

Water Resources Research

RESEARCH ARTICLE

10.1029/2020WR027828

Special Section:

The Quest for Sustainability of Heavily Stressed Aquifers at Regional to Global Scales

Key Points:

- Terrestrial systems modeling gives an interannual probabilistic assessment of subsurface water storage
- The hindsight water storage assessment of the drought years 2011/2012 and 2018/2019 provided useful results
- The assessment can generate valuable information for pronounced dry and also wet anomalies of subsurface water storage

Supporting Information:

- Supporting Information S1

Correspondence to:

C. Hartick,
c.hartick@fz-juelich.de

Citation:

Hartick, C., Furusho-Percot, C., Goergen, K., & Kollet, S. (2021). An interannual probabilistic assessment of subsurface water storage over Europe using a fully coupled terrestrial model. *Water Resources Research*, 57, e2020WR027828. <https://doi.org/10.1029/2020WR027828>

Received 28 APR 2020

Accepted 9 DEC 2020

© 2020. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

An Interannual Probabilistic Assessment of Subsurface Water Storage Over Europe Using a Fully Coupled Terrestrial Model

Carl Hartick^{1,2} , Carina Furusho-Percot^{1,2} , Klaus Goergen^{1,2} , and Stefan Kollet^{1,2} 

¹Agrosphere (IBG-3), Research Centre Jülich, Jülich, Germany, ²Centre for High-Performance Scientific Computing in Terrestrial Systems, Geoverbund ABC/J, Jülich, Germany

Abstract The years 2018 and 2019 were two of the hottest and driest in Mid-Europe, highlighting the need for a comprehensive assessment of available water resources. In this study, we propose a probabilistic, terrestrial water assessment method, which utilizes a terrestrial forward model that closes the coupled water and energy cycles, from groundwater to the top of the atmosphere. In this methodology, the model is initialized with the current state of the water year and forced with a climatologic ensemble of atmospheric forcing to account for atmospheric uncertainty and natural variability. The simulations result in an ensemble of ensuing water years that are analyzed for subsurface water storage anomalies. The methodology was applied to the water years 2011/2012 and 2018/2019 and showed an increased probability of a significant water deficit in regions that had a water deficit in the previous year. This was also observed in an evaluation simulation. The results were compared to simulations with perfect forcing and uncertain initial conditions, and showed predictability at the interannual timescale and beyond, depending on the strength of the anomaly. The methodology was then applied to 2019/2020 to provide an outlook of the evolution of the current anomalies. The results emphasize the importance of accounting for groundwater dynamics in applied terrestrial models to account for long-term memory effects in the terrestrial water cycle in forward simulations, over large spatial scales. This method of probabilistic subsurface water storage assessment may provide crucial information to public and industrial sectors for long-term water resource planning.

Plain Language Summary In central and northern Europe, 2018 was an extremely hot and dry year, raising questions about water availability and security. There is scientific and public interest in determining whether water storage, especially from deeper soil, can recover after 1 year, and how likely a continuous water deficit is. This question could be answered by modeling, which captures the whole water cycle from clouds to groundwater, and simulates the strong memory effect of water storage. Together with the initial state of the terrestrial system in the previous water year, and a large number of possible scenarios from climatologic information capturing the possibilities of how the weather can evolve, we assessed the water deficit at the end of the following year. The results suggest that a substantial deficit from the outset resulted in a significant probability that the water deficit would continue, which was indeed the case. This approach is useful in situations where there is a strong deficit, or an abundance of water, and could contribute to water resource assessment and planning for public and private sectors, on a continental scale.

1. Introduction

Droughts associated with heatwaves are difficult to predict, over the seasonal timescale and beyond, mainly due to the uncertainty of atmospheric forcing (Miralles et al., 2019). However, memory effects, related to the slow dynamics of groundwater-soil water-vegetation systems, mean that the challenge is an initial value problem, which may potentially lead to increased predictive ability over extended periods (Dirmeyer, 2000). The major assumption is that memory effects, including interactions of the groundwater-soil water-vegetation system, under uncertain atmospheric forcing conditions are correctly simulated in the forward model.

In recent years, there have been many efforts to establish a reliable seasonal forecasting system, applying various approaches, for example, using statistical tools such as linear stochastic models (e.g.,

Mishra & Desai, 2005), Bayesian frameworks (e.g., Madadgar & Moradkhani, 2013) or the collection and transformation of multiple datasets (e.g. Hao et al., 2014). These approaches rely on drought indices, derived from as much observational data as possible, to develop probability density functions. Climate models with a dynamical core can also be used (e.g., Yuan et al., 2013; Yuan & Wood, 2013); however, predictions from climate models alone do not generally produce accurate results (Hao et al., 2018; Yuan et al., 2015b). Recent studies, therefore, have used machine learning methods to combine models with observational data (Rhee & Im, 2017; Hosseini-Moghari & Araghinejad, 2015). An experimental combination of hydrological modeling and regional climate forecasts in the Sub-Saharan region has also been successfully used (Sheffield et al., 2013). Sheffield et al. (2013) used reanalysis data to calculate a climatology with a hydrological model. The current state is modeled with remote sensing data, while the forecast is then calculated with a climate model to provide predictions for up to 6 months. In Europe, important research from the European drought observatory (Vogt et al., 2018) and the German drought monitor (Zink et al., 2016), have used observational data to describe the current drought situation. For forecasts over seasonal timescales, the Ensemble Streamflow Prediction (ESP) approach, or seasonal forecasts provided by European Center for Medium-Range Weather Forecasts, ECMWF (Arnal et al., 2018; Turco et al., 2017) can be used. There are even efforts to derive long-term drought indicators from long-term climate simulations (Solaraju-Murali et al., 2019).

In recent years, droughts have frequently occurred in Mid-Europe (ME); 2018 and 2003 were particularly dry and hot anomalies. While in 2003, the ecosystem was able to rapidly recover, the ecosystem effects of 2018 persisted until 2019 (Hari et al., 2020). Both events were triggered by an Omega block in the atmosphere, although they differ in their specific pattern. In 2003, the lack of soil moisture increased the drought conditions, especially in southern Europe, while in 2018, the number of clear sky days was most influential (Liu et al., 2020). Using data from the satellite product GRACE (Tapley et al., 2019), the missing water in 2018/2019 was estimated to be up to 145 Gt (Boergens et al., 2020). In this study, we quantify droughts as a significant anomaly in the subsurface water storage, compared to the individual monthly means obtained by climatology.

The importance of soil moisture and groundwater in the occurrence of extreme temperature events and droughts has been demonstrated (e.g., Hirschi et al., 2011; Seneviratne et al., 2010). However, at the continental scale, weather forecasts or climate models do not provide the complete water cycle and oversimplify the groundwater-soil water-vegetation system, generally neglecting groundwater dynamics.

While hydrological models appear to be a viable alternative, at the continental scale, the groundwater compartment and soil water coupling are also simplified (Zink et al., 2016), which may adversely affect the simulation of memory effects (Lo & Famiglietti, 2010).

To relax simplifying assumptions, and to study the predictability of subsurface water storage anomalies, we propose the application of an integrated terrestrial modeling approach for subsurface water storage anomaly assessment. In this context, integrated terrestrial modeling refers to the representation of the complete terrestrial hydrologic and energy cycle, from groundwater across the land surface into the atmosphere. Following on from the early work of York et al. (2002) and Maxwell et al. (2007), significant advances have been made in coupling subsurface, land surface, and atmospheric models, and in providing physically consistent states and fluxes throughout the terrestrial system. At the continental scale, which is of interest in this study, studies have successfully applied groundwater parameterizations, and simulated the terrestrial cycle over the Amazon, and globally at relatively coarse spatial resolutions (Miguez-Macho & Fan, 2012; Walco et al., 2000). Keune et al. (2016) used the Terrestrial Systems Modeling Platform (TSMP), including a 3D variably saturated groundwater representation, in a continuum approach to show that surface-atmosphere feedbacks are well-captured, and that groundwater actually mitigated the 2003 heatwave over Europe. Keune et al. (2018) also demonstrated that human water use, related to groundwater abstraction and irrigation, may systematically change the distribution of water resources, due to local and non-local subsurface-land surface-atmosphere feedbacks.

In this study, we propose an interannual probabilistic assessment of subsurface water storage, in combination with the TSMP across the European continent, utilizing past meteorological information. With this approach, all subsurface water resources for the entire continent, down to the bedrock, can

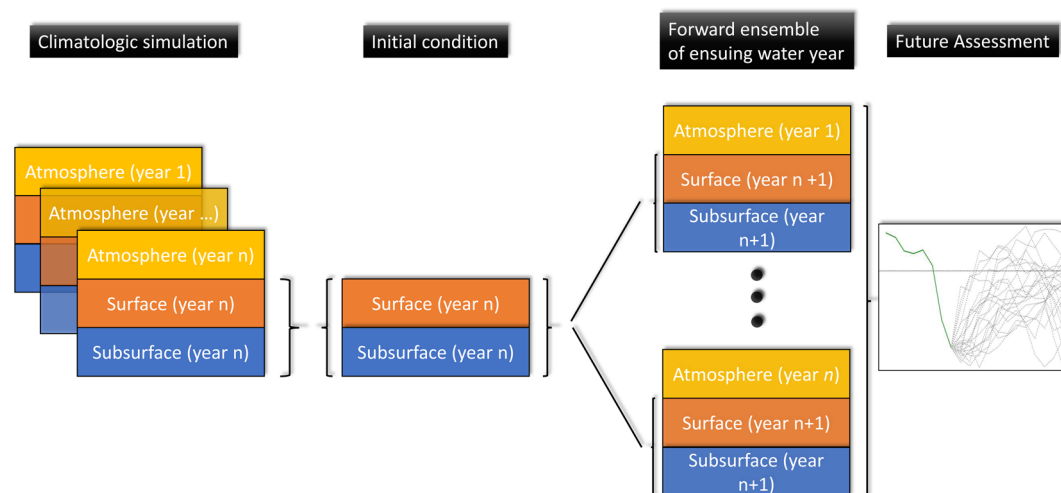


Figure 1. Schematic of the proposed methodology for a probabilistic interannual subsurface water assessment. The initial surface and subsurface conditions are obtained from a climatologic simulation and are used to initialize an ensemble of simulations forced by atmospheric boundary conditions of all available years (n) from climatology data. Individual ensemble members evolve independently until the end of the subsequent water year. The results from the forward ensemble are used for the probabilistic future assessment of subsurface water.

be included in a physically consistent simulation. This methodology may be especially valuable in the assessment of the impact of extreme climatic conditions on subsurface water storage anomalies. In the following sections, the approach is explained in detail, followed by an application at the continental scale. The results are discussed in the context of water scarcity probabilities at the interannual timescale, and the impact of memory effects on the assessment results.

2. Methods

In the following section, we describe the proposed methodology to assess anomalies in subsurface water storage (over the full subsurface column), the relevant components of the terrestrial forward model applied in this study, the experimental setup of the probabilistic assessment and its reverse, to determine the predictability and relative impact of atmospheric and hydrologic initial conditions uncertainty.

2.1. Interannual Probabilistic Assessment of Subsurface Water

Figure 1 provides a schematic overview for the probabilistic interannual subsurface water assessment method. The state of the terrestrial system is simulated with an integrated terrestrial model (forward model) that closes the water and energy cycle, from groundwater across the land surface to the top of the atmosphere. Importantly, groundwater dynamics are incorporated in a PDE-based, 3D continuum approach. In the interannual assessment, an ensemble of forward simulations for the following water year is initialized with the land surface and subsurface conditions from the same forward model at the end of the previous water year. The subsequent water year is then simulated by applying the atmospheric boundary conditions from atmospheric reanalysis information from all available previous years. The result is an ensemble of ensuing water years, based on the available climatology of atmospheric boundary conditions. The assessment accounts for climatologic uncertainty and natural variability, as atmospheric processes may not be predictable at the interannual timescale. In this study, the probabilistic assessment was performed at the interannual time scale, over a complete water year. In other applications, the scale of the lead time and the month of initialization may vary.

2.2. Terrestrial Systems Modeling Platform (TSMP)

In this study, we applied TSMP (Gasper et al., 2014; Shrestha et al., 2014) as the forward model. However, based on the proposed methodology (Section 2.1 and Figure 1), any integrated terrestrial model may be applied that accounts for groundwater dynamics and memory effects, and couples to the land surface and atmosphere. TSMP closes the terrestrial water and energy cycle, from groundwater across the land surface into the atmosphere, by coupling the atmospheric model COSMO v5.01 with the land surface model CLM v3.5 and the groundwater model ParFlow v3.2, via the Ocean Atmosphere Sea Ice Soil Model Coupling Toolkit (OASIS3-MCT). Below, we provide a brief overview of the different components of the TSMP. See Shrestha et al. (2014) and Gasper et al. (2014) for further details.

The nonhydrostatic model COSMO has been developed by a consortium of weather services, under the leadership of the German weather service (Baldauf et al., 2011). In different configurations, COSMO can serve as an operational weather forecast model, or for regional climate simulations. COSMO solves the primitive Euler-equations and includes multiple schemes for precipitation, shallow convection, radiation, and turbulence closure. The atmospheric state at the lowest level, including precipitation, temperature, humidity, and horizontal wind is coupled to the Community Land Model (CLM).

The CLM consists of shallow soil, snow layers, land cover and vegetation, and handles the interaction with the atmosphere (Oleson et al., 2004, 2008). The CLM parameterizes hydrological, biological, and radiation processes, such as evapotranspiration, sensible, and ground heat flux. In the TSMP, the CLM supplies COSMO overland with the boundary conditions of surface albedos, latent and sensible heat flux, zonal momentum flux, and outgoing longwave radiation. Vegetation is described by 16 different Plant Functional Types (PFTs). In the TSMP, the hydrologic component of the CLM is completely replaced by ParFlow.

The hydrological model, ParFlow (Ashby & Falgout, 1996; Jones & Woodward, 2001; Kollet & Maxwell, 2006; Maxwell, 2013) solves the 3D Richards equation with a Newton-Krylov solver to model integrated variably saturated groundwater-surface water flow. In ParFlow, the Richards equation is discretized in space, using finite differences and an implicit backward Euler scheme in time. Overland flow is modeled by solving the kinematic wave equation in a finite volume approach. ParFlow receives the incoming precipitation after canopy interception, as well as the water loss from evapotranspiration from the CLM. In turn, ParFlow provides the CLM with the hydrologic state, in terms of soil moisture and matric potential.

In the TSMP, the coupler OASIS3-MCT (Valcke, 2013) connects the different component models in the form of a linked library, based on a Multiple Process Multiple Data (MPMD) approach. OASIS3-MCT acts as the driver initializing the models, managing the time steps and coupling frequencies, exchanging the coupling data in 2D arrays in memory, and finally terminating the simulation.

2.3. Model Domain and Setup

In this study, the proposed methodology was applied to the European continent. The model domain (Figure 2), was implemented according to the Coordinated Regional Downscaling Experiment (CORDEX) (Giorgi et al., 2009; Gutowski et al., 2016), with subdomains according to Christensen and Christensen (2007). The domain covers all of Europe, based on a rotated latitude-longitude grid, with a horizontal resolution of 0.11° . This results in a lateral resolution of approximately 12.5 km. COSMO has a vertical range of 22 km subdivided into 50 levels. The applied time step size was 60 s. While CLM has 10 soil layers ranging to a depth of 3 m, ParFlow has five extra layers, covering a total depth of 57 m below the surface; the first 10 layers of CLM and ParFlow are identical. In ParFlow (and CLM), there is a variable vertical discretization, ranging from 2 cm at the land surface to ≥ 10 m toward the bottom of the aquifer, based on a terrain-following grid. The CLM and ParFlow use a time step size of 900 s, which also represents the coupling frequency between the subsurface, land surface and atmosphere.

The topographic slopes required by ParFlow were estimated from the USGS GTOPO30 (DAAC, 2004). Lateral boundary conditions along the coast are defined by a constant hydraulic pressure (Dirichlet) with a hydrostatic vertical profile in ParFlow. The soil parameters of the ParFlow model were estimated with the help of the Food and Agricultural Organization (FAO) database (Carballas et al., 1990) and were assumed

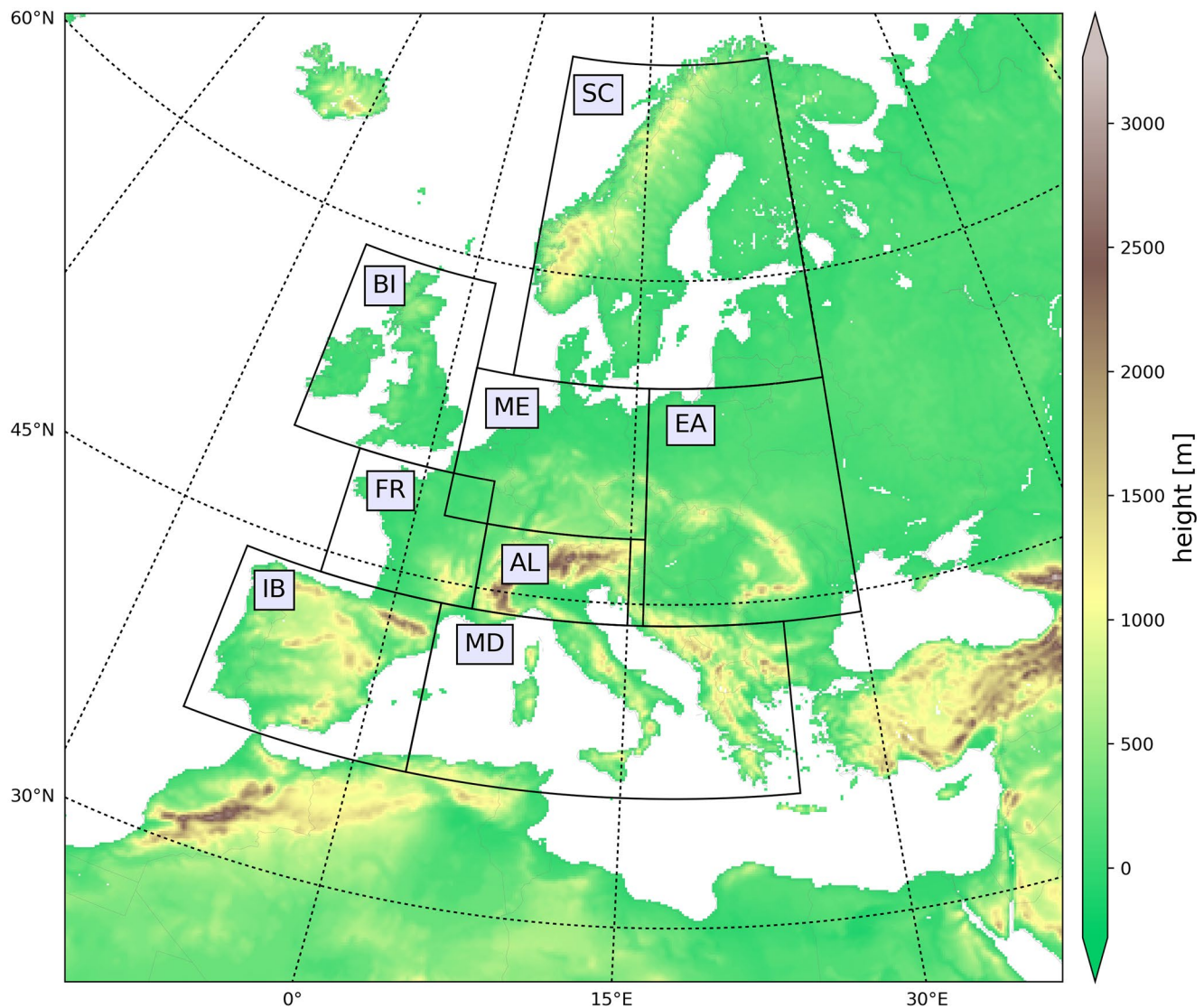


Figure 2. The European CORDEX model domain, including the topographic height. Each box represents a different geographical PRUDENCE region, with BI, British Isles; IB, Iberian Peninsula; FR, France; ME, Mid-Europe; SC, Scandinavia; AL, Alps; MD, Mediterranean; EA, Eastern Europe. CORDEX, Coordinated Regional Downscaling Experiment.

to be vertically homogenous. To achieve this, 15 soil types are defined based on the texture information. To account for information loss due to spatial aggregation and anisotropy, the values of the horizontal permeability were scaled by 1,000. This concept was introduced by Niedda (2004) and improves the results in coupled modeling (Fang et al., 2016). For the CLM, PFTs were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) database (Friedl et al., 2002). The individual values for the stem and leaf area index and the monthly bottom and top heights of the PFT were calculated with the global CLM surface dataset. Additional details are available from Furusho-Percot et al. (2019).

2.4. Subsurface Water Storage Assessments of 2011/2012, 2018/2019, and 2019/2020

The proposed methodology was applied and evaluated in probabilistic hindsight assessments of the water years 2011/2012 and 2018/2019. Additionally, an assessment for 2019/2020 was performed. In the assessments, the terrestrial state at midnight, August 31 of the corresponding water year was used for the initial condition. The initial terrestrial states were obtained from a deterministic climatological Evaluation

Simulation (EVAL), starting in 1989 (Furusho-Percot et al., 2019), based on atmospheric boundary conditions from the ERA-Interim atmospheric reanalysis (Dee et al., 2011). Furusho-Percot et al. (2019) demonstrated that the climatology data showed a good agreement with the E-OBS dataset (Cornes et al., 2018) and ERA-Interim for precipitation and 2 m air temperature anomalies. In addition, the climatology of subsurface water storage showed good agreement with data derived from GRACE (Watkins et al., 2015). We expanded the comparison, using the novel GRACE-REC dataset (Humphrey & Gudmundsson, 2019), which constitutes a historic reconstruction of GRACE from 1901 to the present day, at daily and monthly time scales. GRACE-REC reconstructs historical total water storage (*tws*) changes with a statistical model trained with data from GRACE and meteorological reanalysis products. We compared EVAL with an average of the GRACE-REC datasets that were created with the new ERA5 reanalysis product of the ECMWF.

Based on the initial state on August 31 of the assessment year, ensemble simulations were performed applying the ERA-Interim lateral atmospheric boundary conditions, of the years 1996–2018/2019. The original boundary conditions of the assessment year are left out. Accordingly, the ensemble simulation resulted in 22 possible outcomes for the water year. For the evaluation of the hindsight assessment in 2018/2019, EVAL was continued by applying atmospheric boundary conditions from the ERA-Interim atmospheric reanalysis. In the case of the 2019/2020 assessment, no comparable deterministic run was available, because the EVAL climatology could not be extended due to the cessation of the ERA-Interim by ECMWF.

2.5. Evaluation of Predictability and the Memory Effect

The proposed method assumes that there is a strong memory effect of slow subsurface water dynamics, resulting in predictability, despite strong uncertainty in the atmospheric forcing. This method is similar to a traditional ensemble streamflow prediction (Day, 1985; Wood & Lettenmaier, 2008; Wood et al., 2002). Therefore, we quantify the influence of the initial condition with an additional experiment that is the reverse ensemble prediction or assessment (e.g., Wood & Lettenmaier, 2008): all available initial conditions are forced with the assumed perfect atmospheric forcing of the assessment year. In this study, the reverse ensemble was performed by reducing the TSMP to the coupled CLM and ParFlow and applying the atmospheric forcing of the assessment year to all available initial climatology conditions in offline hydrologic simulations.

While an ensemble assessment provides predictability via the known initial conditions and continuous divergence over time, a reverse assessment offers predictability because of the forcing, resulting in a convergence of all ensemble members with time. To quantify the relative influence of the initial condition and atmospheric forcing on predictability, Wood and Lettenmaier (2008) proposed estimating the error ratio of both ensembles. In this study, the Root Mean Square Error (RMSE) of all individual ensemble members and the Mean Absolute Error (MAE) were used: the RMSE takes outliers into account, while the MAE focuses on the ensemble mean. Thus, with those two errors, we present a possible upper and lower error bound with regard to the impact of outliers in the ensemble. The use of the MAE is justified because this method focuses on drought situations and longer timescales, where the ensemble naturally diverges. For the error ratio, the respective error of the assessment is divided by the error of the reverse assessment. If the error ratio is greater than one, the influence of atmospheric forcing is assumed to be more important; otherwise, the initial condition dominates. Previous studies have shown that the initial soil moisture condition is the dominating factor for up to three months (Shukla et al., 2013), and sometimes up to 7 months in Europe (Arnal, 2018; Staudinger & Seibert, 2014).

2.6. Analyses

Data were extracted as spatially averaged monthly mean values of the following variables: 2 m air temperature, *tas* (K); precipitation, *pr* (L); and total column water storage, *s* (L) from the land surface to the bottom of the aquifer. The latter is an integrated measure of water resources and was calculated from the simulated relative saturation $sat_{i,j,k}$ (–), the porosity $\phi_{i,j,k}$ (–), for a pixel with indices i,j,k in the lateral and vertical direction, respectively, where dz_k (L) is the vertical extent of each model layer that is laterally constant, and nz is the number of grid cells in the vertical direction:

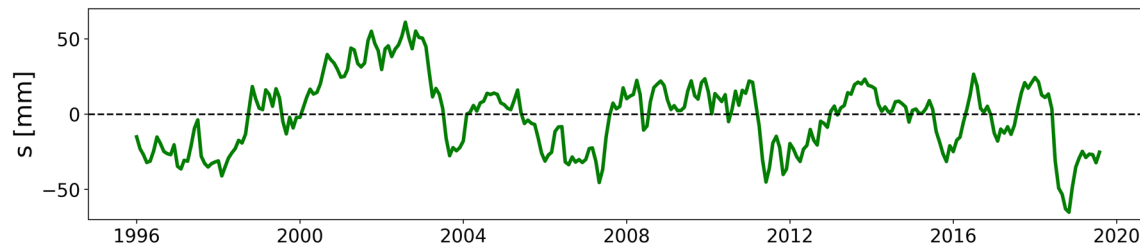


Figure 3. Time series of anomaly of monthly column storage, s , averaged over Mid-Europe (ME) between 1996 and 2019 from the climatology EVAL. EVAL, Evaluation Simulation.

$$S_{i,j} = \sum_k^{nz} sat_{i,j,k} \phi_{i,j,k} dz_k \quad (1)$$

This gives a water storage estimate for every subsurface column in the model domain. For all variables and grid elements over land, reference monthly means were calculated for 1996–2018. The reference values were subtracted from every monthly output to obtain monthly anomalies, which were spatially averaged over the corresponding region. To calculate the probability of a water deficit, the Weibull plotting position formula was used (Hao et al., 2016):

$$P(x < 0) = \frac{n_0}{n + 1} \quad (2)$$

where P is the probability that a negative/positive water storage anomaly exists; n_0 is the number of ensemble members that fulfill this condition; and n is the total number of ensemble members.

3. Results and Discussion

3.1. Climatology of Subsurface Water Storage Anomalies Over Europe

The evolution of the subsurface water storage anomaly, s , from EVAL, which leads to the initial state of the hindsight assessment for ME is shown in Figure 3. Real world events, like the 2002 Elbe-Danube flood, with a high positive anomaly, and drought events like those of 2003 and 2018 are clearly distinguishable. The subsurface water storage anomaly ranges from -60 to 60 mm, while extreme droughts are characterized by an interannual change of up to 80 mm. In terms of a negative anomaly, 2018 is exceptional, although 2003 would probably be on the same level without the Elbe-Danube flood in the previous year. An overview of all PRUDENCE regions is included in the supporting information S1.

Figure 4 provides the results of the spatial analyses of the s anomaly over Europe. In August 2018, in EVAL (Figure 4a), the water deficit is significant, especially in the ME region. In August 2019 (Figure 4b), the situation is more heterogeneous. Many dry spots persisted, while water scarcity increased in southwest Europe and recovered in east Europe. Figure 4 also emphasizes the strong spatial variability from the regional to the continental scale, which depends on multiscale heterogeneity in physical parameters and fluxes, especially related to precipitation. At the smallest spatial scale on the order of the model's resolution, the anomaly patterns must be treated with caution due to the uncertainty, for example, related to local rainfall amounts in the transient, coupled simulations.

Furusho-Percot et al. (2019) compared EVAL to Total Water Storage (tws) derived from GRACE: in general, EVAL is able to capture the evolution of tws anomalies, including subsurface and surface water. We extended the comparison for the whole EVAL period, with the tws anomalies of the GRACE-REC dataset (Humphrey & Gudmundsson, 2019). Figure 5 compares the mean of the dataset derived from GRACE-REC created with ERA5 including the seasonal cycle to the climatology of Furusho-Percot et al. (2019). EVAL well captures the general temporal evolution of water resources, especially in ME and FR. Differences in the Mediterranean region (MD) between 1996 and 2000 could be due to spin-up effects. For the

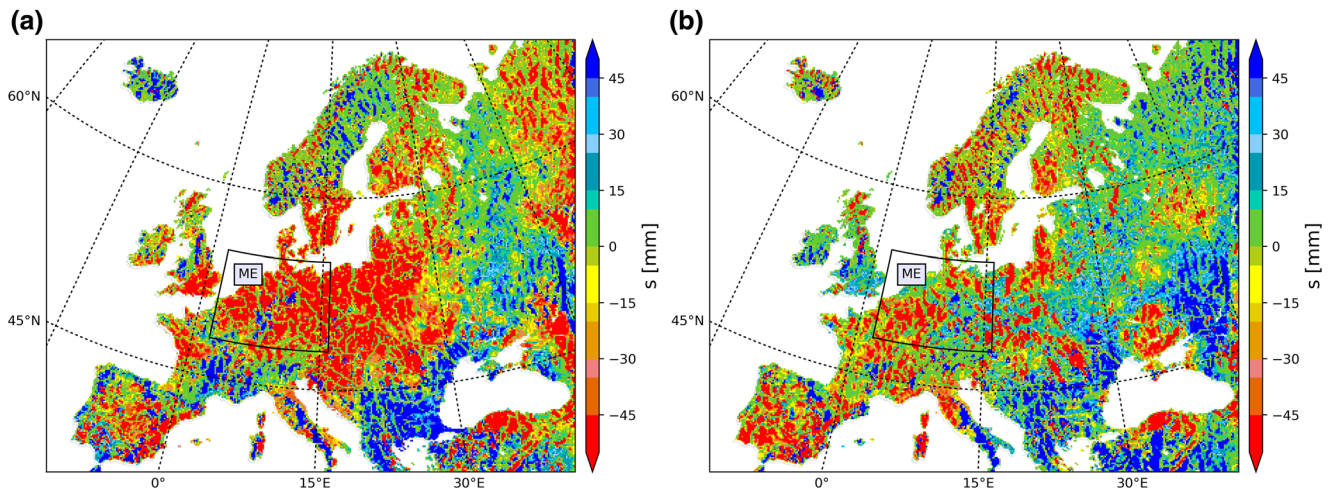


Figure 4. Subsurface monthly water storage anomaly, s , over the European model domain, with a focus on the Mid-Europe (ME) region for EVAL in (a) August 2018, and (b) August 2019. Red symbolizes a strong water deficit (drought), blue water abundance. Regions covered in green show nearly no anomaly compared to the climatology. EVAL, Evaluation Simulation.

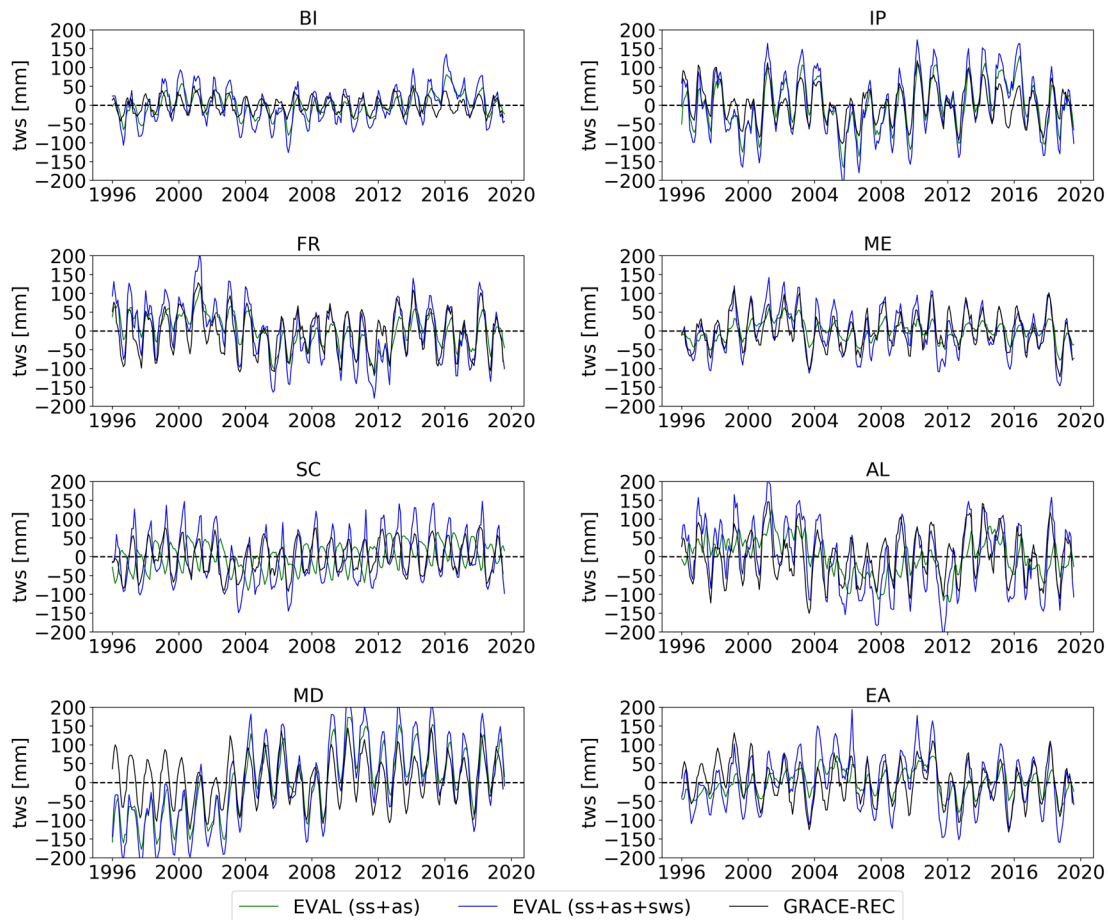


Figure 5. Anomalies of total water storage, tw_s , for the entire soil column including soil water, ss , and aquifer, as , storages ($ss + as$) and surface water storage including snow ($ss + as + s_{ws}$) based on the GRACE-REC datasets created with ERA5 and the EVAL climatology for different PRUDENCE regions (BI, British Isles; IP, Iberian Peninsula; FR, France; ME, Mid-Europe; SC, Scandinavia; AL, Alps; MD, Mediterranean; and EA, Eastern Europe). EVAL, Evaluation Simulation.

Table 1

Pearson Correlation, Bias and RMSE Over the EVAL Period for all PRUDENCE Regions for Total Water Storage With GRACE-REC as Reference

Region	Pearson correlation	Bias (mm)	RMSE (mm)
BI	0.69	0.08	35.64
IP	0.86	−3.17	46.84
FR	0.88	3.83	36.65
ME	0.89	−0.46	26.57
SC	0.86	0.69	36.33
AL	0.82	0.02	50.88
MD	0.75	−5.64	73.87
EA	0.78	−0.65	44.70

regions SC (Scandinavia) and AL (Alps) snow is important for the estimation of the total water storage. EVAL includes values for snow but only at the beginning of each month. Nevertheless, a good agreement is reached. Additionally, the fully coupled simulations were performed with a transient atmosphere, for example without re-initialization or spectral nudging, therefore, while the model results are reasonable and internally consistent, biases are expected. For every region, Table 1 shows the Pearson correlation coefficient, the bias and the RMSE resulting from the comparison with GRACE-REC.

3.2. Hindsight Assessment and the Impact of Interannual Memory Effects

Figure 6 depicts the ME analysis of monthly anomaly time series for pr (Figure 6a), tas (Figure 6b), and s (Figure 6c) from EVAL between January and the end of August 2018, and the resultant probabilistic assessment of the water year 2018/2019. The heatwave is visible with negative precipitation (Figure 6a) and positive temperature anomalies (Figure 6b)

in EVAL, E-OBS, and ERA-Interim (data shown for reference). However, at the end of 2019, EVAL exhibits a wet bias compared to the other datasets. The influence of precipitation and temperature on the subsurface water storage, s , is apparent (Figure 6c). The s anomaly changed from a wet (positive) anomaly to a strongly dry (negative) anomaly in August 2018 over the ME region. The hydrologic state at the end of August 2018 then served as the initial condition for the probabilistic assessment of 2019, based on the ensemble atmospheric boundary conditions of 1996/1997 to 2017/2018. In Figure 6, the individual ensemble members are plotted in dashed lines.

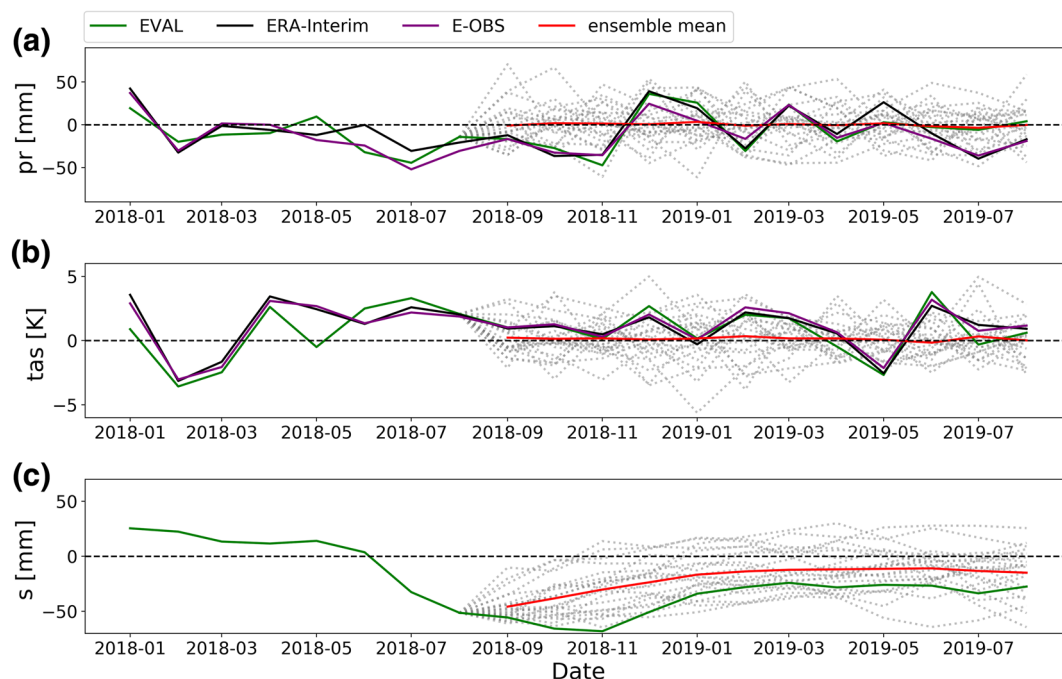


Figure 6. Monthly (a) precipitation anomalies, pr ; (b) temperature anomalies, tas ; and (c) subsurface water storage anomalies, s , from January 2018 to August 2019 in the Mid-Europe (ME) region. The deterministic climatology EVAL (green line) is supported by the evolution of ERA-Interim and E-OBS anomalies for pr and tas (black and purple line). The time period of the hindsight assessment of the water year 2018/2019 includes all ensemble members (dashed lines).

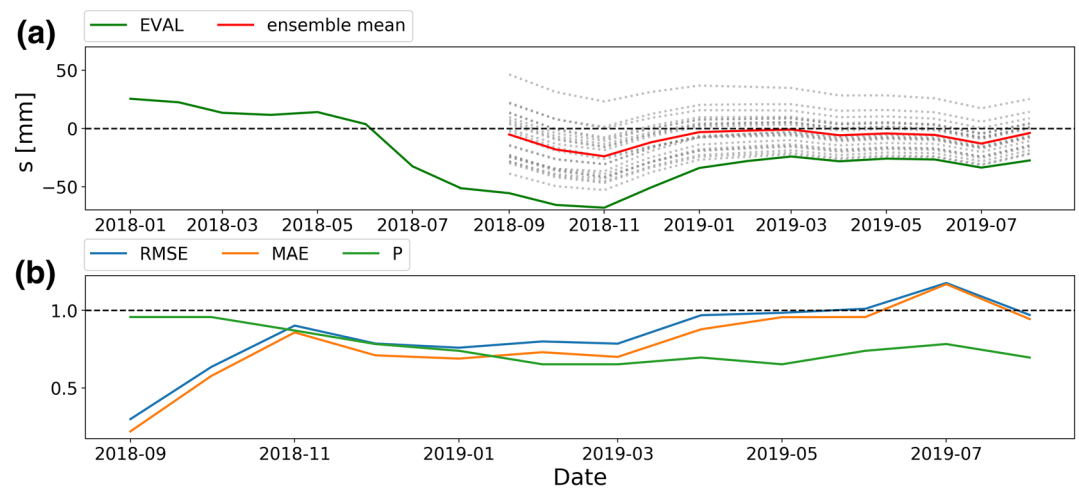


Figure 7. Monthly subsurface water storage anomalies, s , of EVAL and the ensemble members in Mid-Europe (ME) for the reverse assessment (a); and the resulting monthly Mean Absolute Error (MAE) ratio, Root Mean Square Error (RMSE) ratio combining assessment and reverse assessment and Weibull probability P , for a dry anomaly over the assessment period in ME (b).

The ensemble mean of s shows that there is initially a significant water deficit, which then stabilizes in August 2019 (mean -14.96 mm, SD 22.1 mm). While most ensemble members exhibit a significant reduction of the dry anomaly until the beginning of summer 2019, the Weibull probability estimation for s leads to a P of approximately 70% of a continuing water deficit. In contrast, the Weibull probability for pr is approximately 50%, reflecting the independence of precipitation on the initial conditions at the interannual timescale. The difference in P values for s and pr is a measure of the strong impact of memory effects of subsurface hydrodynamics on the temporal evolution of water storage anomalies, even at the interannual time scale.

Comparing the ensemble to EVAL, the actual simulated moisture conditions are in the extreme quantiles of the ensemble in late 2018 and were drier at $s = -27.49$ mm in August 2019. This emphasizes the strength of the 2018 drought; since the start of the analysis in 1996, there have not been such continuously dry conditions due to the continuous lack of precipitation following the initial dry conditions. The ensemble mean of the precipitation and temperature anomaly are close to zero reflecting the negligible impact of the initial condition.

As previously mentioned, a strong memory effect of groundwater can be identified at the interannual timescale, with P being highly dependent on the initial condition. This can be demonstrated by performing the reverse assessment, with a “perfect” forcing, applying all possible initial conditions. Figure 7a illustrates the range of initial conditions and the 55% reduction in the spread during the assessment period, which reduces the error of the ensemble over time. In theory, the error should eventually reach zero, however, in the assessment period the convergence of the ensemble is relatively small, because of the pronounced memory effect of subsurface water dynamics. This is consistent with the findings of Parry et al. (2016), who investigated and reviewed studies on the cessation of long-term droughts, using in situ observations from groundwater observation wells. To quantify the error, and to determine if the initial condition or atmospheric forcing dominates the predictability of the assessment, we calculated the MAE of the ensemble to EVAL, and the RMSE of both experiments, with associated error ratios. In the error ratio, the reverse assessment is the denominator. Figure 7b shows the time series of the error ratios and the evolution of P during the assessment period. While P reduces from nearly 100% to 72%, MAE and RSME fluctuate around one, for the whole assessment period. Thus, the initial condition is more important than the forcing. This finding is in contrast to Shukla et al. (2013); where the RSME ratio over Europe is significantly larger than one, after 6 months lead time, suggesting that important processes leading to memory, and therefore predictability in the assessment, have been neglected. The error ratios of the RSME are more sensitive to outliers, which may occur because the forward simulations were performed with a fully transient atmosphere, without spectral

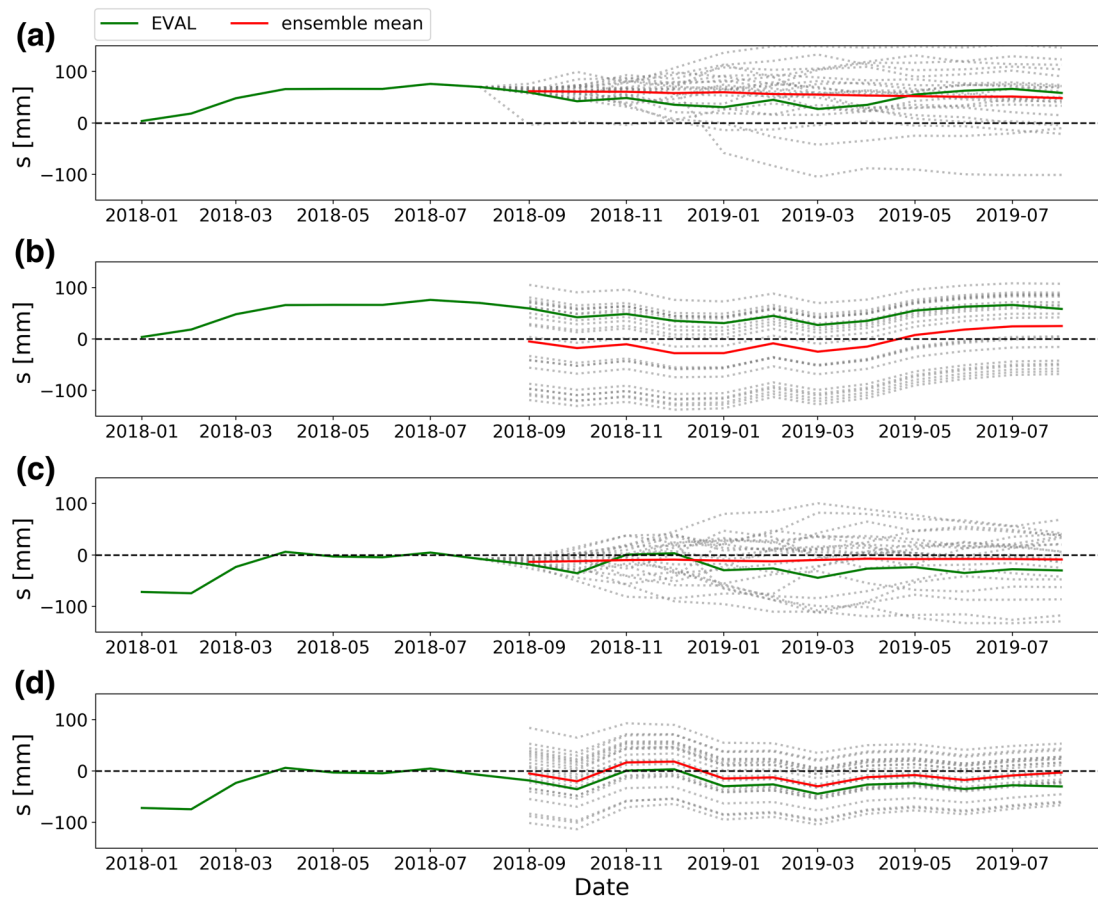


Figure 8. Monthly subsurface water storage anomalies, s , between January 2018 and August 2019 for (a) the assessment in the Mediterranean (MD) region; (b) the reverse assessment in the MD region; (c) the assessment in the Iberian Peninsula (IP); and (d) the reverse assessment in the IP.

nudging and re-initialization. The true error ratio, therefore, probably falls between the RSME and MAE ratios. The RSME and MAE error ratios suggest that the assessment provides useful information at the interannual time scale, and potentially beyond, which represents a major improvement to previous lead times (Hao et al., 2014; Shukla et al., 2013; Yuan et al., 2015a).

Over Europe, there are regions with a positive anomaly, or no anomaly during the assessment period. In the latter case, there should be no predictability in terms of s anomalies, following the above rational. This is seen in the Mediterranean (MD) region and the Iberian Peninsula (IP) for the assessment and the reverse assessment in Figure 8. In the MD (Figures 8a and 8b), there is a large positive anomaly that is quite stable in our simulation. Similar to the case of a strong water deficit, the assessment results in a Weibull probability of $P = 74\%$ that the positive anomaly persists. The error ratios are below one for the whole period (S3), therefore, the initial condition dominates. For the IP (Figures 8c and 8d), there is no anomaly, therefore the assessment exhibits no predictability ($P \sim 50\%$), and the error ratios are larger than one for lead times over 5 months (S3). This suggests that the predictabilities are strongly dependent on the strength of the anomaly at the beginning of the assessment.

A spatial analysis of the whole domain is shown in Figure 9. For every pixel, RMSE and MAE ratios were calculated for 3, 6, 9, and 12 months lead times. Overall there is no significant difference between the two error ratios but there are spatial clusters for both ratios that mostly contribute to regional predictability. Time series of error ratios and P values for all PRUDENCE regions are provided in the supporting information (S2–S4). There is no clear difference in error ratios with respect to climatic regions, because the strength of the anomaly determines the predictability, not the average ambient atmospheric and hydrologic conditions.

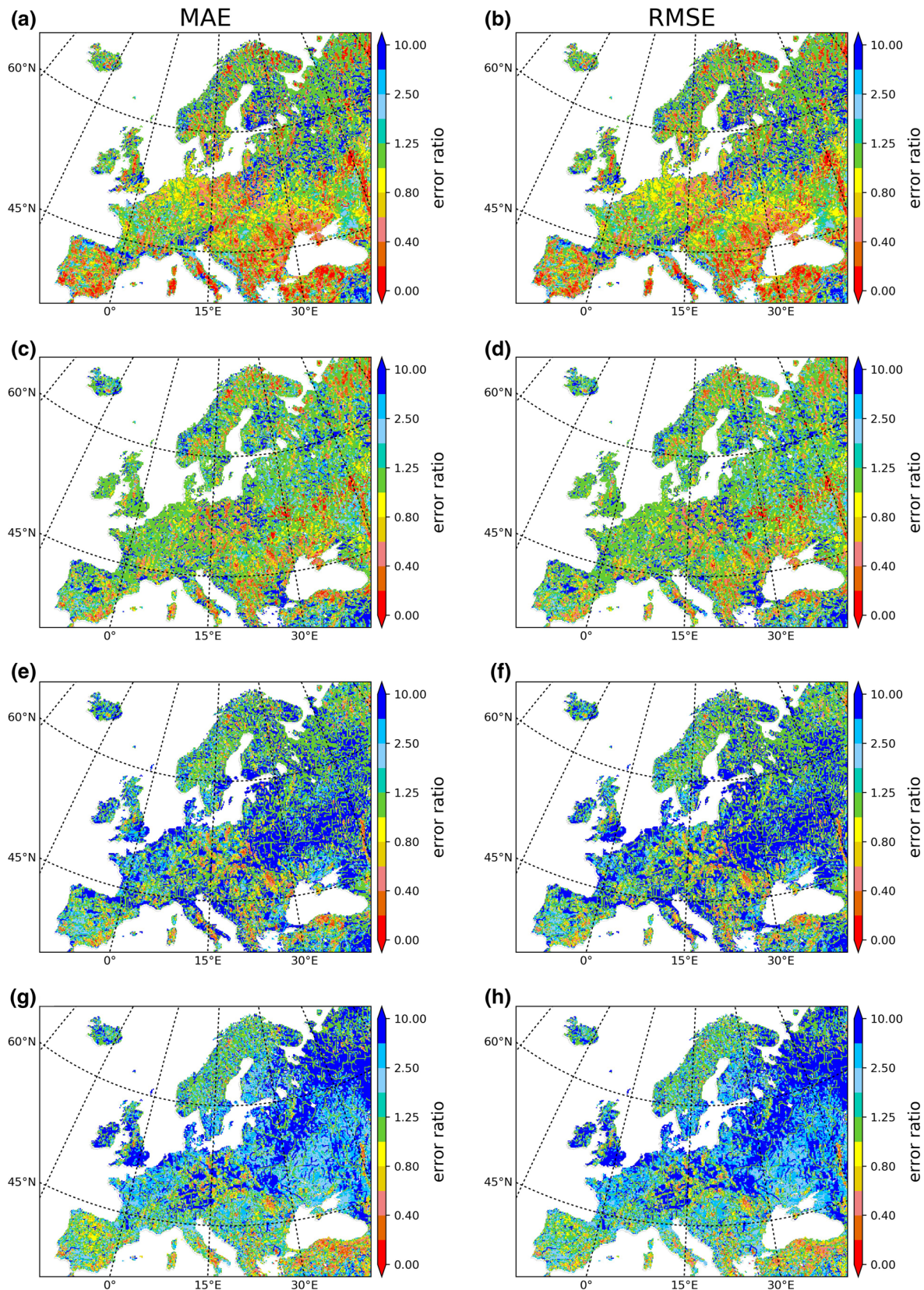


Figure 9. Error ratios of the Mean Absolute Error (MAE, left column) and Root Mean Square Error (RMSE, right column) over the European domain for (a and b) 3 months; (c and d) 6 months; (e and f) 9 months; and (g and h) 12 months lead time.

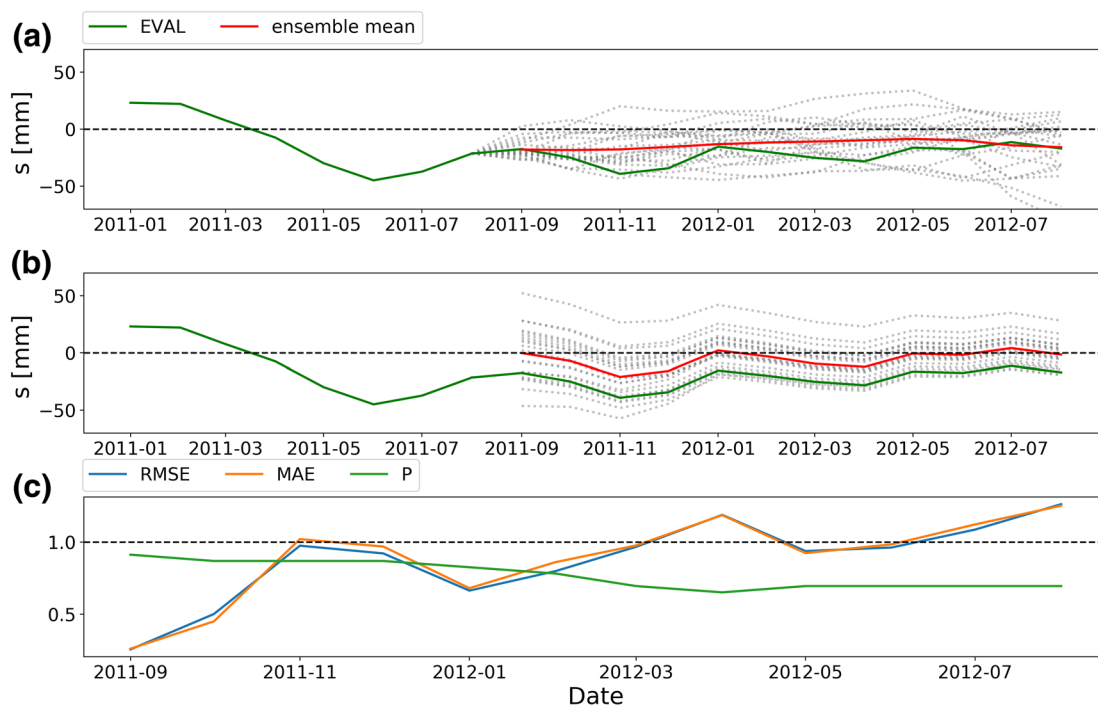


Figure 10. Monthly subsurface water storage anomalies, s , for January 2011 to August 2012 for (a) the assessment; (b) the reverse assessment in Mid-Europe (ME); and (c) the Root Mean Square Error (RMSE) ratio and the Mean Absolute Error (MAE) ratio of the ensemble mean to the reference EVAL are depicted, together with the probability for a water deficit, P . EVAL, Evaluation Simulation.

For validation, the assessment and the reverse assessment were repeated for 2011/2012, with the initial condition of August 2011 retrieved from EVAL. The results are shown in Figure 10. Similar to 2018/2019, ME was under drought conditions. Overall, the assessment (Figure 10a) provides a good indication of the situation in August 2012, resulting in P of approximately 70% for a continuing water deficit. MAE and RMSE ratios are smaller or close to one also at the interannual timescale corroborating the dominating impact of the initial condition (Figure 10c). Maps over Europe of the error ratios for 2011/2012 are provided in the supporting information (S5).

3.3. Future Assessment of 2019/2020

In the final step, the assessment was performed for 2019/2020, with the initial condition August 2019 (Figure 11), to determine the probability of the drought continuing into the future. In contrast to 2018, the continuation of the deterministic climatology EVAL, based on ERA-Interim atmospheric boundary information, shows a continuous water deficit in s throughout 2019 until August (Figures 11c). August 2019, the initial condition, is not as dry as August 2018, due to the underestimation of the pr anomaly in EVAL in July 2019 (Figures 11a). Figures 11a and 11b show an especially hot and dry summer in ERA-Interim, which is less pronounced in EVAL, which might lead to an underestimation of the water deficit in the initial condition. The ensemble of the following year, 2019/2020, suggests a reduction of the water deficit, to an average of -7.46 mm (SD 10.65 mm) in August 2020. The Weibull probability, P , of dry conditions is still 65%, suggesting a high probability of anomalously dry conditions, which were indeed observed over ME (Boergens et al., 2020) suggesting the usefulness of the proposed probabilistic assessment approach.

4. Summary and Conclusions

The study proposes an interannual probabilistic assessment methodology of subsurface water storage of the terrestrial system, by applying an integrated terrestrial modeling approach from groundwater across the land surface into the atmosphere at the continental scale. This approach assumes that the

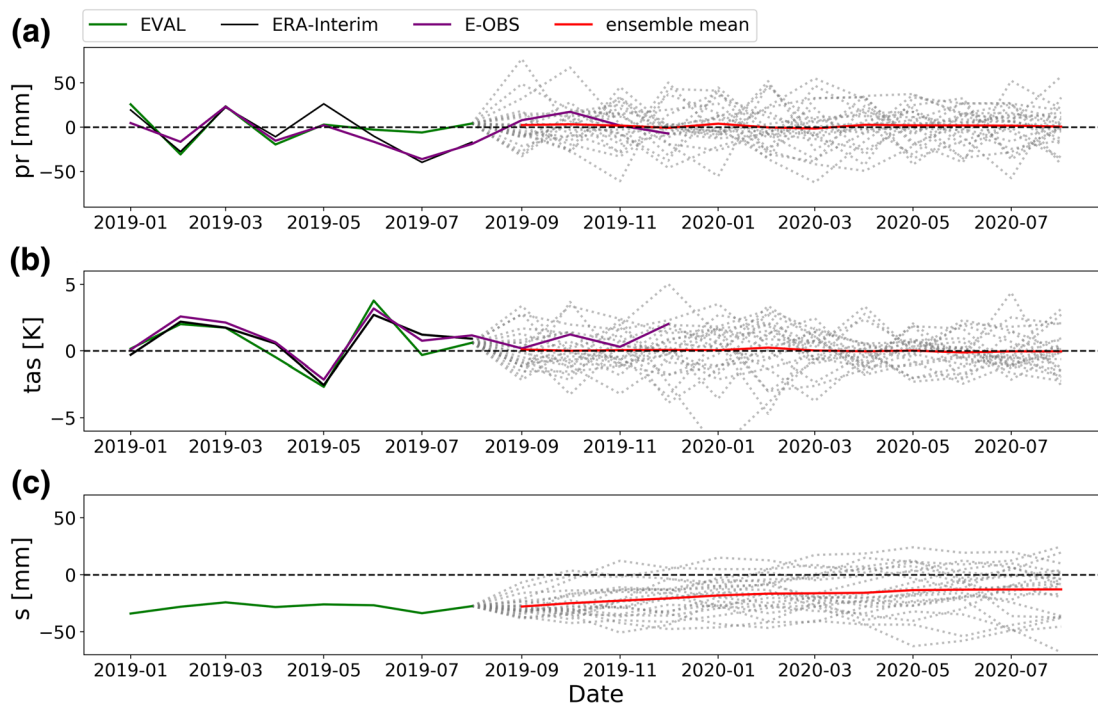


Figure 11. Monthly (a) precipitation anomalies, pr ; (b) temperature anomalies, tas ; and (c) subsurface water storage anomalies, s , from January 2019 to August 2020 in the Mid-Europe (ME) region. The deterministic climatology EVAL (green line) is supported by the evolution of ERA-Interim and E-OBS anomalies for pr and tas (black and purple line). The time period of the future assessment of the water year 2019/2020 includes all ensemble members (dashed lines). EVAL, Evaluation Simulation.

model is able to reproduce the subsurface water dynamics in the terrestrial system, ultimately rendering the assessment an initial value problem, due to strong memory effects. Atmospheric uncertainty is accounted for by utilizing the climatologic atmospheric boundary information in the simulation of the ensuing water year, in a probabilistic approach. The result is an ensemble of possible realizations of subsurface water in the ensuing water year that is analyzed for probabilities of storage anomalies. These realizations cover major parts of the ME climate variability. The approach was applied to the water years 2018/2019, 2011/2012, and 2019/2020 and focused on regions with a strong anomaly. To quantify the influence of the initial condition, a reverse assessment with assumed perfect forcing was performed.

The climatology of subsurface water storage, from a deterministic evaluation simulation over ME, shows that the water year 2017/2018 was potentially the driest since the start of the analysis in 1996. Based on the anomalous August 2018 initial condition, the probabilistic hindsight assessment indicates that there was a high probability of a continued water deficit at the end of August 2019, due to the memory effects of the subsurface system. This corresponded with the actual conditions observed over ME during that time period. The future assessment, based on the initial condition of August 2019, suggested an increased probability for a continuing water deficit at the end of 2020, which, based on current observations, appears to be realistic.

In this model, 22 years of historic atmospheric forcing has been applied to capture the atmospheric uncertainty, i.e., the natural variability. However, this is probably not sufficient to cover the full uncertainty range, therefore, additional years will be added to assess the robustness of the approach. While an additional limitation of this method is the relatively coarse resolution of the model, which does not resolve local hydrologic variability, the study demonstrates the usefulness of the approach over larger regions supported by a favorable agreement with GRACE-REC. In the future, the proposed methodology needs to be verified with observations and evaluated for forecast skill requiring the extraction of, e.g., subsurface water storage from GRACE-REC, which is complicated.

The proposed method is applicable to wet and dry anomalies, yet the information obtained from the assessment is limited in the case of small anomalies. In case of dry anomalies, the error ratio of both experiments shows that there is predictability due to the initial condition until eight months, or if the anomaly is strong enough, at the interannual time scale and beyond. The usefulness of this approach relies on the realistic simulation of memory effects in the terrestrial system, which requires explicit inclusion of the groundwater compartment, at the cost of increased computational demand. The benefit is a modeling framework with the potential to provide a continuous, probabilistic assessment over the course of a water year and beyond, which might be useful for water managers and industrial sectors, such as agriculture. In case of wet anomalies, predictability needs to be studied in the context of floods and flood generating processes considering the impact of model resolution in space and time.

In the assessment, while coupling with the atmosphere is important to dynamically downscale the atmospheric boundary conditions and close the water and energy cycle in a consistent modeling framework, the methodology could also be applied in an offline hydrologic modeling framework, if feedbacks with the atmosphere are negligible. In this study, coupling with the atmosphere was included to provide opportunities to study the potential feedbacks of drought memory effects on atmospheric processes in the future. Exploring these opportunities requires in-depth analyses, which are beyond the scope of this study.

Data Availability Statement

The data are made available at https://datapub.fz-juelich.de/slots/prob_cordex/.

Acknowledgments

The study received funding from the Helmholtz-RSF Joint Research Group and the Initiative and Networking Fund of the Helmholtz Association (HGF) through the project "Advanced Earth System Modeling Capacity (ESM)". The content of the paper is the sole responsibility of the author(s) and does not represent the opinion of the Helmholtz Association, and the Helmholtz Association is not responsible for any use that might be made of the information contained. The authors gratefully acknowledge the ESM project for funding this work by providing computing time on the ESM partition of the supercomputer JUWELS at the Jülich Supercomputing Center (JSC). The authors also gratefully acknowledge the constructive comments and suggestions by Martyn Clark and two anonymous reviewers, which strongly helped in improving the manuscript.

References

- Arnal, L., Cloke, H. L., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., et al. (2018). Skilful seasonal forecasts of streamflow over Europe?. *Hydrology and Earth System Sciences*, 22(4), 2057–2072. <https://doi.org/10.5194/hess-22-2057-2018>
- Ashby, S. F., & Falgout, R. D. (1996). A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. *Nuclear Science & Engineering*, 124(1), 145–159. <https://doi.org/10.13182/NSE96-A24230>
- Baldauf, M., Seifert, A., Foerstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the COSMO Model: Description and sensitivities. *Monthly Weather Review*, 139(12), 3887–3905. <https://doi.org/10.1175/MWR-D-10-05013.1>
- Boergens, E., Güntner, A., Dobslaw, H., & Dahle, C. (2020). Quantifying the Central European droughts in 2018 and 2019 with GRACE follow-on. *Geophysical Research Letters*, 47, 1–9. <https://doi.org/10.1029/2020GL087285>
- Carballas, T., Macías, F., Diaz-Fierros, F., Carballas, M., & Fernandez-Urrutia, J. A. (1990). *FAO-UNESCO soil map of the world. Revised legend*. Informes Sobre Recursos Mundiales de Suelos (FAO).
- Christensen, J. H., & Christensen, O. B. (2007). A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climatic Change*, 81(Suppl. 1), 7–30. <https://doi.org/10.1007/s10584-006-9210-7>
- Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., & Jones, P. D. (2018). An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123, 9391–9409. <https://doi.org/10.1029/2017JD028200>
- DAAC, L. P. (2004). *Global 30 Arc-Second elevation data set GTOPO30*. Land Process Distributed Active Archive Center.
- Day, G. N. (1985). Extended streamflow forecasting using NWSRFS. *Journal of Water Resources Planning and Management*, 111(2), 157–170.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. <https://doi.org/10.1002/qj.828>
- Dirmeyer, P. A. (2000). Using a global soil wetness dataset to improve seasonal climate simulation. *Journal of Climate*, 13(16), 2900–2922. [https://doi.org/10.1175/1520-0442\(2000\)013<2900:UAGSWD>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2900:UAGSWD>2.0.CO;2)
- Fang, Z., Bogen, H., Kollet, S., & Vereecken, H. (2016). Scale dependent parameterization of soil hydraulic conductivity in 3D simulation of hydrological processes in a forested headwater catchment. *Journal of Hydrology*, 536, 365–375. <https://doi.org/10.1016/j.jhydrol.2016.03.020>
- Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, 83(1–2), 287–302. [https://doi.org/10.1016/S0034-4257\(02\)00078-0](https://doi.org/10.1016/S0034-4257(02)00078-0)
- Furusko-Percot, C., Goergen, K., Hartick, C., Kulkarni, K., Keune, J., & Kollet, S. (2019). Pan-European groundwater to atmosphere terrestrial systems climatology from a physically consistent simulation. *Scientific Data*, 6(1), 320. <https://doi.org/10.1038/s41597-019-0328-7>
- Gasper, F., Goergen, K., Shrestha, P., Sulis, M., Rihani, J., Geimer, M., & Kollet, S. (2014). Implementation and scaling of the fully coupled Terrestrial Systems Modeling Platform (TerrSysMP v1.0) in a massively parallel supercomputing environment—A case study on JUQUEEN (IBM Blue Gene/Q). *Geoscientific Model Development*, 7(5), 2531–2543. <https://doi.org/10.5194/gmd-7-2531-2014>
- Giorgi, F., Jones, C., & Asrar, G. R. (2009). Addressing climate information needs at the regional level: The CORDEX framework. *World Meteorological Organization Bulletin*, 58(3), 175.
- Gutowski, J. W., Giorgi, F., Timbal, B., Frigon, A., Jacob, D., Kang, H. S., et al. (2016). WCRP COordinated Regional Downscaling Experiment (CORDEX): A diagnostic MIP for CMIP6. *Geoscientific Model Development*, 9(11), 4087–4095. <https://doi.org/10.5194/gmd-9-4087-2016>
- Hao, Z., AghaKouchak, A., Nakhjiri, N., & Farahmand, A. (2014). Global integrated drought monitoring and prediction system. *Scientific Data*, 1, 140001. <https://doi.org/10.1038/sdata.2014.1>

- Hao, Z., Hao, F., & Singh, V. P. (2016). A general framework for multivariate multi-index drought prediction based on Multivariate Ensemble Streamflow Prediction (MESP). *Journal of Hydrology*, 539, 1–10. <https://doi.org/10.1016/j.jhydrol.2016.04.074>
- Hao, Z., Singh, V. P., & Xia, Y. (2018). Seasonal drought prediction: Advances, challenges, and future prospects. *Reviews of Geophysics*, 56(1), 108–141. <https://doi.org/10.1002/2016RG000549>
- Hari, V., Rakovec, O., Markonis, Y., Hanel, M., & Kumar, R. (2020). Increased future occurrences of the exceptional 2018–2019 Central European drought under global warming. *Scientific Reports*, 10(1), 1–10. <https://doi.org/10.1038/s41598-020-68872-9>
- Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., et al. (2011). Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*, 4(1), 17–21. <https://doi.org/10.1038/ngeo1032>
- Hosseini-Moghari, S. M., & Araghinejad, S. (2015). Monthly and seasonal drought forecasting using statistical neural networks. *Environmental Earth Sciences*, 74(1), 397–412. <https://doi.org/10.1007/s12665-015-4047-x>
- Humphrey, V., & Gudmundsson, L. (2019). GRACE-REC: A reconstruction of climate-driven water storage changes over the last century. *Earth System Science Data*, 11(3), 1153–1170. <https://doi.org/10.5194/essd-11-1153-2019>
- Jones, J. E., & Woodward, C. S. (2001). Newton-Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. *Advances in Water Resources*, 24(7), 763–774. [https://doi.org/10.1016/S0309-1708\(00\)00075-0](https://doi.org/10.1016/S0309-1708(00)00075-0)
- Keune, J., Gasper, F., Goergen, K., Hense, A., Shrestha, P., Sulis, M., & Kollet, S. (2016). Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heat wave in 2003. *Journal of Geophysical Research: Atmospheres*, 121, 13301–13325. <https://doi.org/10.1002/2016JD025426>
- Keune, J., Sulis, M., Kollet, S., Siebert, S., & Wada, Y. (2018). Human water use impacts on the strength of the continental sink for atmospheric water. *Geophysical Research Letters*, 45, 4068–4076. <https://doi.org/10.1029/2018GL077621>
- Kollet, S. J., & Maxwell, R. M. (2006). Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, 29(7), 945–958. <https://doi.org/10.1016/j.advwatres.2005.08.006>
- Liu, X., He, B., Guo, L., Huang, L., & Chen, D. (2020). Similarities and differences in the mechanisms causing the European summer heat waves in 2003, 2010, and 2018. *Earth's Future*, 8(4), 1–11. <https://doi.org/10.1029/2019EF001386>
- Lo, M. H., & Famiglietti, J. S. (2010). Effect of water table dynamics on land surface hydrologic memory. *Journal of Geophysical Research*, 115, D22118. <https://doi.org/10.1029/2010JD014191>
- Madadgar, S., & Moradkhani, H. (2013). A Bayesian Framework for Probabilistic Seasonal Drought Forecasting. *Journal of Hydrometeorology*, 14(6), 1685–1705. <https://doi.org/10.1175/jhm-d-13-010.1>
- Maxwell, R. M. (2013). A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. *Advances in Water Resources*, 53, 109–117. <https://doi.org/10.1016/j.advwatres.2012.10.001>
- Maxwell, R. M., Chow, F. K., & Kollet, S. J. (2007). The groundwater-land-surface-atmosphere connection: Soil moisture effects on the atmospheric boundary layer in fully-coupled simulations. *Advances in Water Resources*, 30(12), 2447–2466. <https://doi.org/10.1016/j.advwatres.2007.05.018>
- Miguez-Macho, G., & Fan, Y. (2012). The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. *Journal of Geophysical Research*, 117, D15113. <https://doi.org/10.1029/2012JD017539>
- Miralles, D. G., Gentile, P., Seneviratne, S. I., & Teuling, A. J. (2019). Land-atmospheric feedbacks during droughts and heatwaves: State of the science and current challenges. *Annals of the New York Academy of Sciences*, 1436(1), 19–35. <https://doi.org/10.1111/nyas.13912>
- Mishra, A. K., & Desai, V. R. (2005). Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment*, 19(5), 326–339. <https://doi.org/10.1007/s00477-005-0238-4>
- Niedda, M. (2004). Upscaling hydraulic conductivity by means of entropy of terrain curvature representation. *Water Resources Research*, 40, W04206. <https://doi.org/10.1029/2003WR002721>
- Oleson, K. W., Dai, Y., Bonan, G. B., Bosilovich, M., Dickinson, R. E., Dirmeyer, P. A., et al. (2004). *Technical description of the community land model (CLM) (NCAR Tech. Note NCAR/TN-46, pp. 186)*. <https://doi.org/10.5065/D6N877R0>
- Oleson, K. W., Niu, G. Y., Yang, Z. L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., et al. (2008). Improvements to the community land model and their impact on the hydrological cycle. *Journal of Geophysical Research*, 113, G01021. <https://doi.org/10.1029/2007JG000563>
- Parry, S., Prudhomme, C., Wilby, R. L., & Wood, P. J. (2016). Drought termination: Concept and characterization. *Progress in Physical Geography*, 40(6), 743–767. <https://doi.org/10.1177/0309133316652801>
- Rhee, J., & Im, J. (2017). Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agricultural and Forest Meteorology*, 237–238, 105–222. <https://doi.org/10.1016/j.agrformet.2017.02.011>
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3–4), 125–161. <https://doi.org/10.1016/j.earscirev.2010.02.004>
- Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., et al. (2013). A drought monitoring and forecasting system for Sub-Saharan African water resources and food security. *Bulletin of the American Meteorological Society*, 95(6), 861–882. <https://doi.org/10.1175/bams-d-12-00124.1>
- Shrestha, P., Sulis, M., Masbou, M., Kollet, S., & Simmer, C. (2014). A scale-consistent terrestrial systems modeling platform based on COSMO, CLM, and ParFlow. *Monthly Weather Review*, 142(9), 3466–3483. <https://doi.org/10.1175/MWR-D-14-00029.1>
- Shukla, S., Sheffield, J., Wood, E. F., & Lettenmaier, D. P. (2013). On the sources of global land surface hydrologic predictability. *Hydrology and Earth System Sciences*, 17(7), 2781–2796. <https://doi.org/10.5194/hess-17-2781-2013>
- Solaraju-Murali, B., Caron, L. P., Gonzalez-Reviriego, N., & Doblas-Reyes, F. J. (2019). Multi-year prediction of European summer drought conditions for the agricultural sector. *Environmental Research Letters*, 14(12), 124014. <https://doi.org/10.1088/1748-9326/ab5043>
- Staudinger, M., & Seibert, J. (2014). Predictability of low flow—An assessment with simulation experiments. *Journal of Hydrology*, 519(PB), 1383–1393. <https://doi.org/10.1016/j.jhydrol.2014.08.061>
- Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., et al. (2019). Contributions of GRACE to understanding climate change. *Nature Climate Change*, 9(5), 358–369. <https://doi.org/10.1038/s41558-019-0456-2>
- Turco, M., Ceglar, A., Prudhomme, C., Soret, A., Toreti, A., & Doblas-Reyes Francisco, J. (2017). Summer drought predictability over Europe: Empirical versus dynamical forecasts. *Environmental Research Letters*, 12(8). <https://doi.org/10.1088/1748-9326/aa7859>
- Valcke, S. (2013). The OASIS3 coupler: A European climate modeling community software. *Geoscientific Model Development*, 6(2), 373–388. <https://doi.org/10.5194/gmd-6-373-2013>
- Vogt, J. V., Naumann, G., Masante, D., Spinoni, J., Cammalleri, C., Erian, W., et al. (2018). *Drought risk assessment and management. A conceptual framework (Eur 29464 EN)*. Publications Office of the European Union. <https://doi.org/10.2760/057223>
- Walko, R. L., Band, L. E., Baron, J., Kittel, T. G. F., Lammers, R., Lee, T. J., et al. (2000). Coupled atmosphere-biophysics-hydrology models for environmental modeling. *Journal of Applied Meteorology*, 39(6), 931–944. [https://doi.org/10.1175/1520-0450\(2000\)039<0931:CABHMF>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<0931:CABHMF>2.0.CO;2)

- Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, 120, 2648–2671. <https://doi.org/10.1002/2014JB011547>
- Wood, A. W., & Lettenmaier, D. P. (2008). An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophysical Research Letters*, 35, L14401. <https://doi.org/10.1029/2008GL034648>
- Wood, A. W., Maurer, E. P., Kumar, A., & Lettenmaier, D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research*, 107(20) ACL 6-1–ACL 6-15. <https://doi.org/10.1029/2001JD000659>
- York, J. P., Person, M., Gutowski, W. J., & Winter, T. C. (2002). Putting aquifers into atmospheric simulation models: An example from the Mill Creek Watershed, northeastern Kansas. *Advances in Water Resources*, 25(2), 221–238. [https://doi.org/10.1016/S0309-1708\(01\)00021-5](https://doi.org/10.1016/S0309-1708(01)00021-5)
- Yuan, X., Roundy, J. K., Wood, E. F., & Sheffield, J. (2015a). Seasonal forecasting of global hydrologic extremes: System development and evaluation over GEWEX basins. *Bulletin of the American Meteorological Society*, 96(11), 1895–1912. <https://doi.org/10.1175/BAMS-D-14-00003.1>
- Yuan, X., & Wood, E. F. (2013). Multimodel seasonal forecasting of global drought onset. *Geophysical Research Letters*, 40, 4900–4905. <https://doi.org/10.1002/grl.50949>
- Yuan, X., Wood, E. F., Chaney, N. W., Sheffield, J., Kam, J., Liang, M., & Guan, K. (2013). Probabilistic Seasonal Forecasting of African Drought by Dynamical Models. *Journal of Hydrometeorology*, 14(6), 1706–1720. <https://doi.org/10.1175/JHM-D-13-054.1>
- Yuan, X., Wood, E. F., & Ma, Z. (2015b). A review on climate-model-based seasonal hydrologic forecasting: Physical understanding and system development. *Wiley Interdisciplinary Reviews: Water*, 2(5), 523–536. <https://doi.org/10.1002/wat2.1088>
- Zink, M., Pommerenke, J., Kumar, R., Thober, S., Samaniego, L., & Marx, A. (2016). The German Drought Monitor. *Environmental Research Letters*, 17, 5625.