

PAPER • OPEN ACCESS

Evaluation of high-resolution river discharge over Europe based on the ICON-CLM climate ensemble simulations

To cite this article: Stefan Hagemann *et al* 2026 *Environ. Res.: Climate* **5** 045011

View the [article online](#) for updates and enhancements.

You may also like

- [Triple Jeopardy: trends in the co-occurrence of extreme heat, drought, and PM2.5 across the conterminous United States](#)
Anuska Narayanan, Mostafa Rezaali, David Keellings et al.
- [Hemispheric stationary-eddy activity of the winter North Atlantic oscillation](#)
M Brotons, J García-Serrano and R J Haarsma
- [Sea-level rise and coastal hazards expose critical infrastructure vulnerabilities across O'ahu](#)
Anna B Mikkelsen, Charles H Fletcher, Richelle U Moskvichev et al.

ENVIRONMENTAL RESEARCH CLIMATE



PAPER

OPEN ACCESS

RECEIVED
2 March 2026

REVISED
10 August 2026

ACCEPTED FOR PUBLICATION
18 September 2026

PUBLISHED
12 October 2026

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the
title of the work, journal
citation and DOI.



Evaluation of high-resolution river discharge over Europe based on the ICON-CLM climate ensemble simulations

Stefan Hagemann^{1,*} , Beate Geyer¹ , Sergei Petrov¹ , Evgenii Churiulin² , Michael Jähn³ ,
Muhammed Muhshif Karadan⁴ , Vera Maurer⁵ , Christopher Purr⁵ , Shweta Singh⁵
and Heimo Truhetz⁶

¹ Institute of Coastal Systems—Analysis and Modelling, Helmholtz-Zentrum Hereon, Max-Planck-Str. 1, 21502 Geesthacht, Germany

² Institute of Meteorology and Climate Research Troposphere Research (IMKTRO), Karlsruhe Institute of Technology, Karlsruhe, Germany

³ Center for Climate Systems Modeling (C2SM), ETH Zurich, Zurich, Switzerland

⁴ Fachbereich für Atmosphärische Prozesse, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus, Germany

⁵ Deutscher Wetterdienst, Offenbach am Main, Germany

⁶ Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

* Author to whom any correspondence should be addressed.

E-mail: stefan.hagemann@hereon.de

Keywords: discharge trends, high-resolution, Europe, historical simulations, ICON-CLM

Supplementary material for this article is available [online](#)

Abstract

High-resolution river discharge provides relevant information for evaluating regional climate models and assessing the hydrological impact of climate change. In this study, we generate high-resolution European river discharges by forcing the hydrological discharge (HD) model with runoff fields from the Icosahedral non-hydrostatic climate limited-area mode (ICON-CLM) regional climate model. We then evaluate the quality of the simulated discharge obtained from a reanalysis-driven hindcast by comparing it with observations. Moreover, we assess how well the historical discharge trends are represented across six simulations in which ICON-CLM was driven by different global climate models (GCMs). A comparison of the results with observed river discharge across Europe shows that ICON-CLM-based simulations exhibit widespread dry biases, which are partly attributable to underestimation of precipitation in these regions, and partly to a too weak surface runoff component in the land-surface scheme of ICON-CLM. Using the HydroPy hydrology model as an alternative representation of land-surface hydrology improves the simulated discharge in many regions. However, biases remain in areas with regulated rivers, as the HD model does not take regulations into account. Analysis of discharge trends from 1960–2014 reveals robust, large-scale patterns in the reference hindcasts, such as an increase in mean and low-flow discharge in Northern Europe and a decrease in Southern Europe. While some historical simulations reproduce these patterns, others show opposing or inconsistent signals, highlighting the strong influence of the driving GCMs. Our results underscore the suitability of ICON-CLM atmospheric forcing for hydrological impact studies and suggest the need for improved model tuning with regard to precipitation, including undercatch corrections in reference observations. They also emphasize the importance of multi-model ensembles to reliably capture climate-induced hydrological changes.

1. Introduction

Currently, there is a joint effort between the climate limited-area modelling (CLM) Community (www.clm-community.eu) and the German Federal Ministry of Research, Technology and Space (BMFTR) project ‘Updating the data basis for adaptation to climate change in Germany’ (UDAG; Früh 2023, Singh *et al* 2026). The aim is to downscale an ensemble of selected climate change simulations from the coupled model intercomparison project Phase 6 (CMIP6; Eyring *et al* 2016). This effort will update the data basis for climate change impact studies and the planning of adaptation measures to climate change in Germany (Kahlenborn *et al* 2021). Thus, its contribution will form part of the basis for implementing the German adaptation strategy (Deutsche Anpassungsstrategie: DAS). This regional climate model (RCM) ensemble is conducted with the same model version and setup of the Icosahedral Non-hydrostatic (ICON) model used in CLM (ICON-CLM; Pham *et al* 2021, Geyer *et al* 2026). Although results from other RCMs will also be included in the DAS, this new ensemble will form a core part of the DAS. Therefore, our study focuses on this ensemble based only on one RCM. This differs from previous studies in which an ensemble of different RCMs was considered (e.g. Sieck *et al* 2026).

ICON-CLM belongs to the common class of regional climate models that represent atmosphere and land processes without considering lateral water flows at the land surface, i.e. usually designated as river runoff or discharge. However, discharge is an important component of global and regional water cycles. Changes in discharge can have a significant impact on the water resources of the respective catchment area (Haddeland *et al* 2013, Hagemann *et al* 2013). Increased discharge extremes can lead to floods and droughts, which pose serious threats to human life and welfare (Kundzewicz and Matczak 2015). Such an increased risk necessitates enhanced adaptation measures. As discharge closes the water cycle between land and ocean, changes in the associated freshwater inflow may affect the thermohaline circulation (Marzeion *et al* 2007, Hordoir *et al* 2008), ocean salinity (Väli *et al* 2013), stratification (Hordoir and Meier 2010), and the functioning of marine ecosystems (e.g. Daewel and Schrum 2017). Therefore, quantifying discharge and its future changes is crucial for climate change impact assessments in coastal regions and sea basins.

In order to fill the gap that discharge is not provided by ICON-CLM, we used a state-of-the-art river runoff model, the hydrological discharge (HD) model (Hagemann *et al* 2020), to generate discharges that are consistent with the ICON-CLM output. It utilizes the 6 h input fields of surface and subsurface runoff that were generated by TERRA (Schrodin and Heise 2001, Schulz *et al* 2016), the land surface scheme of ICON-CLM. This method produces simulated discharges that align with ICON-CLM simulations. Therefore, the discharge can be used as an additional measure for evaluating the model with respect to the terrestrial water cycle at the land surface. Note that this method is different to the approach commonly used in hydrology where full hydrology models are forced by atmospheric climate model fields (see, e.g. Hagemann *et al* 2020). Here, the hydrological model calculates its own land surface water balance and hydrological processes may potentially be better represented than in the climate model. However, in this approach, hydrological feedback to the atmosphere is not considered and the respective land surface water fluxes may be inconsistent with the results yielded by the respective climate model.

The main aim of the present study is to assess the suitability of the ICON-CLM output for its use in climate change impact studies on river runoff. We acknowledge that climate change impact studies typically focus on future projections, where the associated signals are more pronounced than those observed during the historical period. However, we wanted to investigate how the new ensemble behaves in the latter case. Therefore, two objectives were set for this study. First, we want to evaluate the simulated river runoff that is based on the ICON-CLM ensemble. This evaluation will use various metrics in relation to observed discharges. Second, we want to investigate how the impact of the recent climate change on river runoff is represented by analysing consistent trend patterns in several discharge statistics. Please note that it is difficult to quantify climate change-related trends from observed river discharges on a large scale, since most large rivers—especially those with a long-term streamflow record—have been impacted by human influences, such as dam construction or land use (Hartmann *et al* 2013). Therefore, rather than comparing the simulated trends with observations, we compare them with simulated discharges that were generated analogously to Hagemann and Stacke (2023). They applied the HD model to derive consistent, high-resolution discharges from two commonly used, observation-based datasets over Europe.

The manuscript is structured as follows: section 2 provides an overview of the models, data, experimental setup and metrics employed. Section 3 presents an evaluation of the simulated discharges, accompanied by a more detailed consideration of hydrological variables for selected rivers. Section 4

Table 1. ICON-CLM experiments forced by 6 h GCM data (3 h for ERA5).

Experiment	ICON CLM forcing	Period	Forcing reference
ICON-CLM EVAL	ERA5	1950–2024	Hersbach <i>et al</i> (2020)
ICON-CLM ECE	EC-Earth3-Veg	1950–2014	Döscher <i>et al</i> (2022)
ICON-CLM CNRM	CNRM-ESM2-1	1950–2014	Séférian <i>et al</i> (2019)
ICON-CLM MPI	MPI-ESM-1-2-HR	1950–2014	Mauritsen <i>et al</i> (2019)
ICON-CLM MIROC	MIROC6	1950–2014	Tatebe <i>et al</i> (2019)
ICON-CLM CMCC	CMCC-CM2-SR5	1950–2014	Cherchi <i>et al</i> (2019)
ICON-CLM NOR	NorESM2-MM	1950–2014	Seland <i>et al</i> (2020)

provides an analysis of trends and a discussion of the results. Section 5 ends with a summary and conclusions.

2. Data and methods

First, we describe the experimental setup used. Then, we provide information on the evaluation metrics and trend measures that are analysed.

2.1. ICON-CLM simulations

ICON is a flexible and scalable high-performance modelling framework for weather, climate and environmental predictions and projections (Zängl *et al* 2015, Müller *et al* 2025). In the present study, ICON was used in CLM (ICON-CLM; Pham *et al* 2021) to conduct regional climate simulations over the European domain of the coordinated regional downscaling experiment (EURO-CORDEX; Jacob *et al* 2013) of the World Climate Research Programme. The selected EURO-CORDEX domain has a spatial resolution of 0.11° in rotated coordinates (approx. 12 km) and covers Europe and its surrounding regions including the North Atlantic and the Mediterranean Basin as well as parts of North Africa and western Asia. The ICON-CLM model setup was first based on the ICON setup used for operational numerical weather prediction at the Deutscher Wetterdienst, and then determined from optimization exercises through model extensions and a novel parameter tuning strategy (Geyer *et al* 2026). On the one hand, ICON-CLM was used to downscale reanalysis data (see section 2.3) to evaluate the model performance. On the other hand, it was used to downscale various global simulations of the CMIP Phase 6 (CMIP6; Eyring *et al* 2016). Table 1 provides the experiment names as well as information on the reanalysis and the respective CMIP6 models used for downscaling. Please note that the first 10 years of these simulations are considered as a spin-up period and are not taken into account thereafter.

2.2. The HD model

The HD model (Hagemann *et al* 2020) is a river-routing model that is well-established and implemented in a range of global and regional model systems. The HD model is forced by daily or sub-daily time series of surface and subsurface runoff that are provided by the driving climate or hydrology model. It calculates discharge for every land grid box of the considered domain and at the final inflow points into the ocean. In the present study, we applied the HD model v5.2.4 (Hagemann *et al* 2025) over its European domain (land areas between 11° W to 69° E and 27° N to 72° N) at $1/12^\circ$ spatial resolution. Figure 1 shows the domain along with a few rivers particularly considered in this study. We note that conservative remapping was used to interpolate data onto the HD model domain for either model forcing or evaluation purposes.

2.3. Observational data

For the evaluation of simulated discharge, we used data on daily discharges from downstream gauges on many European rivers with catchment areas of around 1000 km^2 or more. We retrieved the respective station data from the Global Runoff Data Centre and various agencies and institutions. The latter are listed in table 2 of Hagemann and Stacke (2023) and additionally French discharge data were obtained from the EU Copernicus Marine Service (<https://marine.copernicus.eu/>).

In order to conduct two additional hydrological hindcast simulations with a hydrology model (see section 2.4), we assessed two independent datasets. On the one hand, we obtained atmospheric forcing data from the ERA5 reanalysis (Hersbach *et al* 2020), which is the fifth generation of atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). It has a spatial resolution of about 0.28° and is available for 1940–2024. We also used ERA5 evapotranspiration (ET)

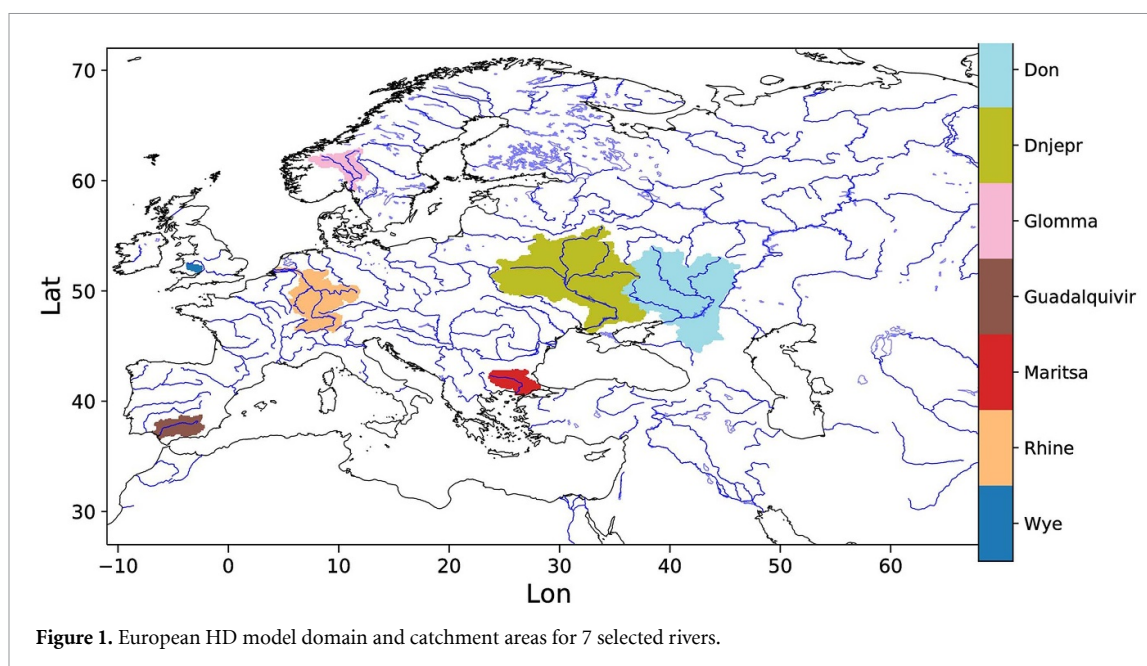


Table 2. HD model experiments forced by 6 h data of ICON-CLM.

Experiment	Type	HD forcing	Period
HD5-EVAL	Hindcast	ICON-CLM EVAL	1950–2024
HD5-ECE	Historical	ICON-CLM ECE	1950–2014
HD5-CNRM	Historical	ICON-CLM CNRM	1950–2014
HD5-MPI	Historical	ICON-CLM MPI	1950–2014
HD5-MIROC	Historical	ICON-CLM MIROC	1950–2014
HD5-CMCC	Historical	ICON-CLM CMCC	1950–2014
HD5-NOR	Historical	ICON-CLM NOR	1950–2014

and vertically integrated water vapour data to evaluate the results for specific rivers. On the other hand, we utilized a recent version of the E-OBS dataset (Cornes *et al* 2018), which comprises several daily gridded surface variables at 0.1° resolution over Europe covering the area 25°N – 71.5°N , 25°W – 45°E . These gridded data were derived from station data using an ensemble simulation procedure that is described in detail by Cornes *et al* (2018). In the present study, we used the mean of the ensemble members for precipitation and 2 m temperature of v. 31 (EOBS31), which covers the period 1950–2024. We note that the quality of the EOBS data is poor for the last five years. However, in the present study we only considered data until 2018.

For the evaluation of ICON-CLM precipitation, we used EOBS31 and the WFDE5 (Cucchi *et al* 2020) dataset that includes bias corrected time series of surface meteorological variables from ERA5. Cucchi *et al* (2020) generated these time series by applying the WATCH Forcing Data methodology (Weedon *et al* 2011) to the respective ERA5 data. The spatial resolution of WFDE5 is 0.5° and it covers the period 1979–2019.

In addition, we considered the simulated monthly climatology of ET in selected river catchments based on ERA5 and the Global Land Evaporation Amsterdam Model V. 4.2a (GLEAM; Miralles *et al* 2011, 2025) because Cai *et al* (2024) identified considerable uncertainties in the ET estimates of reanalysis products. GLEAM provides evaporative flux estimates based on satellite observations and reanalysis data with a spatial resolution of 0.1° for the period 1980–2024.

2.4. Experimental setup and additional hydrological simulations

The HD model was applied to simulate discharge from the ICON-CLM hindcast (HD5-EVAL) and the six historical simulations. In each of the seven experiments, 6 h time series of surface and subsurface runoff were used to force the HD model. Table 2 provides the respective experiment acronyms and summarizes their characteristics.

Table 3. HD model experiments forced by daily data of hydropy.

Experiment	Atmospheric forcing	HydroPy input	HydroPy resolution	Period
HD5-ERA5	ERA5	Precipitation, 2 m temperature, downwelling shortwave and longwave radiation, 2 m specific humidity, surface pressure, 10 m wind.	$\sim 0.28^\circ$	1940–2024
HD5-EOBS31	EOBS31	Precipitation, 2 m temperature	0.1°	1950–2024
HD5-Hpy-EVAL	ICON-CLM EVAL	Precipitation, 2 m temperature, downwelling shortwave and longwave radiation, 2 m specific humidity, surface pressure, 10 m wind.	$\sim 0.08^\circ$ ($1/12^\circ$)	1950–2024

In addition, we conducted three hydrological hindcast simulations that are summarized in table 3. Here, we used the hydrological model HydroPy (Stacke and Hagemann 2021), which is a state-of-the-art global hydrology model and the successor of the Max-Planck Institute for Meteorology's Hydrology Model (MPI-HM; Stacke and Hagemann 2012). For the three hindcast simulations, HydroPy was forced by daily atmospheric forcing data (including, e.g. precipitation and 2 m temperature), and then daily time series of surface and sub-surface runoff from HydroPy were used to simulate daily discharges with the HD model. In HD5-HPy-EVAL, daily atmospheric data of the ICON-CLM hindcast were interpolated to the HD European domain (figure 1) and then used to force HydroPy. Here, a restart state of 1950 was taken from a century-long simulation (Hagemann *et al* 2024) and used at the start of the simulation in 1950.

To support the evaluation of results, two hindcast simulations were set up analogously to those analysed by Hagemann and Stacke (2023). First, we forced HydroPy globally with daily atmospheric data of ERA5 at the ERA5 resolution. For Europe and the period 1979–2018, the resulting discharge simulation should match the simulation considered by Hagemann and Stacke (2023). In this period, smaller differences between the two simulations may occur in catchments that cross the boundary of the European HD model domain. In addition, differences may occur due to the different water storage contents in HydroPy in the beginning of 1979. While Hagemann and Stacke (2023) conducted a 50 years spin-up simulation by repeatedly using forcing of the year 1979, a transient simulation starting in 1940 was conducted in the present study. The latter was preceded by a 52 years spin-up simulation that consecutively repeated the year 1940. By comparing these two hindcasts, we noticed only minor improvements for Eastern European rivers in HD5-ERA5 (supplementary figure S1). Therefore, the evaluation of the respective simulation conducted in Hagemann and Stacke (2023) is also representative for HD5-ERA5.

Second, we forced HydroPy with daily precipitation and temperature data of EOBS31 over the EOBS 0.1° domain. Here, we conducted a 50 years spin-up simulation by repeatedly using forcing of the year 1950 before the actual simulation was started. Noticeable differences to the EOBS22-based simulation considered by Hagemann and Stacke (2023) occur over Spain and Turkey. Over Turkey, this can be considered as an improvement as the data gaps of EOBS v. 22 over Turkey, which were identified by Hagemann and Stacke (2023), are not present in EOBS31. In Spain, new gauge data were added, particularly from regions that were previously under-represented (van den Besselaar 2025). This changed the interpolation and therefore the gridded precipitation fields. As Spain is known to have strong spatial precipitation gradients, adding or removing stations often causes significant differences in the spatial distribution of precipitation between versions (Vicente-Serrano *et al* 2025). We note that, for the majority of Europe, the evaluation of EOBS22-based discharge in Hagemann and Stacke (2023) should also be representative of HD5-EOBS31. However, to demonstrate the aforementioned differences, we have included the supplementary figure S1 that can be directly compared with the figures 6 and 7 in Hagemann and Stacke (2023).

2.5. Evaluation metrics

The discharge evaluation of the HD5 simulations was conducted for those grid boxes that correspond to the station locations within the river network of the respective discharge observations. All evaluation metrics were calculated using simulated and observed time series of daily discharge for the respective periods (1979–2018 or 1960–2014), thereby covering only those days where observed data are available. We used the mean bias, the Pearson correlation coefficient, the Kling–Gupta-efficiency (KGE; Gupta *et al* 2009, Kling *et al* 2012) and the variability ratio defined by the ratio of the simulated and observed coefficients of variation. The KGE is a quality metric that combines bias, correlation, and coefficient of variation. A $KGE > -0.41$ means that a simulated discharge time series better represents the observations

Table 4. Discharge related trend statistics.

Variable	Description
Annual maximum, $Q_{\max}$	Annual maximum of daily discharge [%/decade]
Annual mean, Q_{mean}	Annual mean of daily discharge [%/decade]
Annual minimum, $Q_{\min}$	Annual minimum of daily discharge [%/decade]
No. of low flow days	Annual number of days with flows below the environmental water requirement [d/decade]

than the observed long-term mean discharge (Knoben *et al* 2019). These metrics are described, e.g. in Hagemann *et al* (2020).

In addition, we calculated the monthly mean climatology from the monthly mean time series of several variables in the ICON-CLM simulations and the respective reference data. For all the rivers considered, we averaged these variables spatially over the upstream catchment area of the discharge measurement gauge before calculating the climatology. For precipitation, biases were calculated from the aggregated catchment means.

2.6. Trend statistics

For each river, various trend statistics (see table 4) were calculated for the period 1960–2014 at the respective most downstream station with available observations. As in Hagemann and Stacke (2023), the number of low-flow days is the amount of days each year where the flow is below the environmental water requirement. The latter is defined as 30% of the long-term annual mean discharge (Hagemann *et al* 2013) of the respective catchment. This definition is based on results of Smakhtin *et al* (2004). In addition, we calculated annual and seasonal mean trends for several variables of the ICON-CLM simulations. Using monthly mean time series of these variables, we calculated trends for each ICON-CLM simulation and the respective reference data. For all the rivers considered, we averaged these variables spatially over the upstream catchment of the respective discharge measurement gauge. For the calculation of trends, a linear regression formulation was used. Note that even if the true trend is non-linear, such as for temperature, the linear regression slope provides a best-fit estimate of the average rate of change over the entire period (Shea 2014). For the CMIP6 downscaling, we considered trends for the specific ensemble member simulations as well as ensemble mean trends. The latter were calculated by averaging over the trend distributions of the different ensemble members.

In our analysis of trends, we follow the approach of Hagemann and Stacke (2023) by considering and comparing robust spatial trend patterns, i.e. spatial clusters of rivers with trends in the same direction, instead of calculating the significance of trends per river. The large interannual or decadal variability in discharge among the various ensemble members leads to high noise levels, making it difficult to separate weaker signals (i.e. trends) from the noise when considering only time series for specific rivers. In such cases, a significance test may incorrectly classify some of these trends as insignificant, even if they form part of a large-scale change pattern. Therefore, examining robust spatial patterns is analogous to reducing the noise level obtained from discharge time series at specific rivers.

3. Evaluation of discharge quality metrics

In the following, various metrics (see section 2.5) were calculated at discharge measurement station locations or from spatial averages integrated over their respective upstream catchments. However, for the graphical representation, these measures were allocated to the respective total catchment areas. White areas on the 2D maps indicate ocean areas and catchments for which either:

- (1) insufficient river gauge data was available to the authors
- (2) they were classified as too small (with an area of less than about 1500 km²)
- (3) they are not fully covered by the respective atmospheric forcing data (i.e. ICON-CLM or EOBS31).

Please note that, when evaluating the hindcast simulations (see section 3.2), we considered the period from 1979 to 2018. This is because WDFE5 precipitation data are only available from 1979 onwards. However, for the evaluation of historical discharges (section 3.3), we used the period 1960–2014 to be consistent with the period used to evaluate trends in section 4. For the detailed analysis of the results for the selected rivers in section 3.3, we considered the overlapping part of these periods (i.e. 1979–2014).

3.1. Evaluation of the hindcast

To evaluate the hindcast simulation HD5-EVAL against observations we considered several metrics shown in figure 2. HD5-EVAL shows a general dry bias that is present for most of the catchments and particularly pronounced over Western Europe and Turkey. Correlations are mostly in a mediocre range between 0.4 and 0.7, and the variability ratio is largely underestimated in Western Europe but tends to be overestimated in Northern Europe. The latter is likely related to the fact that many northern European rivers are regulated which is usually reducing the variability of discharge (e.g. Hagemann *et al* 2024). Consequently, higher KGE values (>0.4 , bluish colours) are mainly obtained for catchments in Northeast and Central Europe. We also note positive discharge biases for the Don and Dnepr rivers, which are accompanied with low metric values. This is due to the fact that both rivers are heavily regulated and affected by human water withdrawals, a factor not considered by the HD model. Withdrawals from the Don may range from around 20% (Dzhamalov *et al* 2013) to 50% or more (Khublaryan 2009), whereas withdrawals from the Dnepr comprise around 20% (Vasylenko and Koshkina 2020).

In order to find out how much the land surface scheme of ICON-CLM, TERRA, is responsible for this behaviour of the simulated discharge, we compared the simulated discharge of HD5-EVAL to the discharge obtained by using the HydroPy model (figure 2, right column). In HD5-Hpy-EVAL, the bias is reduced over Central Europe and France, but comparable to HD5-EVAL over most other regions, with a somewhat increased dry bias over Northern Europe and a strongly increased dry bias over the Iberian Peninsula (figure 2(a)). However, the correlation improves for the majority of catchments, pointing towards a better timing of surface runoff and a better separation of the total runoff into its slow (subsurface runoff) and fast flow (surface runoff) components (figure 2(b)). The latter is supported by a considerably improved variability ratio over Western Europe (figure 2(c)). The improved variability ratio, in turn, is caused by the generation of substantially more surface runoff and less subsurface runoff in HydroPy than in TERRA, especially over Western, Central and Southern Europe (figure 3). Consequently, the KGE also gets generally improved with HD5-HPy-EVAL compared to HD5-EVAL, especially in Western, Central and Eastern Europe (figure 2(d)). Over Northern Europe, KGE gets worse for some rivers, which is likely related to the fact that the reduced surface runoff and variability in TERRA are similar to the impact of human regulation on the discharge in these catchments. Here, human regulation often eliminates peak discharges and keeps discharges above certain levels during low-flow periods, thereby reducing discharge variability (Hagemann *et al* 2024). The relatively high KGE values of HD5-HPy-EVAL suggest that the respective atmospheric forcing variables (see table 3) of the ICON-CLM simulation are rather well suited to force hydrological impact simulations.

Also, when using HydroPy to represent the land surface hydrology, we still yield the pattern of dry biases over Northern Europe, British Isles, Iberian Peninsula and Turkey. Hence, we compare the underlying ICON-CLM precipitation with EOBS31 and WFDE5 data in figure 4. Compared to EOBS31, EVAL shows rather low precipitation biases for most of the catchments and tends to have larger precipitation than EOBS31 over Northern and Central Europe. However, EOBS31 is based on station observations that are not corrected for gauge undercatch of the measurement devices (Corney *et al* 2018). The precipitation measurements may have an error of up to 10%–50%, which depends on wind speed, air temperature and state of aggregation (see, e.g. Rudolf and Rubel 2005). The undercatch can be especially large in mountainous areas and areas with snow fall. In WFDE5, such an undercatch correction (Weedon *et al* 2011) was (Becker *et al* 2013) incorporated, which leads to an increase of precipitation. Consequently, the comparison of EVAL precipitation with WFDE5 indicates a dry bias over most catchments in Europe except for the most northern rivers. This dry bias in precipitation contributes to the widespread dry bias in HD5-EVAL and HD5-HPy-EVAL discharge, which is strongest over the Iberian Peninsula, western UK and Turkey. Here, it can be noted that the parameter optimization of ICON-CLM was using EOBS data as a reference precipitation dataset (Petrov *et al* 2025, Geyer *et al* 2026). Based on our results, we recommend that precipitation products corrected for undercatch are also taken into account in future parameter optimization exercises.

3.2. Evaluation of historical simulations

Figure 5 shows the discharge biases for the historical simulations for the period 1960–2014. The biases for HD5-ECE are very similar to those for HD5-EVAL. This similarity appears to be due to the fact that the lateral boundary conditions for the respective ICON-CLM simulations exhibit comparable physical and dynamic behaviour. This is because the atmospheric component of the ECMWF-IFS system utilized to produce the ERA5 data (Hersbach *et al* 2020) is similar to that of EC-Earth3-Veg (Döscher *et al* 2022). This assumption is supported by the similarity of the precipitation biases, and therefore their distributions, in both simulations (supplementary figure S2).

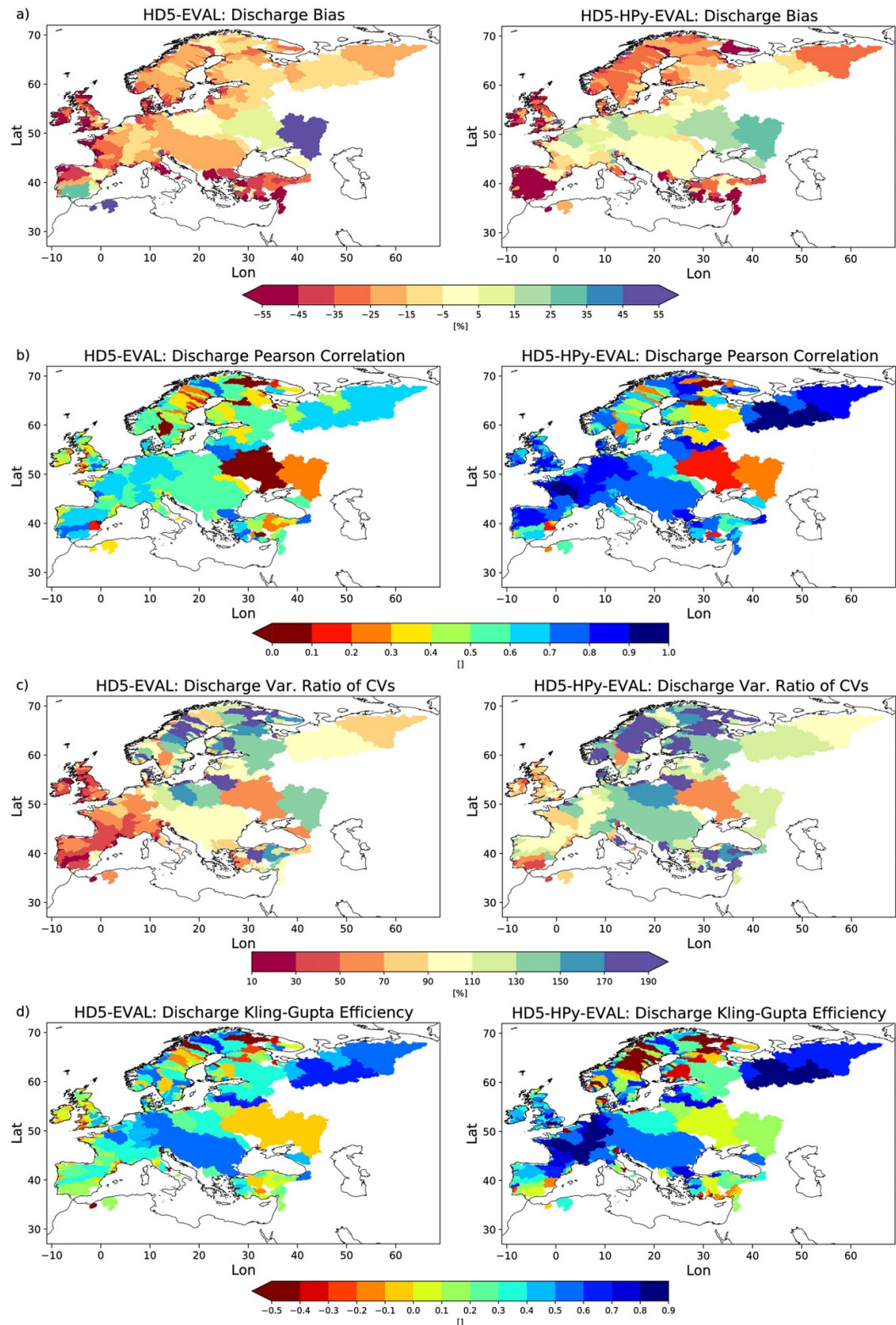


Figure 2. Evaluation of discharge against observations from HD5-EVAL (left) and HD5-HPy-EVAL (right) for 1979–2018: (a) Bias [%], (b) Pearson correlation, (c) Variability ratio, and (d) KGE.

The other five simulations possess only a few bias features that are common among all historical simulations. These comprise dry biases over the British Isles except for southeast England, and southern Scandinavia as well as the wet bias for the rivers Dnepr and Don (see section 3.1). HD5-EVAL and HD5-ECE, HD5-CMCC and HD5-NOR also feature a dry bias over the Iberian Peninsula and southern France, while the other three simulations share an opposing wet bias over the Iberian Peninsula. The dry bias over Turkey is present in HD5-MPI, HD5-CMCC and HD5-NOR but replaced by a prominent

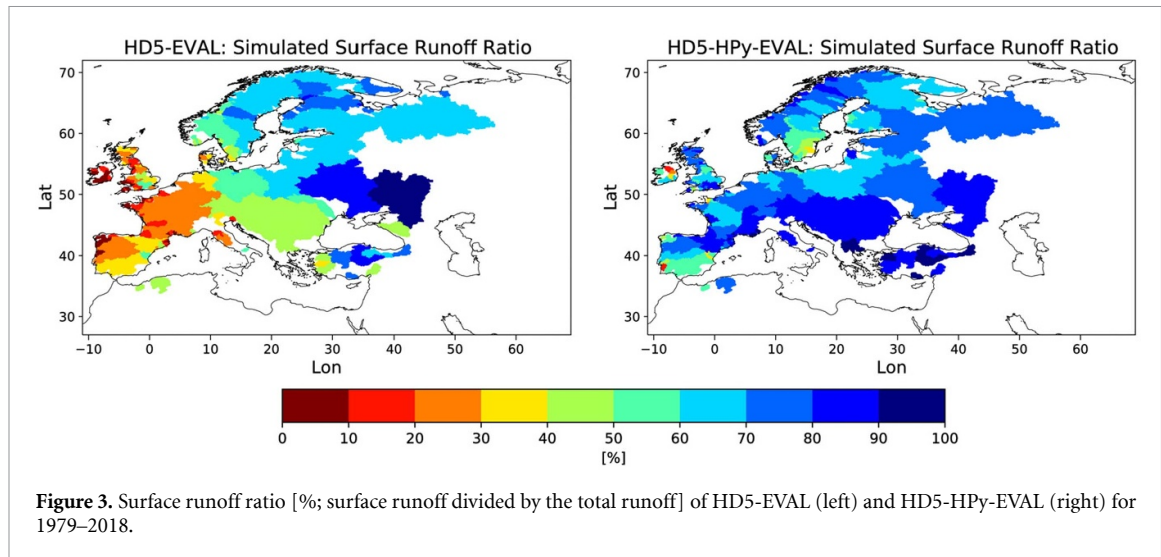


Figure 3. Surface runoff ratio [%; surface runoff divided by the total runoff] of HD5-EVAL (left) and HD5-HPy-EVAL (right) for 1979–2018.

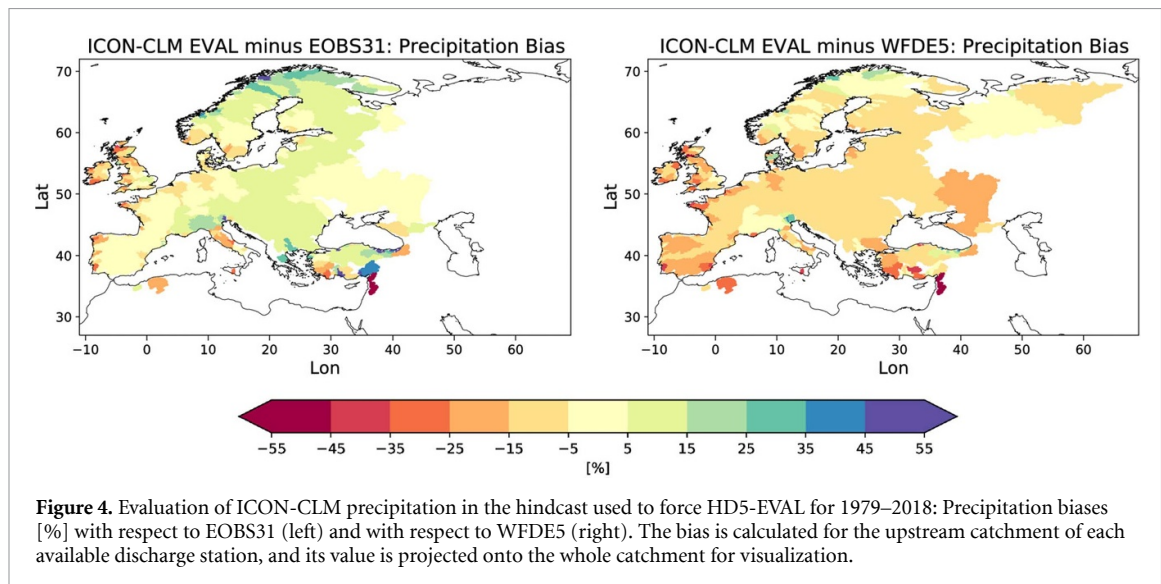


Figure 4. Evaluation of ICON-CLM precipitation in the hindcast used to force HD5-EVAL for 1979–2018: Precipitation biases [%] with respect to EOBS31 (left) and with respect to WFDE5 (right). The bias is calculated for the upstream catchment of each available discharge station, and its value is projected onto the whole catchment for visualization.

wet bias in HD5-CNRM and HD5-MIROC. The latter two also share smaller wet biases over France and Central Europe, while MPI-ESM shows larger wet biases extending from France to Germany and the Southern Baltic Sea catchments.

All historical simulations show variability ratio pattern that are generally similar to the one of HD5-EVAL for the same period (not shown), whereas the latter matches well with the period 1979–2018 shown in figure 2(c). Here, only HD5-CNRM sticks out with having much lower variability in Eastern Europe and Turkey than the observations.

The different behaviour of the historical simulations can be attributed to the characteristics of the global climate model (GCM) forcing used by ICON-CLM. A major driver of the discharge bias pattern is the associated ICON-CLM precipitation. Considering the deviations of the simulated precipitation from the WFDE5 data for the period 1979–2014 (figure S2) shows that the distribution of positive and negative biases resembles the discharge bias pattern over southern and northern Europe. However, noticeable differences occur over Central Europe and the British Isles.

3.3. Analysis of results for selected rivers

In order to take a closer look at some of the biases that strongly deviate from HD5-EVAL, we selected several rivers to be representative for the respective issues (see table 5 and figure 1). To investigate the large discrepancies over southern Europe with hot spots over the Iberian Peninsula and Turkey, the rivers Guadalquivir and Maritsa, respectively, are considered in more detail. Although both catchments are characterized by a pronounced dry season in the summer and a rainy season peaking in the end of

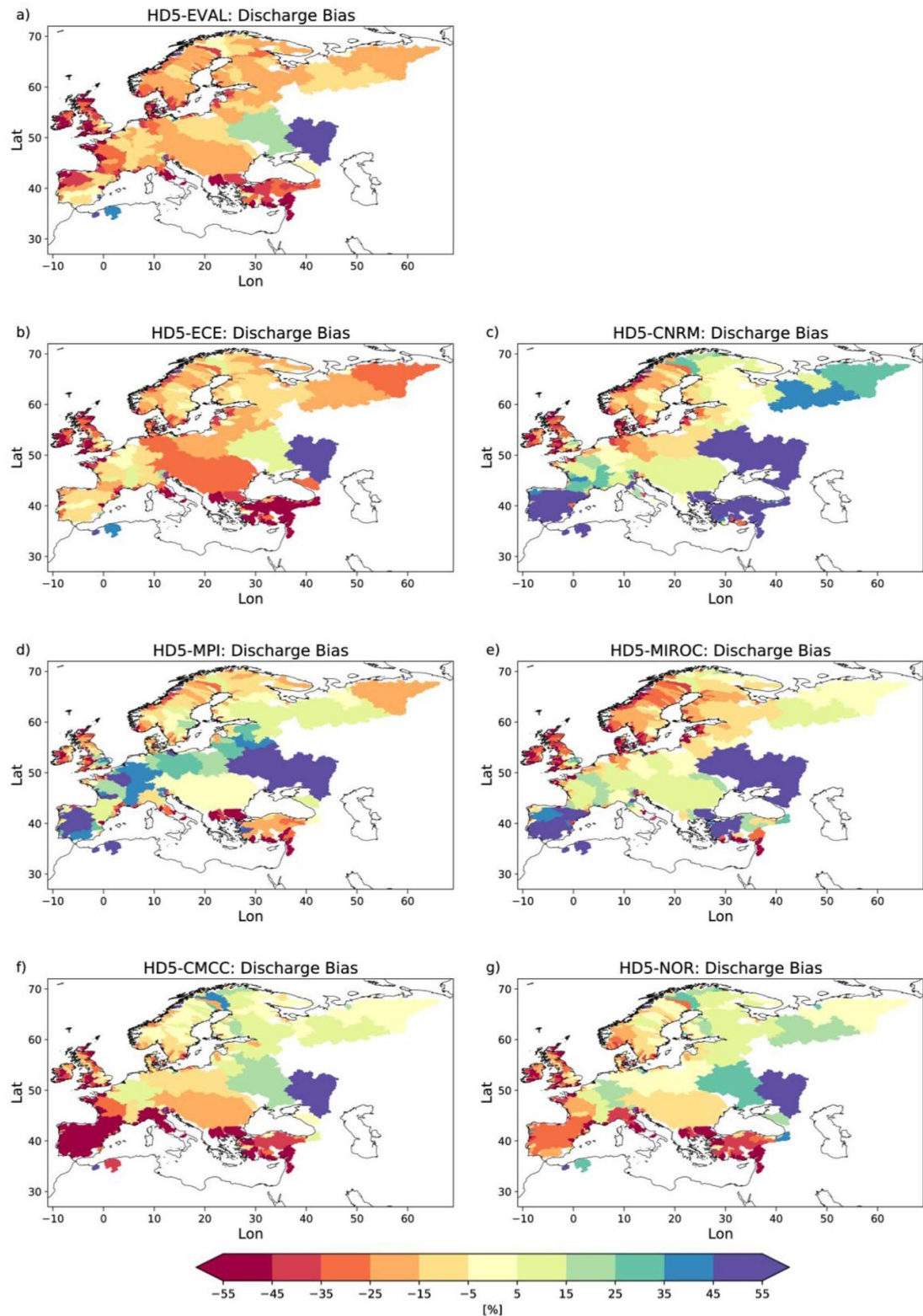


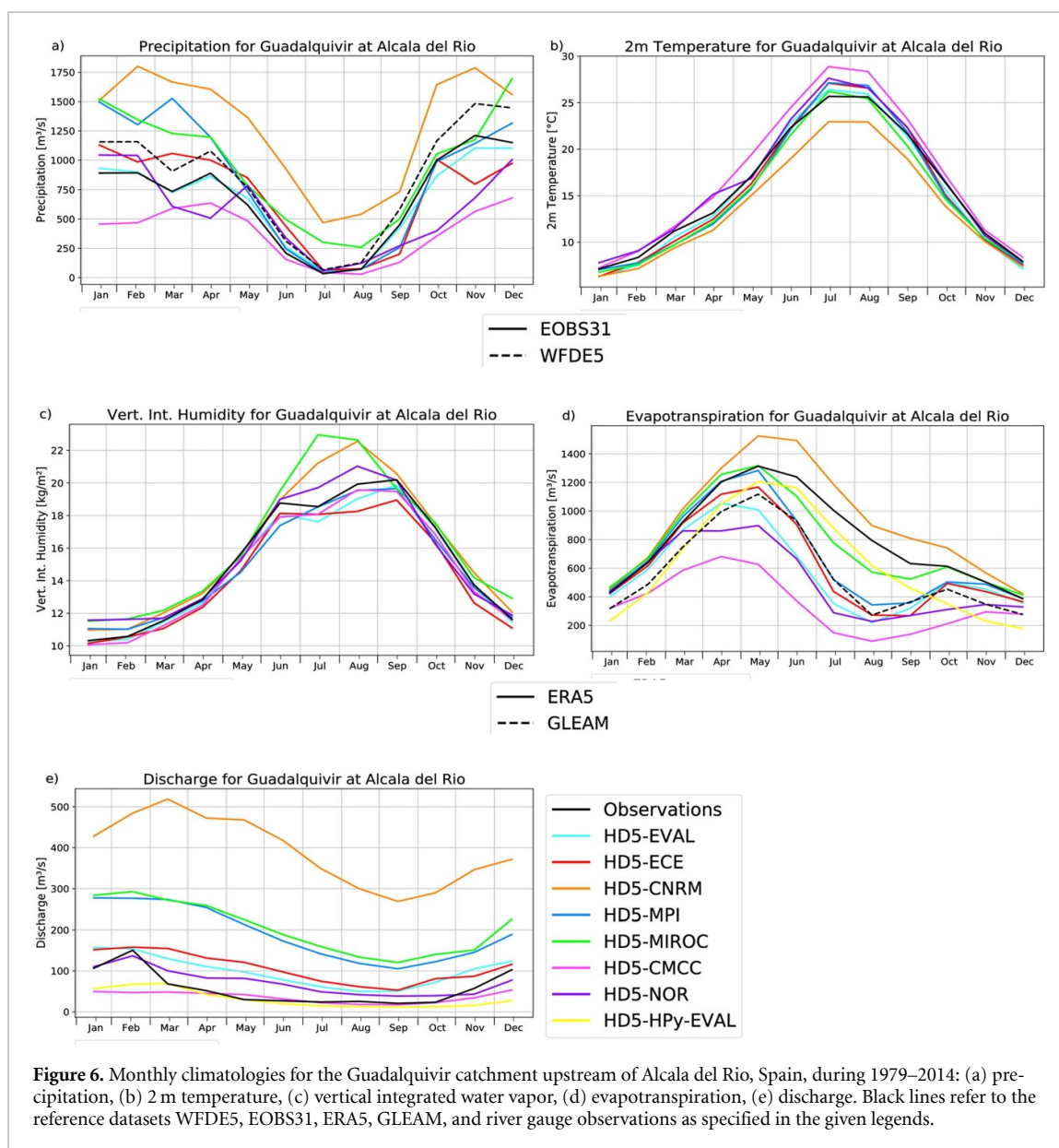
Figure 5. Mean discharge bias [%] during 1960–2014 for (a) HD5-EVAL, (b) HD5-ECE, (c) HD5-CNRM, (d) HD5-MPI, (e) HD5-MIROC, (f) HD5-CMCC, and (g) HD5-NOR.

the year, the intra-annual variability is much larger for the Guadalquivir with precipitation in the driest month amounting to less than 10% of the precipitation in the wettest month (figures 6(a) and 7(a)).

The mean monthly discharge climatology of the Guadalquivir River (figure 6(e)) clearly shows the large wet biases for HD5-CNRM, HD5-MPI and HD5-MIROC that are representative of their widespread wet biases over the Iberian Peninsula. These wet biases are directly related to an overestimation of precipitation (figure 6(a)) that is largest for HD5-CNRM all year long. For HD5-MPI and

Table 5. Catchment characteristics for selected rivers considered in sections 3.3 and 4.2.

River	Catchment area [km ²]	Discharge station	Obs. upstream area [km ²]	HD upstream area [km ²]
Glomma	41 918	Langnes	40 540	40 182
Guadalquivir	56 978	Alcala del Rio	46 134	46 460
Maritsa	53 000	Ferre Kupuru	50 140	51 055
Rhine	165 300	Lobith	160 800	158 980
Wye	4136	Redbrook	4010	4112

**Figure 6.** Monthly climatologies for the Guadalquivir catchment upstream of Alcala del Rio, Spain, during 1979–2014: (a) precipitation, (b) 2 m temperature, (c) vertical integrated water vapor, (d) evapotranspiration, (e) discharge. Black lines refer to the reference datasets WFDE5, EOBS31, ERA5, GLEAM, and river gauge observations as specified in the given legends.

HD5-MIROC, this overestimation is mainly confined to winter and spring so that the resulting positive discharge biases are lower than for HD5-CNRM. It can be noted that all simulations except for the very wet HD5-CNRM simulate less ET than ERA5 (figure 6(d)). This might partially contribute to the positive discharge bias. However, although HD5-ECE has relatively small biases in discharge and precipitation, it yields substantially less ET than ERA5 ET. This suggests deficits in the ERA5 ET, which depends more on its land surface model than on data assimilation. It should be noted that the ET of HD5-ECE simulated by ICON-CLM is relatively close to the GLEAM data, indicating that both datasets provide adequate ET estimates. In contrast to the other historical simulations, HD5-NOR and especially

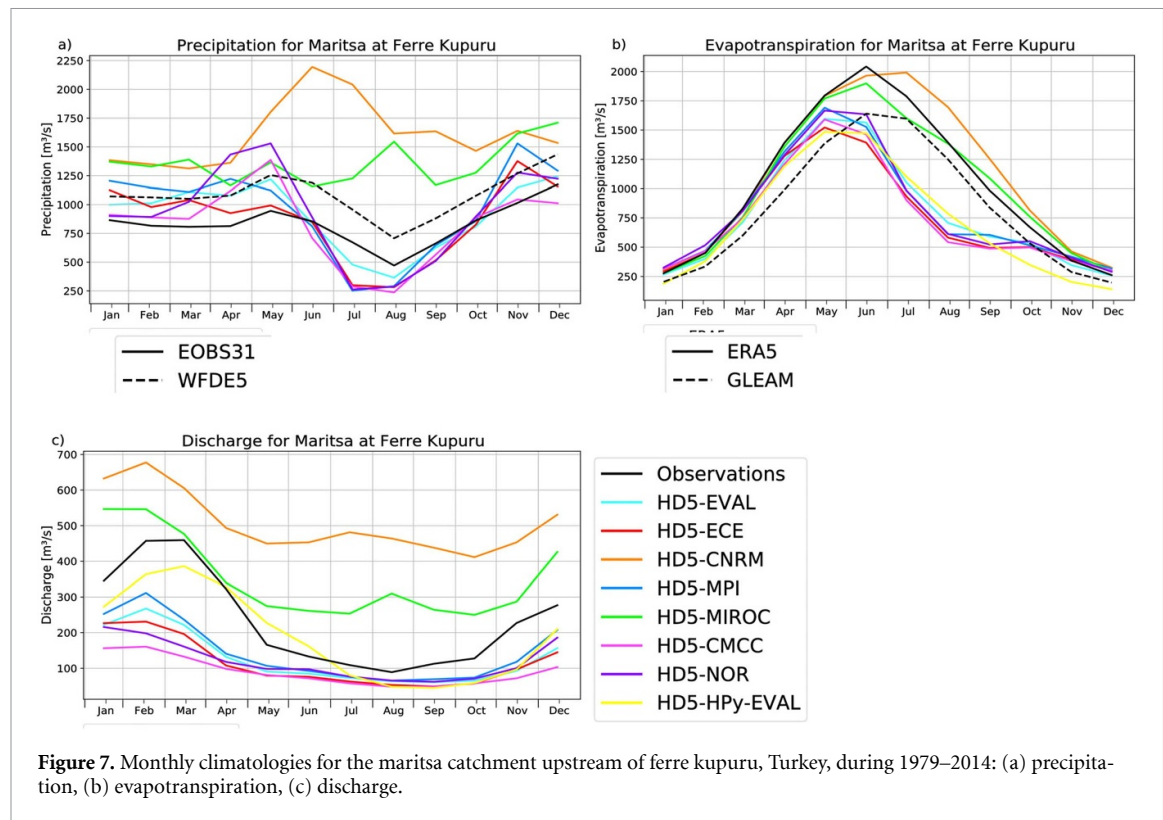


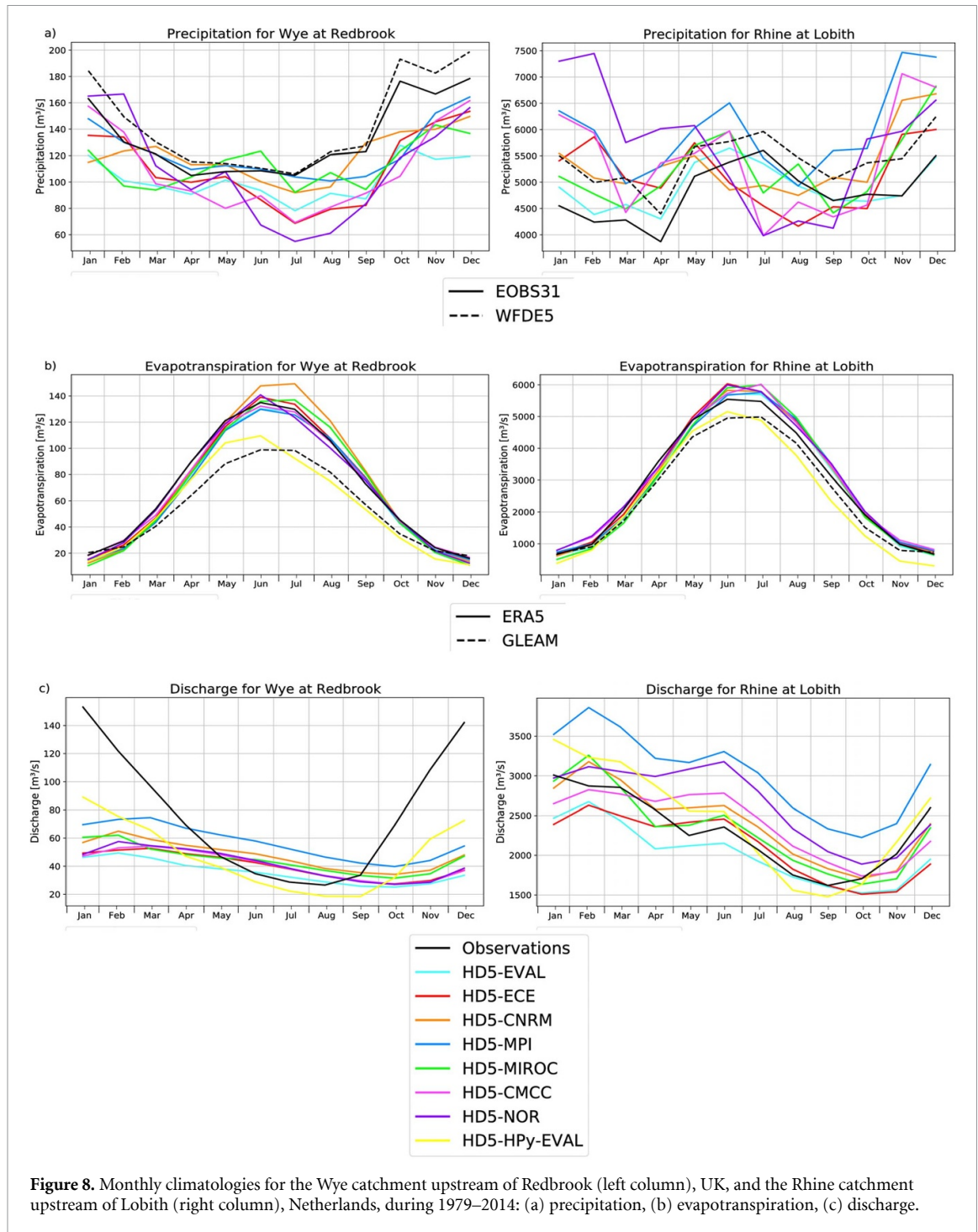
Figure 7. Monthly climatologies for the maritsa catchment upstream of ferre kupuru, Turkey, during 1979–2014: (a) precipitation, (b) evapotranspiration, (c) discharge.

HD5-CMCC underestimate the Guadalquivir discharge during the winter and early spring that is indicative for their widespread dry biases over the Iberian Peninsula. This is directly related to the underestimated ICON-CLM precipitation from late autumn, winter (only HD5-CMCC) and early spring. Due to this dry bias, also ET is much lower in these two simulations than for the other ensemble members. We also note that the discharge curve of HD5-HPY-EVAL is lower than for HD5-EVAL but with a similar shape (figure 6(e)). The observed February peak is almost not present for both simulations. This is likely caused by the negative precipitation bias (compared to WFDE5) of ICON-CLM EVAL during late autumn and winter, which leads to too dry land surface conditions so that sufficient surface runoff cannot be generated to yield the respective discharge peak.

For the Maritsa, HD5-CNRM and HD5-MIROC yield a similar large positive discharge bias (figure 7(c)) as for the Guadalquivir, which is also caused by a general overestimation of precipitation (figure 7(a)). For both catchments, the overestimation of precipitation is accompanied by an overestimation of atmospheric water vapour content (IWV) in summer compared to ERA5 (figure 6(c), supplementary figure S3). The large positive biases in precipitation and IWV of ICON-CLM CNRM are accompanied by a negative 2 m temperature bias (figure 6(b), supplementary figure S3), indicating a negative short wave radiation bias at the surface caused by overestimated cloud cover (not shown). Interestingly, the even larger summer IWV in ICON-CLM MIROC is associated with a smaller negative short wave radiation bias than for ICON-CLM CNRM compared to WFDE5 data (not shown), and it is not accompanied by a temperature bias. This is likely a compensating effect induced by the positive SST bias over the Mediterranean Sea in the MIROC6 forcing data (Tatebe *et al* 2019).

With the exception of the overly wet HD5-CNRM and HD5-MIROC simulations, all others consistently underestimate the discharge while following the seasonal curve of the observations (figure 7(c)). For HD5-MPI, this is in contrast to its performance in the Guadalquivir catchment. Additionally, these simulations consistently show lower ET than ERA5 and GLEAM during summer and autumn (figure 7(b)). The too wet HD5-CNRM and HD5-MIROC simulations, however, show larger ET than GLEAM but similar ET curves as ERA5, again indicating deficits in the ERA5 ET. We note that HD5-HPy-EVAL improves especially the discharge peak at the beginning of the year compared to HD5-EVAL. Its ET is relatively close to the ET of ICON-EVAL except during the last three months of the year. This is likely related to land surface being too dry during summer and early autumn when the dry bias in precipitation of HD5-EVAL is most pronounced.

As mentioned in section 3.2, all simulations show substantial negative discharge biases for numerous rivers in the British Isles, although precipitation biases are relatively weak (except for ICON-CLM EVAL,



ECE and MIROC, which exhibit larger dry biases, particularly along the western coastlines). We selected the Wye River as a representative river from this area. Here, all ICON-CLM simulations tend to underestimate precipitation compared to both the EOBS31 and WFDE5 datasets, particularly in the second half of the year (figure 8(a)). This leads to a general underestimation of discharge (figure 8(c)). The simulated flow of HD5-EVAL and the historical simulations also lag behind the observed discharge curve, due to the erroneous separation of total runoff into surface and subsurface runoff. HD5-HPy-EVAL does not exhibit this lag, suggesting a more accurate separation in the HydroPy model (see also section 3.1). Considering the simulated ET (figure 8(b)), we can see that HydroPy ET agrees well with GLEAM ET, but is much lower than ERA5 ET in summer. This is a reaction to the overly dry conditions caused by the underestimation of precipitation in the ICON-CLM hindcast simulation. However, the ICON-CLM ET does not seem to be affected by dryness due to the underestimated precipitation as only slight deviations from ERA5 are present during the summer; ICON_CLM CNRM and MIROC show even larger ET than ERA5.

As a central European river, we selected the river Rhine. Here, HD5-MPI and HD5-NOR largely overestimate the discharge for most parts of the year (figure 8(c)), which is related to an overestimation of precipitation between late autumn and early spring that is most prominent in winter and long lasting snow melt in the beginning of the year from the alpine part of the catchment (figure 8(a)). The discharge from the other simulations is much closer to the observed climatology, thereby adequately representing its monthly variations. Most simulations exhibit negative precipitation biases during the summer, especially if compared to WFDE5 data, which are less pronounced for ICON-CLM MPI, MIROC, and CNRM. ICON-CLM EVAL yields precipitation that is generally close to EOBS31 data, but lower than WFDE5 all year round except during spring. This actual dry bias in precipitation causes the discharge being underestimated during spring and summer. Similar but less intense as for the Wye River, we note that the ICON-CLM based discharge lags behind the observations by about one month, which is corrected in HD5-HPy-EVAL. Moreover, all simulations yield similar summer ET values that are larger than ERA5 data (figure 8(b)), even those with a large negative precipitation bias during the summer (ICON-CLM ECE, CMCC and NOR). As for the Wye river, the ICON-CLM ET does not show any noticeable sensitivity to biases in precipitation. In the contrary, HydroPy reacts sensitive to the drier conditions imposed by the ICON-CLM EVAL precipitation and yields less ET than ERA5 during the summer. While the HydroPy ET is generally close to the GLEAM data, it shows some underestimation in the second half of the year compared to the GLEAM data. This is probably due to the negative precipitation bias in ICON-CLM EVAL, which has led to drier soil conditions.

The results found in section 3.1, which are analysed in more detail for the Wye and Rhine rivers, clearly indicate that ICON-CLM tends to generate too little surface runoff. If the precipitation has no substantial dry bias, this usually means that too much water is infiltrating into the soil, thereby making the soil wetter than the actual runoff-generating event would do. This results in more subsurface runoff, but this process is much slower than surface runoff generation, meaning this wetter soil state persists for much longer than in reality. Consequently, the wetter soil can sustain ET in phases of drying, making ET rather insensitive to fluctuations in precipitation. Therefore, ICON-CLM ET does not appear to be sensitive to precipitation biases in central European catchments or the British Isles.

4. Evaluation of historical trends

Understanding long-term trends in simulated river discharge is essential for evaluating the ability of regional climate models to capture climate-driven hydrological changes. In this section, we analyse trends for various discharge statistics (see section 2.6) for the period 1960–2014 based on the ICON-CLM ensemble and compare them with two hydrological reference hindcasts. First, we evaluate how well the historical simulations reproduce robust, large-scale patterns found in the reference datasets (section 4.1). This evaluation provides insight into the influence of the driving GCMs on the simulated hydrological response, highlighting where the model behaviour consistently reflects or deviates from the climate-related signals in the reference hindcasts. Section 4.2 then investigates noticeable deviations from robust trend patterns in more detail.

4.1. Trends in simulated discharge

Figures 9(a)–(d) illustrate the annual mean discharge trends for the three hindcasts and the ensemble mean of the six historical simulations. HD5-EVAL captures the general trend patterns indicated by both HD5-ERA5 and HD5-EOBS31. This comprises wetting in Northern Europe and the British Isles, and drying in Southern Europe, particularly over the Iberian Peninsula, Italy and Turkey. However, it is surprising that there are large differences between HD5-EVAL and HD5-ERA5, despite both originating from the same ERA5 data. These differences are particularly evident over Central and North Eastern Europe. In these areas, the patterns of HD5-EVAL patterns align much better with those of HD5-EOBS31. The ensemble mean trend exhibits a pattern similar to that of HD5-ERA5, but with much weaker trend values. However, the strong drying over the Iberian Peninsula evident in the three hindcasts is only hinted at along the eastern coast. Conversely, some wetting is observed for several rivers flowing into the Atlantic on the west coast.

The ensemble mean trends are generally lower because the six historical simulations do not exhibit robust trends; instead, they yield conflicting signals inherited from the various GCM forcings (figures 9(e)–(j)). HD5-ECE and HD5-CNRM capture the wetting pattern of the hindcast, but HD5-MPI and HD5-MIROC show opposite trends, with widespread drying over the respective areas (Scandinavia and British Isles). HD5-NOR only captures the wetting pattern over northern Europe but shows the opposite trend over the British Isles and southern Scandinavia, while the behaviour of HD5-CMCC

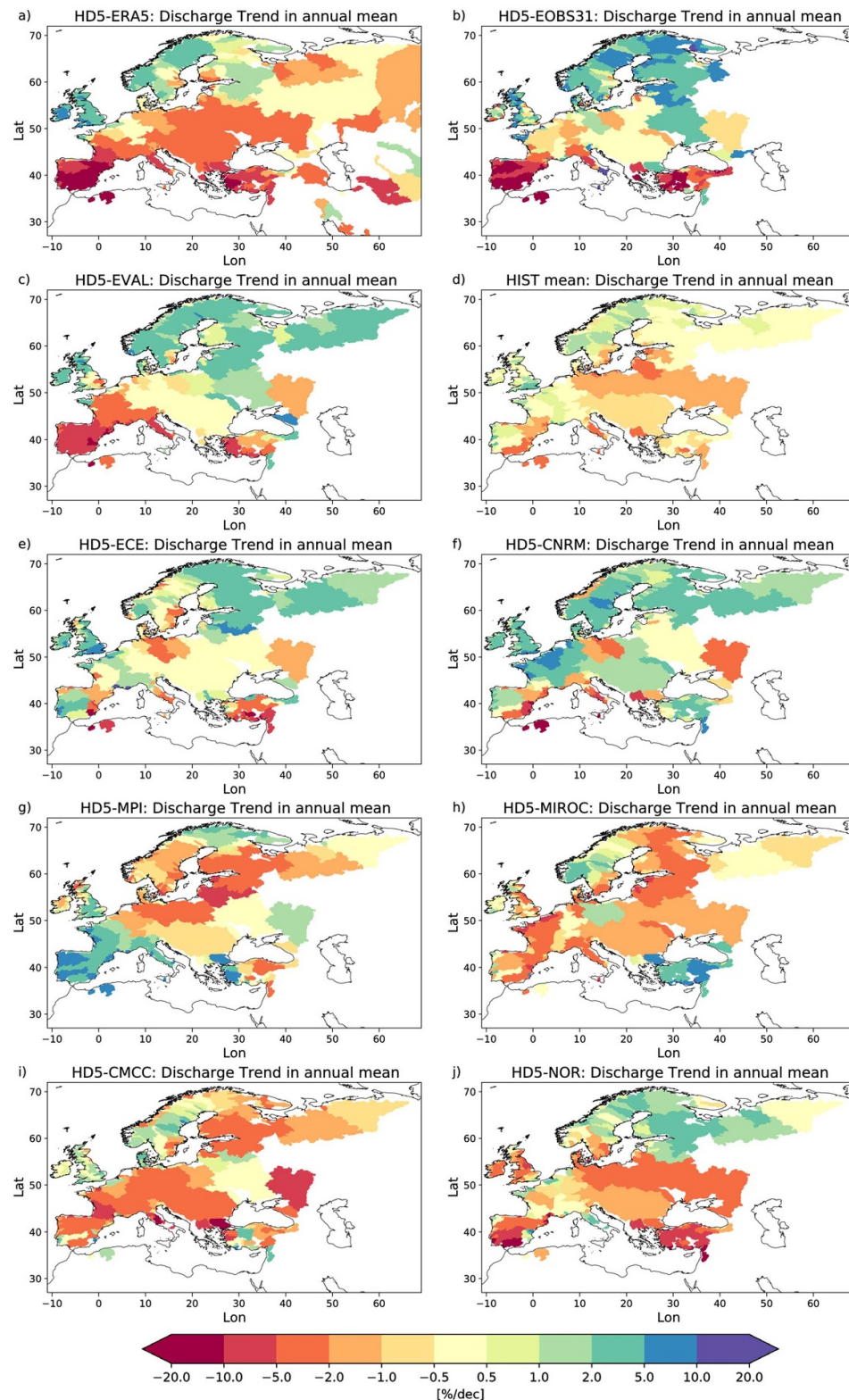


Figure 9. Trends in annual mean discharge [%/decade] for (a) HD5-ERA5, (b) HD5-EOBS31, (c) HD5-EVAL, (d) historical ensemble mean (HIST mean), (e) HD5-ECE, (f) HD5-CNRM, (g) HD5-MPI, (h) HD5-MIROC, (i) HD5-CMCC, and (j) HD5-NOR for 1960–2014.

is vice versa, showing weak increases over the UK and southwest Scandinavia, but drying over most other parts of Northern Europe. The drying over the Iberian Peninsula is evident in HD5-NOR, HD5-CMCC and HD5-MIROC, albeit with lower values for the latter two. HD5-ECE and HD5-CNRM only show drying in eastern Spain, whereas HD5-MPI shows wetting across the entire peninsula. This wetting extends to Italy and western Turkey—the opposite of the signal from the hindcasts. The other historical

simulations, except for HD5-NOR, show the drying in Italy. Meanwhile, the drying in Turkey is present in HD5-ECE, HD5-CMCC and HD5-NOR. HD5-CNRM and HD5-MIROC show an opposing signal there, while HD5-MPI indicates drying only in eastern Turkey.

For the annual maximum discharge ($Q_{\max}$), the two reference hindcasts (HD5-ERA5 and HD5-EOBS31) exhibit similar trends to those found in annual mean discharge (Q_{mean}), except for the tendency of $Q_{\max}$ to decrease in southern Scandinavia, in contrast to the increase of Q_{mean} (figures 10(a) and (b)). However, HD5-EVAL and several of the historical simulations show a different behaviour (figure 10). For HD5-EVAL (figure 10(c)), $Q_{\max}$ trends are similar to Q_{mean} trends only in southern Europe, but widespread opposing negative $Q_{\max}$ trends occur in northern and central Europe, and the British Isles. This different behaviour is surprising given that the lateral boundary and SST forcing for ICON-CLM in HD5-EVAL has the same origin as the atmospheric forcing of HD5-ERA5. The ensemble mean trends (figure 10(d)) are similar to those in HD5-EVAL. However, positive trends seen in some northern and mid-European rivers in the three hindcasts are almost completely absent in the ensemble mean. It can be noted that except for the western part of Central Europe and the British Isles, the general pattern of decreasing $Q_{\max}$ in HD5-EVAL and the historical simulations is relatively consistent with the trend pattern shown in HD5-ERA5.

For the historical simulations (figures 10(e)–(j)), HD5-CMCC, HD5-MIROC, HD5-MPI and HD5-NOR show similar trends in $Q_{\max}$ and Q_{mean} , although the latter two exhibit some exceptions. HD5-MPI shows decreasing $Q_{\max}$ trends in the northernmost Scandinavian rivers and the southern UK, whereas HD5-NOR exhibits a mixed pattern with a tendency for negative $Q_{\max}$ trends in place of the widespread positive Q_{mean} trends seen in the reference hindcasts over Northern Europe. HD5-ECE and HD5-CNRM contrast more sharply with the reference hindcasts than HD5-EVAL, exhibiting lower or decreasing $Q_{\max}$ trends in many catchments, contrary to the increasing or neutral Q_{mean} trends. Similar trend patterns only occur in southern Europe (for both simulations) and north-eastern Europe (for HD5-ECE). The general pattern of decreasing $Q_{\max}$ shown in the ensemble mean (see above) is present in each historical simulation, albeit with varying regional details.

The three hindcasts largely concur on the general trend pattern for the annual minimum discharge ($Q_{\min}$), showing a decrease in southern Europe and an increase elsewhere (figures 11(a)–(c)). An exception is seen for HD5-ERA5, which indicates negative $Q_{\min}$ trends for several rivers in eastern Central Europe. The ensemble mean trends (figure 11(d)) also capture the general trend pattern shown in HD5-EVAL, but have generally lower trend magnitudes. In line with the three hindcasts, both HD5-ECE and HD5-CNRM (figures 11(e) and (f)) show large-scale increases in $Q_{\min}$, though only a limited number of southern European rivers experience decreases. HD5-CMCC and HD5-NOR (figures 11(i) and (j)) indicate decreases in southern Europe, but only partially capture the respective increases. HD5-CMCC shows opposite decreasing trends over eastern Scandinavia, while HD5-NOR shows them over the British Isles and some southern Baltic rivers. The widespread negative Q_{mean} trends in Central Europe and the positive trends in southern Europe are also visible in the $Q_{\min}$ trend pattern for HD5-MPI (figure 11(g)), whereby the latter oppose the Q_{mean} and $Q_{\min}$ trends indicated by the three hindcasts. HD5-MIROC (figure 11(h)) shows a $Q_{\min}$ trend pattern that only partly follows its Q_{mean} trends. It captures the negative trends in southwestern Europe and the positive trends in southern Scandinavia and northern Germany indicated by the three hindcasts. However, it exhibits negative trends in the Eastern Baltic Sea catchment and positive trends in Turkey that oppose these trends. These latter trends are evident, at least in part, in all historical simulations except those for HD5-CMCC and HD5-NOR.

With the exception of Central Europe, trends in the annual number of low flow days are mostly opposite to the Q_{mean} trend pattern (figure 12). Wetter Q_{mean} are usually associated with a reduction in the number of low flow days, and vice versa. The dominant pattern in the three hindcasts (figures 12(a)–(c)) is a trend towards fewer low flow days in northern Europe and more in southern Europe. The ensemble mean trend pattern (figure 12(d)) broadly matches that from HD5-EVAL (figure 12(c)), albeit with lower magnitudes. The different historical ensemble members mainly yield similar patterns with some exceptions (figures 12(e)–(j)). HD5-MPI does not fully capture the decreasing low flow days in northern Europe (there are no decreases in the southern Baltic Sea catchments and even increases for some eastern Baltic rivers), nor do HD5-MIROC (there are increases over Northern Scandinavia and eastern Baltic rivers) or HD5-CMCC (there are mixed trends, mostly weak, in both directions). In southern Europe, HD5-CNRM does not exhibit a discernable trend pattern whereas HD5-MPI shows decreasing trends. HD5-ECE and HD5-MIROC also show decreasing trends in Turkey but only partially capture the increases over the Iberian Peninsula. HD5-ECE, for example, does not capture these increases in the south-west).

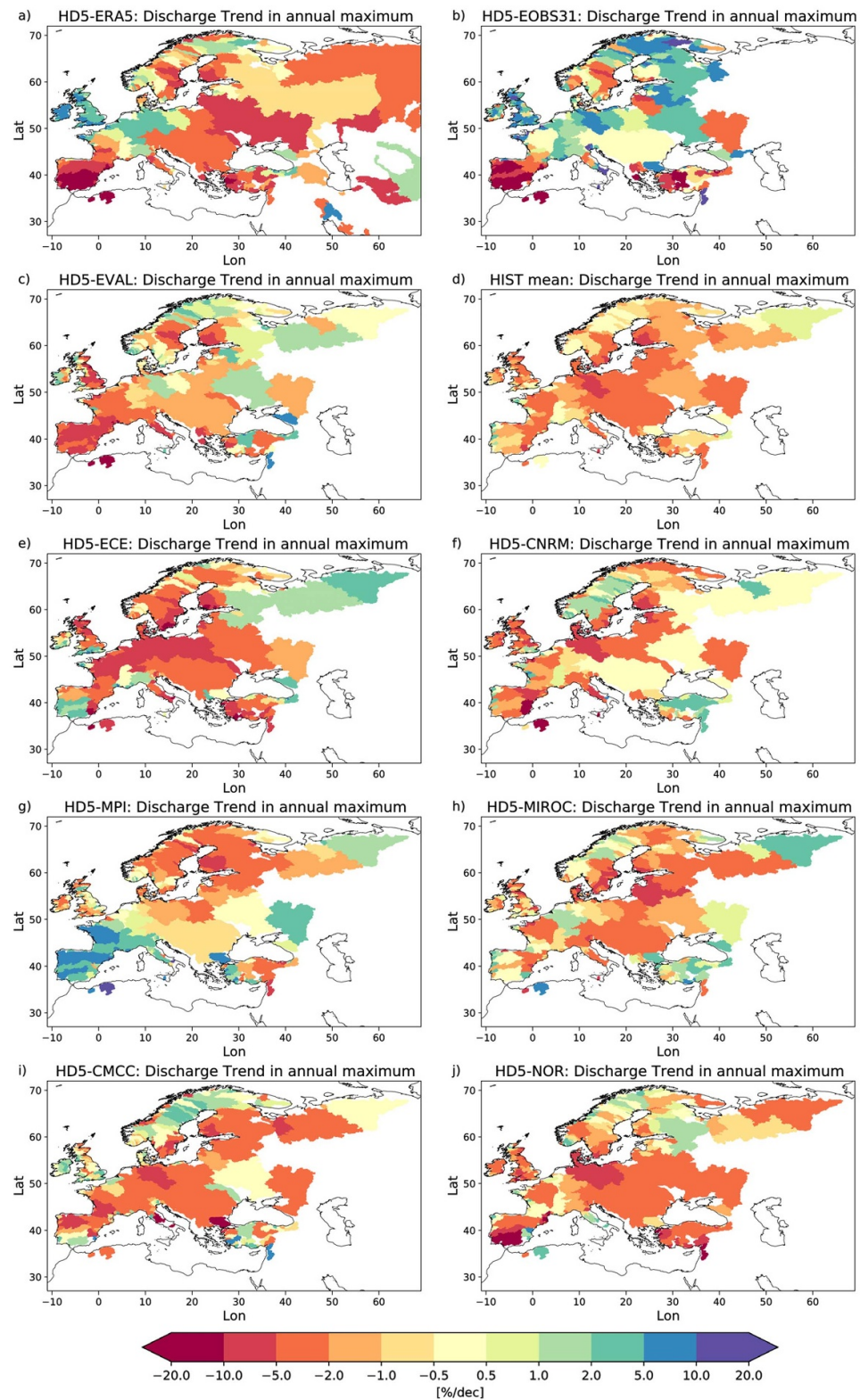


Figure 10. Trends in annual maximum discharge [%/decade] for (a) HD5-ERA5, (b) HD5-EOBS31, (c) HD5-EVAL, (d) historical ensemble mean (HIST mean), (e) HD5-ECE, (f) HD5-CNRM, (g) HD5-MPI, (h) HD5-MIROC, (i) HD5-CMCC, and (j) HD5-NOR for 1960–2014.

4.2. Inconsistent discharge trends in ERA5 driven simulations

In section 4.1, significant discrepancies in the discharge trends of HD5-EVAL and HD5-ERA5 were found, particularly over Central and North Eastern Europe. This occurred despite both simulations being

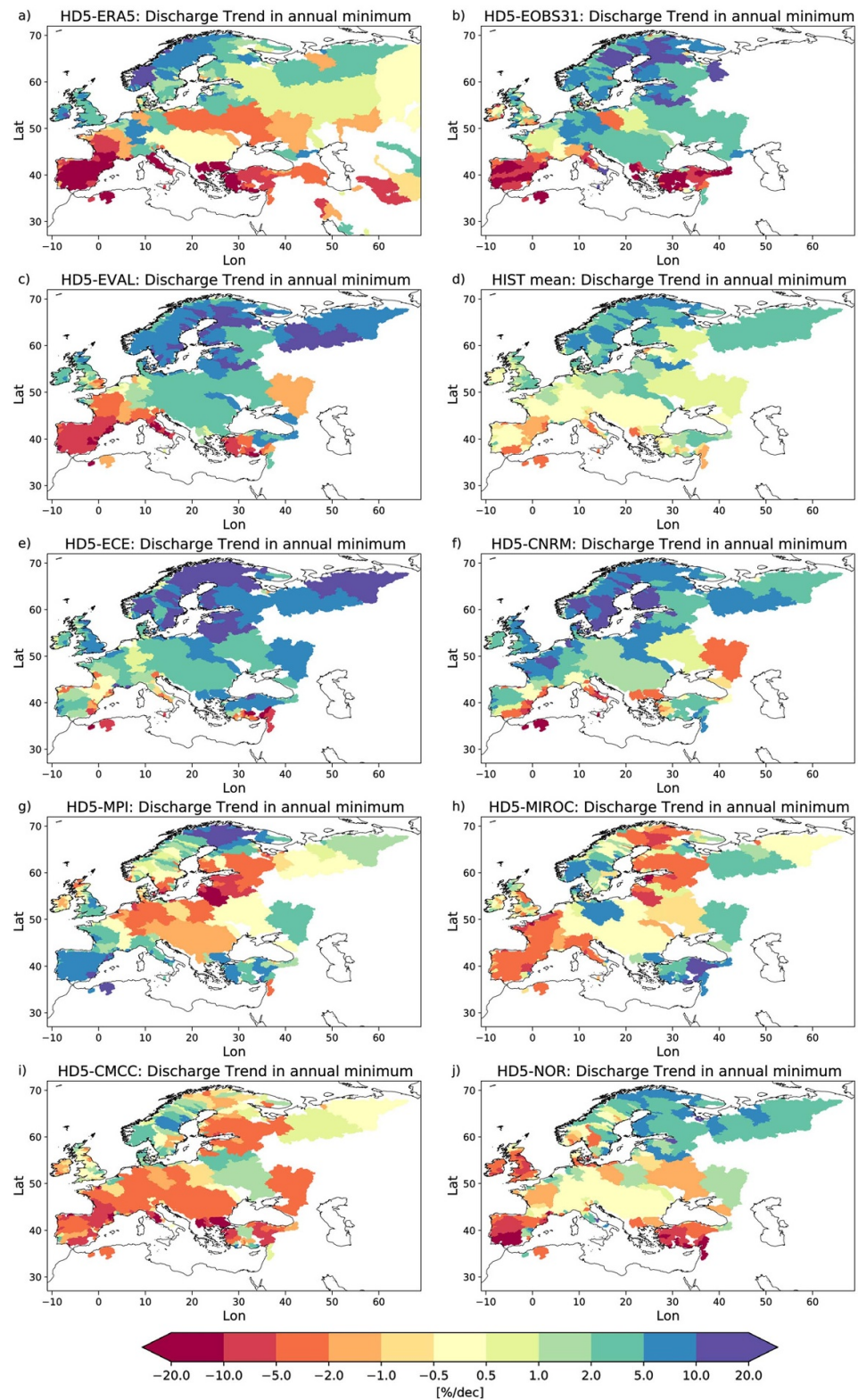


Figure 11. Trends in annual minimum discharge [%/decade] for (a) HD-ERA5, (b) HD-EOBS31, (c) HD-EVAL, (d) historical ensemble mean (HIST mean), (e) HD5-ECE, (f) HD5-CNRM, (g) HD5-MPI, (h) HD5-MIROC, (i) HD5-CMCC, and (j) HD5-NOR for 1960–2014.

forced by data from the same ERA5 dataset. These differences are mainly due to the contrasting precipitation trends in these areas, which also differ between ICON-CLM EVAL and ERA5 (figure 13). While ICON-CLM indicates consistent increasing trends, ERA5 shows mostly negligible or even decreasing trends. This deficient behaviour of ERA5 is likely related to the fact that ERA5 precipitation is affected

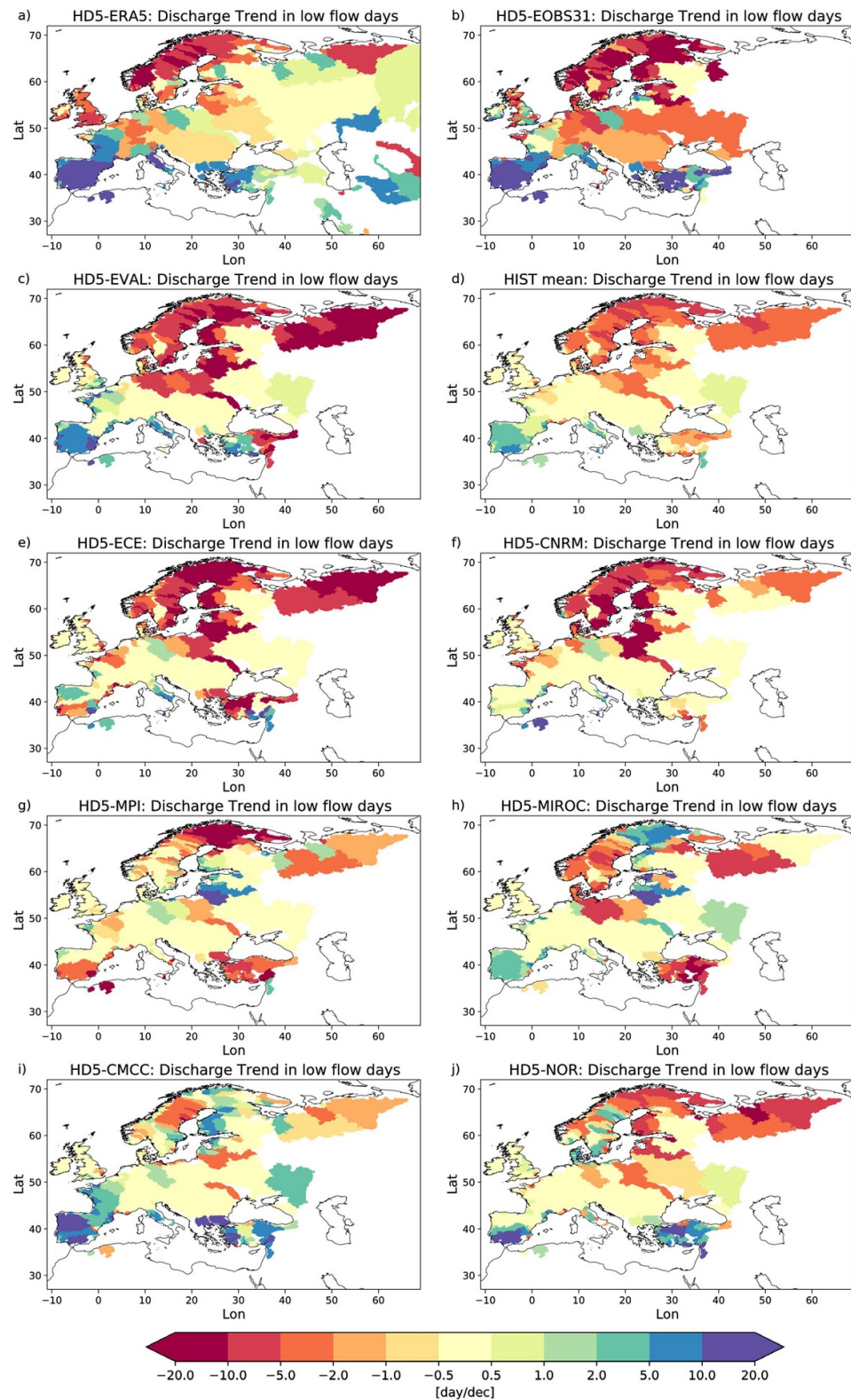
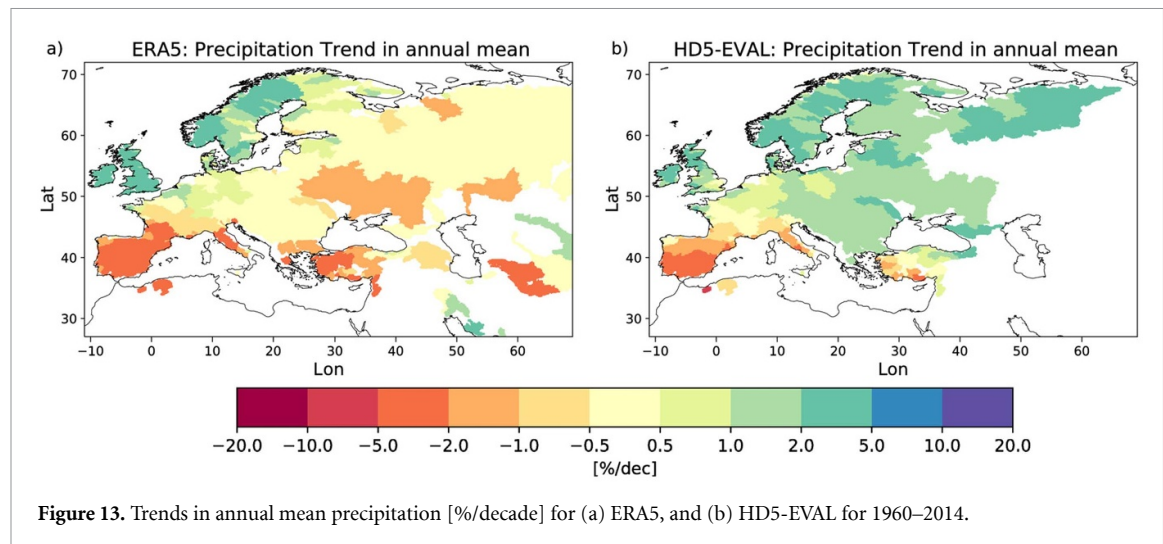


Figure 12. Trends in annual number of low flow days [day/decade] for (a) HD-ERA5, (b) HD-EOBS31, (c) HD-EVAL, (d) historical ensemble mean (HIST mean), (e) HD5-ECE, (f) HD5-CNRM, (g) HD5-MPI, (h) HD5-MIROC, (i) HD5-CMCC, and (j) HD5-NOR for 1960–2014.

by the amount and type of data available for data assimilation during different time periods. Such an impact of the observing system on integrated water vapour and precipitation was also found in the ERA40 reanalysis (Uppala *et al* 2006), a predecessor of ERA5 (Bengtsson *et al* 2004, Hagemann *et al* 2005).



4.3. On the robustness of hindcast trends in the historical simulations

We found that some of the historical simulations do not capture several robust trend patterns of the hindcasts. On the one hand, the increases in Q_{mean} and Q_{min} in Northern Europe are only partially present in HD5-MPI, HD5-MIROC and HD5-CMCC. These increases are directly related to warming in the recent decades, which is associated with increased winter precipitation, lower snowfall ratios, and earlier and lower snowmelt-related discharge peaks (Hagemann and Stacke 2023). Therefore, one would expect these signals to be present in historical simulations. Consequently, we analysed in more detail why these simulations do not fully capture this signal. We calculated annual and seasonal mean trends for each ICON-CLM simulation and the respective reference data using monthly mean time series of several variables (see section 2.6).

Both HD5-MIROC and HD5-CMCC show a largely confined pattern of increased winter precipitation (figure S4), which is mostly limited to western Scandinavia and a few southern Baltic rivers in the case of HD5-MIROC, or localized to central and southern Baltic rivers in the case of HD5-CMCC. Unlike EOBS31 and the hindcast HD5-EVAL, they do not show increases in precipitation in the eastern Baltic (figure 14). As temperature rises in this region (figure 15) are accompanied by the expected increase in ET (not shown), runoff decreases, leading to the deviating trends in the annual mean discharge (figure 9).

For HD5-MPI, the mean warming trend in the 2 m temperature is lower than that observed and simulated in the other simulations. In particular, observed winter warming is almost absent or much lower than in the other simulations (figure S5). Annual mean precipitation (figure 14), and winter precipitation (figure S4) in particular show an opposite pattern in south-western Europe and the Baltic Sea catchment. Additionally, winter ET decreases over western Scandinavia and the southern Baltic, contrary to the expected effects of warming (figure S6).

The reason for the atypical behaviour of the HD5-MPI trend lies in the selection of the corresponding GCM simulation used for the downscaling. This was taken from a 10-member initial condition ensemble conducted with MPI-ESM-1-2-HR (see, e.g. Tebaldi *et al* 2021). Regarding the global mean 2 m temperature, the selected member was one of the warmest during the 1960s and 1970s. After this, it entered a cold phase at the end of the 1990s and the beginning of 21st century (figure S6). Consequently, the warming gradient from 1960 to 2014 is lower than the ensemble mean. Therefore, the unusual temporal evolution of HD5-MPI is not erroneous, but merely an expression of the natural variability represented by MPI-ESM-1-2-HR.

A peculiar characteristic of HD5-MPI during the winter is that it shows an increase in shortwave downward radiation in the southern Baltic whereas the other experiments indicate a general decrease in northern Europe. This increase coincides with reductions in precipitation and integrated water vapour. These changes seem contradictory to the decrease in ET, which is energy-limited in this region (Teuling *et al* 2009). However, these changes suggest a reduction in cloud cover, which leads to lower values of longwave downward radiation. Since incoming shortwave radiation is low in winter, the reduction in longwave downward radiation exceeds the increase in incoming shortwave radiation, resulting in less

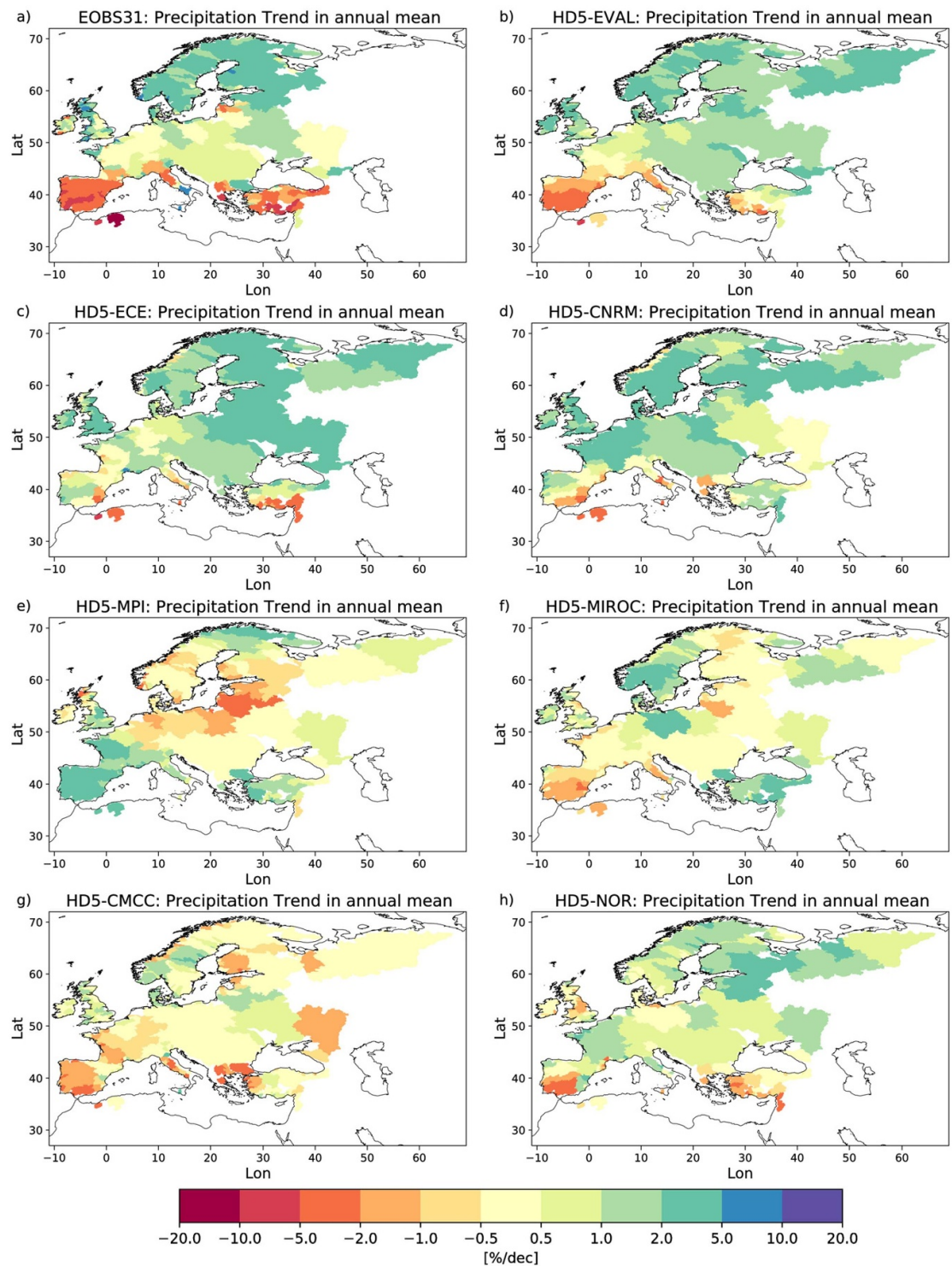
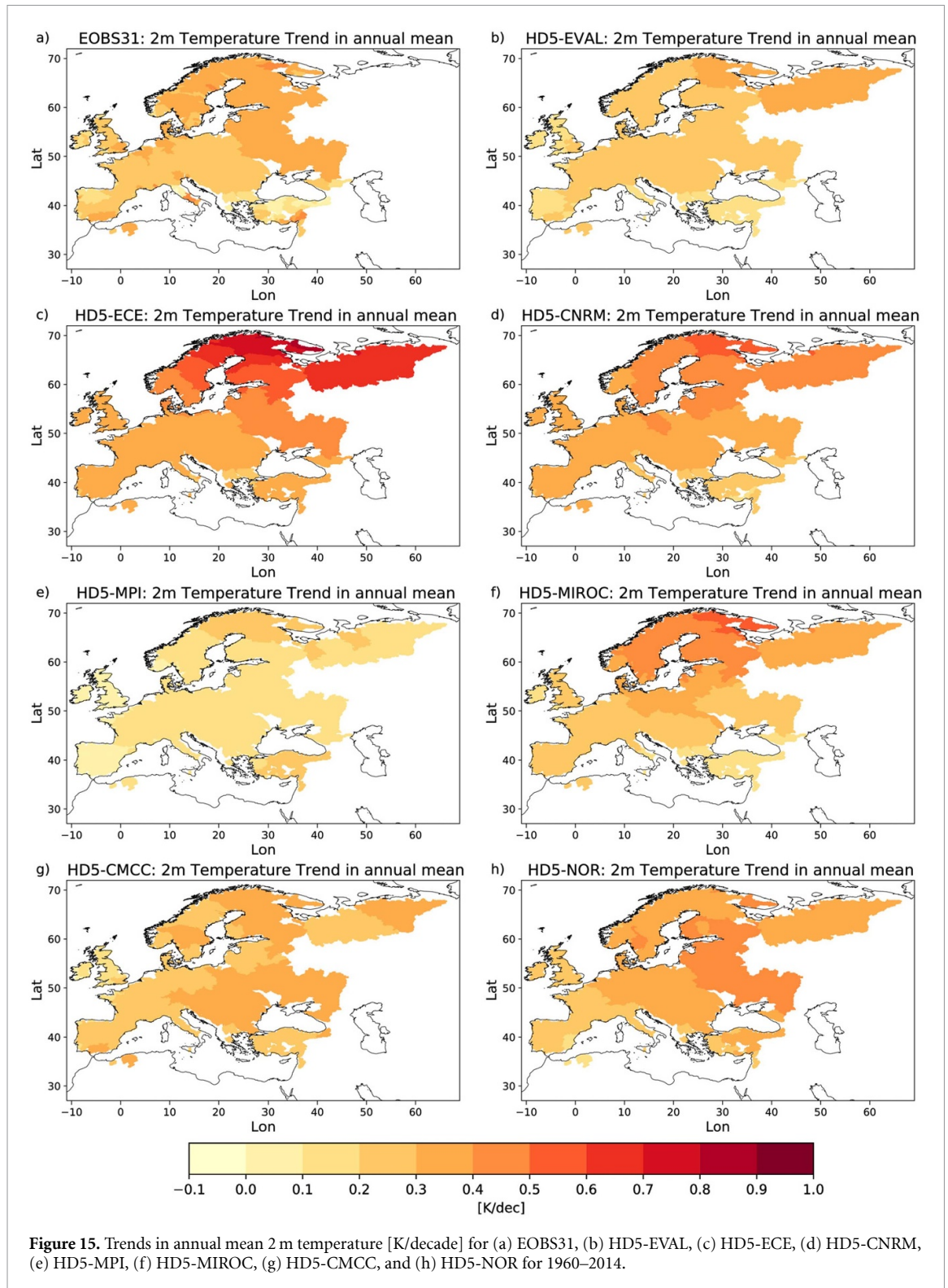


Figure 14. Trends in annual mean precipitation [%/decade] for (a) EOBS31, (b) HD5-EVAL, (c) HD5-ECE, (d) HD5-CNRM, (e) HD5-MPI, (f) HD5-MIROC, (g) HD5-CMCC, and (h) HD5-NOR for 1960–2014.

energy being available to warm the surface and drive ET. Table 6 provides corresponding values for the Glomma catchment to illustrate these effects.

On the other hand, the general drying in the mean, maximum and minimum discharges in Southern Europe is either absent or only partially present in some historical simulations. HD5-MPI stands out again here, as it shows widespread wetting in most of the respective catchments. The same is true for HD5-MIROC in Turkey. This is directly related to increasing precipitation trends in these areas, which are opposite to the trends shown in the EOBS31 data (figure 14). Note that the negative precipitation trends over Turkey in HD5-EVAL are less pronounced than in the EOBS31 data, resulting in lower magnitude negative discharge trends than in the two reference hindcasts (figure 9).



5. Summary and conclusions

This study examines the suitability of the ICON-CLM regional climate model for hydrological impact studies by generating high-resolution river discharge data for Europe using the HD model. Because ICON-CLM does not simulate lateral water flow, the HD model is used to route its surface and sub-surface runoff fields and obtain consistent river discharge estimates. These estimates are based on a hindcast simulation driven by ERA5 and six historical simulations driven by different CMIP6 global models. Additional hydrological hindcasts based on ERA5 and E-OBS data serve as quasi-observational

Table 6. Trends and mean values for the period 1960–2014 in the glomma catchment upstream of langnes, Norway. SW and LW down refer to shortwave and longwave downward radiation. The rightmost column shows the absolute trend of the sum of shortwave and longwave downward radiation.

Data	ET	SW down		LW down		LW + SW
	Trend [%/dec]	Trend [%/dec]	Mean [W m ⁻²]	Trend [%/dec]	Mean [W m ⁻²]	abs. Trend [W m ⁻² /dec]
ERA5	5.2	−0.6	19.6	1.0	235.4	2.24
ICON-CLM EVAL	13.8	−1.0	19.0	0.9	231.7	1.90
ICON-CLM MPI	−7.4	0.9	18.6	−0.1	229.6	−0.06

references. The study evaluates discharge biases, hydrological behaviour, and trend patterns for the period 1960–2014 using observed data from European river gauges.

Across Europe, the ICON-CLM-based simulations show widespread dry discharge biases, primarily caused by underestimated precipitation and an inadequately low surface runoff fraction simulated by the TERRA land surface scheme. These biases are particularly pronounced in western Europe, on the Iberian Peninsula, and in Turkey, while some eastern European rivers exhibit wet biases due to missing human water withdrawals in the HD model. Using HydroPy instead of TERRA improves the partitioning between surface and subsurface runoff, increases correlations and KGE scores, and reduces biases in many central and western European catchments. However, the performance worsens where strong river regulation reduces natural variability, as in the case for many Scandinavian rivers, since regulation is not taken into account by the HD model. Case studies for the Guadalquivir, Maritsa, Wye, and Rhine catchments confirm that precipitation errors dominate discharge biases, and that ICON-CLM ET shows limited sensitivity to deficits in precipitation due to overly wet soil states.

Trend analyses of the hindcasts reveal robust climate change signals, such as an increase in mean and minimum discharge in northern Europe and a decrease in discharge, particularly in terms of low flows, in southern Europe. However, the historical simulations show inconsistent and often contradictory trend patterns that are strongly influenced by their driving GCMs. While some models reproduce the reference pattern (e.g. ICON-CLM ECE and CNRM), others (e.g. ICON-CLM MPI, and MIROC) show opposite signals. This highlights the substantial spread caused by natural variability and GCM uncertainty in historical simulations. The global MPI-ESM simulation reflects the uncertainty introduced by natural variability well, as its behaviour deviates substantially from that of the other members of the MPI-ESM initial condition ensemble. However, a more typical member of this MPI-ESM ensemble would have been more appropriate for impact assessment studies. This would make it easier to communicate the results to stakeholders and the general public. Using the selected ensemble member will make it more difficult to assess the climate change signal in the continued scenario simulations. Furthermore, this will affect users who use the corresponding ICON-CLM data or derived discharges as a driving force for their impact model, for example for freshwater input in coupled simulations.

For the trend analysis, it is assumed that model biases remained stationary between 1960–2014. However, this assumption may not be entirely accurate because climatic changes during this period may have altered the characteristics of large-scale circulation and land-atmosphere interactions. As these processes are often non-linear, associated model biases may also evolve over time (see, for example, Seneviratne *et al* 2010, Shepherd 2014, Maraun 2016), which could affect the magnitude of simulated trends. However, a detailed analysis of such behaviour is beyond the scope of this study.

In principle, to comprehensively address model uncertainty, an ensemble comprising only one RCM using the forcing of different GCMs is insufficient. Ideally, different RCMs should also be used to perform historical and future scenario simulations. However, Déqué *et al* (2007) stated that the uncertainty arising from the use of different GCMs is generally greater than uncertainty arising from different RCMs. Christensen and Kjellström (2020) also documented that the selection of GCMs generally exerts a more substantial influence on the climate change signal than the selection of RCM. Nevertheless, Evin *et al* (2021) reported that, by the end of the 21st century, the uncertainty arising from RCMs could be comparable to that arising from GCMs for changes in total precipitation and mean temperature in winter and summer. They found that both sources of uncertainty are mainly dominated by a small number of individual models. It should be noted in this respect that the predecessor version of ICON-CLM used in the German NUKLEUS project (Sieck *et al* 2026) did not exhibit any apparent outlier behaviour (Cusinato *et al* 2026). To assess the uncertainty stemming from natural variability represented by the GCM, it is desirable to downscale multiple realizations of the same GCM, as Paquin *et al* (2025) did for

North America. This would also help to avoid confusion when communicating results relating to the selection of an unusual ensemble member, as seen with HD5-MPI.

One might assume that the historical simulation with the smallest biases is the most reliable for future projections. However, as Hagemann and Jacob (2007) have shown, this is not necessarily the case. They considered an ensemble of ten RCMs driven by one GCM and found that the hydrological values (precipitation, ET and runoff) were close to the corresponding observations for the RCM HadRM3P. However, it showed outlier behaviour in its predicted changes, which deviated significantly from the RCM multi-model ensemble mean. Therefore, in addition to having low biases when simulating today's climate, it is necessary for interactions and feedback to be accurately represented in a specific GCM-downscaling chain simulation. For instance, limited sensitivity to precipitation deficits (see section 3.3) could impact the adequate representation of land-atmosphere feedback for regions experiencing future spatial shifts in transitional soil moisture regimes. This issue should be addressed in future studies on scenario simulations conducted with the ICON-CLM ensemble that will be released in 2026.

Overall, we conclude that ICON-CLM provides suitable atmospheric forcing for hydrological impact studies. This indicates a satisfactory simulation of precipitation patterns and their temporal development. However, improvements are needed in terms of model tuning with regard to precipitation. In particular, undercatch correction in the observational reference data must be taken into account. Furthermore, the land surface scheme requires further development, particularly with respect to surface runoff generation. To conduct robust climate change assessments, we recommend combining the results of multiple GCMs and RCMs rather than relying on a single model's downscaling chain. Additionally, downscaling multiple realizations of the same GCM would help assess the uncertainty introduced by natural variability. For Europe, such assessments could be coordinated within the Euro-CORDEX framework.

Acknowledgments

This work was supported from funds of the German Federal Ministry of Research, Technology and Space in the project 'Updating the data basis for adaptation to climate change in Germany (UDAG)' (Grant No. 01LP2326A-E). We used DeepL and Copilot for improving the English. Note that no content was added by these two tools. We acknowledge the DKRZ for the use of resources in terms of computing time and storage capacity. HT is specifically thankful for the computational resources granted by the John von Neumann Institute for Computing (NIC) on the supercomputer JURECA at the Jülich Supercomputing Centre (JSC) through the Grant JJSC39.

Data availability statement

Most of the observed daily discharge data used can be obtained from the Global Runoff Data Centre, 56 068 Koblenz, Germany (https://grdc.bafg.de/data/data_portal/, last access: 7 November 2024). Other data were retrieved from public websites associated with the sources listed in table 2 of Hagemann and Stacke (2023). Additionally, French discharge data were retrieved from the EU Copernicus Marine Service information portal. ERA5 data were obtained from the German Climate Computing Centre (DKRZ) who received them from ECMWF's MARS archive. E-OBS vs. 31 data can be downloaded from the ECA&D website (<https://www.ecad.eu/download/ensembles/download.php>; last access: 7 August 2025). WFDE5 data were retrieved from the Copernicus Climate Data Store (<https://cds.climate.copernicus.eu>; last access: 7 November 2024). Monthly ET data from GLEAM v4.2a were obtained from the Integrated Climate Data Center (ICDC) at the University of Hamburg. The ICDC last accessed the data from www.GLEAM.eu on 9 July 2025 (last access: 19 December 2025). Within the CORDEX-Europe framework, the UDAG project and contributing members from the CLM community made the ICON-CLM data available via the ESGF, where they can be accessed at <https://metagrid-ceda.east.esgf.io/search?project=CORDEX-CMIP6>. An additional publication via the World Data Centre for Climate at DKRZ is planned. HD model discharge data and related HydroPy data are accessible via the World Data Centre for Climate at DKRZ. These data comprise:

- HD5-EVAL and HD5-Hpy-EVAL discharges, monthly ET and daily surface runoff and subsurface runoff data from HydroPy (Hagemann and Geyer 2026).
- HD model discharges for the six historical simulations (Hagemann 2026b).
- HD5-ERA5 discharge and related HydroPy runoffs (Hagemann 2025).

- HD5-EOBS31 discharge and related HydroPy runoffs (Hagemann 2026a).

Supplementary figures are available at: <https://doi.org/10.1088/2752-5295/aea9bc/data1>.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

Stefan Hagemann  0000-0001-5444-2945

Conceptualization (lead), Data curation (lead), Formal analysis (lead), Funding acquisition (equal), Investigation (lead), Methodology (lead), Software (lead), Validation (lead), Visualization (lead), Writing – original draft (lead), Writing – review & editing (lead)

Beate Geyer  0000-0001-8017-3136

Funding acquisition (equal), Investigation (equal), Methodology (supporting), Writing – review & editing (supporting)

Sergei Petrov  0009-0006-8880-9390

Investigation (equal), Writing – review & editing (supporting)

Evgenii Churiulin  0000-0003-4527-2125

Investigation (equal), Writing – review & editing (supporting)

Michael Jähn  0000-0003-0886-8245

Investigation (equal), Writing – review & editing (supporting)

Muhammed Muhshif Karadan  0000-0002-6666-9028

Investigation (equal), Writing – review & editing (supporting)

Vera Maurer  0000-0002-4064-7357

Investigation (supporting), Methodology (supporting), Writing – review & editing (supporting)

Christopher Purr  0000-0002-9376-432X

Investigation (equal), Methodology (supporting), Visualization (supporting), Writing – review & editing (supporting)

Shweta Singh  0000-0003-0752-1952

Investigation (supporting), Writing – review & editing (supporting)

Heimo Truhetz  0000-0002-1255-302X

Investigation (equal), Writing – review & editing (supporting)

References

- Becker A, Finger P, Meyer-christoffer A, Rudolf B, Schamm K, Schneider U and Ziese M 2013 A description of the global land-surface precipitation data products of the global precipitation climatology Centre with sample applications including centennial (trend) analysis from 1901–present *Earth Syst. Sci. Data* **5** 71–99
- Bengtsson L, Hagemann S and Hodges K I 2004 Can climate trends be calculated from reanalysis data? *J. Geophys. Res. Atmos.* **109** D11111
- Cai Y *et al* 2024 Reconciling global terrestrial evapotranspiration estimates from multi-product intercomparison and evaluation *Water Resour. Res.* **60** e2024WR037608
- Cherchi A *et al* 2019 Global mean climate and main patterns of variability in the CMCC-CM2 coupled model *J. Adv. Model. Earth Syst.* **11** 185–209
- Christensen O B and Kjellström E 2020 Partitioning uncertainty components of mean climate and climate change in a large ensemble of European regional climate model projections *Clim. Dyn.* **54** 4293–308
- Cornes R C, van der Schrier G, van den Besselaar E J M and Jones P D 2018 An ensemble version of the E-OBS temperature and precipitation data sets *J. Geophys. Res. D: Atmos.* **123** 9391–409
- Cucchi M, Weedon G P, Amici A, Bellouin N, Lange S, Schmied H M, Hersbach H and Buontempo C 2020 WFDE5: bias-adjusted ERA5 reanalysis data for impact studies *Earth Syst. Sci. Data* **12** 2097–120
- Cusinato E, Feldmann H, Geyer B, Ludwig P, Trachte K and Pinto J G 2026 From global to regional climate models: consistency assessment of the hydrological cycle *Clim. Dyn.* (submitted)

- Daewel U and Schrum C 2017 Low-frequency variability in North Sea and Baltic Sea identified through simulations with the 3-D coupled physical–biogeochemical model ECOSMO *Earth Syst. Dyn.* **8** 801
- Déqué M, Rowell D P, Lüthi D, Giorgi F, Christensen J H, Rockel B, Jacob D, Kjellström E, de Castro M and van den Hurk B 2007 An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections *Clim. Change* **81** 53–70
- Döscher R et al 2022 The EC-Earth3 Earth system model for the coupled model intercomparison project 6 *Geosci. Model Dev.* **15** 2973–3020
- Dzhamalov R G, Frolova N L and Kireeva M B 2013 Current changes in river water regime in the don river basin *Water Resour.* **40** 573–84
- Evin G, Somot S and Hingray B 2021 Balanced estimate and uncertainty assessment of European climate change using the large EURO-CORDEX regional climate model ensemble *Earth Syst. Dyn.* **12** 1543–69
- Eyring V, Bony S, Meehl G A, Senior C A, Stevens B, Stouffer R J and Taylor K E 2016 Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization *Geosci. Model Dev.* **9** 1937–58
- Früh B 2023 Updating the data basis for adaptation to climate change in Germany (UDAG) (available at: www.dwd.de/EN/research/projects/udag/udag.html, access: 27.01.2026)
- Geyer B et al 2026 Optimisation of ICON-CLM for the EURO-CORDEX domain: developments, sensitivities *Geosci. Model Dev.* **19** 5439–90
- Gupta H V, Kling H, Yilmaz K K and Martinez G F 2009 Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling *J. Hydrol.* **377** 80–91
- Haddeland I et al 2013 Global water resources affected by human interventions and climate change *Proc. Natl Acad. Sci. USA* **111** 3251–6
- Hagemann S et al 2013 Climate change impact on available water resources obtained using multiple global climate and hydrology models *Earth Syst. Dyn.* **4** 129–14
- Hagemann S 2025 Global ERA5-based high-resolution river runoff (Version 1). World data center for climate (WDCC) at DKRZ (https://doi.org/10.26050/WDCC/HDERA5_v1)
- Hagemann S 2026a European EOBS-based high-resolution river runoff (Version 1). World data center for climate (WDCC) at DKRZ (https://doi.org/10.26050/WDCC/EU-EOBS_v1)
- Hagemann S 2026b European historical river runoff based on ICON-CLM. World data center for climate (WDCC) at DKRZ (<https://doi.org/10.26050/WDCC/EU-hist>)
- Hagemann S, Arpe K and Bengtsson L 2005 Validation of the hydrological cycle of ERA-40, ECMWF, Reading, ERA-40 *Project Report Series 24*
- Hagemann S and Geyer B 2026 *European Hindcast of River Runoff Based on ICON-CLM* (World Data Center for Climate (WDCC) at DKRZ) (<https://doi.org/10.26050/WDCC/EU-hind>)
- Hagemann S, Ho-Hagemann H T M and Hanke M 2025 *The Hydrological Discharge Model—a River Runoff Component for Offline and Coupled Model Applications (5.2.4)* (Zenodo) (<https://doi.org/10.5281/zenodo.15004456>)
- Hagemann S and Jacob D 2007 Gradient in the climate change signal of European discharge predicted by a multi-model ensemble *Clim. Change* **81** 309–27
- Hagemann S, Nguyen T T and Ho-Hagemann H T M 2024 M.: a three-quantile bias correction with spatial transfer for the correction of simulated European river runoff to force ocean models *Ocean Sci.* **20** 1457–78
- Hagemann S and Stacke T 2023 Complementing ERA5 and E-OBS with high-resolution river discharge over Europe *Oceanologia* **65** 230–48
- Hagemann S, Stacke T and Ho-Hagemann H T M 2020 High resolution discharge simulations over Europe and the Baltic sea catchment *Front. Earth Sci.* **8**
- Hartmann D L et al 2013 Observations: atmosphere and Surface *The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker, D Qin, G K Plattner, M Tignor, S K Allen, J Boschung, A Nauels, Y Xia, V Bex and P M Midgley (Cambridge University Press)
- Hersbach H et al 2020 The ERA5 global reanalysis *QJR. Meteorol. Soc* **146** 1999–2049
- Hordoir R and Meier H E M 2010 Freshwater fluxes in the Baltic Sea: a model study *J. Geophys. Res.* **115** C08028
- Hordoir R, Polcher J, Brun-cottan J C and Madec G 2008 Towards a parametrization of river discharges into ocean general circulation models: a closure through energy conservation *Clim. Dyn.* **31** 891–908
- Jacob D et al 2013 EURO-CORDEX: new high-resolution climate change projections for European impact research *Reg Environ. Change* **14** 563–78
- Kahlenborn W, Porst L, Voß M, Fritsch U, Renner K, Zebisch M, Wolf M, Schönthaler K and Schauser I 2021 *Klimawirkungs- und Risikoanalyse 2021 für Deutschland (Kurzfassung)* (Umweltbundesamt, Climate Change) pp 127
- Khublaryan M G 2009 *Water Resources for Sustainable Development, with Particular Reference to Russia* ed N P Laverov (EOLSS Publications) pp 81–94
- Kling H, Fuchs M and Paulin M 2012 Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios *J. Hydrol.* **424–425** 264–77
- Knoben W J M, Freer J E and Woods R A 2019 Technical note: inherent benchmark or not? comparing nash-sutcliffe and kling-gupta efficiency scores *Hydrol. Earth Syst. Sci.* **23** 4323–31
- Kundzewicz Z W and Matczak P 2015 Hydrological extremes and security, *Pro. Int. Association of Hydrological Sciences* vol 366 pp 44–53
- Maraun D 2016 Bias correcting climate change simulations—a critical review *Curr. Clim. Change Rep.* **2** 211–20
- Marzeion B, Levermann A and Mignot J 2007 The role of stratification-dependent mixing for the stability of the atlantic overturning in a global climate model* *J. Phys. Oceanogr.* **37** 2672–81
- Mauritsen T et al 2019 Developments in the MPI-M earth system model version 1.2 (MPI-ESM1.2) and its response to increasing CO₂ *J. Adv. Model. Earth Syst.* **11** 998–1038
- Miralles D G et al 2025 GLEAM4: global land evaporation and soil moisture dataset at 0.1° resolution from 1980 to near present *Sci. Data* **12** 416
- Miralles D G, Holmes T R H, De Jeu R A M, Gash J H, Meesters A G C A and Dolman A J 2011 Global land-surface evaporation estimated from satellite-based observations *Hydrol. Earth Syst. Sci.* **15** 453–69
- Müller W A et al 2025 ICON: toward vertically integrated model configurations for numerical weather prediction, Climate predictions, and Projections *Bull. Am. Meteorol. Soc.* **106** E1017–31
- Paquin D, McCray C D, Gauthier C B, Giguère M, Asselin O, Bourgault P, Labonté M-P and Matte D 2025 The ouranos CRCM5-CMIP6 ensemble: a dynamically downscaled ensemble of CMIP6 simulations over North America *Sci. Data* **12** 1984

- Petrov S, Will A and Geyer B 2025 Linear meta-model optimization for regional climate models (LiMMo version 1.0) *Geosci. Model Dev.* **18** 6177–94
- Pham T V, Steger C, Rockel B, Keuler K, Kirchner I, Mertens M, Rieger D, Zängl G and Früh B 2021 ICON in climate limited-area mode (ICON release version 2.6.1): a new regional climate model *Geosci. Model Dev.* **14** 985–1005
- Rudolf B and Rubel F 2005 11 Global precipitation (Part 1/3) *Observed Global Climate, Landolt-Börnstein—Group V Geophysics* (Springer) pp 1–24
- Schrodin R and Heise E 2001 The multi-layer version of the DWD soil model TERRA_LM Consortium for Small-Scale Modelling *Tech. Rep.* **2** 16
- Schulz J-P, Vogel G, Becker C, Kothe S, Rummel U and Ahrens B 2016 Evaluation of the ground heat flux simulated by a multi-layer land surface scheme using high-quality observations at grass land and bare soil *Meteorol. Z.* **25** 607–20
- Séférian R et al 2019 Evaluation of CNRM earth system model, CNRM-ESM2-1: role of earth system processes in present-day and future climate *J. Adv. Model. Earth Syst.* **11** 4182–227
- Seland Ø et al 2020 Overview of the norwegian earth system model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations *Geosci. Model Dev.* **13** 6165–200
- Seneviratne S I, Corti T, Davin E L, Hirschi M, Jaeger E B, Lehner I, Orlowsky B and Teuling A J 2010 Investigating soil moisture–climate interactions in a changing climate: a review *Earth Sci. Rev.* **99** 125–61
- Shea D 2014 Trend analysis. NCAR climate data guide (available at: <https://climatedataguide.ucar.edu/climate-tools/trend-analysis>, access: 28.07.2026)
- Shepherd T G 2014 Atmospheric circulation as a source of uncertainty in climate change projections *Nat. Geosci.* **7** 703–8
- Sieck K et al 2026 NUKLEUS—A first kilometre scale multi-model climate ensemble for Germany: evaluation, EGU sphere [preprint] (<https://doi.org/10.5194/egusphere-2026-1024>)
- Singh S et al 2026 Updating the data basis for adaptation to climate change in Germany using the regional climate model ICON-CLM *Int. J. Climatol.* (submitted)
- Smakhtin V, Revenga C and Döll P 2004 A pilot global assessment of environmental water requirements and scarcity *Water Int.* **29** 307–17
- Stacke T and Hagemann S 2012 Development and evaluation of a global dynamical wetlands extent scheme *Hydrol. Earth Syst. Sci.* **16** 2915–33
- Stacke T and Hagemann S 2021 HydroPy (v1.0): a new global hydrology model written in Python *Geosci. Model Dev.* **14** 7795–816
- Tatebe H et al 2019 Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6 *Geosci. Model Dev.* **12** 2727–65
- Tebaldi C et al 2021 Climate model projections from the scenario model intercomparison project (scenariomip) of CMIP6 *Earth Syst. Dyn.* **12** 253–93
- Teuling A J et al 2009 I.: a regional perspective on trends in continental evaporation *Geophys. Res. Lett.* **36** L02404
- Uppala S M et al 2006 The ERA-40 re-analysis *QJR. Meteorol. Soc.* **131** 2961–3012
- Väli G, Meier H E M and Elken J 2013 Simulated halocline variability in the Baltic Sea and its impact on hypoxia during 1961–2007 *J. Geophys. Res. Oceans* **118** 6982–7000
- van den Besselaar E 2025 E-OBS daily gridded observations for Europe from 1950 to present: product user guide (available at: <https://confluence.ecmwf.int/display/CKB/E-OBS+daily+gridded+observations+for+Europe+from+1950+to+present%3A+Product+user+guide>, access: 12.02.2026)
- Vasylenko Y and Koshkina O 2020 Description of the characteristics of dnipro river basin. summary, version 2.0, umweltbundesamt gmbh; office international de l'eau (iow), Vienna, austria; paris, france, european union water initiative plus for Eastern Partnership Countries (EUWI+): results 2 and 3 ENI/2016/372-403
- Vicente-Serrano S M et al 2025 Characteristics of widespread extreme precipitation events in Peninsular Spain and the Balearic Islands: spatio-temporal dynamics and driving mechanisms *Clim. Dyn.* **63** 340
- Weedon G P, Gomes S, Viterbo P, Shuttleworth W J, Blyth E, Österle H, Adam J C, Bellouin N, Boucher O and Best M 2011 Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century *J. Hydrometeorol.* **12** 823–48
- Zängl G, Reinert D, Rípodas P and Baldauf M 2015 The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: description of the non-hydrostatic dynamical core *QJR. Meteorol. Soc.* **141** 563–79