

Water Resources Research®



RESEARCH ARTICLE

10.1029/2022WR033655

Key Points:

- Earth observations (EO) and in situ data assimilation (DA) improve hydrological forecasting quality depending on season and assimilated data; effects persist for long periods
- DA is relatively more important than meteorological forcing to improve the hydrological forecast quality during all seasons except autumn
- Assimilating EO-based snow water equivalent in winter impacts seasonal forecasts for 3 months+, hence it's beneficial to have high-quality snow information

Correspondence to:

J. L. Musuza,
jude.musuza@smhi.se

Citation:

Musuza, J. L., Crochemore, L., & Pechlivanidis, I. G. (2023). Evaluation of earth observations and in situ data assimilation for seasonal hydrological forecasting. *Water Resources Research*, 59, e2022WR033655. <https://doi.org/10.1029/2022WR033655>

Received 9 SEP 2022

Accepted 1 APR 2023

Author Contributions:

Conceptualization: Louise Crochemore, Ilias G. Pechlivanidis

Data curation: Jude L. Musuza

Formal analysis: Jude L. Musuza,

Louise Crochemore, Ilias G. Pechlivanidis

Funding acquisition: Ilias G.

Pechlivanidis

Investigation: Jude L. Musuza, Louise

Crochemore, Ilias G. Pechlivanidis

Methodology: Jude L. Musuza, Louise

Crochemore, Ilias G. Pechlivanidis

Project Administration: Ilias G.

Pechlivanidis

Resources: Ilias G. Pechlivanidis

Software: Jude L. Musuza, Louise

Crochemore

Supervision: Ilias G. Pechlivanidis

© 2023. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Evaluation of Earth Observations and In Situ Data Assimilation for Seasonal Hydrological Forecasting

Jude L. Musuza¹ , Louise Crochemore² , and Ilias G. Pechlivanidis¹ 

¹Swedish Meteorological and Hydrological Institute, Norrköping, Sweden, ²Université Grenoble Alpes, CNRS, IRD, Grenoble-INP, IGE, Grenoble, France

Abstract Earth observations (EOs) are a valuable complement to in situ measurements in hydrology because they provide information in locations where direct measurements are unavailable or prohibitively expensive to make. Recent advances have enabled the assimilation of data sets of different physical variables into hydrological models to better estimate states and fluxes. Along with the meteorological forcings, the assimilated data exert controls on forecasts; it is therefore important to apportion the contributions from the forcing and assimilated data. Quality assessment is sometimes conducted in the pseudo-reality, which neglects observations and model limitations. Here we introduce a diagnostic framework that accounts for observations to assess the sources of skill and infer the seasonal importance of assimilated and forcing data. We test the framework with a forcing data set from a downscaled Global Circulation Model and assimilate four EO and two in situ data sets to initialize the forecasts. We investigate the hydrological response and seasonal predictions over a Swedish snowmelt-dominated catchment using the HYPE model over 2001–2015. The framework allows assessing the improvement in seasonal skill due to the different assimilated data and meteorological forcing. For the studied catchment, all EO and in situ data sets add information to the final forecast. The lead times during which data assimilation influences forecast skill also differ between data sets and seasons for example, assimilating snow water equivalent impacts the forecast for more than 20 weeks during winter. Lastly, assimilated data sets are generally more important to streamflow forecasting skill than meteorological forcing in the studied snow-dominated catchment.

1. Introduction

Hydrological forecasting is the science and practice of estimating and predicting the future state of a river system. It is essential for water resources managers to plan for and mitigate natural disasters (i.e., droughts and floods) and to develop strategies to adapt to variations in the climate. Forecasts are broadly classified according to their future time horizons, that is, short- (up to 2 days), medium- (up to 15 days), extended- or sub-seasonal (2 weeks to a month) or long-range (up to 1 year; also known as seasonal forecasts); and can be either a single value (deterministic) or a range of possible values (probabilistic) (Roundy et al., 2019). Due to uncertainties from model errors and physical meteorological processes which increase with lead time, hydrologic forecast performance generally decreases with lead time. Long-range forecasts are inherently uncertain but nonetheless useful to decision makers in for example, planning long-term strategies for reservoir storage (Anghileri et al., 2016; Essénfelder et al., 2020) and supporting adaptation to climate change (White et al., 2021; Winsemius et al., 2014).

Seasonal forecasting methods in hydrology can be broadly divided into: (a) statistical, which use stochastic modeling techniques (Jenicek et al., 2016; Macian-Sorribes et al., 2020), and (b) dynamical, which use seasonal meteorological forecasts from Global Circulation Models (GCMs) as input to a hydrological model (Tuteja et al., 2011; X. Yuan et al., 2015). Hybrid methods also exist where the two methods are combined (Wang et al., 2021). Ensemble forecasting (Buizza et al., 2019), where multiple realizations for a time and location are generated to account for future uncertainties, is part of the state-of-the-art early warning services. Current ensembles are commonly obtained by applying different perturbations to the control (unperturbed) forecast, and should preferably be based on multiple models (Sahai et al., 2021). Ensemble forecasting is used to predict flooding (Alfieri et al., 2014; Ramos et al., 2007), droughts (Colossi & Tucci, 2020); and in hydro-climate services, for example, for water resources management applications (Macian-Sorribes et al., 2020; Pechlivanidis et al., 2020), agriculture (Villani et al., 2021) and hydropower (Contreras et al., 2020).

Validation: Jude L. Musuuza, Ilias G. Pechlivanidis
Visualization: Jude L. Musuuza, Ilias G. Pechlivanidis
Writing – original draft: Jude L. Musuuza, Louise Crochemore, Ilias G. Pechlivanidis
Writing – review & editing: Jude L. Musuuza, Louise Crochemore, Ilias G. Pechlivanidis

The predictability (source of skill) of hydrological variables lies in two main factors: (a) meteorological forcings (MF), and (b) initial states (IHC) of the hydrological model (Arnal et al., 2017; Giuliani et al., 2020; Wood & Lettenmaier, 2008; Wood et al., 2016; H. Yuan et al., 2019). Various techniques are used to improve forecasts: data assimilation, multi-modeling, and pre- and post-processing, that is, bias-adjustment of meteorological and/or hydrological forecasts (see e.g., Macian-Sorribes et al. (2020)). For streamflow forecasts, it is important to understand the spatiotemporal patterns of predictability and the drivers behind them (Girons Lopez et al., 2021) because the patterns can be mapped onto physical catchment characteristics. Initial conditions of soil moisture and snow in spring have been identified as significant sources of skill over Europe (Wanders et al., 2019), whilst Harrigan et al. (2018) concluded that the baseflow index is a hydrological signature strongly correlated with forecasting skill over the United Kingdom. In a pan-European investigation, Pechlivanidis et al. (2020) highlighted that the hydrological forecast skill at early lead times depends on the hydrological regime of the river system with flashy and long river memory catchments having limited and high forecast skill, respectively. de Lavenne et al. (2022) additionally showed that long-term memory in France is driven by the presence of aquifers and is particularly observed in catchments with a high aridity index.

A modeling framework producing reasonable forecasts of a single variable for example, streamflow can have inconsistencies with respect to other variables for example, groundwater level. Vogel et al. (2021) forecast several variables relevant to extreme events related to droughts and bush fires for lead times of up to 3 months. The consideration of both soil moisture and actual evapotranspiration ensured the consistent partitioning of precipitation into its main constituent fluxes. Such a verification of internal model fluxes is however not commonly done, and can be informed by earth observations (EO) or in-situ data, and corrected through data assimilation (DA).

DA is a common post-processing methodology used to incorporate observations into a modeling framework while accounting for different sources of uncertainty. This minimizes model errors (defined as the difference between model estimates and observations) that arise from inadequate conceptualization (process representation), non uniqueness of parameters and inaccurate initialization (Moradkhani et al., 2019). DA is one of the main methods external to the hydrological model used to improve the model's initial states (DeChant & Moradkhani, 2015; Liu & Gupta, 2007; Poletti et al., 2019; Wang et al., 2021; Yucel et al., 2015) thereby producing more reliable hydrological forecasts than without DA implementation. DA reduces the uncertainty in initial forecast states by improving model parameters (Pathiraja et al., 2018) and structure (Abbaszadeh et al., 2018) although the parameter improvement is not common in operational hydrology forecasting (Moradkhani et al., 2019). Reduction of model uncertainty through DA has been investigated in Abbaszadeh et al. (2019); DeChant and Moradkhani (2014), among others.

DA is becoming a well-established hydrological modeling practice especially in prediction services where correcting initial model states has been shown to improve the forecasts at medium-range time scales (DeChant & Moradkhani, 2011; X. Yuan et al., 2016). The process requires high-quality observations which can be a limiting factor in remote and sparsely monitored areas or for entities that are expensive or difficult to measure. EO fill this gap by offering consistent observations over wide spatial extents and at reasonable spatiotemporal resolution. Assimilation of in situ observations, for example, streamflow, is a common practice with justified added value but the impact of assimilating earth observations is still under scientific exploration (T. L. Du et al., 2022; Nijzink et al., 2018). The time lag between acquisition and retrieval of many EO data sets limits their use to historical studies. Nevertheless, Musuuza et al. (2020) recently explored the impact of EO data assimilation in a mountainous river system and highlighted the corresponding improvements in the modeled streamflow and reservoir inflow. Assimilation of EO into models has also been done in for example, Dechant and Moradkhani (2011), Leisenring and Moradkhani (2011), and Magnusson et al. (2014).

In addition to providing spatially continuous fields, EO offer a large variety of hydrologically useful variables, for example, soil moisture, snow cover, water temperature, etc. EO data sets can assist in assessing the performance of hydrological models, which would help to: (a) understand how and against which observations model estimates can be most appropriately compared, (b) point toward model structural inconsistencies (systematic differences, which can subsequently lead to improvements in model structure and/or parametrization), (c) add value to the development of data assimilation techniques, and (d) achieve a “right for the right reasons” model.

Forecast performance is known to decrease with lead time and yet, to the best of our knowledge, no study has explicitly quantified the lead times over which improved initial states acquired from DA impact the forecasts and their performance. The objective of this study is therefore to introduce and apply a framework to quantify

such critical lead times and determine the relative importance of GCM meteorological forcings and assimilation of different data sets in a seasonal forecasting system. The framework allows answering the following scientific questions:

1. How does the assimilation of different data sets impact seasonal forecasts and how long do the impacts last for in different seasons?
2. What is the relative importance of meteorological forcings and assimilated data in a seasonal forecasting system?

The framework is illustrated here on a Swedish catchment, for which the scientific questions are answered. We next present the study area and data in Section 3. Section 4 provides a description of the experimental setup including the hydrological model, the ensemble Kalman filter and the methodology used to evaluate the performance of the seasonal forecast systems. We present and interpret the results in Section 5, followed by a discussion in Section 6. Finally, Section 7 states the conclusions.

2. A Diagnostic Framework to Detect Sources of Forecast Skill

The added performance (skill) of a “new” forecast system is commonly assessed against a benchmark system using statistical metrics that characterize the forecast performance with regard to the overall forecast distribution, extremes, temporal variability, reliability, accuracy or sharpness (Anctil & Ramos, 2019). However, the benchmark selection can also help isolate and attribute the forecast skill to MF and/or IHC (Crochemore et al., 2020). The performance of both the “new” forecast and benchmark systems is assessed using a reference data set that is selected to represent “reality” (Pappenberger et al., 2015). While forecasts are sometimes evaluated with respect to their performance against the historical model “reference” simulation (also known as “pseudo-observations” or “perfect forecast”), model limitations (i.e., structural and parametric) that lead to simulation errors are neglected (Girons Lopez et al., 2021). Alternatively forecasts can be evaluated against the historical “real” observations, in which case the forecast performance is subject to model limitations. This latter case is particularly relevant when assessing the performance of data assimilation techniques whose aim is to approach “real” observations, rather than model-based “pseudo-observations.”

Sensitivity analysis methods can quantify the increase in streamflow predictability as a result of decreasing the uncertainty in one of the dominant sources (Arnal et al., 2017; Wood et al., 2016). However, these theoretical methods are not tailored for quantifying the importance of a unique (or combined) forcing or assimilated data set(s), which is needed when making data set choices to improve (operational) forecasting services. Here, we overcome this challenge by taking advantage of various forecast systems that are commonly applied in (operational) seasonal streamflow forecasting. Understanding each system (hypothesis and limitations) and further inter-comparing them, allows setting up a framework that diagnoses the added performance of the different sources of skill (Figure 1). We note that the framework is not restrictive to a certain type of catchment conditions and/or model setup, and hence can be applied to any catchment, hydrological model, data set used for model initialization or assimilation method. The following systems are available where “streamflow climatology” means the climatology of historical streamflow (Crochemore et al., 2017):

- *Observed streamflow climatology*: Is a simplified system that only requires historical streamflow observations. The method does not depend on either MF or quantification of the IHC. An ensemble of streamflow observations is generated by resampling past scenarios from the observed climatology based on their start date. The uncertainty represented in both IHC and MF is that of the natural hydrologic system.
- *Simulated streamflow climatology*: Is similar to the one above, yet the system is subject to historical MF driving a hydrological model to generate historical streamflow simulations (“reference” simulation). The quality of the MF and the hydrological model error influence the quality of the streamflow climatology (Harrigan et al., 2018). The uncertainty in MF thus represented is that of historical MF climatology, and the uncertainty in IHC is that of the hydrological model climatology in IHC.
- *Ensemble Streamflow Prediction (ESP)*: Assumes that the resampled ensemble of historical meteorological data contains adequate information to predict the future, which can also lead to forecast overconfidence since this reliance on historical data limits its ability to capture responses to unprecedented events (Buizza et al., 2019; Day, 1985). In comparison to the system above, ESP requires model initialization (generated from historical MF) which is a source of skill, and hence the updated catchment conditions in the initialization

makes the ESP framework advantageous over the two methods above (Wood & Lettenmaier, 2006). The ESP approach has been shown to provide accurate forecasts with an added value to decision making (Girons Lopez et al., 2021).

- *GCM-based streamflow forecasts*: Require dynamic meteorological forcing from a Global Circulation Model (GCM). Such meteorological forecasts are subject to regional biases indicating the need for bias-adjustment to reduce their propagation into the hydrological predictions (Crochemore et al., 2016). Particularly for hydrological modeling, initialization is required, and hence the difference in skill between the GCM-based predictions and ESP is attributed to the improved (or not) dynamic MF.

Moreover, some of the seasonal streamflow forecast systems can be further enhanced through EO and in situ data assimilation that improves the initial hydrological conditions. Among the four systems introduced earlier, only the ESP and GCM-based streamflow forecast systems can take advantage of the improved IHC through DA; named in Figure 1 as *ESP with data assimilation* and *GCM-based streamflow forecasts with data assimilation* respectively.

Applying a system inter comparison, one can diagnostically quantify the importance of the different sources of skill. However, attention is needed on the benchmark since the calculated skill in some system inter comparisons could be driven by both MF and IHC (e.g., comparing ESP or GCM-based streamflow forecasts against observed streamflow climatology). In practice, the skill of any of the four systems that do not consider data assimilation can be benchmarked against one of the remaining three systems, consequently pointing toward the sources of skill (the system characteristics that differ between the two systems). A comparison of the two systems that consider data assimilation against their counterparts without data assimilation allows quantifying the added value from improved model initialization due to DA against a model initialization driven by climatology-based forcing.

3. Study Area and Data

3.1. Study Area

The Umeälven catchment is located in northern Sweden with an area of 26,000 km² and a mean elevation of 510 m above mean sea level. The regime is dominated by snowmelt in spring and summer (Figure 2).

Annual precipitation ranges from 542 to 874 mm with a mean of 691 mm during the period 2001–2015. For the same period, daily mean temperature oscillated between −12° and 20°C with a mean value close to 0°C. The dominant land uses are coniferous forest and open land with vegetation, covering 38% and 32% of the area, respectively. Moraine, the dominant soil, makes up almost 60% of the area. The catchment is extensively used for hydropower production with 19 installed hydropower plants. The temporal power production needs are almost the reverse of the flow regime. Consequently, the water storage from the snowmelt season to the next winter in large headwater reservoirs is the most important reservoir management objective. The hydropower sector would greatly benefit from high-quality forecasts, which are particularly crucial during extreme years (Crochemore et al., 2021).

3.2. Earth Observations and In Situ Data

Four EO data sets (Table 1) of fractional snow cover (FSC), snow water equivalent (SWE) and the actual and potential evapotranspiration (AET and PET) and the two in situ data sets highlighted in the table were available from different providers and spatiotemporal resolutions.

FSC is obtained from a daily pan-European product by the CRYOLAND initiative (Malnes et al., 2015) of the Finnish Meteorological Institute at a 500 m grid from 2000 to near real time. It is retrieved from the optical sensors on the NASA TERRA Y AQUA satellite. The data represents the area fraction from 0% to 100% of either the forested canopy or the ground covered with snow. Optical sensors are affected by clouds but not high latitudes.

SWE is obtained from a global (Northern Hemisphere) daily product provided by the European Space Agency Climate Change Initiative (ESA-CCI) (Kouki et al., 2021), at a 12.5 km grid. The product is retrieved from passive microwave sensors and based on Scanning Multichannel Microwave Radiometer (SMMR) and Special Sensor Microwave/Imager (SSM/I) data from a number of Defense Meteorological Satellite Program (DMSP) F-series satellites over the period 1978–2019. Such passive microwave-based products are uncertain

Table 1

The EO and In Situ (Highlighted) Data Sets Used in This Study With Their Sources, Time Coverages and Spatiotemporal Resolutions

Hydrologic variable (abbreviation; unit)	Provider	Availability window	Spatial resolution [m]	Temporal resolution
Fractional snow cover (FSC; -)	Finnish Meteorological Institute (FMI)	2000–2021	500	Daily
Snow water equivalent (SWE; mm)	European Space Agency—Climate Change Initiative (ESA CCI)	1979–2019	12500	Daily
Actual evapotranspiration (AET; mm)	National Aeronautics and Space Administration (NASA)	2001–2021	500	8-day aggregate
Potential evapotranspiration (PET; mm)	National Aeronautics and Space Administration (NASA)	2001–2021	500	8-day aggregate
Streamflow at sub catchment outlet (QOB; m ³ · s ⁻¹)	Swedish Meteorological and Hydrological Institute (SMHI)	1900–2020	-	Daily
Reservoir inflow (INFLOW; m ³ · s ⁻¹)	Vattenregleringsföretaget (VRF)	1961–2019	-	Daily

and erroneous in complex terrains (highlands), deep snow conditions, during snowmelt and along the coast (Magnusson et al., 2014). However, the data still provides useful information in areas with moderate terrain and cold/dry climate.

The evapotranspiration products are global, 8-day aggregations at a 500 m resolution from NASA (Mu et al., 2011). The data is retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Terra and Aqua satellites in 36 spectral bands. The data covers the period from 2001 to near real time.

The streamflow data at the sub catchment outlet stations are collected by SMHI from different Swedish institutions for example, water councils, municipalities, water authorities and hydropower companies and is highly quality-checked. It is measured directly or inferred from fitting river stage observations to rating curves. The data set covers 1900 to near real time and is published on the Vattenwebb portal <https://vattenwebb.smhi.se/station/#/> (SMHI, 2022).

The local reservoir inflow is a daily product from the Swedish water regulation authority (VRF) covering the period 1961–2020. It is the sum of snowmelt and runoff from the area between the outlet of an upstream reservoir (if existing) and the outlet of the most downstream reservoir in the forecast area and the net precipitation on the reservoirs within the catchment. This local reservoir inflow is computed as the residual between the reservoir outflow and inflow, and the change in reservoir storage volume. The links to all data sets are presented in Appendix C2.

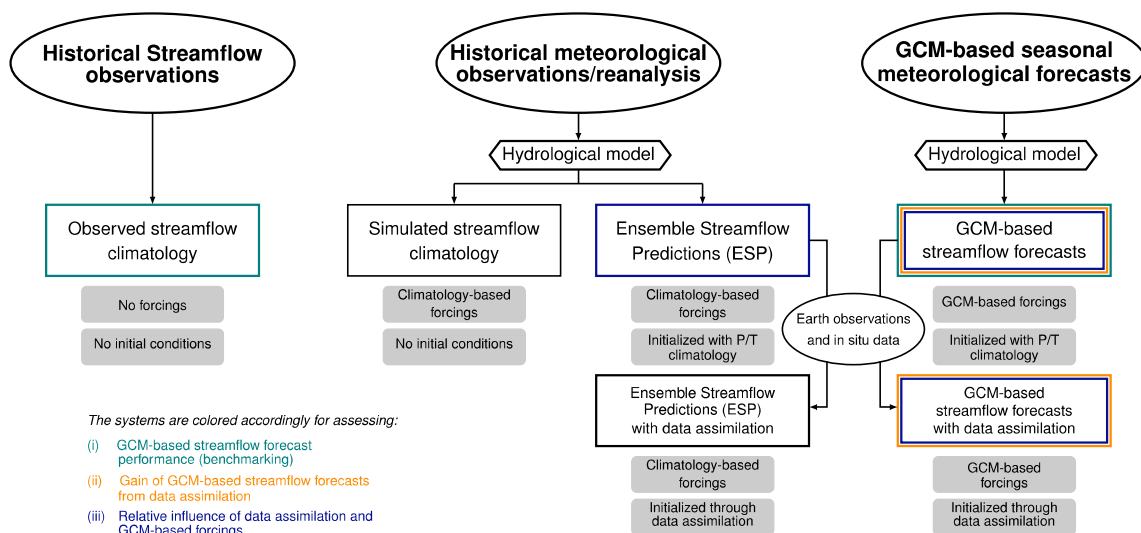


Figure 1. The diagnostic framework guiding toward decomposition of the sources of skill in seasonal streamflow forecasting. The gray boxes describe the characteristics of each forecast system. Some forecast systems are colored accordingly depending on their prerequisite to address the paper's tasks (i)–(iii) presented in Section 4.4.

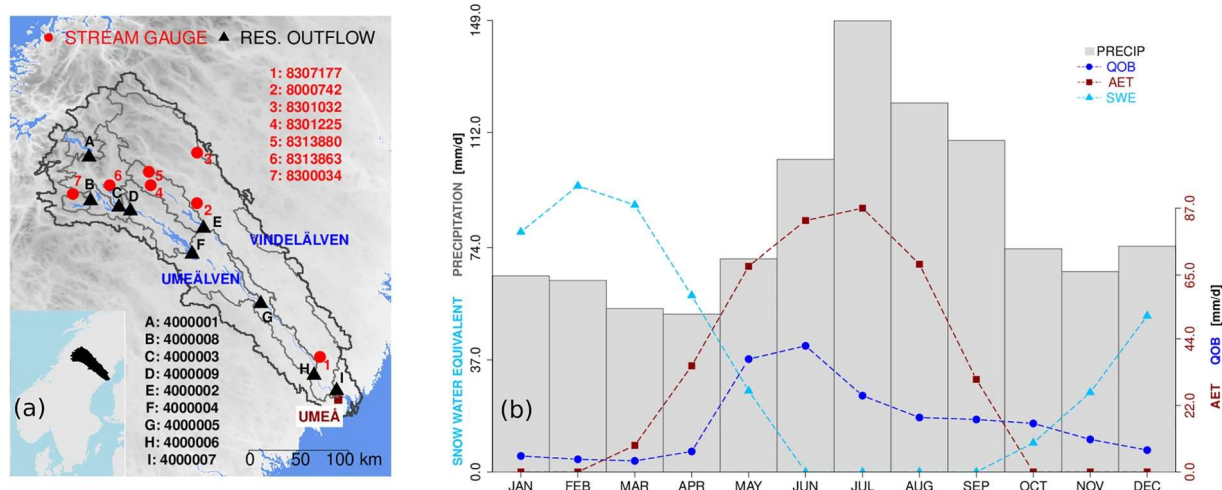


Figure 2. (a) The Umeälven study area showing the managed Umeälven and unmanaged Vindelälven tributaries. The red circles are the unregulated sub catchment outflow stations and the black triangles the reservoir outflow stations. The monitoring point IDs are listed in red and black, respectively. (b) The monthly snow water equivalent (light blue), normalized streamflow at the outlet (blue), the basin's mean actual evapotranspiration (maroon) and the precipitation (histogram), all in $\text{mm} \cdot \text{day}^{-1}$.

3.3. Meteorological Historical Forcing and Seasonal Forecasts

Here, we make use of the Hydrological Global Forcing Data version 2.0 (HydroGFD v2.0) product, an observation-corrected reanalysis data set providing historical meteorological information of precipitation and mean temperature for 1971–2018 at a 0.5° resolution (Berg et al., 2018). HydroGFD is used to drive the simulated streamflow climatology and ESP (with and without data assimilation) forecast systems in Figure 1.

Seasonal predictions of daily mean precipitation and temperature taken from the ECMWF SEAS5 system (Johnson et al., 2019) constitute the GCM-based forcing data set introduced in Section 2. The SEAS5 system is an example of data obtained after assimilating current weather observations into a GCM framework to forecast future weather. The system generates operationally an ensemble of 51 members initialized every month with a seven month lead time (Johnson et al., 2019). The SEAS5 re-forecasts (hindcasts) used here consist of 25 ensemble members available at a resolution of approximately 36 km.

GCM outputs are subject to regional biases which are further propagated into the hydrological predictions (Crochemore et al., 2016). We therefore bias-adjusted the data using a modified version of the Distribution Based Scaling (DBS) method (Yang et al., 2010) that accounts for drifting. Using the data for the whole analysis period, the bias-adjustment parameters are conditioned on the lead month and the forecast issue date. DBS was originally developed to adjust biases in climate projections and is a quantile mapping method adapted here, for seasonal forecasting. Bias adjustment was conducted on all monthly initialized forecasts using the HydroGFD v2.0 (Berg et al., 2018) as reference data set. After bias adjustment, the cumulative distribution of daily meteorological forecasts follows closely the one of HydroGFD. See the results of bias-adjusting SEAS5 toward HydroGFD v2.0 in Pechlivanidis et al. (2020).

4. Experimental Setup

4.1. Hydrological Model Description

The Hydrological Predictions for the Environment (HYPE) model is a semi-distributed process-based model capable of simulating the hydrological processes from a single basin to global scale (Lindström et al., 2010). The parameters in the model are defined for the entire domain or dependent on land use or soil classes. The model has conceptual routines for most of the major land surface and subsurface processes. The snow accumulation and melt processes are modeled using the degree-day method with land use dependent parameters. HYPE simulates the water flow paths in soil, which is divided into three layers with a fluctuating groundwater table. A fraction of rainfall or snowmelt infiltrates into the top soil, which is limited by a soil-type dependent maximum infiltration

rate. If the soil moisture in the upper soil layer exceeds a threshold for macropore flow, part of the remaining water forms macropore flow. Potential evapotranspiration (PET) is estimated using several possible models, for example, the modified Jensen-Haise (Oudin et al., 2005) among others. PET is achieved only if either the actual soil moisture exceeds a large portion of the soil field capacity or the sub catchment is defined as a water body. For soil moisture below this limit in non-water-body areas, the actual evapotranspiration, computed following the crop coefficient method in Allen et al. (1998), decreases linearly to zero at the wilting point. Runoff from the soil zone is computed when the soil moisture exceeds field capacity and it percolates to lower soil layers when the soil moisture in the upper layers exceeds field capacity. The ground water level is estimated based on the level in the soil zone where the pore space is filled.

The hydrological model for the Umeälven study area (see Figure 2) is a subset of the pan-European E-HYPE model (Hundecha et al., 2016) ran at a daily time step. It consists of 157 sub catchments with nine reservoir inflow regions and seven gauging stations considered free from anthropogenic influences. The calibration of the hydrological model was done with streamflow as the main objective, which resulted in a high performance with respect to streamflow and reservoir inflow. The Kling and Gupta Efficiency metric (KGE; (Gupta et al., 2009)), whose optimum value is 1 was close to 0.8 for all variables, except for SWE and PET at the mountainous and coastal areas (see details in Musuza et al. (2020)). The poor performance in the SWE in those regions is due to the failure of the passive microwave sensors stated in Section 3.2.

4.2. Data Assimilation With the Ensemble Kalman Filter

The Ensemble Kalman Filter (Evensen, 2003) is a sequential algorithm that uses a finite number of ensemble members to represent the error statistics. The error covariance matrices are computed from an ensemble and propagated forward by the dynamic model. This setup requires: (a) an infinite number of members in the ensemble, (b) a linear dynamic model, and (c) Gaussian errors. However, these requirements are never satisfied, since large ensembles quickly become computational bottlenecks in large domains or with long simulation runs. The normally-distributed error assumption is also invalid for physically bounded variables such as FSC (Mishra & Coulibaly, 2009; Pathiraja et al., 2018). Here, we used the logit-normal error transformation for bounded variables, for example, FSC, and the lognormal transformation for semi-bounded variables, for example, SWE. See Bertino et al. (2003) for details.

At each time step in the assimilation, we introduced an ensemble of finite, random perturbations in the forcing and other observation data that resulted in randomly generated model trajectories. The perturbations in forcing are spatially correlated fields that help propagate information to locations where it may be missing (Reichle & Koster, 2003). The spatial correlation is controlled by the localization, which suppresses superfluous covariances that can arise in mountainous regions (Magnusson et al., 2014).

The ensemble of model states and fluxes was then propagated through the dynamic model and transformed to the observation space, where the *predicted observation* and observation error terms were obtained. The localization can be included by modifying the error covariance matrix terms present in the Kalman gain matrix. These are multiplied by two scalar matrices that restrict the influence of observations on updated variables using element-wise products. One ensures that the distance-dependent functions used to define the scalar matrices diminish to zero at appropriate separation lengths. See Evensen (2003) for a detailed treatment of the Ensemble Kalman Filter.

4.3. Seasonal Hydrological Forecasting

To generate initial hydrologic states, the model was initially forced with the HydroGFD v2.0 meteorological data, whilst in the case of data assimilation the model states were generated with the four EO datasets (FSC, SWE, AET, PET) and the two in situ data sets of streamflow (QOB) and reservoir inflow (INFLOW), as well as the combination (COMBI) of FSC, QOB and INFLOW (i.e., a total of eight possible initializations, seven exploiting DA). In Musuza et al. (2020), a total of 62 possible combinations of the six data sets in Table 1 were generated and the constituent data sets assimilated individually or simultaneously (called combinations) using the Ensemble Kalman filter technique over the period 2001–2015. COMBI was among the several combinations that improved the model performance in streamflow and reservoir inflows. The modeled streamflow and reservoir inflow were used as target variables during the previous assimilation experiment due to their importance to the hydropower sector in the area.

The hydrological model thus initialized was forced with the bias-adjusted SEAS5 (hereafter named GCM-MF) and HydroGFD-based climatological forcing datasets to generate seasonal hydrological forecasts initialized with and without DA at the start of each month with a 7-month lead time.

We extracted modeled streamflow and reservoir inflow in all gauged sub catchments of the region to assess the model's predictive accuracy at seasonal time scales. Seasonal re-forecasts are evaluated with respect to their forecast performance against the historical observations per initialization month and lead time. Even if the hydrological model is run at a daily time scale, the generated forecasts are aggregated to weekly means for the 2001–2015 period, during which both re-forecasts and earth observations are available. It is important to note that the verification of modeled re-forecasts against real observations allows the estimation of the actual forecast accuracy, which depends not only on model initialization and meteorological forcings, but also on the model's structural and parametric inadequacies to represent reality (numerically described by the historical model performance).

4.4. The Forecast System Inter Comparison

In this paper, the forecast systems are inter-compared in order to assess:

1. The quality of GCM-MF seasonal hydrological forecasts for each initialization month and lead time;
2. The gain in performance from data assimilation based on the change in skill when different in situ and EO data sets are assimilated during different seasons; and the persistence of data assimilation impacts that is, the number of lead weeks over which DA influences the skill;
3. The relative importance of DA and GCM-based forcing on hydrological forecasting skill during different seasons.

Following the framework in Section 2, task (i) is addressed by assessing the GCM-MF forecasts initialized without DA against the observed climatology (note that these systems are presented in light blue boxes in Figure 1). Task (ii), the impact of data assimilation on forecast skill is addressed by assessing the GCM-MF forecasts initialized with DA against its counterpart system without DA (see yellow boxes in Figure 1). Finally, task (iii) is addressed in two steps: first by using the skill calculated in task (ii) and then assessing the skill of GCM-MF forecasts benchmarked against ESP forecasts, both initialized without DA (see dark blue boxes in Figure 1). The difference in the two skills would indicate the relative importance of the two sources of skill (MF and DA).

4.5. Forecast Evaluation

We first assess the quality of the forecasts in terms of their accuracy, that is the performance of the temporally aggregated streamflow and reservoir inflow forecasts in comparison to observations, which is expressed in $\text{m}^3 \cdot \text{s}^{-1}$. For this purpose, the daily model outputs were aggregated to weekly mean values and the Continuous Ranked Probability Score (CRPS; (Hersbach, 2000)) was calculated for each sub catchment, initialization month and lead week. Given that SEAS5 provides forecasts for 7 months ahead, the evaluations here are conducted for lead weeks 0–27. CRPS is a common measure of overall ensemble forecast performance and is equivalent to the mean absolute error when applied to deterministic forecasts (Buizza et al., 2019; Hersbach, 2000). The time averaged CRPS (Schepen et al., 2019) is formulated as:

$$CRPS = \frac{1}{N} \sum_{t=1}^N \int_{-\infty}^{\infty} [F_t^f(x) - F_t^o(x)]^2 dx, \quad (1)$$

where $F_t^f(x)$ is the cumulative distribution of the forecast at time step t , $F_t^o(x)$ is the step function corresponding to the observation at time step t (0 if x is smaller than the observation value at time step t and 1 otherwise), and N is the total number of time steps in the temporal period. For each initialization month, there is the temporal period (e.g., week, month, etc.) where the mean forecast value is calculated and the lead time.

We next assess the Continuous Ranked Probability Skill Score (CRPSS), as a measure of forecasting skill, by comparing the CRPS value of the forecast system to that of the selected benchmark. The skill metric is non-dimensional and takes values on $[-\infty, 1]$, with 1 the perfect forecast skill. Positive (negative) skill scores

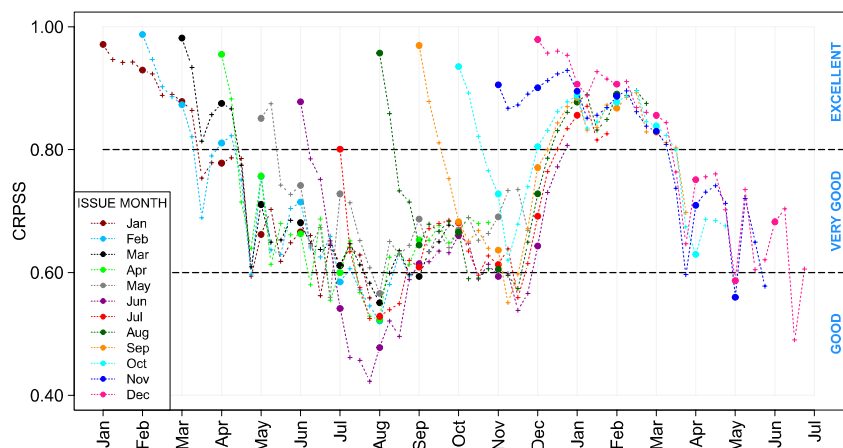


Figure 3. The median weekly CRPSS of the GCM-MF system initialized without data assimilation and benchmarked against real observation climatology. The forecasts were issued at the beginning of every month with a 7-month lead time and aggregated to the median for all stations. Black dashed lines demarcate the “good,” “very good” and “excellent” skill levels.

indicate that the forecast system performs better (worse) than the benchmark in terms of CRPS. Values close to 0 indicate that the forecast system has comparable performance to the benchmark. CRPSS is defined as:

$$CRPSS = 1 - \frac{CRPS_S}{CRPS_B}, \quad (2)$$

where $CRPS_S$ and $CRPS_B$ are the respective CRPS of the forecasting and benchmark systems. Finally for better communication of the results, we classified the skill into “good” ($0.4 \leq CRPSS < 0.6$), “very good” ($0.6 \leq CRPSS < 0.8$), and “excellent” ($CRPSS \geq 0.8$) ranges.

5. Results

5.1. Benchmarking Seasonal Streamflow Forecast Quality

Here we assess the skill of the GCM-MF streamflow forecasts initialized without DA using real observations as reference and the observed streamflow climatology as benchmark. The forecasts at each gauge in the study area were issued at the start of each month and the daily values are aggregated to weekly for the 7 months ahead, while the skill was computed for each station. However for better visualization, the median skill from all stations was used for each initialization month and lead week. Figure 3 shows the median CRPSS from all gauges for all lead weeks and initialization months. The skill is positive for all initialization months and lead weeks but as expected drops with increasing lead time, with the exception of forecasts targeting January and February whose quality is excellent regardless of the lead time. Both hydrological and meteorological initializations play a significant role since the skill at lead week 0 is always excellent with CRPSS greater than 0.8 (except for the July initialization). Depending on the initialization month, the skill can decrease significantly with values reaching very good ranges (CRPSS between 0.6 and 0.8) or even good ranges (CRPSS between 0.4 and 0.6), especially when targeting the months of July, August and November. Forecast skill is comparatively lower during the spring flood period (May–August) than in other months. Most of the streamflow volume is generated during May–August (see Figure 2) and hence, volumetric errors in streamflow measurements and inaccuracies in determining the start of snowmelt could be affecting the CRPSS values.

To better estimate the onset of snowmelt and improve the streamflow generation processes, it might be necessary to introduce snow (FSC, SWE) and temperature-related information (PET, AET) during the months preceding the spring flood, as well as streamflow information (QOBS, INFLOWS) directly. Additionally, seasonal streamflow peaks in the summer months with the hydrological response being a combination of snowmelt and rainfall (Figure 2). Consequently, the forecast skill depends on both the quantification of snowpack and predicted rainfall. It is also interesting to note that the decreasing trend of skill during the summer is sharper than for the other months, which is further proof of the low predictability during this season.

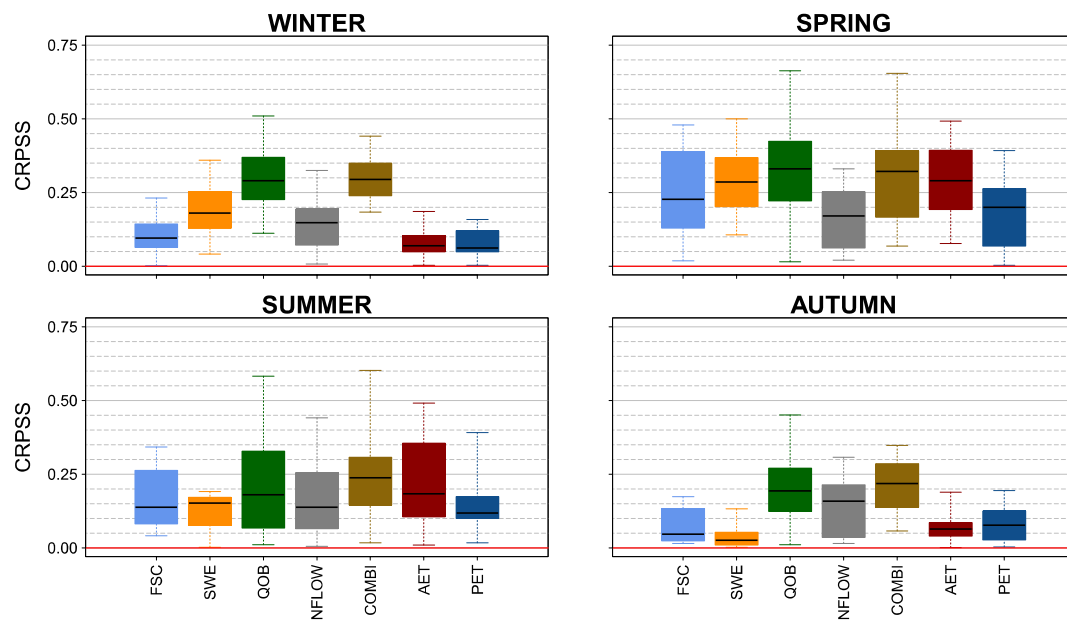


Figure 4. The gain in CRPSS for lead week 0 due to assimilation of different data sets. The red line corresponds to the skill when the “new” forecast (GCM-MF streamflow forecasts with DA) and benchmark (GCM-MF streamflow forecasts without DA) systems have identical long-term performances.

5.2. Impact of Data Assimilation on Hydrological Forecasting

In this section we assess the impact of assimilating different data sets in different seasons using the GCM-MF streamflow forecasts initialized with DA as the “new” forecasting systems and their counterpart system without DA as common benchmark. The CRPSS values for lead week 0 were computed after assimilating different data sets during model initialization and extracted separately for each station during the 3 months of the season (i.e., December–February for winter; March–May for spring, etc.). The forecasting skill was computed for each station and season as the 3-month median. Figure 4 shows the CRPSS for lead week 0 from all stations and during the four seasons.

The high skills during winter and spring highlight the importance of some initial hydrological conditions during these seasons. In autumn, the performance across all data sets is similar to that in winter yet comparatively low, probably due to the transient states of the snowpack and soil moisture storage from near total depletion during the summer to full capacity during the winter. The onset of polar nights also means that the summer control of evaporation gets gradually reduced until it becomes irrelevant in winter.

The skill variation between the different assimilation experiments and seasons can be attributed to the seasonal variability of both the accuracy of individual data sets used in the assimilation and the control of each variable on the hydrological response (Parajka & Blöschl, 2008). Most assimilated data sets attain their highest median skill during spring and lowest during summer/autumn. The QOB, COMBI, SWE and AET data sets have the top three median skills during the high flood months (spring and, to some extent, summer). There is relatively higher gain in the summer than in the autumn possibly due to the easy-to-beat low performance of the benchmark system in the summer.

The impact of assimilating in situ (streamflow and reservoir inflow) data on the median skill shows a limited seasonal pattern with the highest gain during spring and the lowest during summer. The skill from assimilating streamflow information is always high, because streamflow gauges aggregate the catchment's natural response from the entire upstream area and the E-HYPE model was originally calibrated using streamflow data. In addition, the skill from assimilating streamflow information is higher than that from reservoir inflow. Although we are currently not assessing the information content in the different types of observations and how this is extracted through DA, we note that reservoir inflow is a post-processed product that is derived by the hydropower companies after water balance assessments in the reservoirs. Consequently, reservoir inflow data could be subject to higher uncertainty and bias

than streamflow measurements, which can affect the forecast skill after assimilation. The artificial manipulation of inflows, which is sometimes abrupt, can lead to loss of information over time that is however preserved in streamflow.

The EO snow products (FSC and SWE) have the highest median skills in the spring and the lowest in the autumn; and SWE is more skillful than FSC, especially in spring and summer. In fact, SWE always outperforms reservoir inflow except in the autumn. The SWE, along with AET, has the highest skill after QOB and COMBI in spring, which further emphasizes the importance of snow information in estimating the spring flood. The high SWE skill in spring is consistent with previous findings in Greuell et al. (2019), Shukla et al. (2013), and Wanders et al. (2019), where the spring snow was one of the identified drivers of skill in snow-dominated European river systems. The low skills of snow data during the autumn might be due to the nearly depleted snowpack.

The EO evapotranspiration products (AET and PET) have the highest median skills during spring followed by summer and the lowest in winter and autumn for AET and PET, respectively. AET is clearly more skillful than PET in the spring and summer. Evapotranspiration affects the runoff generation in spring and summer through its control on snowmelt and soil moisture storage, which are particularly relevant for the high flood events. The low skills in autumn and winter are due to the low solar angles at the high latitude of the study area.

Moreover, the assimilation of the combination data set (COMBI) shows high potential, since the median improvements are always higher than the constituent data sets in the combination (except during spring when the QOB median gain is slightly higher), while the improvement is relatively more consistent (less variability in skill between stations and seasons) than for the other datasets. This first highlights the complementary roles the different data sets have. In addition, it underpins the value of assimilating combined datasets, which is receiving increased attention from the scientific community (see e.g., Gao et al. (2011)), although challenges still remain (MacBean et al., 2016).

Finally, we note that the spread in the CRPSS values is generally highest in spring and summer than in the other two seasons. This is possibly due to first the river's hydrological seasonal variability characterized by high streamflow during these seasons and second the quality of the EO and in situ data that might vary seasonally; although the latter is not investigated here. As mentioned in Section 5.1, the likely volumetric errors during the high-flow months and the fact that we merge results from different gauges, which are subject to different data quality, can affect the forecasts resulting into such spreads in the CRPSS values.

5.3. The Persistence of Data Assimilation at Seasonal Time Scales

Here we assess the “persistence” of assimilating different data sets in streamflow forecasting. We define this persistence as the number of lead weeks over which the performance of GCM-MF streamflow forecasts initialized with DA remains higher than that of its counterpart without DA (benchmark). It is thus the time taken for the IHC improvements to dissipate from the system. Figure 5 presents the persistence for different assimilated data sets across seasons. The persistence differs between the assimilated products and seasons, with the shortest persistences generally observed in summer.

The persistence of assimilating snow products (FSC and SWE) during winter and, in most locations, in spring persists for over 3 months. The snow products also show a relatively long persistence at some stations in the highlands during autumn due to the snowpack that starts to form. This highlights the dominant role of snow processes in controlling the seasonal hydrological regime in snow-dominated catchments, such as Umeälven, and also the potential gain from accurately estimating and integrating snow information in the hydrological model during these seasons.

The assimilation of in situ data sets shows its longest persistence in winter, and its shortest in summer and spring. The short summer persistence might be caused by the volumetric errors in the streamflow that affect the forecast skill. Persistence is often longer at the sub catchment streamflow outlets than at the reservoir inflow stations except in the summer and, at a few stations, autumn and winter. This could be due to the fact that the reservoir inflow data are subject to model structural errors.

Assimilating evapotranspiration (AET and PET) has long persistence particularly during winter, spring and autumn. The evapotranspiration data sets are based on temperature which can exert control on the timing of snow accumulation and snowmelt, and hence on the runoff volume from winter to spring. The study area experiences its highest rainfall volumes in the summer months, making precipitation exert more control on the soil storage dynamics than evapotranspiration. When the rain ceases at the beginning of autumn, the control on soil

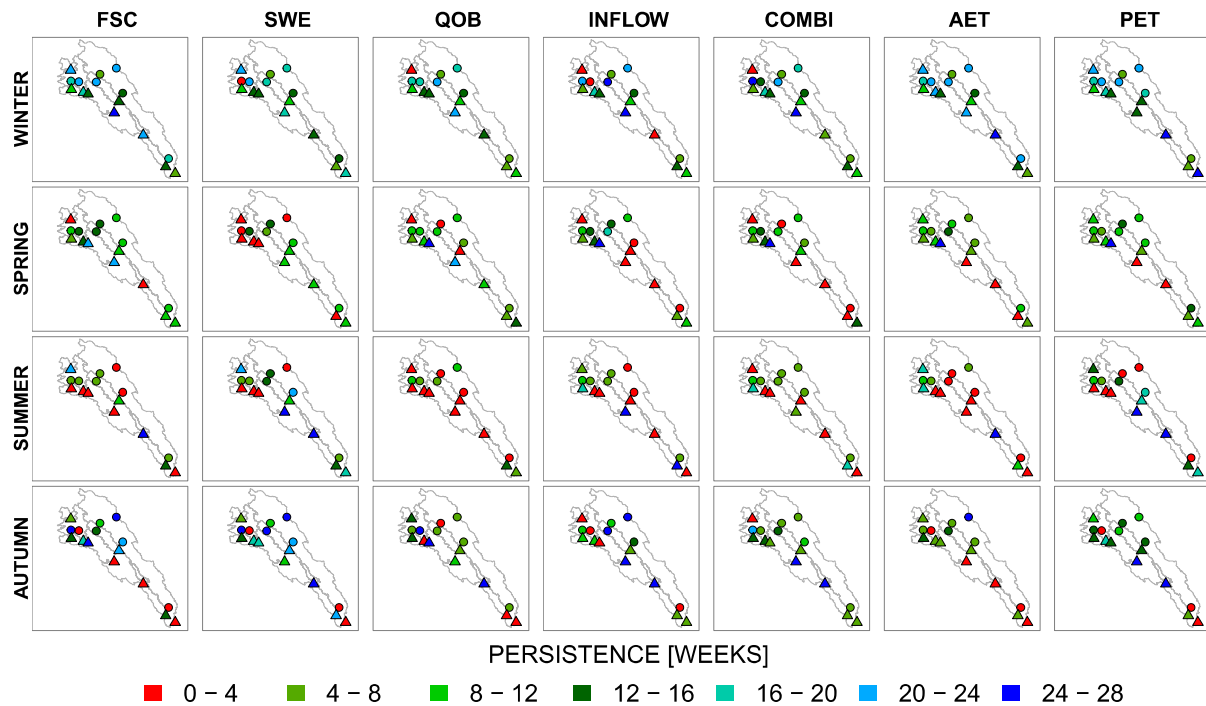


Figure 5. The persistence of DA (in weeks) in the streamflow forecast skill at sub catchment outflow (circles) and reservoir inflow (triangles) stations resulting from assimilating data sets during different seasons of the year. The forecast and benchmark systems are the GCM-MF streamflow forecasts with and without DA, respectively.

water storage and runoff generation reverts back to temperature. The description of physical process interactions presented here is supported by the numerical modeling exercise.

The persistence from assimilating COMBI is longer in the cold (autumn and winter) seasons than spring and summer due to the presence of the two in situ products that showed a similar behavior. Apart from FSC in spring and winter, assimilation of COMBI always persisted for longer than the individual constituent data sets, which again highlights the value of assimilating multiple data sets simultaneously. We also note that the combination returned the longest overall persistence during the autumn, which might be due to different data sets complementing one another. The persistence of different assimilations for all seasons at individual observation points is presented in Appendix A1.

5.4. Relative Importance of Data Assimilation and GCM-Based Meteorological Forcing on Hydrological Forecasting Performance

Here we recognize that the hydrological forecasting skill is controlled by both the initial hydrological conditions (whose improvement is achieved through DA) and meteorological forcing (whose improvement is achieved through the dynamic GCM-based data set). In particular, we aim at providing diagnostics and guidance on which and when the investigated products should be assimilated in order to improve the initial hydrological conditions, or when the focus should be on improving GCM-based meteorological data sets. We first compute the skill of the GCM-MF streamflow forecasts initialized with DA benchmarked against their counterpart without DA. This skill is solely due to the improvement of initial hydrological conditions through DA. Next we compute the skill of the GCM-MF streamflow forecast system benchmarked against the ESP system, in both cases considering no DA. The latter comparison quantifies the gain from the GCM-based MF. We infer the relative importance of DA and MF from the difference between these two intercomparisons: DA is more important if the difference is positive and vice versa.

Figure 6 shows the apportionment of the control on forecasting skill improvements from assimilating different data sets and from forcing the model with GCM-based MF across seasons and for lead week 0. In winter, DA is more important in most of the highland outflow stations (circles) and for most assimilated data sets except FSC, AET and PET. The marginal role of evapotranspiration might be due to the nearly non-existent evaporative fluxes at high latitudes, while the marginal role of the fractional snow cover might be explained by a maximum snow

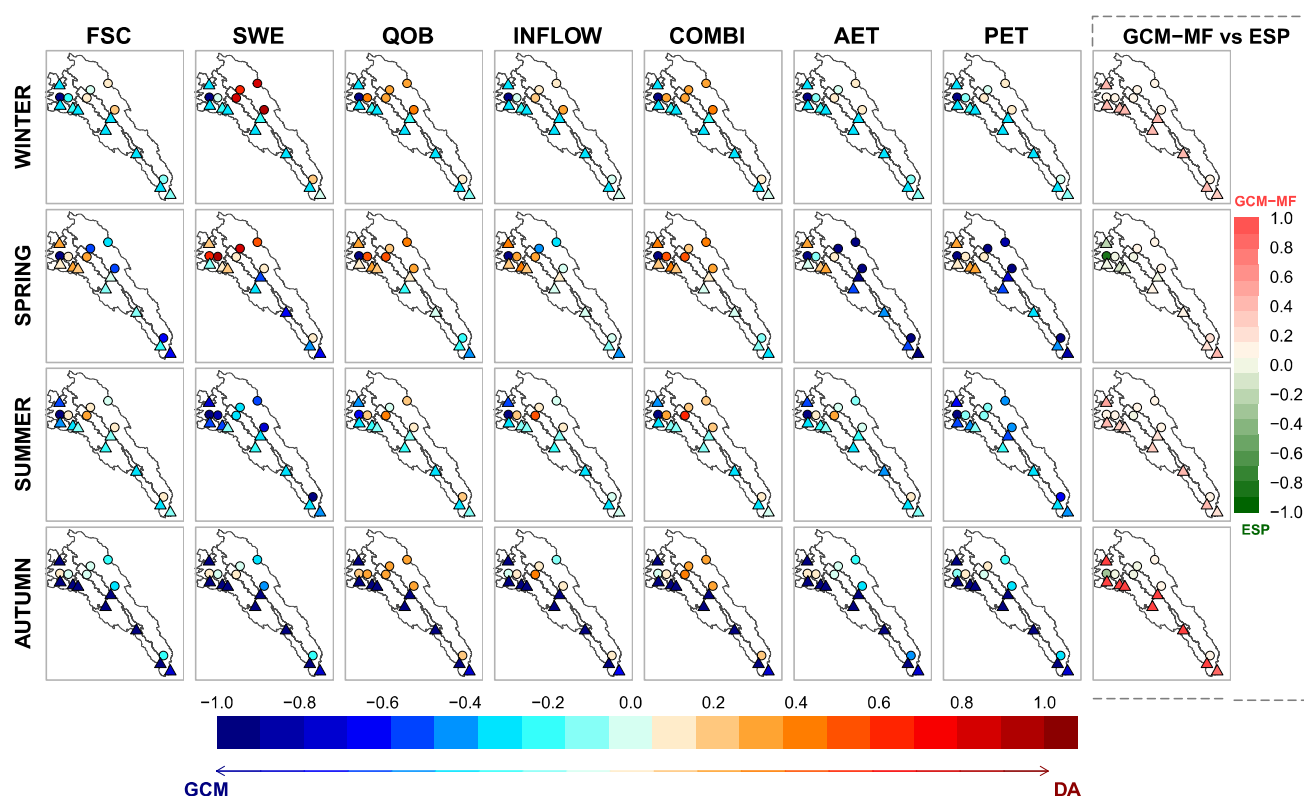


Figure 6. The relative importance of data assimilation and GCM-based meteorological forcing for different assimilations and seasons for lead week 0. The triangles and circles are the reservoir inflow and sub catchment outflow stations, respectively. Red and blue colors indicate importance of data assimilation and the meteorological forcing, respectively. The last column “GCM-MF versus ESP” shows the skill of GCM-based streamflow forecasts benchmarked against ESP-based streamflow forecasts, both without data assimilation. Red (green) is when the GCM-based (ESP) streamflow forecasts are more skillful.

cover in winter, leading to little information content coming from the FSC data set. However, DA has very low importance for reservoir inflows (triangles) everywhere with all datasets. This might be due to water being held in the snowpack during winter with very little melting to generate local runoff into reservoirs. Therefore, the reservoir inflow cannot be improved by assimilating any data set.

During spring, DA is again more important in the highlands for all assimilated data sets including FSC, and to a limited extent, AET and PET. This might be explained by the contribution of the data sets to the snowmelt and runoff processes. The GCM-MF shows high importance at some headwater stations for ET products possibly due to the dependence of ET on temperature, hence its role in snowmelt. The QOB, SWE and COMBI data sets returned the top three gains in skill in Figure 4, which is consistent with the current results.

Precipitation and temperature are the main constituents of forcing datasets. The annual rainfall peaks during May–August (see Figure 2), and accordingly, the gain from GCM-MF on seasonal streamflow forecasting exceeds that from data assimilation during these months. Nevertheless, the influence of data assimilation prevails in highland sub catchments for QOB, INFLOW and COMBI. This might be due to the relevance of the assimilated datasets’ information content to the aggregated response of the system. For the catchment under investigation, this emphasizes the importance of initial hydrological conditions and potentially the bigger role of snowmelt compared to rainfall in streamflow forecasting.

For streamflow forecasting in autumn, DA is more important than meteorological forcing in sub catchments when the in situ and combination of data sets are assimilated. The low importance of assimilating snow products might be due to the nearly depleted snowpack, although the importance is slightly higher in the highlands where the snowpack is either starting to form or did not completely melt during the summer. The moderate importance of assimilating ET in the highlands might be attributed to the end of the rainy season and subsequent resumption of the evaporation process control on soil storage and runoff generation. Meteorological forcing is more

important for streamflow forecasts (reservoir inflow stations) during the autumn, which can be attributed to the local temperature and precipitation conditions driving the reservoir inflow flux generation processes. Thus, forecasts of sub catchment outflow and reservoir inflows are more sensitive to DA and GCM-based forcings, respectively. This is in spite of the widely held belief in the community that at short lead times, meteorological forcing is always less important to seasonal streamflow forecasts than initial conditions.

The skill of the GCM-based streamflow forecasts benchmarked against the ESP-based forecasts, both initialized without data assimilation was used in the apportionment of the relative importance of DA and meteorological forcing presented in the preceding paragraphs. The column “GCM-MF versus ESP” in Figure 6 shows the skill between the two forcing data sets during different seasons. For the high-latitude, snow-dominated catchment under consideration, the streamflow forecasts based on the dynamic GCM-based meteorological forcing is always superior to ESP, which is derived from historical data, except during the spring in the highlands (both streamflow and reservoir inflow) and in the autumn at the highland streamflow gauges.

The evolution of the importance with lead time for all seasons and assimilations is presented in Appendix B1. For streamflow in winter and spring, DA is more important than the meteorological forcing (MF) for most assimilations at early lead times. Even where MF is more important, the importance reduces with time. SWE shows an abrupt change in importance from DA to MF between lead weeks 0 and 4. At lead week 0 in the summer, DA is slightly more important at most sub catchment outlets for all assimilations except SWE and PET but MF dominates from lead week 4 for all assimilations. In the autumn, the importance gradually changes from DA at early times to MF at some outlets and vice versa for others.

For reservoir inflow, MF is generally more important than DA during all seasons and most assimilations. During winter and spring, DA is more important at longer lead times while MF is always more important at all lead weeks during the summer. At early lead weeks in the autumn, MF has a very high importance everywhere for all assimilations but the importance reduces gradually and eventually changes to DA at lead week 20. The importance of MF for reservoir inflow might be due to the role of temperature and precipitation to the local runoff generation processes. Importance results at longer lead weeks 4, 8, 12, 16 and 20 are presented in Appendices B2–B6.

We investigated the relative importance of DA and meteorological forcing on the forecast quality further by comparing the forecast hydrographs of the 25 ensemble members to the observations at stream gauge 8313863 (see location in Figure 2). Figure 6 showed DA to be more important for both FSC and SWE assimilations during winter and spring, with SWE having a higher importance. From the same figure, both the GCM and ESP-based streamflow forecast systems performed equally well at the station during both seasons.

Moving to the local single station investigation, we present in Figure 7 the ESP and GCM-based (using the SEAS5 MF) streamflow forecasts without DA. Since both sets of forecasts are initialized without data assimilation, we are essentially comparing the impact of meteorological forcing. In winter, ESP forecasts outperform GCM-based streamflow forecasts in the first lead month, as its ensemble is more accurate (closer to observations), while in spring the GCM-based forecasts better anticipate the slow rising limb in April. The GCM-based forecasts estimate the peak discharge slightly better and produce a better ensemble spread around observations. In spring, the peaks are improved further when SWE data are assimilated, while in winter little difference can be observed from assimilating FSC (third column). The higher importance of DA in SWE assimilation (second row, third column) is evident from the very accurate prediction of the observed peak at lead week 12. This is consistent with the box plots in Figure 4. The error between the initializations and observations shown by the space between the solid black line and the maroon markers in the lower panel reduces from the two first sub figures to the last. This supports the conclusion that assimilating SWE information further improved initial states.

6. Discussion

6.1. Applicability of Insights in Other Domains

The diagnostic framework presented here allows detecting the sources of skill in operational seasonal streamflow forecasts. It was tested on a snow-dominated catchment using GCM-based forcing data sets and a single algorithm (i.e., Ensemble Kalman Filter) to assimilate a set of observations. Here we built upon the modeling work of Musuza et al. (2020), who achieved a high performance during model identification over the region indicating the acceptability of the model setup (structure and parameterization). Although one can argue that the

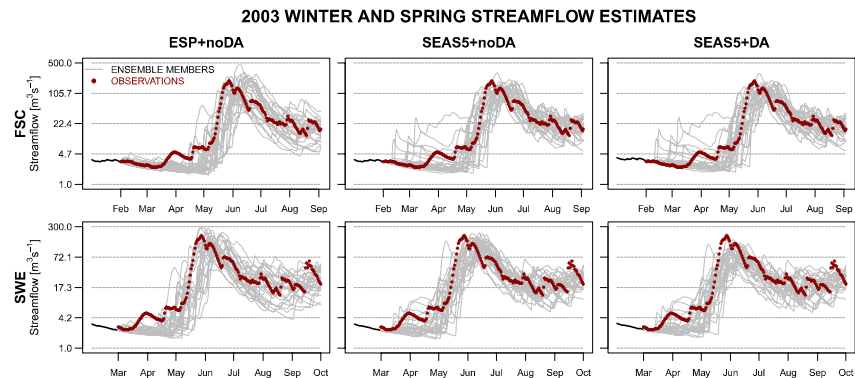


Figure 7. The forecast streamflow for winter (top row, issued at the start of February) and spring (bottom row, issued at the start of March). The first two columns show the forecasts using the ESP and GCM-based meteorological forcing without data assimilation, respectively. The third column is the GCM-based MF with assimilation of FSC (top) and SWE (bottom).

results presented here are valid for this specific experimental setup (i.e., catchment, hydrological model, forecast systems etc.), we highlight that our diagnostic framework can be extended to other impact models, river systems (controlled by different physical processes), assimilation methods and forcing datasets. The deployment of this framework in other hydro-climatic domains will broaden the knowledge of forecast skill attribution. Furthermore, to inform operational systems improvement, it should be deployed at geographical and temporal scales relevant to the system operations.

In a European investigation, Pechlivanidis et al. (2020) concluded that the seasonal hydrological predictability at lead months from 0 to 2 is strongly linked to the hydrological regime of the underlying domain. For instance, hydrological forecasts in flashy catchments (e.g., due to urbanization, impervious soil etc.) are characterized by low skill, while forecasts in catchments with long river memory (e.g., due to baseflow domination, snow contribution, dampening from lakes and wetlands) can achieve high skill. The latter statement was supported by the results here. Our findings are also consistent with previous studies relating to key drivers of seasonal hydrological forecasting skill (Crochemore et al., 2020; Giron Lopez et al., 2021; Harrigan et al., 2018). Therefore, our conclusions may be regionalized to other snow-dominated domains with comparable EO data quality. For applications at other snow-dominated catchments, deviations from the current results might arise, yet their exact magnitude is difficult to quantify. However, we do not expect our insights to change significantly due to the strong domination of seasonal snow melting/accumulation and temperature changes.

Finally, we note that even if the Swedish catchment's intra-annual response does not fully represent the snow accumulation/melting regimes over the globe, our diagnostic framework is still useful with the current insights potentially transferable to other similar river systems. In particular, systems with similar climatological and hydrological conditions over the globe can be of high relevance to the energy sector. Hence the results presented here can guide decision making in hydropower management and risk reduction (Contreras et al., 2020; Wörmann et al., 2020) by pointing out the seasons and lead times for which implementing ESP and/or GCM-MF hydrological forecasts with and without data assimilation can be strategic.

6.2. Recommendations for Moving Forward

We have extensively highlighted that the assimilation of different EO products increases the skill of hydrological forecasts. Although data assimilation primarily improves the initial state of the hydrological forecasting system, we note that the success of assimilation depends on the quality of the assimilated data as well. Therefore, we recommend efforts to further improve the quality of EO, for instance, through improvements in the spatiotemporal resolution of the products and/or the acquisition of in situ measurements of snow (transects) and evapotranspiration (eddy flux towers) to constrain and validate the respective EO products and to refine the post-processing algorithms. Future research should investigate the assimilation of the same variable from different providers, for example, evapotranspiration from GLEAM (available online: <https://www.gleam.eu>) and LSA-SAF (available online: <https://landsaf.ipma.pt/en>) (both accessed on 17.05.2022) to ascertain if the poor/good performances in the assimilation are related to the specific NASA MODIS products used here.

Here, we assimilated snow, evapotranspiration, streamflow and reservoir inflow data into the model, although hydrological forecasts in many river systems are known to be driven by antecedent soil storage conditions (J. Du et al., 2019). Soil moisture information was not considered in this investigation due to the strong control exerted by snow accumulation and melting on the seasonal hydrological signal in the catchment. Still, data on soil moisture and total water storage anomalies are available from different EO providers (e.g., ESA-CCI, NASA, etc.) at acceptable spatiotemporal resolutions. Soil moisture was reported as a key control variable on the hydrological predictions especially in high-latitude domains (Wanders et al., 2019). We therefore recommend investigating the impact of soil moisture on hydrological forecasting skill in snow-dominated catchments as well.

Uncertainty quantification is still missing in the EO products, although it could be used in the data assimilation; for instance, in the computation of the error covariance matrices to improve model performance in domains where EO data are uncertain (e.g., high-latitude autumn/winter evapotranspiration, complex terrain snow water equivalent, etc). Moreover most of the EO data sets are not available in (near) real time. This consequently limits their use to historical investigations and validation of hindcasts. In order to use the data in real time forecasting services, we recommend the development of platforms such as the CRYOLAND project (<http://www.cryoland.eu/>) that have short acquisition/retrieval/processing times to facilitate assimilation of the EO data in operational forecasting services.

Finally, we note that the E-HYPE model was originally calibrated against streamflow, and we showed here that its assimilation had the highest impact on forecast skill. However, the observed streamflow at the outlet of a large snow-dominated catchment represents past headwater snowmelt events. This is due to the delay between the temperature conditions triggering the melting and the propagation of the hydrological response through the river system. Such “memory” effects should be integrated into streamflow observations (e.g., as metadata) in regions with long concentration times, while the data assimilation algorithms should be modified accordingly to take this information into consideration, in order to further improve the hydrological predictions.

7. Conclusions

A scientific experiment was conducted that allowed, for the first time to our knowledge, the quantification of the gain in skill of seasonal streamflow forecasts from assimilation of both individual and combined EO and in situ data. A methodological framework was proposed to diagnose the relative importance of meteorological forcing and data assimilation of different hydrological variables, that is, four EO (fractional snow cover, snow water equivalent, actual and potential evapotranspiration); two in situ (streamflow and reservoir inflow) and the combination of data sets COMBI. We investigated the seasonal hydrological predictions in the snowmelt-driven Umeälven catchment in northern Sweden, which is heavily regulated for hydropower production, and hence seasonal prediction of streamflow and reservoir inflow is of high economic benefit. The results from this investigation are very promising for enhancing hydro-climate services for snow-dominated river systems.

Assimilating different in situ and EO data sets improves the forecast skill, but the improvement varies depending on the season, lead time and assimilated data. The persistence, that is the lead weeks during which data assimilation influences the forecast skill (here quantified as the number of successive weeks when the skill from the system initialized with DA remains above that of the counterpart system without DA), is also different for different data sets and seasons. Overall the assimilated data sets exert more control on the forecast quality than the meteorological forcing data, highlighting the importance of accurately quantifying the initial hydrological conditions for this snow-dominated river system.

In particular, the results show a generally high skill at short lead times for the forecasting system initialized without DA, although it is comparatively lower in the summer months. The skill was further improved by assimilating different in situ and EO products during all seasons of the year. After data assimilation, the skill was highest in winter and spring, while for most of the assimilation experiments, it showed a large spatial variability in spring and summer. We attribute this large spread to the errors related to estimating the spring and summer floods and mixing performances from several observation points.

The results show a high median skill improvement with the assimilation of snow and evapotranspiration during spring at the onset of the spring flood, which is of high interest to hydropower reservoir operators. Seasonal forecasts can hence be used in the regulation of the reservoirs resulting, for instance, in storing the water or using it for power generation even when energy demand/prices might not be optimal. Otherwise if the reservoirs fill up

before the end of the spring flood, water would need to be spilled which is of no economic value. Improvements in the forecast skill through DA can consequently result in improved and highly certain decisions. Nevertheless, this further highlights the potential for in situ snow and evapotranspiration measurements, which can be used for validating and adjusting the EO datasets.

The improvements from DA persisted in the forecasting system for long periods with the persistence depending on the assimilated data and season. In particular, the impact of assimilating SWE during winter persisted for more than 3 months, which again underpins the value of accurate snow information in space and time. Moreover, at short lead weeks, data assimilation is relatively more important than meteorological forcing to improve the hydrological forecast skill during all seasons except autumn. This not only emphasizes the benefit of improved data sets for assimilation, but also the benefit of improved and efficient assimilation algorithms.

Appendix A: Seasonal Persistences

The persistence at each observation point for different seasons and assimilations.

Table A1

The Persistence in Weeks at the Stream Gauges and Reservoir Outlets (See Figure 2 for the Location IDs) for Different Assimilations and Seasons

		Location	FSC	SWE	QOB	INFLOW	COMBI	AET	PET
Winter	Stream gauges	8300034	18	26	0	10	12	9	9
		8301225	22	22	5	14	5	14	28
		8313863	22	14	13	6	5	5	0
		8313880	6	6	13	11	4	4	12
		8301032	21	20	0	7	11	0	25
		8000742	15	15	10	8	4	19	16
		8307177	19	7	11	11	4	3	4
	Reservoir outlets	4000001	21	2	0	9	2	13	2
		4000002	15	12	10	8	3	18	7
		4000003	17	17	0	10	3	0	10
		4000004	26	26	12	2	4	28	28
		4000005	23	5	10	2	4	28	28
		4000006	13	13	0	0	14	14	5
		4000007	8	10	10	7	6	19	7
		4000008	10	8	0	5	0	0	10
		4000009	16	15	2	28	2	2	4
Spring	Stream gauges	8300034	0	18	10	11	12	27	23
		8301225	17	22	11	10	5	16	15
		8313863	22	22	10	6	5	0	5
		8313880	5	5	3	13	5	12	7
		8301032	17	21	9	10	4	25	8
		8000742	15	13	7	10	4	21	10
		8307177	14	22	8	13	6	0	7

Table A1
Continued

		Location	FSC	SWE	QOB	INFLOW	COMBI	AET	PET
Summer	Reservoir outlets	4000001	21	21	2	10	5	8	2
		4000002	11	11	3	8	2	23	7
		4000003	15	21	11	10	4	17	14
		4000004	20	21	22	2	28	0	26
		4000005	13	25	10	2	4	0	27
		4000006	6	13	5	7	28	14	5
		4000007	18	5	16	10	3	4	7
		4000008	9	12	6	5	18	14	14
		4000009	14	16	28	28	2	26	8
	Stream gauges	8300034	18	18	11	8	12	27	7
		8301225	24	22	18	5	5	27	15
		8313863	19	22	14	5	5	0	0
		8313880	7	6	14	5	5	9	8
		8301032	19	21	10	4	5	27	25
		8000742	14	17	2	4	5	21	8
		8307177	7	5	2	6	6	0	0
	Reservoir outlets	4000001	4	21	2	21	2	8	5
		4000002	11	15	3	12	3	22	7
		4000003	15	17	14	0	4	17	8
		4000004	21	16	2	4	5	9	0
		4000005	13	25	2	25	4	26	0
		4000006	8	8	6	16	17	24	7
		4000007	9	26	11	3	3	4	0
		4000008	12	12	6	0	18	15	15
		4000009	16	16	28	2	2	20	8
Autumn	Stream gauges	8300034	24	10	10	8	10	7	16
		8301225	26	13	11	14	4	8	11
		8313863	3	13	14	5	5	28	0
		8313880	6	13	3	13	4	3	14
		8301032	21	10	9	0	5	6	12
		8000742	14	10	8	24	3	8	15
		8307177	6	11	2	6	3	6	0
	Reservoir outlets	4000001	2	2	2	22	20	14	10
		4000002	10	10	11	10	3	7	14
		4000003	17	13	14	0	2	3	17
		4000004	28	22	2	28	0	9	28
		4000005	4	2	2	28	25	26	28
		4000006	13	10	3	14	12	3	7
		4000007	10	10	16	19	2	2	0
		4000008	8	6	6	0	18	16	15
		4000009	15	22	28	2	2	27	14

Appendix B: Supplementary Figures

The figure shows the relative importance between DA and meteorological forcing at individual observation points for the different assimilations and seasons.

The relative importance between data assimilation and meteorological forcing at longer lead weeks.

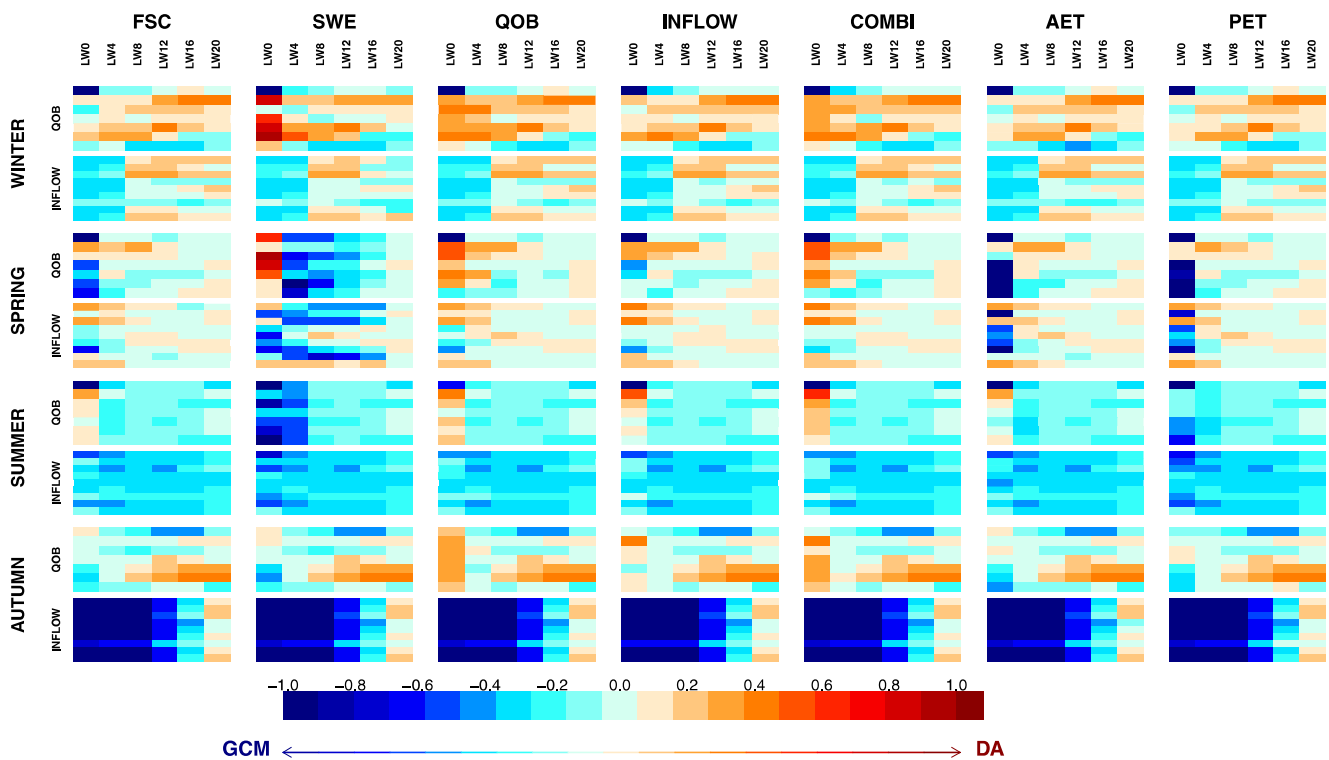


Figure B1. The variation of the importance between DA and meteorological forcing with lead time. Lead weeks (LW) 0, 4, 8, 12, 16, and 20 are shown for each assimilated data set and season. The streamflow, QOB is plotted on top the reservoir inflow, INFLOW for each season.

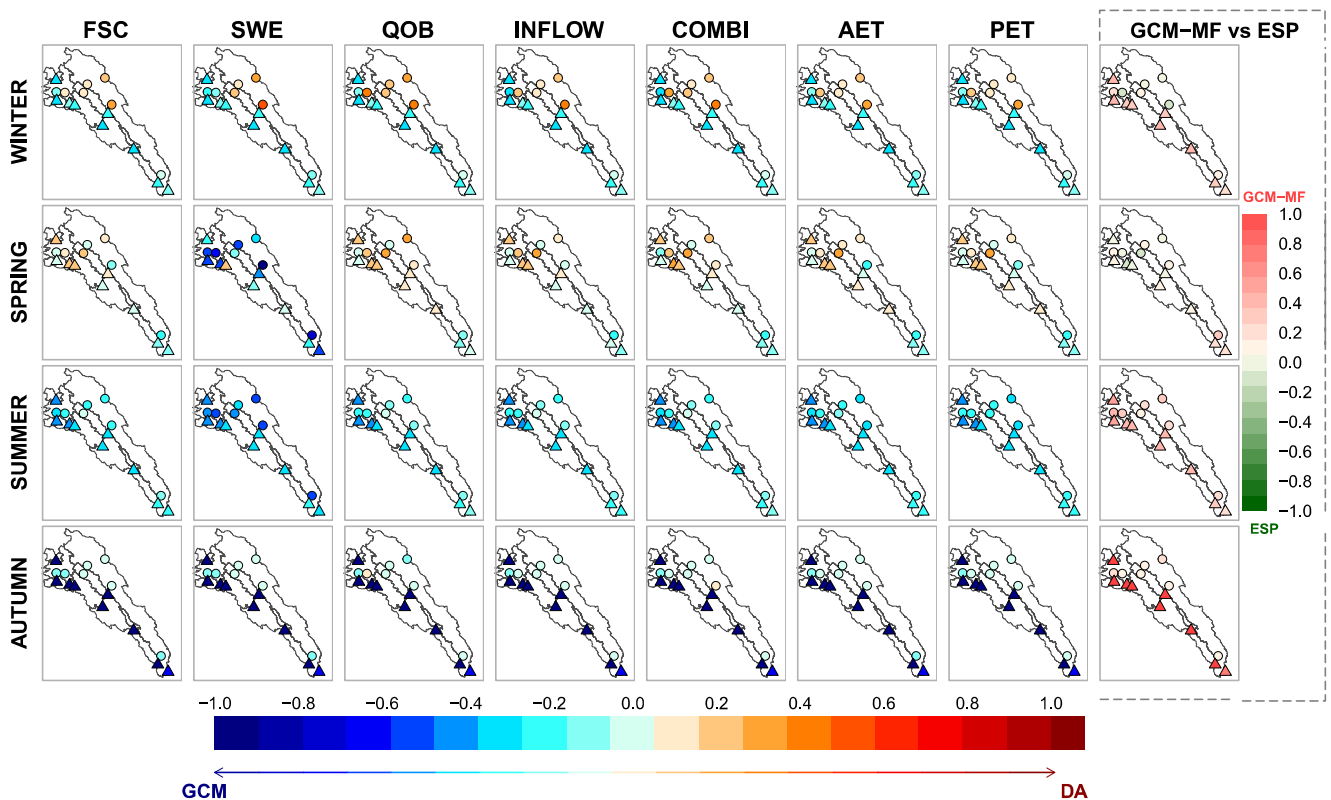


Figure B2. The relative importance of DA and GCM-based MF for different assimilations and seasons for lead week 4. The triangles and circles are the reservoir inflow and sub catchment outflow stations, respectively. Red and blue colors indicate importance of data assimilation and the meteorological forcing, respectively. The last column “GCM-MF versus ESP” shows the skill of GCM-based streamflow forecasts benchmarked against ESP-based streamflow forecasts, both without data assimilation. Red (green) is when the GCM-based (ESP) streamflow forecasts are more skillful.

See Figure B3

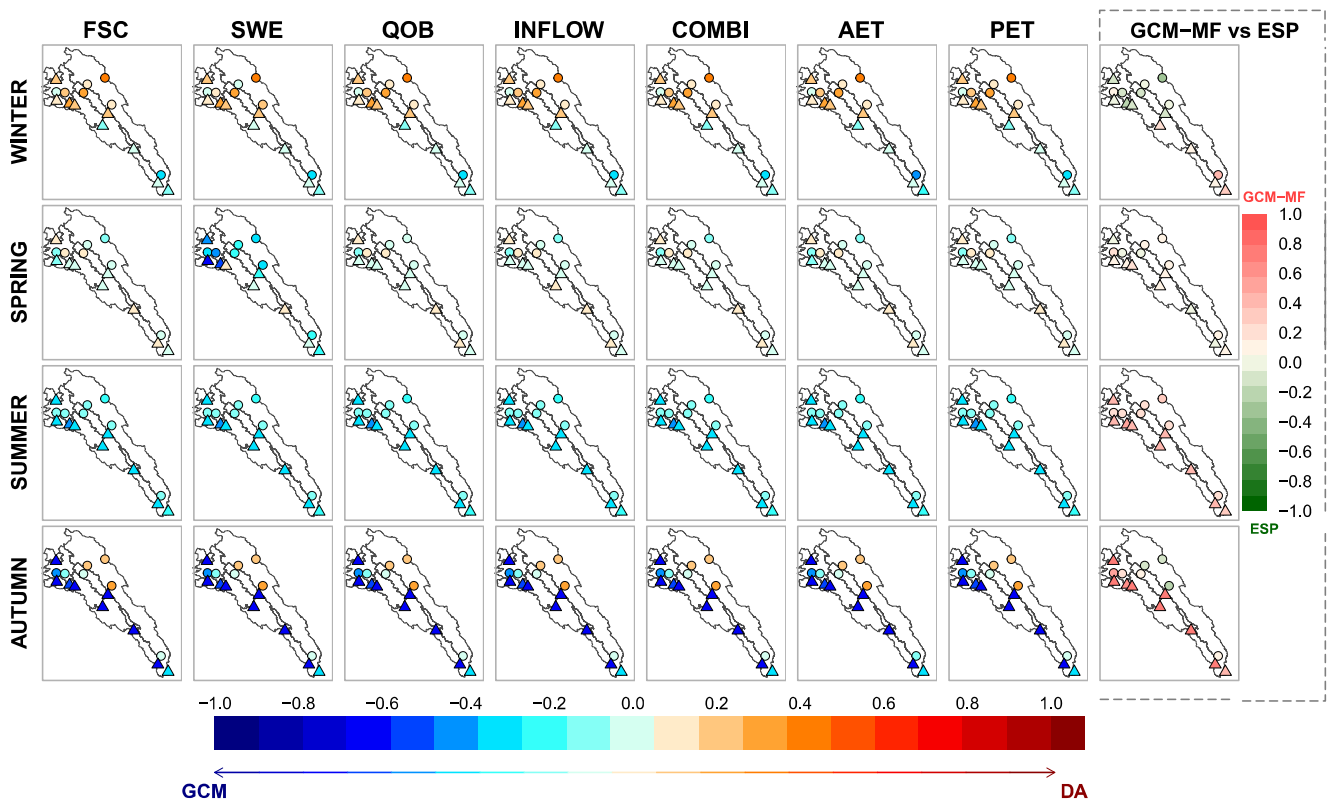


Figure B3. The relative importance of DA and GCM-based MF for different assimilations and seasons for lead week 8. The triangles and circles are the reservoir inflow and sub catchment outflow stations, respectively. Red and blue colors indicate importance of data assimilation and the meteorological forcing, respectively. The last column “GCM-MF versus ESP” shows the skill of GCM-based streamflow forecasts benchmarked against ESP-based streamflow forecasts, both without data assimilation. Red (green) is when the GCM-based (ESP) streamflow forecasts are more skillful.

See Figure B4

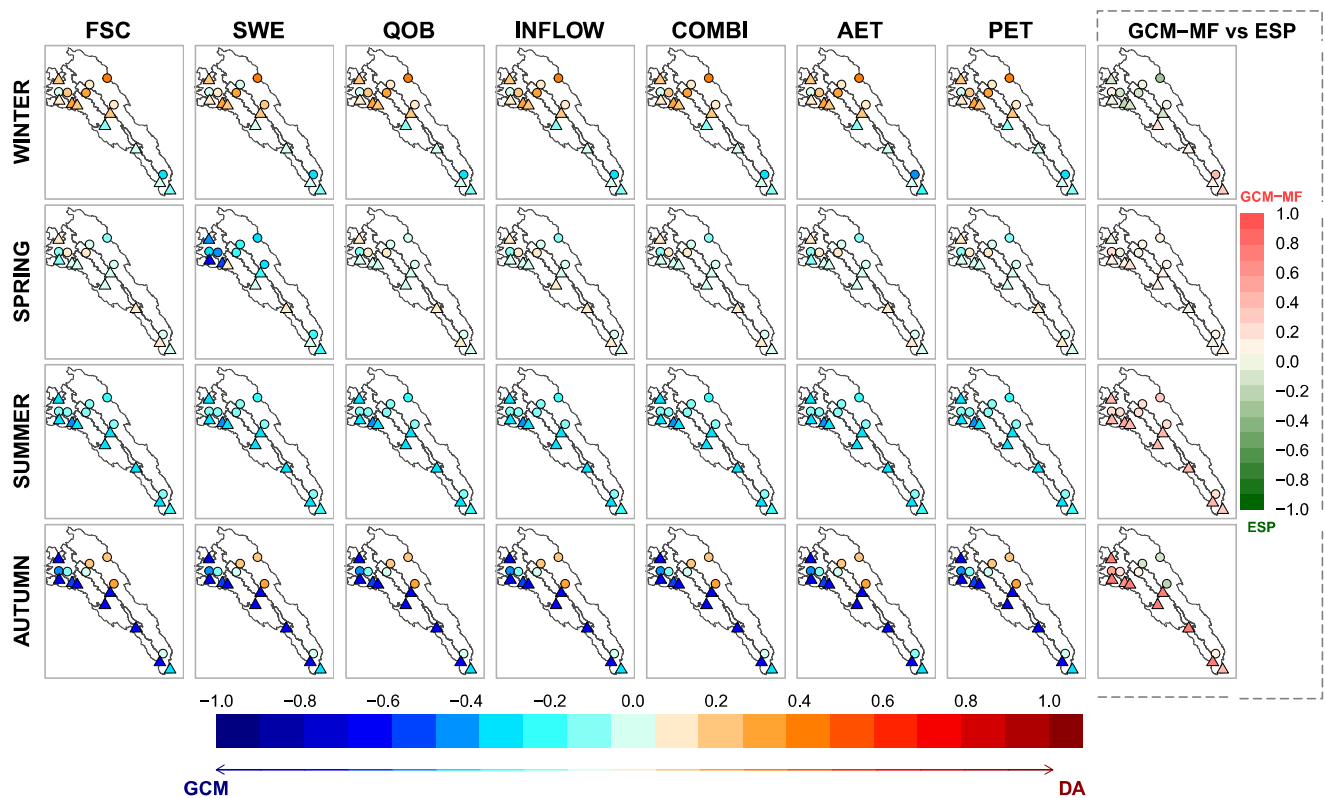


Figure B4. The relative importance of DA and GCM-based MF for different assimilations and seasons for lead week 12. The triangles and circles are the reservoir inflow and sub catchment outflow stations, respectively. Red and blue colors indicate importance of data assimilation and the meteorological forcing, respectively. The last column “GCM-MF versus ESP” shows the skill of GCM-based streamflow forecasts benchmarked against ESP-based streamflow forecasts, both without data assimilation. Red (green) is when the GCM-based (ESP) streamflow forecasts are more skillful.

See Figure B5

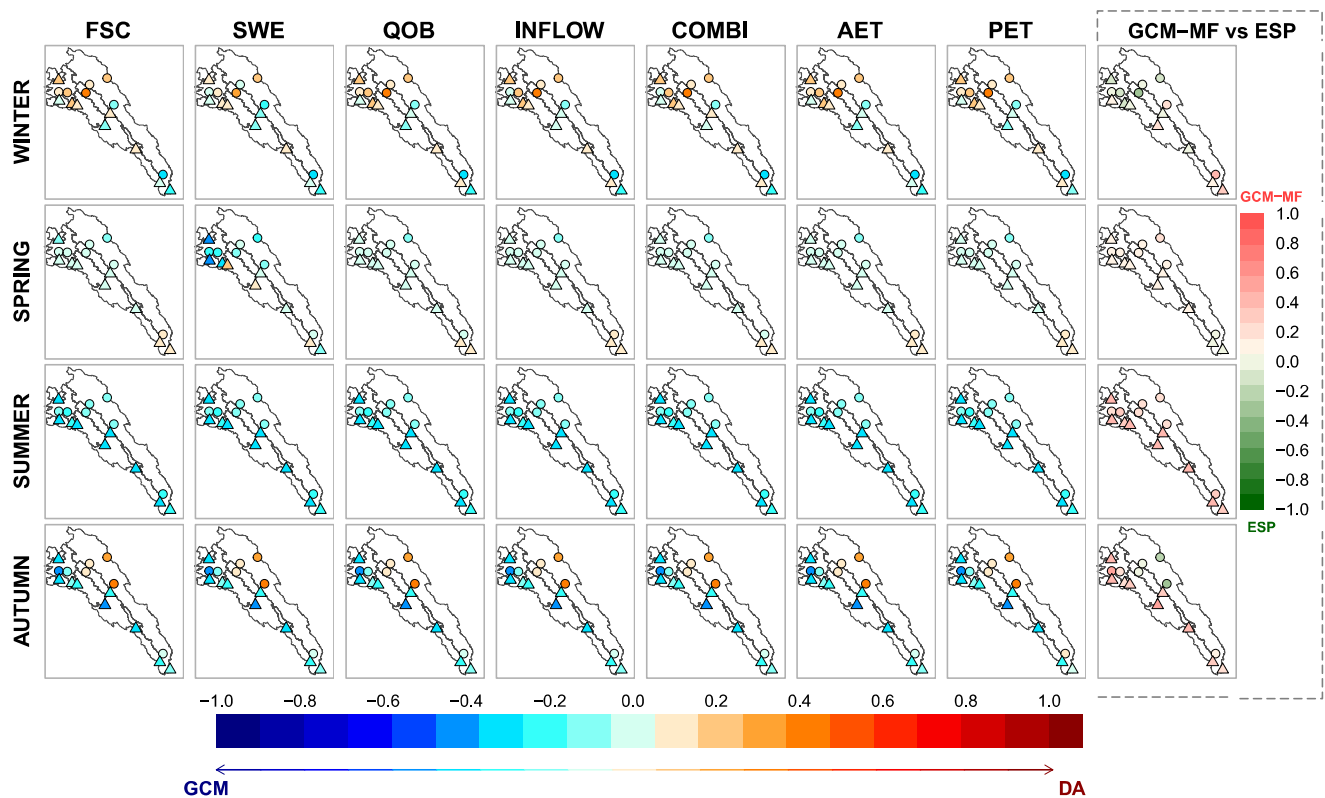


Figure B5. The relative importance of DA and GCM-based MF for different assimilations and seasons for lead week 16. The triangles and circles are the reservoir inflow and sub catchment outflow stations, respectively. Red and blue colors indicate importance of data assimilation and the meteorological forcing, respectively. The last column “GCM-MF versus ESP” shows the skill of GCM-based streamflow forecasts benchmarked against ESP-based streamflow forecasts, both without data assimilation. Red (green) is when the GCM-based (ESP) streamflow forecasts are more skillful.

See Figure B6

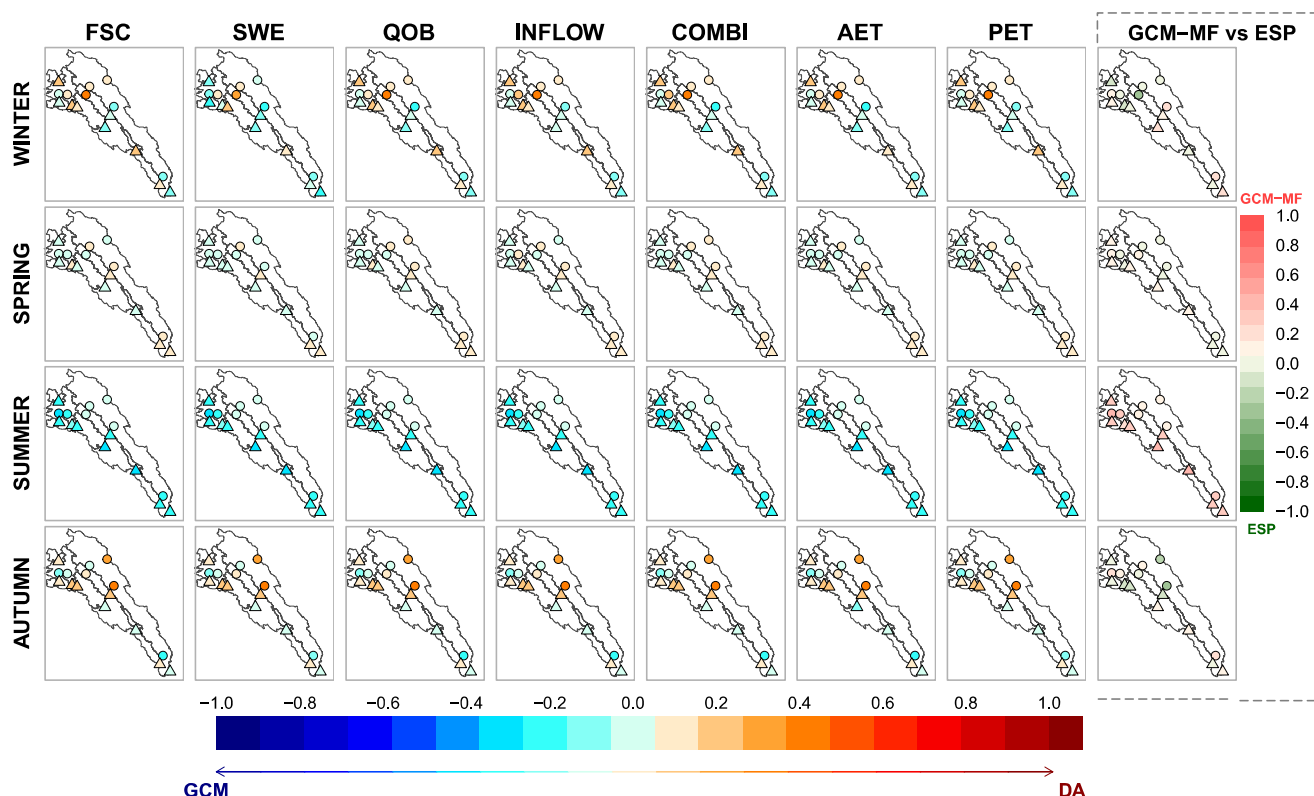


Figure B6. The relative importance of DA and GCM-based MF for different assimilations and seasons for lead week 20. The triangles and circles are the reservoir inflow and sub catchment outflow stations, respectively. Red and blue colors indicate importance of data assimilation and the meteorological forcing, respectively. The last column “GCM-MF versus ESP” shows the skill of GCM-based streamflow forecasts benchmarked against ESP-based streamflow forecasts, both without data assimilation. Red (green) is when the GCM-based (ESP) streamflow forecasts are more skillful.

Appendix C: General Information

C1. Acronyms

Acronyms

SMHI	Swedish Meteorological and Hydrological Institute
HydroGFD	Hydrological Global Forcing Dataset
ESP	Ensemble Streamflow Prediction
ECMWF	European Centre for Medium-range Weather Forecasts
DA	Data assimilation
EO	Earth Observations
MF	Meteorological Forcing
SF	Seasonal Forecast(ing)
SWE	Snow Water Equivalent
FSC	Fractional Snow Cover
AET	Actual Evapotranspiration
PET	Potential Evapotranspiration
QOB	Streamflow Observation(s)
CRPS	Continuous Ranked Probability Score
CRPSS	Continuous Ranked Probability Skill Score
GCM	Global Circulation Model

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

All data sets used in this research except the reservoir inflows are open and freely available. They are available from the websites and citations listed below.

SWE: Kouki et al. (2021), <https://data.ceda.ac.uk/neodc/esacci/snow/data/swe>

FSC: Malnes et al. (2015), <http://neso.cryoland.enveo.at/cryoclient>

AET: Mu et al. (2011), <https://lpdaac.usgs.gov/products/mod16a2v006/>

PET: Mu et al. (2011), <https://lpdaac.usgs.gov/products/mod16a2v006/>

QOB: SMHI (2022), Vattenwebb portal <https://vattenwebb.smhi.se/station/#/> or on request from the SMHI.

References

- Abbaszadeh, P., Moradkhani, H., & Daescu, D. N. (2019). The quest for model uncertainty quantification: A hybrid ensemble and variational data assimilation framework. *Water Resources Research*, 55(3), 2407–2431. <https://doi.org/10.1029/2018WR023629>
- Abbaszadeh, P., Moradkhani, H., & Yan, H. (2018). Enhancing hydrologic data assimilation by evolutionary particle filter and Markov chain Monte Carlo. *Advances in Water Resources*, 111, 192–204. <https://doi.org/10.1016/j.advwatres.2017.11.011>
- Alfieri, L., Pappenberger, F., Wetterhall, F., Haiden, T., Richardson, D., & Salamon, P. (2014). Evaluation of ensemble streamflow predictions in Europe. *Journal of Hydrology*, 517, 913–922. <https://doi.org/10.1016/j.jhydrol.2014.06.035>
- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration - guidelines for computing crop water requirements-crop evapotranspiration - guidelines for computing crop water requirements*. (Tech. Rep. No. 5-AGWA-8). Food and Agriculture Organization of the United Nations.
- Anctil, F., & Ramos, M.-H. (2019). Verification metrics for hydrological ensemble forecasts. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 893–922). Springer.
- Anghileri, D., Voisin, N., Castelletti, A., Pianosi, F., Nijssen, B., & Lettenmaier, D. P. (2016). Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments. *Water Resources Research*, 52(6), 4209–4225. <https://doi.org/10.1002/2015WR017864>
- Arnal, L., Wood, A. W., Stephens, E., Cloke, H. L., & Pappenberger, F. (2017). An efficient approach for estimating streamflow forecast skill elasticity. *Journal of Hydrometeorology*, 18(6), 1715–1729. <https://doi.org/10.1175/JHM-D-16-0259.1>
- Berg, P., Donnelly, C., & Gustafsson, D. (2018). Near-real-time adjusted reanalysis forcing data for hydrology. *Hydrology and Earth System Sciences*, 22(2), 989–1000. <https://doi.org/10.5194/hess-22-989-2018>
- Bertino, L., Evensen, G., & Wackernagel, H. (2003). Sequential data assimilation techniques in oceanography. *International Statistical Review*, 71(2), 223–241. <https://doi.org/10.1111/j.1751-5823.2003.tb00194.x>
- Buizza, R., Du, J., Toth, Z., & Hou, D. (2019). Major operational ensemble prediction systems (EPS) and the future of EPS. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 151–174). Springer.
- Colossi, B. R., & Tucci, C. E. M. (2020). *Ensemble long-term soil moisture forecast using hydrological modeling* (Vol. 25). RBRH. <https://doi.org/10.1590/2318-0331.252020190036>
- Contreras, E., Herrero, J., Crochemore, L., Aguilar, C., & Polo, M. J. (2020). Seasonal climate forecast skill assessment for the management of water resources in a run of river hydropower system in the Piqueira River (southern Spain). *Water-Sui*, 12(8), 2119. <https://doi.org/10.3390/w12082119>
- Crochemore, L., Cantone, C., Pechlivanidis, I. G., & Photiadou, C. S. (2021). How does seasonal forecast performance influence decision-making? Insights from a serious game. *Bulletin of the American Meteorological Society*, 102(9), E1682–E1699. <https://doi.org/10.1175/BAMS-D-20-0169.1>
- Crochemore, L., Ramos, M.-H., & Pappenberger, F. (2016). Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts. *Hydrology and Earth System Sciences*, 20(9), 3601–3618. <https://doi.org/10.5194/hess-20-3601-2016>
- Crochemore, L., Ramos, M.-H., Pappenberger, F., & Perrin, C. (2017). Seasonal streamflow forecasting by conditioning climatology with precipitation indices. *Hydrology and Earth System Sciences*, 21(3), 1573–1591. <https://doi.org/10.5194/hess-21-1573-2017>
- Crochemore, L., Ramos, M.-H., & Pechlivanidis, I. G. (2020). Can continental models convey useful seasonal hydrologic information at the catchment scale? *Water Resources Research*, 56(2), e2019WR025700. <https://doi.org/10.1029/2019WR025700>
- Day, G. N. (1985). Extended streamflow forecasting using NWSRFS. *Journal of Water Resources Planning and Management*, 111(2), 157–170. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1985\)111:2\(157\)](https://doi.org/10.1061/(ASCE)0733-9496(1985)111:2(157))
- Dechant, C., & Moradkhani, H. (2011). Radiance data assimilation for operational snow and streamflow forecasting. *Advances in Water Resources*, 34(3), 351–364. <https://doi.org/10.1016/j.advwatres.2010.12.009>
- DeChant, C. M., & Moradkhani, H. (2011). Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrology and Earth System Sciences*, 15(11), 3399–3410. <https://doi.org/10.5194/hess-15-3399-2011>
- DeChant, C. M., & Moradkhani, H. (2014). Toward a reliable prediction of seasonal forecast uncertainty: Addressing model and initial condition uncertainty with ensemble data assimilation and sequential Bayesian combination. *Journal of Hydrology*, 519, 2967–2977. <https://doi.org/10.1016/j.jhydrol.2014.05.045>
- DeChant, C. M., & Moradkhani, H. (2015). Analyzing the sensitivity of drought recovery forecasts to land surface initial conditions. *Journal of Hydrology*, 526, 89–100. <https://doi.org/10.1016/j.jhydrol.2014.10.021>
- de Lavenne, A., Andréassian, V., Crochemore, L., Lindström, G., & Arheimer, B. (2022). Quantifying multi-year hydrological memory with catchment forgetting curves. *Hydrology and Earth System Sciences*, 26(10), 2715–2732. <https://doi.org/10.5194/hess-26-2715-2022>
- Du, J., Berner, J., Buizza, R., Charron, M., Houtekamer, P., Hou, D., et al. (2019). Ensemble methods for meteorological predictions. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 103–117). Springer.

Acknowledgments

This study was partially funded by the EU Horizon 2020 project PrimeWater (Delivering advanced predictive tools from medium to seasonal range for water dependent industries exploiting the cross-cutting potential of EO and hydroecological modeling) under the Grant Agreement 870497, and from the EU Horizon 2020 project E-SHAPE (EuroGEOSS Showcases: Applications Powered by Europe) under Grant Agreement 820852. In addition, this study was partially funded by the EU Horizon 2020 project I-CISK (Innovating climate services through integrating scientific and local knowledge) under Grant Agreement 101037293, and from the project “The role of hydropower as a regulating resource in a renewable energy system with climate impact and increased internationalization of electricity markets” granted by the Swedish Energy Agency (Grant 52095-1). The investigation was performed at the SMHI Hydrology Research unit, where much work benefits from joint efforts in developing models and concepts by the whole team. The seasonal meteorological forecasts SEAS5 from the European Centre for Medium-Range Weather Forecasts are freely accessible from the Copernicus Climate Data Store (<https://cds.climate.copernicus.eu/>). The HYPE model code is available from the HYPEweb portal (<https://hypeweb.smhi.se/model-water/>). Real-time seasonal forecasts obtained through E-HYPE are openly available on the HYPEweb portal (<https://hypeweb.smhi.se/explore-water/forecasts/seasonal-forecasts-europe/>). Our funders had no role in the choice of research project; design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. Author contributions: JLM contributed to the code development, model runs, result analysis and figures, interpretation of results, and writing of the manuscript. LC contributed to the basic idea, study design, code development, result analysis and figures, interpretation of results, and writing of the manuscript. IGP was responsible for the project management and funding acquisition and contributed to the basic idea, study design, result analysis and figures, interpretation of results, and writing of the manuscript.

- Du, T. L., Lee, H., Bui, D. D., Graham, P. L., Darby, S. D., Pechlivanidis, I. G., et al. (2022). Streamflow prediction in highly regulated, trans-boundary watersheds using multi-basin modeling and remote sensing imagery. *Water Resources Research*, 58(3), e2021WR031191. <https://doi.org/10.1029/2021WR031191>
- Essenfelder, A. H., Larosa, F., Mazzoli, P., Bagli, S., Broccoli, D., Luzzi, V., et al. (2020). Smart climate hydropower tool: A machine-learning seasonal forecasting climate service to support cost-benefit analysis of reservoir management. *Atmosphere*, 11(12), 1305. <https://doi.org/10.3390/atmos11121305>
- Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4), 343–367. <https://doi.org/10.1007/s10236-003-0036-9>
- Gao, C., Wang, H., Weng, E., Lakshminarayanan, S., Zhang, Y., & Luo, Y. (2011). Assimilation of multiple data sets with the ensemble Kalman filter to improve forecasts of forest carbon dynamics. *Ecological Applications*, 21(5), 1461–1473. <https://doi.org/10.1890/09-1234.1>
- Girons Lopez, M., Crochemore, L., & Pechlivanidis, I. G. (2021). Benchmarking an operational hydrological model for providing seasonal forecasts in Sweden. *Hydrology and Earth System Sciences*, 25(3), 1189–1209. <https://doi.org/10.5194/hess-25-1189-2021>
- Giuliani, M., Crochemore, L., Pechlivanidis, I., & Castelletti, A. (2020). From skill to value: Isolating the influence of end user behavior on seasonal forecast assessment. *Hydrology and Earth System Sciences*, 24(12), 5891–5902. <https://doi.org/10.5194/hess-24-5891-2020>
- Greuell, W., Franssen, W. H. P., & Hutjes, R. W. A. (2019). Seasonal streamflow forecasts for Europe – Part 2: Sources of skill. *Hydrology and Earth System Sciences*, 23(1), 371–391. <https://doi.org/10.5194/hess-23-371-2019>
- Gupta, H., Kling, H., Yilmaz, K., & Martinez, G. (2009). Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Harrigan, S., Prudhomme, C., Parry, S., Smith, K., & Tanguy, M. (2018). Benchmarking ensemble streamflow prediction skill in the UK. *Hydrology and Earth System Sciences*, 22(3), 2023–2039. <https://doi.org/10.5194/hess-22-2023-2018>
- Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5), 559–570. [https://doi.org/10.1175/1520-0434\(2000\)015<0559:DOTCRP>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2)
- Hundecha, Y., Arheimer, B., Donnelly, C., & Pechlivanidis, I. (2016). A regional parameter estimation scheme for a pan-European multi-basin model. *Journal of Hydrology: Regional Studies*, 6, 90–111. <https://doi.org/10.1016/j.ejrh.2016.04.002>
- Jenicek, M., Seibert, J., Zappa, M., Staudinger, M., & Jonas, T. (2016). Importance of maximum snow accumulation for summer low flows in humid catchments. *Hydrology and Earth System Sciences*, 20(2), 859–874. <https://doi.org/10.5194/hess-20-859-2016>
- Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12(3), 1087–1117. <https://doi.org/10.5194/gmd-12-1087-2019>
- Kouki, K., Räisänen, P., Luojus, K., Luomajarva, A., & Riihelä, A. (2021). Evaluation of northern hemisphere snow water equivalent in CMIP6 models with satellite-based SNOWCCI data during 1982–2014. *The Cryosphere Discussions*, 2021, 1–32. <https://doi.org/10.5194/tc-2021-195>
- Leisenring, M., & Moradkhani, H. (2011). Snow water equivalent prediction using Bayesian data assimilation methods. *Stochastic Environmental Research and Risk Assessment*, 25(2), 253–270. <https://doi.org/10.1007/s00477-010-0445-5>
- Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., & Arheimer, B. (2010). Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research*, 41(3–4), 295–319. <https://doi.org/10.2166/nh.2010.007>
- Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, 43(7). <https://doi.org/10.1029/2006WR005756>
- MacBean, N., Peylin, P., Chevallier, F., Scholze, M., & Schürmann, G. (2016). Consistent assimilation of multiple data streams in a carbon cycle data assimilation system. *Geoscientific Model Development*, 9(10), 3569–3588. <https://doi.org/10.5194/gmd-9-3569-2016>
- Macian-Sorribes, H., Pechlivanidis, I., Crochemore, L., & Pulido-Velazquez, M. (2020). Fuzzy postprocessing to advance the quality of continental seasonal hydrological forecasts for river basin management. *Journal of Hydrometeorology*, 21(10), 2375–2389. <https://doi.org/10.1175/JHM-D-19-0266.1>
- Magnusson, J., Gustafsson, D., Hüsler, F., & Jonas, T. (2014). Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods. *Water Resources Research*, 50(10), 7816–7835. <https://doi.org/10.1002/2014WR015302>
- Malnes, E., Buane, A., Nagler, T., Bippus, G., Gustafsson, D., Schiller, C., et al. (2015). User requirements for the snow and land ice services – Cryoland. *The Cryosphere*, 9(3), 1191–1202. <https://doi.org/10.5194/tc-9-1191-2015>
- Mishra, A. K., & Coulibaly, P. (2009). Developments in hydrometric network design: A review. *Reviews of Geophysics*, 47(2). <https://doi.org/10.1029/2007RG000243>
- Moradkhani, H., Nearing, G. S., Abbaszadeh, P., & Pathiraja, S. (2019). Fundamentals of data assimilation and theoretical advances. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 676–695). Springer.
- Mu, Q., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, 115(8), 1781–1800. <https://doi.org/10.1016/j.rse.2011.02.019>
- Musuza, J. L., Gustafsson, D., Pimentel, R., Crochemore, L., & Pechlivanidis, I. (2020). Impact of satellite and in situ data assimilation on hydrological predictions. *Remote Sensing*, 12(5), 811. <https://doi.org/10.3390/rs12050811>
- Nijzink, R. C., Almeida, S., Pechlivanidis, I. G., Capell, R., Gustafsson, D., Arheimer, B., et al. (2018). Constraining conceptual hydrological models with multiple information sources. *Water Resources Research*, 54(10), 8332–8362. <https://doi.org/10.1029/2017WR021895>
- Oudin, L., Michel, C., & Anctil, F. (2005). Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 1: Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? *Journal of Hydrology*, 303(1), 275–289. <https://doi.org/10.1016/j.jhydrol.2004.08.025>
- Pappenberger, F., Ramos, M., Cloke, H., Wetterhall, F., Alfieri, L., Bogner, K., et al. (2015). How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. *Journal of Hydrology*, 522, 697–713. <https://doi.org/10.1016/j.jhydrol.2015.01.024>
- Parajka, J., & Blöschl, G. (2008). Spatio-temporal combination of MODIS images – Potential for snow cover mapping. *Water Resources Research*, 44(3). <https://doi.org/10.1029/2007WR006204>
- Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., & Geenens, G. (2018). Data-driven model uncertainty estimation in hydrologic data assimilation. *Water Resources Research*, 54(2), 1252–1280. <https://doi.org/10.1002/2018WR022627>
- Pechlivanidis, I. G., Crochemore, L., Rosberg, J., & Bosshard, T. (2020). What are the key drivers controlling the quality of seasonal streamflow forecasts? *Water Resources Research*, 56(6), e2019WR026987. <https://doi.org/10.1029/2019WR026987>
- Poletti, M. L., Silvestro, F., Davolio, S., Pignone, F., & Reborja, N. (2019). Using nowcasting technique and data assimilation in a meteorological model to improve very short range hydrological forecasts. *Hydrology and Earth System Sciences*, 23(9), 3823–3841. <https://doi.org/10.5194/hess-23-3823-2019>
- Ramos, M. H., Bartholmes, J., & del Pozo, J. T. (2007). Development of decision support products based on ensemble forecasts in the European flood alert system. *Atmospheric Science Letters*, 8(4), 113–119. <https://doi.org/10.1002/asl.161>

- Reichle, R. H., & Koster, R. D. (2003). Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *Journal of Hydrometeorology*, 4(6), 1229–1242. [https://doi.org/10.1175/1525-7541\(2003\)004<1229:ATIOHE>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1229:ATIOHE>2.0.CO;2)
- Roundy, J. K., Duan, Q., & Schaake, J. C. (2019). Hydrological predictability, scales, and uncertainty issues. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 4–29). Springer.
- Sahai, A. K., Kaur, M., Joseph, S., Dey, A., Phani, R., Mandal, R., & Chattopadhyay, R. (2021). Multi-model multi-physics ensemble: A futuristic way to extended range prediction system. *Frontiers in Climate*, 3. <https://doi.org/10.3389/fclim.2021.655919>
- Schepen, A., Wang, Q. J., & Robertson, D. E. (2019). Verification metrics for hydrological ensemble forecasts. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 893–922). Springer.
- 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>
- SMHI. (2022). SMHI Vattenwebb. Retrieved from <https://vattenwebb.smhi.se/>
- Tuteja, N. K., Shin, D., Laugesen, R., Khan, U., Shao, Q., Wang, E., et al. (2011). *Experimental evaluation of the dynamic seasonal streamflow forecasting approach*. (Tech. Rep.). Bureau of Meteorology, Melbourne.
- Villani, G., Tomei, F., Pavan, V., Pirola, A., Spisni, A., & Marletto, V. (2021). The ICOLT climate service: Seasonal predictions of irrigation for Emilia-Romagna, Italy. *Meteorological Applications*, 28(4), e2007. <https://doi.org/10.1002/met.2007>
- Vogel, E., Lerat, J., Pipunic, R., Frost, A., Donnelly, C., Griffiths, M., et al. (2021). Seasonal ensemble forecasts for soil moisture, evapotranspiration and runoff across Australia. *Journal of Hydrology*, 601, 126620. <https://doi.org/10.1016/j.jhydrol.2021.126620>
- Wanders, N., Thober, S., Kumar, R., Pan, M., Sheffield, J., Samaniego, L., & Wood, E. F. (2019). Development and evaluation of a pan-European multimodel seasonal hydrological forecasting system. *Journal of Hydrometeorology*, 20(1), 99–115. <https://doi.org/10.1175/jhm-d-18-0040.1>
- Wang, W., Liu, J., Li, C., Liu, Y., & Yu, F. (2021). Data assimilation for rainfall-runoff prediction based on coupled atmospheric-hydrologic systems with variable complexity. *Remote Sensing*, 13(4), 595. <https://doi.org/10.3390/rs13040595>
- White, C. J., Domeisen, V. D. I., Acharya, N., Adefisan, E. A., Anderson, M. L., Aura, S., et al. (2021). Advances in the application and utility of subseasonal-to-seasonal predictions. *Bulletin America Meteorology Social*, 1–57(6), E1448–E1472. <https://doi.org/10.1175/BAMS-D-20-0224.1>
- Winsemius, H. C., Dutra, E., Engelbrecht, F. A., Archer Van Garderen, E., Wetterhall, F., Pappenberger, F., & Werner, M. G. F. (2014). The potential value of seasonal forecasts in a changing climate in southern Africa. *Hydrology and Earth System Sciences*, 18(4), 1525–1538. <https://doi.org/10.5194/hess-18-1525-2014>
- Wood, A. W., Hopson, T. M., Newman, A. J., Brekke, L. D., Arnold, J. R., & Clark, M. P. (2016). Quantifying streamflow forecast skill elasticity to initial condition and climate prediction skill. *Journal of Hydrometeorology*, 17(2), 651–668. <https://doi.org/10.1175/jhm-d-14-0213.1>
- Wood, A. W., & Lettenmaier, D. P. (2006). A test bed for new seasonal hydrologic forecasting approaches in the western United States. *Bulletin America Meteorology Social*, 87(12), 1699–1712. <https://doi.org/10.1175/BAMS-87-12-1699>
- Wood, A. W., & Lettenmaier, D. P. (2008). An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophysical Research Letters*, 35(14), L14401. <https://doi.org/10.1029/2008GL034648>
- Wörman, A., Uvo, C. B., Brandimarte, L., Busse, S., Crochemore, L., Lopez, M. G., et al. (2020). Virtual energy storage gain resulting from the spatio-temporal coordination of hydropower over Europe. *Applied Energy*, 272, 115249. <https://doi.org/10.1016/j.apenergy.2020.115249>
- Yang, W., Andréasson, J., Phil Graham, L., Olsson, J., Rosberg, J., & Wetterhall, F. (2010). Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrology Research*, 41(3–4), 211–229. <https://doi.org/10.2166/nh.2010.004>
- Yuan, H., Toth, Z., Peña, M., & Kalnay, E. (2019). Overview of weather and climate systems. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 35–45). Springer.
- Yuan, X., Ma, F., Wang, L., Zheng, Z., Ma, Z., Ye, A., & Peng, S. (2016). An experimental seasonal hydrological forecasting system over the Yellow River basin – Part I: Understanding the role of initial hydrological conditions. *Hydrology and Earth System Sciences*, 20(6), 2437–2451. <https://doi.org/10.5194/hess-20-2437-2016>
- Yuan, X., Wood, E. F., & Zhuguo, M. (2015). A review on climate-model-based seasonal hydrologic forecasting: Physical understanding and system development. *WIREs Water*, 2(5), 523–536. <https://doi.org/10.1002/wat2.1088>
- Yucel, I., Onen, A., Yilmaz, K. K., & Gochis, D. J. (2015). Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall. *Journal of Hydrology*, 523, 49–66. <https://doi.org/10.1016/j.jhydrol.2015.01.042>