as REW values in topsoil reached 0.2-0.4. On 15th October, the minimum soil water content was observed. The upper 45 cm of soil was almost completely dry, while deeper horizons under the croplands, beech and opened forests still contained plant-available soil moisture (although not accessible to the crops due to root depth). As November and December brought new precipitation, soil gained enough moisture for plant water supply (REW values 0.4-0.8 for topsoil and 0.5-1 for full column). Recovery of soil moisture under the tall canopies was not as noticeable as under the short ones. This can be explained with the harvesting of cultivated areas and withering on grasslands. Thus, almost no soil moisture was used for transpiration, which is typically the most consumable part of water balance in this climate. Moreover, the
The SEAS5 data used as a forcing in Global BROOK90 allows getting daily ensemble predictions of soil moisture with a lead time of seven months and monthly updates. Figure 6 shows three different hindcast-forced soil moisture simulations for the Natzschung catchment. Start of each hindcast is indicated with the red dot, ensemble mean is shown with red and single ensemble runs with grey colours. In the pre-drought winter period of 2017-2018, the September hindcast ensemble mean showed approximately 10-15% underestimation of water content compared to ERA5 forcing until February. In fact, this is rather a 1-1.5 month lag due to delayed prediction of the rainy season, since the slope of the growing moisture curve as well as the saturation-plateau values look consistent between both forcings. In May 2018, when the soil moisture decline was clear according to ERA5 data, the hindcast demonstrated even a steeper depletion of water content in the first 3 months, thereafter, however, it quickly flattened out and soil moisture refill began due to significant precipitation input. Thus, SEAS5 forcing not only predicted the drought recovery point two months earlier and severely overestimated soil water content by more than 25%. Finally, a hindcast started in January 2019 on the upward ‘recovery’ soil moisture curve showed a decent agreement with ERA5 forcing until a new seasonal decline started in April-May marking the beginning of the 2019 drought. For all the three presented hindcasts, ensemble band (especially lowest members) covers the variability of ERA5 forced soil moisture through all lead times, however, a general overestimation of precipitation for the drought period and its variability in general in SEAS5 hindcast is clear.
Both land cover and soil texture affect performance of the soil moisture forecast. Figure 7 reveals that for ERA5 forcing, HRUs with croplands on loams resulted in a higher soil moisture content and lower seasonal variations during the 2018 drought months than forests on sandy soils. This can be explained by in general lower evaporation rates from herbaceous vegetation, slower soil water movement and higher holding capacity of loamy soils. However, winter moisture recovery was found to be with the same rate for all HRUs. SEAS5 hindcast started in the beginning of drought period (July 2018) showed a good agreement within HRUs until beginning of August, thereafter displaying moisture increase, while according to ERA5 forcing, the drought continued. SEAS5-forced soil moisture content got in line with ERA5-forced data only in winter recovery time (January-February 2019). The absolute error of the forecast was found to be significantly different for various HRUs. It was lower for croplands (-10+18 mm), while for forests it was much higher (-10+50 mm).
Taking a closer look on the driest month during the 2018-2019 drought in the catchment (October), it could be seen, how SEAS5 forcing underestimates precipitation dynamics with an increase of lead time during the drought season (Figure 8). While for one-month lead time the differences in soil moisture between two ERA5-forced and hindcasted soil moisture are minor (up to 50 mm) even in forested areas of the catchment, already starting from 3-month lead time overestimations are up to 100 mm. Lastly, the 7-month forecast showed almost twice as much soil moisture as it was found in the reanalysis runs for the forests and more than +50% for cultivated areas.

Spatial distribution of daily KGE values between hindcasted and reanalysis soil moisture for the whole study period are presented on Figure 9. Best performance was found in the cultivated areas, where KGE values were starting from 0.95 for up to 1-month and were reduced to 0.5 for up to 7-month lead time. Grassland and deciduous broadleaf forest showed slightly lower values. Worst performance was demonstrated by spruce forests, especially on loamy soils (northeastern part of the catchment). Here, already from 3-month lead time difference in hindcast quality to cropland is noticeable (KGE values 0.6-0.8) and for 7-month lead time values dropped below 0. Disaggregation of KGE into its components reveal, that the highest influence on the hindcast quality have overestimated mean and underestimated variance ratios. This means, that long-term planning with regard to drought adaptation and mitigation procedures for the forest ecosystems and agriculture in this case are limited and severely affected by uncertainties of seasonal meteorological forecast.

3.3 Effect of the forecast lead time on the soil moisture predictions in different conditions

The effect of a forecast lead time in predicting the monthly soil moisture for the root zone and the full soil column in twelve selected catchments is shown in Figure 10 and Figure 11. A lead time of one month resulted in a relatively small discrepancy between hindcast and reanalysis forcings for all catchments. The relative and absolute errors rarely exceeded 10% and 20 mm respectively (for full soil column) even in drought periods and no
noticeable correlation with seasonality was observed. However, already a lead time of two (for some catchments – three) months showed much higher differences (both positive and negative) between hindcast- and ERA5-forced soil moisture (up to 30% relative and 150 mm absolute errors for full soil column). For the topsoil, absolute errors are in a range of 5-10 mm for 1-month lead time and increasing to 10-40 mm for 3 months lead time. However, the relative errors were found higher, than for the full column. These results are consistent with similar research (Wanders et al., 2019) and indicate imperfection of meteorological input, namely increasing uncertainties of the seasonal forecasts with the growing lead time. Moreover, for the majority of the catchments (except catchment 4) positive anomalies were found within identified drought periods (REW<0.3). This accounts for general overestimation of small-scale precipitation in autumn forecasts compared to ERA5. On the other hand, a big negative anomaly in August-September 2017 for catchments 2, 4, 5, 6, 10, 12 symbolises a general issue of SEAS5 forecast system, which met some general problems in the whole European domain for these months. With the further increase of a lead time, the differences increase as well, however, not so drastic compared to the differences between 1 and 3 months. Thus, based on results from twelve study sites, it can be concluded that the predictability of the soil moisture using the SEAS5 seasonal forecasts can be successfully accomplished with a lead time up to 2-3 months. Results for both full soil column and root zone look similar and consistent, although the difference in two datasets is more prominent for the latter one. Furthermore, it was noticed, that regardless lead time, SEAS5-forced simulations tend to show highest overestimations of the soil moisture near the end of REW-declared drought periods, thus forecasting the end of a water shortage too early, which is clearly visible for the topsoil (catchments 1, 2, 3, 5, 6, 9, 10, 11, 12). Finally, comparing the relative errors between the catchments in different conditions, it was found, that the smallest errors were achieved for catchment 8 (10-15%), located on a humid alpine forested foothills in the south of Germany. Systematic overestimations of more than 30% were found in catchments 1, 3, 10 and 12 during the drought periods, which sometimes were already significant at a lead time of two months (up to 20%). Underestimations of more than 30% were observed in catchments 6, 9, 10, 12 and appeared mostly in non-drought conditions during autumn. However, no clear pattern between behaviour of forecasted soil moisture and catchment characteristics (i.e. size, orography, dominated land cover and soil type) was found. Thus, it could be assumed with caution that inaccuracy (and its patterns) of the soil moisture forecasts are more likely to be explained by the uncertainties of the meteorological forecast, rather than different responses of catchments with various characteristics.
Figure 10. Absolute and relative difference between monthly soil moisture simulated with SEAS5 hindcasts (mean of ensemble runs) and ERA5 forcings for root zone (0.45 m)
Figure 11. Absolute and relative difference between monthly soil moisture simulated with SEAS5 hindcasts (mean of ensemble runs) and ERA5 forcing for full soil column (up to 2 m)
3.4 Advantages of using probabilistic weather forecasts in drought prediction

Multiple uncertainties of a weather forecast could be compensated by the advantages of using its members instead of only considering an ensemble mean. Here, the quantile predictions of monthly soil moisture simulated from 51 members of SEAS5 meteorological hindcast with different lead times are presented (Figure 12). One month lead time hindcast provides a good fit with ERA5 forcing. Here the width of an ensemble band is relatively narrow due to the small uncertainty band of the meteorological conditions within a short prediction range. Minor inconsistencies in soil moisture predictions for the hindcast and ERA5 forcings (summer 2017 and 2020) probably resulted from the difference in spatial resolution between two datasets. However, already by three month lead time, the spread between ensemble mean and quantiles becomes considerable, especially in the summer period due to increased uncertainty of the meteorological forecast. Here the drought development and propagation is better depicted by lower hindcast quantiles (10-20%), while for drought attenuation in the wet season all probabilities need be used due to a delay of the drought peak in the hindcast forcing dataset. Using a 7 month lead time, thus staying on the edge of seasonal forecast predictability, will bring even a higher spread in quantile hindcasts, however with close developments as for 3 month lead time. Here, the magnitude of soil moisture drought in the 2018 cannot be captured even with 1% quantile for 5 months in a row, meaning a significant overestimation of precipitation with increasing leading time for this year.

Figure 12. Monthly averages of a catchment-mean full-column (up to 2 m) soil moisture simulated with ERA5 and SEAS5 hindcasts forcings in Natzschung catchment
4. Conclusion and outlook

In this article, we present a new version of an automatic framework for water balance modelling on a local scale for any location on the globe called Global BROOK90. The special focus is given on the ability of the new version to provide predictions of water balance components up to seven months in advance by incorporation of SEAS5 seasonal meteorological forecasts. We showed how a combination of global land cover and soil open-source datasets, large-scaled meteorological forecasts and a physically-based model can predict a soil moisture drought with high spatial resolution.

Soil moisture simulations were conducted for twelve small catchments with various landscapes located in Europe within the 2017-2020 time period, which cover the well-known continent-scale European drought of 2018-2019. We used ERA5 reanalysis and SEAS5 hindcasts as a meteorological forcing. The runs with ERA5 forcing served as a reference for the soil moisture to assess the added value and quality of the forecast performance, thus comparing the meteorological data uncertainty. We illustrated spatio-temporal advantages and quality of small-scale modelling in monitoring and forecasting a drought using one of the study sites as an example.

Additionally, the ERA5-forced model runs were compared with SMAP data, which represents a satellite brightness temperature assimilated in a land surface model. Validation resulted in a good agreement of both datasets on a monthly scale, especially for the correlation coefficient. However, Global BROOK90 runs ended up with lower mean and variance of soil moisture. Highest KGE values were found for the topsoil and the goodness of fit declined with a profile depth.

Comparison of HRU-based and catchment-mean monthly soil moisture showed that similarly to large and medium scales, forecasts of soil moisture on a local scale are skilful up to 2-3-month lead-time, which is valid for drought conditions as well. Thereafter, the difference between two forcings could result in more than 100 mm or around 30% of the mean annual soil moisture content, which is a considerable amount, especially for the drought periods. It was found that for the majority of the study sites, SEAS5 forecast resulted in overestimation of the soil moisture in identified plant-water-stress periods due to the general overestimations of precipitation amounts. Furthermore, the absence of patterns in the seasonal forecasts with regard to the catchment characteristics indicates, that a forecast generalization to catchment scale is not recommended for water management purposes. The large deviations of HRU forecasts mirror the catchment heterogeneity and hence the complex interactions of different landscapes and soil profiles and imply the need for a local forecast.

Finally, possible advantages of probabilistic instead of only deterministic soil moisture forecasts using quantiles of ensemble runs were assessed. It was shown that for the lead times up to three months the method could be advantageous as the band of quantile forecast could cover the variability of soil moisture produced with ERA5 forcing. However, as the SEAS5-forced runs in the drought period generally tend to overestimate soil moisture, for three and seven months lead time even 10% quantile of hindcast was found insufficient to reach ERA5-forced soil moisture values.

Overall, Global BROOK90 combined with SEAS5 seasonal ensemble forecasts showed good results for a range of 1-3 months lead time and can serve as a decent basis for drought monitoring and forecasting on a local scale. Moreover, one of the definitive advantages of the framework is that it does not require big computational resources or programming background and could be run on a normal computer or laptop by a non-professional user.
Data and Code availability

Authors fully support open-source and reproducible research. Global BROOK90 framework is available under https://github.com/hydrovorobey/Global_BROOK90 repository (CC BY-NC-ND 4.0). Simulation results, initial and simulation datasets and R-scripts to reproduce figures and tables for the manuscript are available under the following HydroShare composite resource https://doi.org/10.4211/hs.d882e83bae95438881c7b47f003e7a3c.

Author contribution

Conceptualization VI and KR; data curation VI and LTT, formal analysis VI, methodology VI, LTT and KR; supervision KR; visualization VI; writing: original draft preparation VI and LTT, writing: review KR.

Competing interests

The authors declare that they have no conflict of interest.

Funding

Open Access Funding by the Publication Fund of the TU Dresden. This research was also funded by the German Federal Ministry of Education and Research (FKZ 01LR 2005A—funding measure “Regional Information on Climate Action” (RegIKlim), section (a) Model Regions.

References


Copernicus Climate Change Service Information: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. ERA5 hourly data on single levels from 1959 to present, 10.24381/cds.adbb2d47, 2018a.

Copernicus Climate Change Service Information: Copernicus Climate Change Service (C3S): Seasonal forecast daily and subdaily data on single levels, https://doi.org/10.24381/cds.181d637e, 2018b.


Vorobevskii, I.: Supplement materials for publication: Seasonal forecasting of local-scale soil moisture droughts with Global BROOK90, https://doi.org/10.4211/hs.d882e83bae9538881c7b47f003e7a3c, 2023b.


