Journal of Hydrology 620 (2023) 129415

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Assessing uncertainty in hydrological projections arising from local-scale internal variability of climate

Qifen Yuan [a,b,*], Thordis L. Thorarinsdottir [c], Stein Beldring [a], Wai Kwok Wong [a], Chong-Yu Xu [b]

[a] Norwegian Water Resources and Energy Directorate (NVE), Oslo, Norway
[b] Department of Geosciences, University of Oslo (UiO), Oslo, Norway
[c] Norwegian Computing Center (NR), Oslo, Norway

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief.

Keywords:
Local-scale internal variability
Uncertainty
Gridded weather generator
Spatially distributed HBV-model
ANOVA
Climate change impact

ABSTRACT

Hydrological impact assessments are increasingly performed at fine spatial and temporal resolutions in order to resolve local-scale changes under a future climate. Apart from the uncertainty represented by different climate models, emission scenarios and post-processing methods, the local-scale internal variability of the climate can be a major source of uncertainty for hydrological projections. To assess the latter at the catchment scale, this paper presents a methodology which is particularly suitable for spatially distributed hydrological models. An ensemble of daily precipitation and daily mean temperature realizations on a high-resolution grid is simulated from stochastic weather generators (WGs) trained on historical data and equipped with climate change information obtained from a regional climate model. Based on the resulting simulated daily runoff data, the significance of changes in the runoff regime is assessed using a statistical hypothesis test, and the variability contributed by the two input variables is quantified using the analysis of variance (ANOVA). As a proof of concept, simulations on a 1-km grid over a period of 19 years are carried out for nine catchments in central Norway. Significant changes in runoff regimes are found, indicating that the trends introduced in the WGs are not overwhelmed by the local-scale internal variability. Variability in the runoff simulations varies substantially throughout the year; it is highest in periods with potential snowmelt and lowest during winter. Temperature is the dominant source of variability in the colder months (November–March) due to its influence on rainfall and snowmelt. High variability in May–June is contributed comparably by both temperature and precipitation. In summer and early autumn the runoff variability is precipitation dominated. The results are in line with findings in the literature where the runoff variability is driven by the large-scale internal climate variability. This indicates that ignoring the local-scale internal variability may yield an underestimation of the overall variability in runoff projections and projected changes.

1. Introduction

In order to develop adaptation strategies relevant for e.g. agriculture, civil water use, hydropower production and reservoir operation, information on hydrological changes at regional to local scales under a changing climate can be required at natural catchments or catchments that are subject to extensive human interventions (Barros et al., 2014; Hanssen-Bauer et al., 2017b; Abhishek and Kinouchi, 2021; Abhishek et al., 2021). In Northern Europe, the projected warming rate is amongst the highest in the world (Collins et al., 2013), which, combined with the projected changes in precipitation, is expected to considerably impact evaporation, rainfall and snow, leading to significant changes in future runoff regime, floods and droughts (von Storch et al., 2015). In Norway, large variations in topography, climate and land cover are commonly characterized within a catchment, where hydrological processes can be highly sensitive to small-scale climate variability (Beldring et al., 2003). For an accurate quantification of the local-scale changes, hydrological impact assessments have optimally been performed at fine spatial and temporal scales (e.g. Wong et al., 2011; Hanssen-Bauer et al., 2017a).

Progress in high-resolution gridded observational data products such as the seNorge dataset (Lussana et al., 2019) presents new opportunities for hydrologists to carry out reliable impact assessments at required fine scales. However, climate projections at equivalent scales are currently not widely available. Future climate information derives from coupled atmosphere–ocean general circulation models (GCMs) that are too coarse. Regional climate models (RCMs), typically at a spatial resolution of 10–15 km (e.g. Jacob et al., 2014), provide valuable information on mesoscale changes over a region with new

* Corresponding author at: Norwegian Water Resources and Energy Directorate (NVE), Oslo, Norway.
E-mail address: qiyu@nve.no (Q. Yuan).

https://doi.org/10.1016/j.jhydrol.2023.129415

Available online 25 March 2023

advances in convection-permitting climate modeling (Lind et al., 2020; Prein et al., 2020). To bridge the scale gap and remove systematic biases in climate model outputs, various post-processing methods have been performed in hydrological impact studies (Maraun and Widmann, 2018).

It has been found that the uncertainty in hydrological model projections is primarily caused by the uncertainty in future climatic forcing (e.g. Wilby and Harris, 2006; Dobler et al., 2012; Bosshard et al., 2013; Addor et al., 2014; Wang et al., 2020). In a multi-model framework, uncertain climatic forcing is represented by using projections from different climate models and/or under different emission scenarios post-processed with different methods, and often assessed against the uncertainties arising from different hydrological models and/or parameter sets, or furthermore different analysis methods e.g. flood frequency analysis (Lawrence, 2020; Meresa et al., 2021). Less intensively explored is the effect of the internal variability of the climate, partly because of the limited number of model runs available for any particular GCM or RCM (Peel et al., 2015). The natural, internal variability of the climate stems from its intrinsic chaotic nature, and cannot be reduced with our increasing knowledge in climate science (Yip et al., 2011; Ghil and Lucarini, 2020). It thus determines a lower limit of our uncertainties associated with the climate system (Gelfan et al., 2015) and potentially represents a major source of uncertainty in hydrological projections.

Various methods have been suggested to account for the large-scale internal variability of the climate in hydrological impact studies. Seiller and Anctil (2014) found that the uncertainty arising from a five-member GCM/RCM ensemble is larger than the uncertainties contributed by different lumped hydrological models. Peel et al. (2015) proposed to approximate the within-GCM uncertainty using stochastic replicates of available projections, where they found the within-GCM uncertainty is amplified from mean annual precipitation to the simulated mean annual runoff and reservoir yield. Gelfan et al. (2015) simulated a 45-member initial condition ensemble of a GCM and found higher uncertainty in streamflow simulations for flood periods and in catchment with stronger nonlinearity of runoff generation mechanisms. One common issue with these studies is that the climate projections are bias-corrected and downscaled using *model output statistics* (MOS) methods such as simple mean adjustment and empirical quantile mapping, which are problematic for capturing spatial and temporal variability at the finer hydrological modeling scale (Maraun and Widmann, 2018).

To better represent the local-scale internal variability of the climate, stochastic ensembles are commonly generated using statistical methods. Lafaysse et al. (2014) empirically downscaled GCM daily outputs to 100 possible realizations of local-scale surface variables using K-nearest neighbors (K-nn) resampling. They found the sign of change in future annual discharge is uncertain mainly as a result of the internal variability, and the small-scale internal variability accounts for 15%–25% of the total uncertainty in projections given the uncertainties from GCM and downscaling method. Gao et al. (2020) generated 1000 replicates of bias-corrected, downscaled GCM rainfall series using a stochastic daily rainfall model. Their results highlight the dominant effect of the precipitation internal variability in projections of future extreme flows. In Fatichi et al. (2014), hourly precipitation and temperature data were generated using a multi-site weather generator (WG) approach. They found that internal climate variability overwhelms climate change signals in streamflow mean, frequency and seasonality. However, these methods have limited strength to represent the internal variability of a changing climate in a colder region on a high-resolution grid, because (1) temporal sequence was not considered in the resampling procedure (Lafaysse et al., 2014); (2) only the rainfall variability was represented in Gao et al. (2020), whereas in regions with seasonal snow cover the variability of temperature is not negligible (Dobler et al., 2012; Gelfan et al., 2015); and (3) climate change information derived from climate model outputs was added assuming piece-wise stationarity in each future period (Fatichi et al., 2014), whereas a changing climate could preferably be described as a continuously evolving space–time process.

In this paper, we present a methodology for exploring the hydrological uncertainty simulated when local-scale internal variability is represented on a high-resolution grid. Specifically, we employ the WG methods developed for daily mean temperature (Yuan et al., 2019) and daily precipitation (Yuan et al., 2021) to simulate an ensemble of gridded time series. A generator for each variable is trained on a high-resolution observational data product for Norway, where the fine-scale spatial and temporal marginal variations and the spatio-temporal dependences are described using parametric statistical models. The climate change signals obtained from an RCM are added to equivalent components of the generator including the trend. The ensemble of temperature and precipitation data is then used to simulate daily runoff series with a spatially distributed hydrological model, which has been used to perform hydrological impact assessments for the governmental report "Climate in Norway 2100" (Hanssen-Bauer et al., 2017a) issued by the Norwegian Centre for Climate Service (NCCS).

Based on the simulated runoff, we first perform a statistical hypothesis test to assess whether there are significant changes in runoff regime given the trends in the WGs and the simulated variability in the input data. Assessing the significance of regime change is a common task in many studies. Some found that climate change signals in river streamflow are largely masked by internal climate variability (e.g. Lafaysse et al., 2014; Fatichi et al., 2014), while others found robust expected changes in runoff cycle despite projection uncertainty (e.g. Finger et al., 2012; Addor et al., 2014). Our assessment can contribute to this ongoing discussion.

Next, we quantify the variability contributions of the two input variables. Variability in simulations arising from various sources are often estimated using analysis of variance (ANOVA) techniques. Some use a time series based approach called the Quasi-Ergodic ANOVA (Hingray and Saïd, 2014; Lafaysse et al., 2014; Vidal et al., 2016; Hingray et al., 2019), which assumes, among other things, that the uncertainty from the internal climate variability is constant or roughly constant across projection lead times. Here, we adopt the traditional ANOVA approach as it requires fewer assumptions. This approach has seen applications in climate projections (Yip et al., 2011) and hydrological impact studies (Finger et al., 2012; Bosshard et al., 2013; Addor et al., 2014; Lawrence, 2020; Meresa et al., 2021). Quantifying sources of variability related to input variables is supposed to help focus attention on the process tractable components of the total variability.

In a case study, we consider nine different catchments in central Norway. The catchments have different runoff regimes, climatic types and areas varying between 144 km^2 and 3084 km^2. The remainder of the paper is organized as follows. Section 2 introduces the study area. Methods for generating climate input data, simulating daily runoff data and analysis of the simulated runoff data are given in Section 3. The results are presented in Section 4, followed by a discussion in Section 5. The paper then concludes with a brief summary in Section 6.

2. Study area

The study area is located in Trøndelag in central Norway, a region with large climatic variation and abundant hydropower resources. As shown in Fig. 1, it covers the entire Trøndelag and a small part of neighboring Sweden. We focus on nine catchments in this region (shaded in gray in Fig. 1) with sizes ranging from 144 to 3084 km^2 and median elevations varying between 295 and 734 m.a.s.l (see Table 1). They are natural catchments, i.e. direct human interventions have negligible impacts on the catchment hydrology. Most of these catchments have daily discharge time series available for decades long period, maintained and archived in the station network at the Norwegian Water Resources and Energy Directorate. Referring to the Nordic runoff regime map presented by Gottschalk et al. (1979), there is a gradual change of the regime as we move from east to west. The catchments

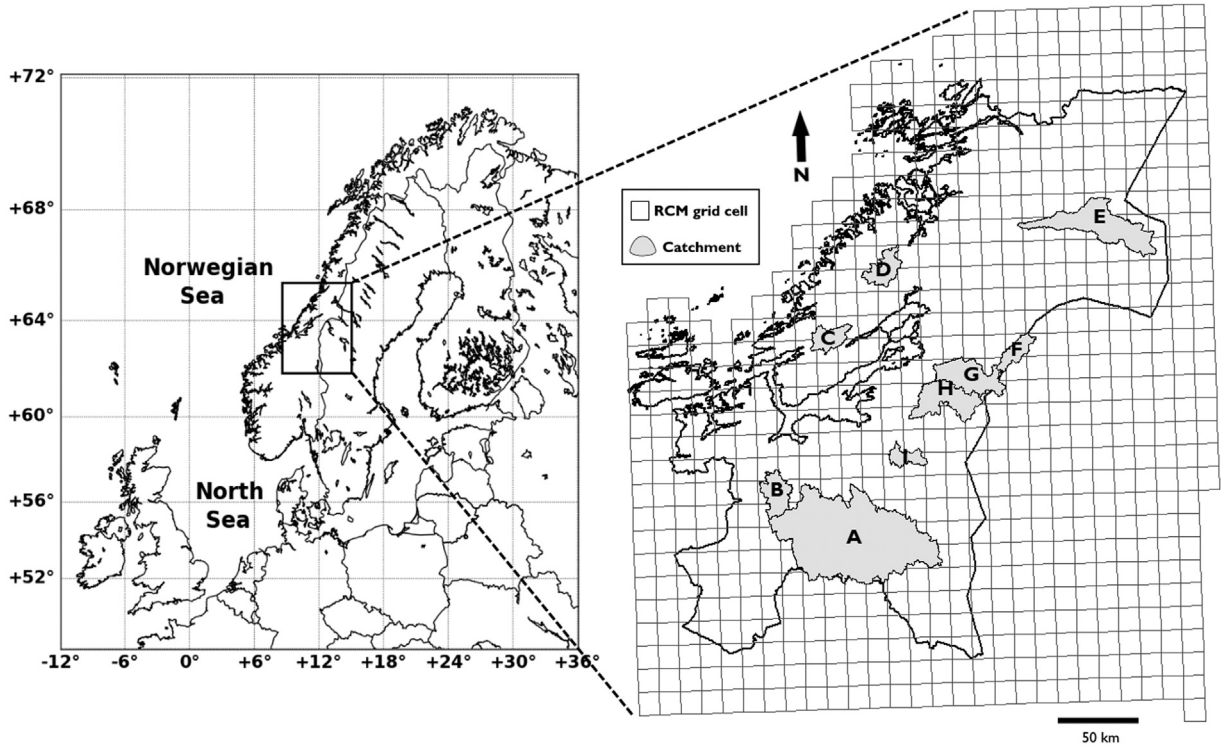

Fig. 1. The study area in central Norway. The nine catchments considered are shaded in gray and indicated by the catchment ID given in Table 1.

Table 1
Characteristics of the nine catchments considered in the case study in Trøndelag, Norway. Annual mean temperature and annual precipitation in the simulation period 1987–2005 are given by the sample mean computed based on the WG-simulated time series averaged over the catchment area, see Section 3.1.1.

Catchment	ID	Size (km^2)	Median elevation (m.a.s.l)	Discharge record (date)	Runoff regime	Temperature (°C)	Precipitation (mm)
Gaulfoss	A	3084	734	1958-05-19	mountain	0.4	1037
Aamot	B	286	460	1987-07-31	mountain	2.2	1239
Krinsvatn	C	206	349	1915-12-19	Baltic	3.1	2379
Oeyungen	D	245	295	1916-09-26	Baltic	2.9	2169
Trangen	E	852	558	1934-10-10	mountain	0.2	1304
Veravatn	F	176	514	1966-10-07	transition	1.1	1561
Dillfoss	G	484	506	1973-01-01	transition	1.7	1161
Hoeggaas	H	491	505	1912-04-25	transition	1.6	1248
Kjeldstad	I	144	578	1913-01-01	transition	1.4	1914

Gaulfoss, Aamot and Trangen (A, B and E) have the mountain type regime, with dominant high flows in spring due to snowmelt and dominant low flows in winter due to snow accumulation. Located in a transition zone between the mountain type and the inland type regime, the four catchments Veravatn, Dillfoss, Hoeggaas and Kjeldstad (F-I) tend to have a second or third highest runoff occurring in autumn in addition to the snowmelt high flows. The catchments Krinsvatn and Oeyungen (C and D) have the Baltic type regime and usually experience a summer low flow in addition to the lowest winter flows. According to the Köppen–Geiger climate classification presented by Kottek et al. (2006), Krinsvatn and Oeyungen have a maritime climate while the other catchments have a humid continental climate.

3. Methods and data

3.1. Climate input data and hydrological model simulation

3.1.1. Generating gridded daily mean temperature and daily precipitation

High-resolution gridded climate input data required for hydrological simulations are stochastically simulated using weather generator (WG) approaches proposed for daily mean temperature (Yuan et al., 2019) and daily precipitation (Yuan et al., 2021). The WGs are trained on a 1-km gridded observational data product, seNorge_2018 version 18.12, provided by the Norwegian Meteorological Institute (Lussana et al., 2019), and combined with climate change information obtained from an RCM in the EURO-CORDEX-11 ensemble (Jacob et al., 2014) operating at a 12 km spatial resolution. The RCM combines the COSMO model in CLimate Mode (CCLM; Steger and Bucchignani, 2020) with boundary conditions from the MPI Earth system model (Giorgetta et al., 2013). As shown in Fig. 1, our study area comprises 695 RCM grid cells (rectangular-like polygons) and 109 514 seNorge grid cells (within the polygons, not shown). The RCM outputs are simulated with different types of forcing before and after the year 2005, i.e. recorded and projected emissions respectively. We thus take only the historical run to avoid possible inconsistency in the model behavior due to the shift in the forcing. Both seNorge and RCM data are available from 1957. We train the WG for each variable and catchment using the seNorge data from 1957 to 1986, and derive climate change information on the RCM grid by comparing the RCM data from the two time periods 1957–1986 and 1987–2005. Simulations of each variable are then performed for the period 1987–2005 for each catchment.

The WGs employed here comprise marginal models with parameters varying smoothly across grid cells and over time, and models describing spatial and temporal dependences for realistic simulations

of daily temperature and precipitation fields. Specifically, the WG for temperature uses a Gaussian distribution with a mean and a variance component, and a stationary and separable spatio-temporal model for residuals (Yuan et al., 2019). We independently sample 10 realizations of spatial residuals and 10 realizations of temporal residuals to sufficiently account for the randomness in the model, and by adding the RCM simulated changes in the mean while keeping the variance unchanged, we obtain 100 realizations of temperature simulations (in °C).

The WG for precipitation uses a discrete-continuous distribution with precipitation occurrence described by a probit regression model, amount on wet days by a gamma distribution, and a single non-stationary Gaussian random field driving the spatio-temporal correlations in both the occurrences and the amounts (Yuan et al., 2021). The RCM simulated changes in the probability of occurrence, the amount on wet days, and the temporal correlation in the underlying spatial random field are used to update the corresponding model components of the WG. From the single random field model, we sample 10 realizations which are then transformed into 10 realizations of precipitation data (in mm day^{-1}).

In Table 1, we list the annual mean temperature and the annual precipitation averaged over the simulation period 1987–2005 for each catchment. The two larger mountain catchments, Gaulfoss and Trangen, have a relatively colder and dryer climate, while the two smaller coastal catchments, Krinsvatn and Oeyungen, are considerably warmer and wetter.

3.1.2. Simulating daily runoff with a distributed HBV-model

For hydrological simulations, we consider the spatially distributed HBV-model developed by Beldring et al. (2003) for Norway. The HBV (Hydrologiska Byråns Vattenbalansavdelning) model has undergone decades of developments and applications in Nordic countries and over the world (Bergström, 1976; Sælthun, 1996; Seibert and Bergström, 2022). The model version employed in this study has a 1 km spatial resolution and is run with a daily time step. It has an interception routine, a snow routine where snowmelt is calculated based on a simple degree-day method, a soil moisture routine where the potential evapotranspiration is estimated using a temperature index method (e.g. Sælthun, 1996; Xu and Singh, 2001), and a dynamical part where runoff is generated from an upper zone and a lower zone. Two land use classes (forests and open lands) and two soil types (wetlands and others) are used to specify the parameters of the processes above and beneath the ground surface (Beldring et al., 2003).

We calibrate and validate the model based on the seNorge data. The nonlinear parameter estimation method PEST (Doherty et al., 1994; Lawrence et al., 2009) is used to find an optimal set of parameters separately for individual catchments, where model simulated discharge is compared with the observed discharge, using the Nash–Sutcliffe efficiency (Nash and Sutcliffe, 1970) and relative volume bias (Yapo et al., 1996) as performance measures. Initial spin-up periods of six months are used to adjust the model before the calibration and validation procedures are started. The calibration period,1988–2002, shows a mean Nash–Sutcliffe efficiency of 0.80 with a mean volume bias of −0.7%, while the validation period,2003–2018, yields a mean Nash–Sutcliffe efficiency of 0.74 with a mean volume bias of 7%.

We then run the model for all possible combinations of the 100 simulated temperature realizations and the 10 simulated precipitation realizations. This yields 1000 time series of simulated daily runoff for each catchment over the period 1987–2005, i.e. time series of daily discharge averaged over the catchment area (in mm day^{-1}). As an example, Fig. 2 shows the simulated mean annual runoff cycles over the period 1987–2005. On average, the runoff is highest in the two smaller coastal catchments Krinsvatn and Oeyungen, followed by the smallest catchment Kjeldstad, consistent with the simulated precipitation amounts (cf. Table 1). In general, low flows in January–March are followed by high flows in April–June with a peak simulated in May. The flows then become more stable or have a mild increase in autumn and decline again in November–December. Krinsvatn and Oeyungen have a summer low flow preceding an autumn high flow as is characteristic for the Baltic type regime. The internal variability of the input results in considerable variability in the simulated runoff, particularly in the catchments with the highest runoff and in the snowmelt season. In addition, there appears to be a pattern of clustering between July and October in some catchments. This will be investigated further in Section 3.2.3.

3.2. Analysis of simulated runoff data

In the following, we focus on the *annual cycle* of runoff, that is, how runoff volumes vary throughout the year. We follow, for example, Finger et al. (2012) and Addor et al. (2014) and describe the *annual cycle of a year* by monthly values which are computed as the average over the simulated daily values in the respective months. Alternatively, it can be estimated by m-day moving averages (e.g. Bosshard et al., 2013) to avoid separation of the data by month. The *mean annual cycle* refers to the average annual cycle over a period of years, represented by (multi-year) mean monthly values (e.g. Fig. 2), while the *change* in the mean annual cycle refers to the difference in the mean monthly values between two periods. Here, we divide the simulations for the period 1987–2005 into two non-overlapping periods,1987–1995 and 1997–2005. We consider the mean annual cycle for each period as well as changes between the two periods. The corresponding runoff statistics (in mm day^{-1}) are computed for each of the 1000 simulations, and the following analyses are performed separately for each catchment.

3.2.1. Testing changes in the mean annual runoff cycle

To assess the significance of changes in the mean annual runoff cycle, we perform a statistical hypothesis test for each calendar month. Denote by X the mean monthly runoff in a given month and by D the simulated change between the two time periods, specifically,

$$D = X^{1997-2005} - X^{1987-1995}. \tag{1}$$

For each month, we have $n = 1000$ independent realizations of the simulated change, $D_1, \ldots, D_n$. We assume the simulated change D follows a normal distribution with mean μ_D and standard deviation σ_D. Then, the mean value $\bar{D}$ follows a normal distribution with mean μ_D and standard deviation $\sigma_D/\sqrt{n}$; formally,

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i \sim \mathcal{N}(\mu_D,\ \sigma_D/\sqrt{n}). \tag{2}$$

The objective is now to test the null hypothesis $H_0 : \mu_D = 0$ versus the alternative $H_a: \mu_D \neq 0$. Denote by S_D the sample standard deviation,

$$S_D = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(D_i - \bar{D})^2}. \tag{3}$$

We use the test statistic of the one-sample t-test, $t = \frac{\bar{D}}{S_D/\sqrt{n}}$, which is approximately standard normal for large n. We reject H_0 at the 5% level if $t \geq 1.96$ or $t \leq -1.96$.

Alternatively, it is common in the literature to use the signal-to-noise ratio (S/N) for assessing the significance of change signal. The signal S is usually the ensemble mean, however, it varies somewhat how the noise N is defined. For example, N has been calculated based on the standard deviation of a multi-member ensemble from a different climate model (e.g. Zhuan et al., 2018; Gu et al., 2019), the observed discharge records (Addor et al., 2014), the interannual variability of the sample averaged future projections (Finger et al., 2012), and based on the control scenario ensemble generated for each climate model (Fatichi et al., 2014). Moreover, the significance has been assessed at different levels. Using our notation, we can write $\frac{S}{N} = \frac{\bar{D}}{cS_D}$, where c is a constant. The constant c has been set to e.g. 1.65, 2 or 1 (Hawkins and Sutton, 2009; Zhuan et al., 2018; Gu et al., 2019).

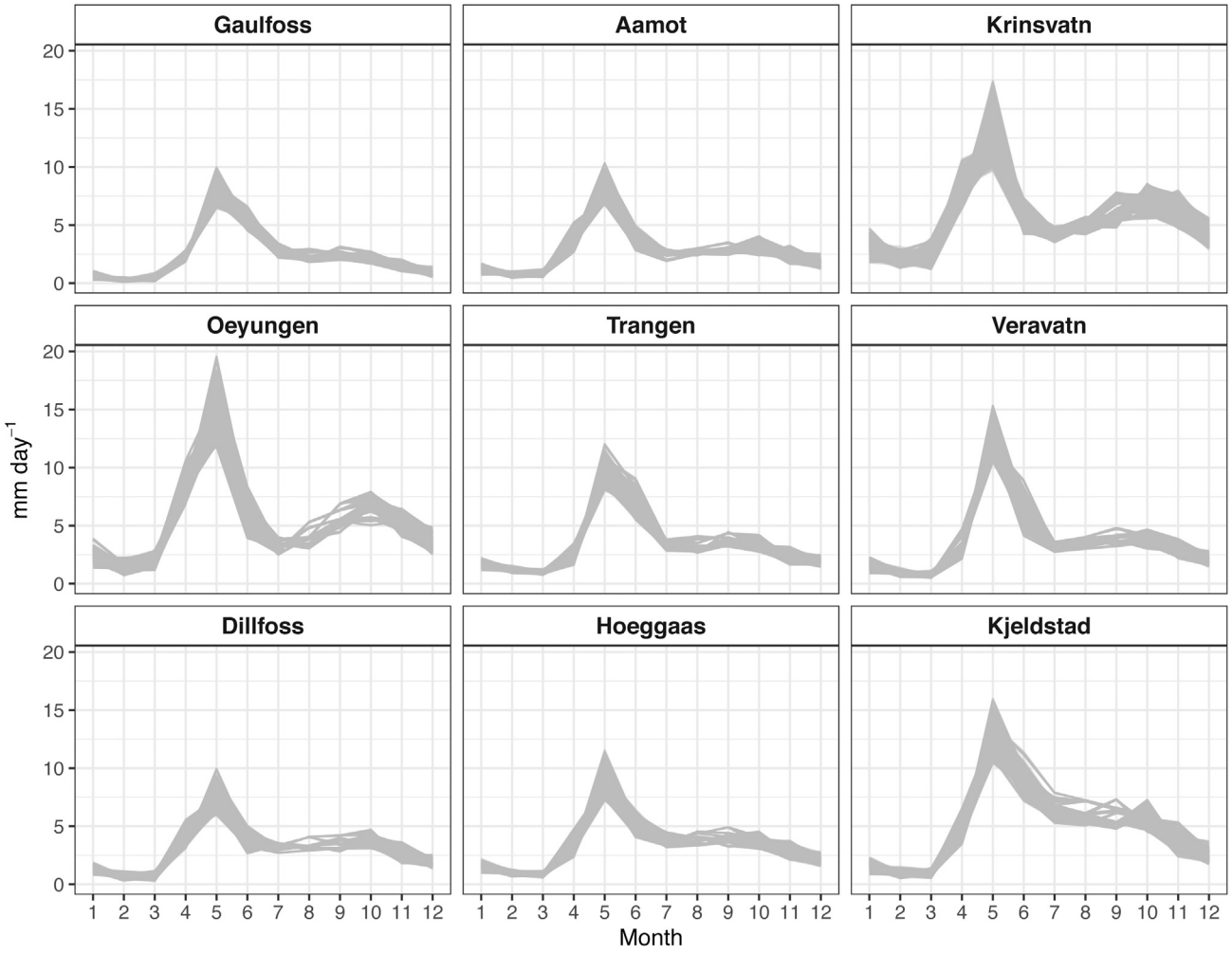

Fig. 2. 1000 simulations of the mean annual runoff cycle for the period 1987–2005. Each simulated runoff series is summarized by multi-year mean monthly averages and indicated by a thin gray line. Results are shown for all 9 catchments considered in this study.

3.2.2. *Quantifying variability due to climate inputs using ANOVA*

Variability in the runoff simulations is analyzed using analysis of variance (ANOVA). The ANOVA method has been applied to quantify different sources of variability in future climate projections (e.g. Yip et al., 2011) hydrological projections of the runoff cycle (e.g. Finger et al., 2012; Bosshard et al., 2013; Addor et al., 2014), extreme flows (e.g. Gao et al., 2020) and flood hazards (e.g. Lawrence, 2020; Meresa et al., 2021).

Here, ANOVA is applied to quantify sources of variability related to temperature and precipitation input data in the runoff simulations. For a given month, denote by X_{ij} a response driven by temperature simulation $i \in \{1, \dots, 100\}$ and precipitation simulation $j \in \{1, \dots, 10\}$. The ANOVA model equation for X_{ij} is given by

$$X_{ij} = \mu + \tau_i + \eta_j + \varepsilon_{ij}, \tag{4}$$

with $\sum_{i=1}^{100} \tau_i = 0$, $\sum_{j=1}^{10} \eta_j = 0$, and the error terms ε_{ij} are assumed independent and identically distributed. The overall mean response is given by μ, τ_i is the effect of the ith temperature simulation and η_j is the effect of the jth precipitation simulation, both given as the deviation from the overall mean.

The total variability in X_{ij} and variability due to various factors can be quantified by *sums of squares*, an approach adopted by e.g. Finger et al. (2012), Bosshard et al. (2013), Addor et al. (2014), and Gao et al. (2020). Here, we follow Yip et al. (2011) and define all sources of variability in terms of *variance*:

$$\mathrm{VT} = \frac{1}{100} \sum_{i=1}^{100} [\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}]^2, \tag{5}$$

$$\mathrm{VP} = \frac{1}{10} \sum_{j=1}^{10} [\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}]^2, \tag{6}$$

$$\mathrm{VE} = \frac{1}{1000} \sum_{j=1}^{10} \sum_{i=1}^{100} [X_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X}_{\cdot\cdot}]^2, \tag{7}$$

$$\mathrm{Vtotal} = \frac{1}{1000} \sum_{j=1}^{10} \sum_{i=1}^{100} [X_{ij} - \bar{X}_{\cdot\cdot}]^2 = \mathrm{VT} + \mathrm{VP} + \mathrm{VE}, \tag{8}$$

where e.g. $\bar{X}_{i\cdot}$ denotes the average over the index j, $\bar{X}_{i\cdot} = \frac{1}{10} \sum_{j=1}^{10} X_{ij}$. That is, the *total variability* is given by the sample variance Vtotal. It decomposes into variability due to temperature simulations, VT, precipitation simulations, VP, and an unexplained part, VE.

We apply the ANOVA framework to three types of responses related to the mean annual runoff cycle: (1) for each month, the mean monthly runoff for 1987–1995 and 1997–2005, respectively, where the sample variance is considered a measure of *natural variability* in the corresponding period; (2) for each month, the monthly runoff for each year, to study the *interannual variation* of the variability throughout the entire simulation period; and (3) for each month, the change in the mean monthly runoff between 1987–1995 and 1997–2005. Variance decomposition results will be presented in unit (mm day^{-1})^2, i.e. the square of the runoff statistics' unit.

Table 2
Average simulated changes in the mean annual runoff cycle from 1987–1995 to 1997–2005 for each catchment, given as change in the mean monthly runoff for each calendar month. Values where the null hypothesis of no change cannot be rejected are indicated in italic. For each catchment, the largest positive change and the largest negative change are indicated in bold.

Catchment	ID	1	2	3	4	5	6	7	8	9	10	11	12
Gaulfoss	A	*0.00*	−0.16	−0.23	**−0.88**	0.81	**1.21**	−0.04	0.17	0.15	−0.15	−0.19	−0.37
Aamot	B	−0.12	−0.28	−0.30	**−1.04**	**2.74**	1.60	0.38	0.17	0.24	−0.07	−0.19	−0.13
Krinsvatn	C	−0.15	−0.78	**−1.19**	−1.07	**4.06**	1.94	0.77	1.37	0.70	0.14	*−0.02*	−0.42
Oeyungen	D	−0.40	−0.75	−0.96	**−2.07**	**5.83**	3.69	0.72	1.30	0.53	0.89	*0.04*	−0.34
Trangen	E	−0.12	−0.25	−0.18	**−0.89**	0.62	**2.50**	0.75	−0.10	*0.02*	−0.29	−0.22	−0.05
Veravatn	F	−0.21	−0.34	−0.29	**−1.68**	2.09	**3.30**	0.38	0.39	0.89	0.23	−0.10	−0.14
Dillfoss	G	−0.24	−0.47	−0.46	**−1.12**	**1.95**	0.97	0.04	−0.17	0.10	−0.24	−0.60	−0.22
Hoeggaas	H	−0.28	−0.34	−0.38	**−1.47**	1.36	**1.48**	0.32	*−0.04*	*−0.01*	−0.05	−0.42	−0.33
Kjeldstad	I	−0.31	−0.45	−0.51	**−1.76**	2.54	**3.34**	1.28	0.81	0.51	−0.41	−0.41	−0.34

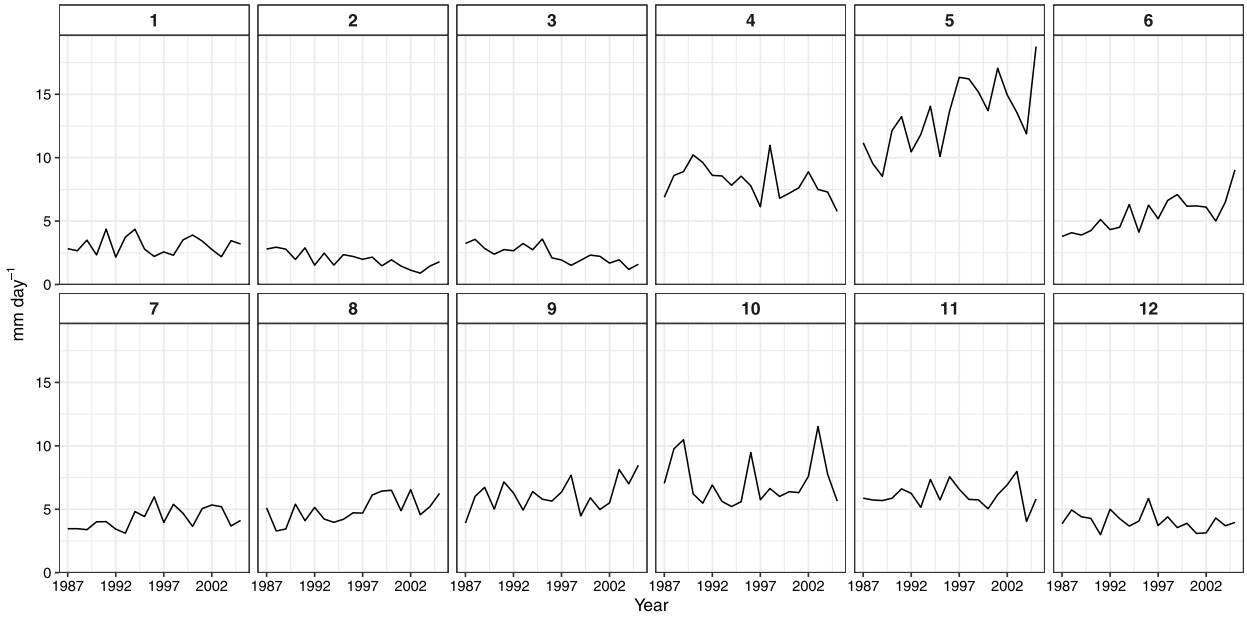

Fig. 3. Interannual variation of the sample averaged monthly runoff across the simulation period 1987–2005 shown for the catchment Krinsvatn.

3.2.3. Assessing model assumptions using bootstrap

To assess the model assumptions of the procedures described above, we additionally apply a bootstrap procedure similar to that employed by Addor et al. (2014) to assess natural variability based on observed discharge records. Specifically, we perform the following: (1) for our set of $n = 1000$ simulations, randomly select a sample of size about 10% of the total number of $n(n-1)/2$ pairs; (2) compute the difference within each selected pair, and then for each month, compute the sample standard deviation based on the difference values. These two steps are repeated 100 times to account for sampling variability, and we subsequently take the 95% bootstrap confidence interval based on the results.

Denote by $D_k = X_{1k} - X_{2k}$ for $k = 1, \ldots, m$ the differences computed in step (2) for the m randomly selected pairs $(X_{1k}, X_{2k})_{k=1}^{m}$ of step (1). The sample variance of the differences is given by

$$\begin{aligned} \mathrm{V}_D &= \frac{1}{m}\sum_{k=1}^{m}[D_k - \bar{D}]^2 \\ &= \frac{1}{m}\sum_{k=1}^{m}[(X_{1k} - \bar{X}_{1.})^2 + (X_{2k} - \bar{X}_{2.})^2 - 2(X_{1k} - \bar{X}_{1.})(X_{2k} - \bar{X}_{2.})] \qquad (9) \\ &\approx 2\,\mathrm{Vtotal}, \end{aligned}$$

where the approximation in the last step holds when m is large and the realizations in the sample are independent. The bootstrap results can thus be used to investigate possible correlations in the data. Specifically, if the 95% bootstrap confidence interval lies below $2\,\mathrm{Vtotal}$, the realizations in the sample may be positively correlated, causing e.g. systematic clustering in the runoff simulations. We perform this assessment for both the mean monthly runoff and the change in the mean monthly runoff.

4. Results

4.1. Changes in the mean annual runoff cycle

The average simulated changes in the mean annual runoff cycle from 1987–1995 to 1997–2005 are listed in Table 2, annotated with conclusions from t-tests testing the significance of the changes. The null hypothesis of no change is generally rejected except for six instances scattered among different months and catchments: January in Gaulfoss, November in Krinsvatn and Oeyungen, September in Trangen, August and September in Hoeggaas.

In addition, the pattern of changes is similar across the catchments: there is generally less runoff from October to April in the more recent time period with the largest decline mostly occurring in April, while the largest increases are simulated in May or June followed by smaller increases of runoff in July–September. Furthermore, Fig. 3 shows that in the catchment Krinsvatn, despite the interannual variation, the sample averaged monthly runoff seems to have underlying trends across the simulation period that generally reflects the change values shown in Table 2. Similar correspondence is found in the other catchments (results not shown), which means in general the changes are driven by an evolving process rather than by randomness.

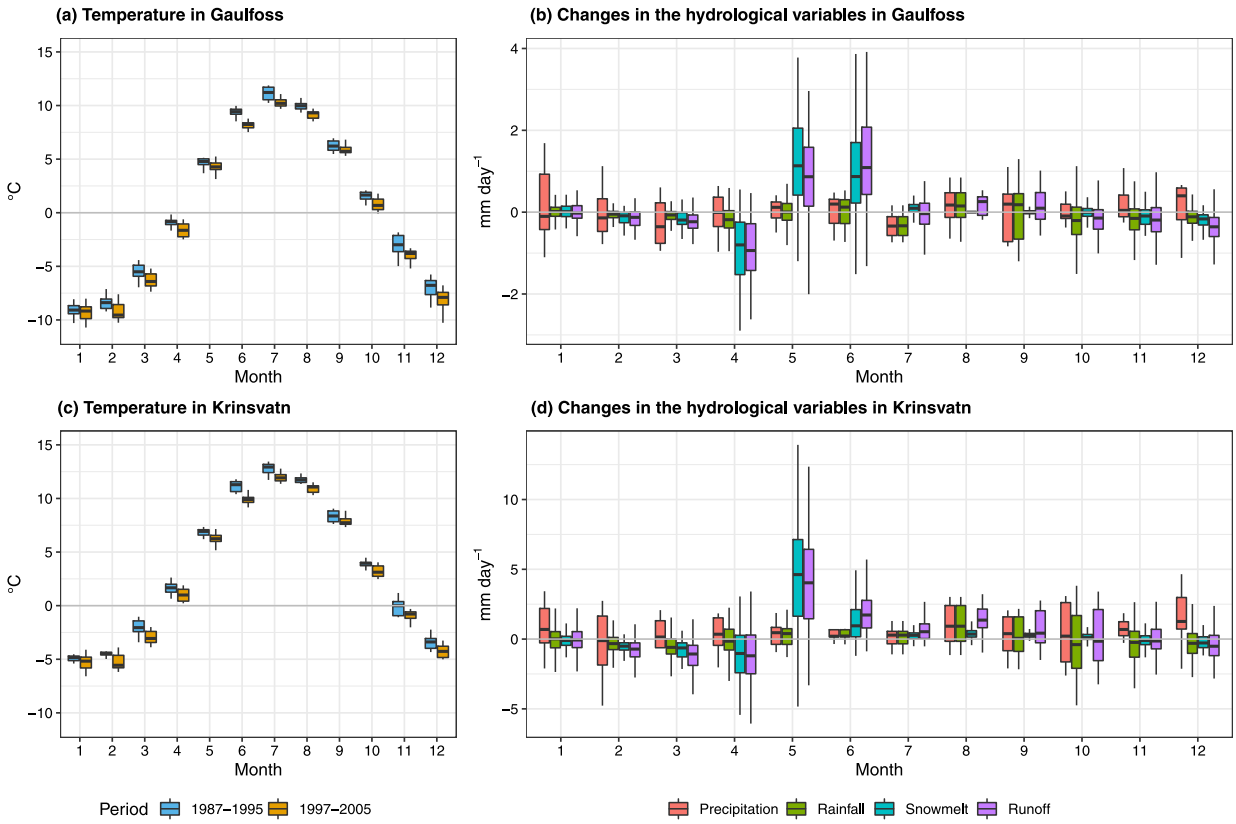

Fig. 4. Left: the simulated mean annual temperature cycle in the two periods 1987–1995 and 1997–2005 with each boxplot representing 100 values. Right: the simulated *changes* in the mean annual cycle of precipitation (not phase differentiated), rainfall (liquid), snowmelt and the resulting runoff across the two periods; for precipitation, each boxplot represents 10 values, while for the other hydrological variables each represents 1000 values. Horizontal gray lines are drawn at zero for ease of interpretation. Results are shown for the catchments Gaulfoss (top) and Krinsvatn (bottom).

To better understand the changes in the mean runoff for each month, we look at the corresponding changes in the hydroclimatological variables shown in Fig. 4 for two representative catchments. Gaulfoss has a continental climate, mountain type runoff regime and is the largest catchment (3084 km^2), whereas Krinsvatn has a maritime climate, Baltic type regime and is much smaller in size (206 km^2). For the temperature generator, the RCM indicates a negative trend over the simulation period (Yuan et al., 2019), and therefore on average a lower temperature is obtained in all months for the period 1997–2005 compared to 1987–1995 (see Fig. 4a and c). Precipitation changes are less consistent between the months and catchments (Fig. 4b and d) because the long-term trends as well as the seasonal variations in the probability of occurrence and amount on wet days are adjusted by the RCM-simulated changes which again vary slightly between the catchments (Yuan et al., 2021).

In the catchment Gaulfoss, the reduced runoff from October to April is associated with the decreasing rainfall and snowmelt (Fig. 4b), where the largest runoff decline in April is mostly driven by the snowmelt. As temperature rises above zero degree in May–June, runoff increases mainly with snowmelt, changes of which exhibit a wide spread. From July to September, the effect of snowmelt diminishes and runoff tends to vary more closely with rainfall. A similar pattern can be found in the catchment Krinsvatn, while the key difference is that Krinsvatn has a warmer climate. Temperature rises above zero during April (Fig. 4c) which triggers an earlier snowmelt than Gaulfoss, causing a larger amount of snowmelt in May with relatively less snowpack left in June, and therefore the largest increase of runoff is in May (Fig. 4d). We found that the runoff changes in the other catchments are similarly related to the changes in rainfall and snowmelt (not shown). Effects of warming on snowmelt flood timing and intensity, as well as rainfall governed changes in autumn discharge are widely discussed based on projections over a longer lead time in the future (e.g., Seiller and Anctil, 2014; Lafaysse et al., 2014).

4.2. Sources of variability in simulated runoff

ANOVA results on the variability in the simulated mean annual runoff cycle for the periods 1987–1995 and 1997–2005 are shown in Fig. 5 for the two catchments Gaulfoss and Krinsvatn. Overall, the total variability varies substantially between calendar months, it is highest around the snowmelt season in spring, followed by autumn and summer, and lowest in February. That is, the physical mechanisms of snowmelt and rainfall high flows are more sensitive to climate variability than the more inertial low flow generation mechanisms (Gelfan et al., 2015). The highest mean runoff in May–June (cf. Fig. 2) is associated mainly with snowmelt, the amount of which is highly uncertain due to uncertainty in temperature changes in that month, as well as changes in snowpack volume which has been adjusted by the temperature and precipitation from previous months until the onset of snowmelt. Therefore, the runoff variability in May–June has comparably large contributions from both temperature and precipitation, as well as other, unexplained variability.

Sources of variability follow a clear seasonal pattern. Variability in temperature is a major source of overall variability in runoff during snowmelt season and in early spring/late autumn when temperatures above or below freezing cause precipitation to fall as rain or snow, respectively. The former echos the findings of Dobler et al. (2012) that the spread of temperature projections by different GCMs leads to uncertain projections of winter and spring runoff. During summer and early autumn, rainfall becomes the primary source of runoff and variability in runoff is thus tightly linked to that in precipitation. The

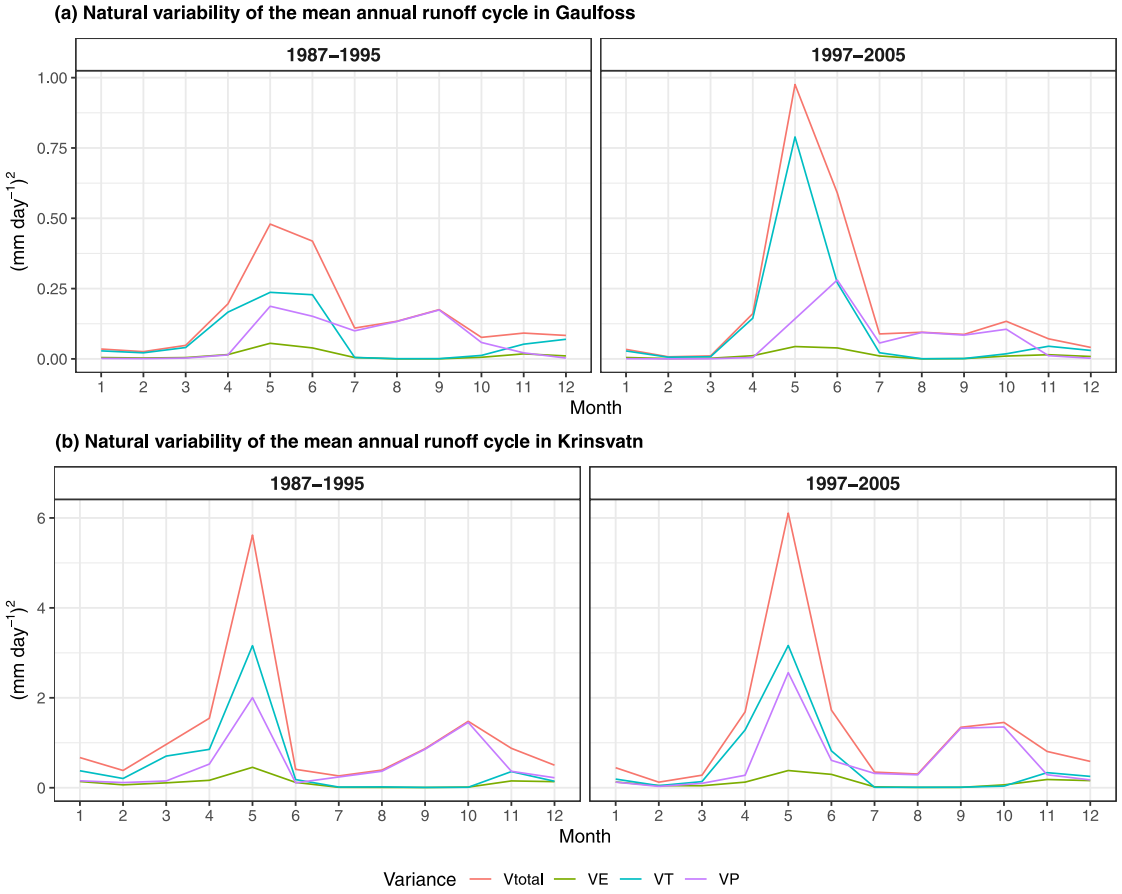

Fig. 5. Sample variance of the simulated mean annual runoff cycles analyzed using ANOVA for the two periods 1987–1995 (left) and 1997–2005 (right). The plots show total variability (Vtotal; red lines), the variability due to temperature simulations (VT; teal lines) and precipitation simulations (VP; violet lines), and the unexplained variability (VE; green lines). Results are shown for the catchments Gaulfoss (top) and Krinsvatn (bottom). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

unexplained variability is generally low compared to the other two factors.

As Fig. 5 shows, there are substantial changes in the total variability between the two time periods in Gaulfoss, represented by a large increase in May due to the temperature simulations and a smaller increase in June contributed mostly by the precipitation simulations. By contrast, the seasonal pattern in Krinsvatn seems more consistent across the two time periods, except for increases in June and September. In both catchments, the low variability in February–March further decreases. The results shown here and found in the other catchments (not shown) suggest that natural variability and the underlying sources may vary somewhat between the two time periods.

To better understand the changes in the variability over the two time periods, interannual variation of the variability in the annual runoff cycle at Krinsvatn is shown in Fig. 6. For most months, the total variability varies across the years without showing an apparent trend, which is in line with the results shown in Fig. 5. However, the increasing trend in June and September as well as the decreasing trend in February and March seem to have good correspondence with the changes found in Fig. 5. In addition, trends in the individual sources of variability match the changes shown in Fig. 5, for instance the increasing dominance of temperature in April and the increased variability due to both temperature and precipitation in June.

A distinct feature shown in Fig. 6 is that the variability in the simulations can be very high in October for individual years, although the variability in the decadal mean runoff in October is much lower than that for May (cf. Fig. 5). This is because the average variability is higher in May. In the other Baltic-type catchment Oeyungen and the catchments with a transition-type runoff regime, the simulated runoff variability in August–October fluctuates substantially between the years. It can be as high as the variability in April but always lower than variability in the snowmelt high flow season (results not shown).

Variability in the simulated changes in the mean annual runoff cycle from 1987–1995 to 1997–2005 are shown in Fig. 7 for the two catchments Gaulfoss and Krinsvatn. The patterns in variability correspond well with the patterns shown in Fig. 4. Note that the total variability in the simulated changes roughly corresponds to the sum of the total variability over the two time periods, cf. Fig. 5. This indicates that there is a low correlation between the mean runoff values of the two time periods, and stresses that ignoring the local-scale internal variability has a compound effect when projected changes are considered rather than projected values.

4.3. Assessment of model assumptions

Fig. 8 shows the results of the bootstrap procedure described in Section 3.2.3 to assess model assumptions of independent simulations of runoff. The results show that the sample standard deviation $\sqrt{\mathrm{Vtotal}}$ of the ANOVA procedure given in Eq. (8) and the standard deviation estimate $\sqrt{\mathrm{V}_D}$ based on a set of pairwise differences given in Eq. (9) possess identical seasonal patterns. Furthermore, the ratio between the two measures is generally close to $\sqrt{2}$, the expected value for independent simulations.

Specifically, the 95% bootstrap confidence interval for $\sqrt{\mathrm{V}_D}$ includes $\sqrt{2\,\mathrm{Vtotal}}$ in all months for Gaulfoss (see Fig. 8a and b). For Krinsvatn, the confidence interval is slightly lower than the expected

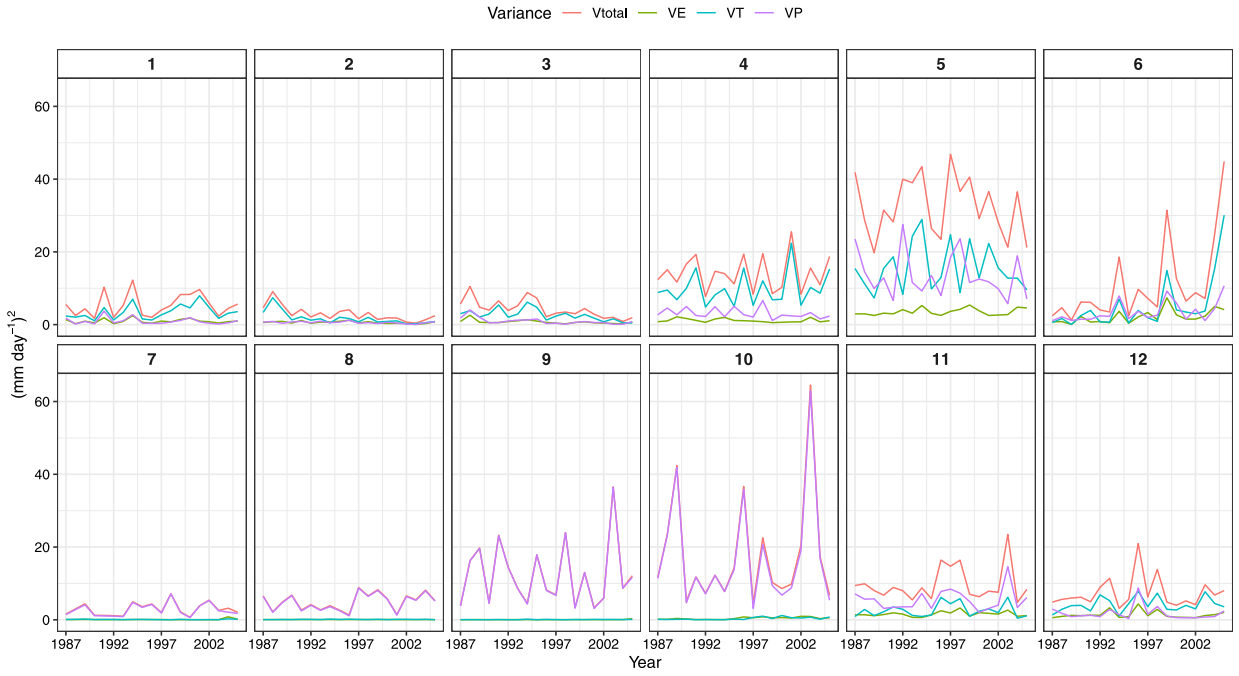

Fig. 6. Interannual variation of the variability in the annual runoff cycle over the period 1987–2005 for the catchment Krinsvatn. For each calendar month, the total variability in the monthly runoff for each year as measured by the sample variance (Vtotal; red lines) is decomposed using ANOVA into the variability due to temperature simulations (VT; teal lines) and precipitation simulations (VP; violet lines), and the unexplained variability (VE; green lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

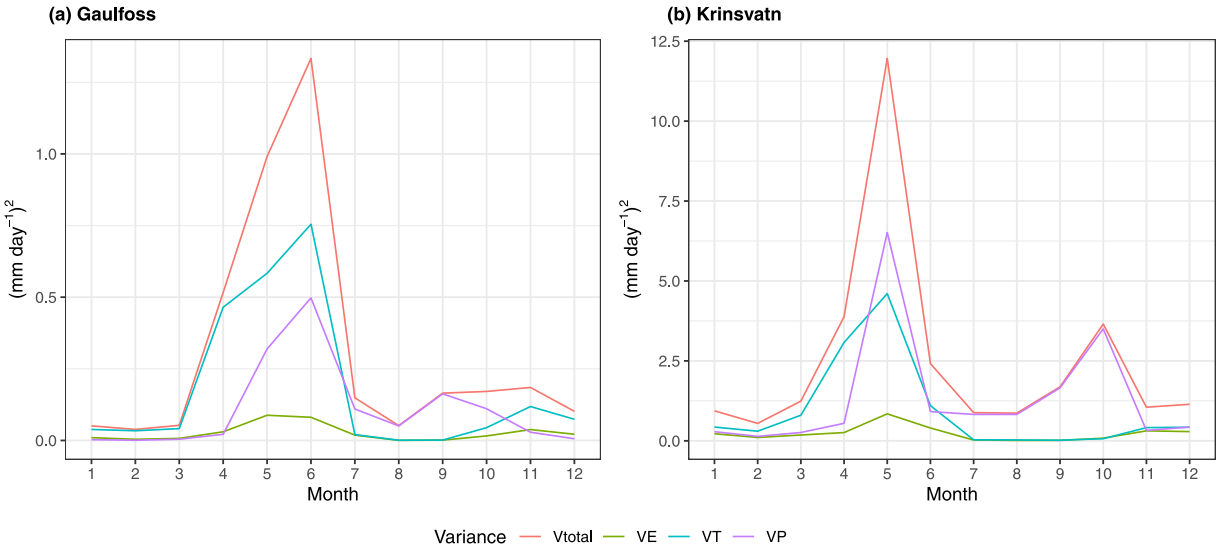

Fig. 7. Total variability in the simulated changes from 1987–1995 to 1997–2005 in mean monthly runoff as measured by the sample variance (Vtotal; red line), decomposed using ANOVA into the variability due to temperature simulations (VT; teal line) and precipitation simulations (VP; violet line), and the unexplained variability (VE; green line). Results are shown for the catchments Gaulfoss and Krinsvatn. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

value in July, September and October for the period 1987–1995 (Fig. 8d) and August and September for 1997–2005 (Fig. 8e). We found that except for the catchment Trangen (E), the confidence interval is slightly below the expected value in individual months typically between July and October (not shown), which coincides with the months exhibiting a pattern of clustering in Fig. 2. As indicated by the ANOVA results, the mean runoff of these months probably cluster around the 10 precipitation simulations. This coincides with the results of Fig. 5 that the total variability in the mean monthly runoff during this time of the year is largely dominated by the variability in the precipitation simulations.

The bootstrap procedure was additionally applied to the sample of changes in the mean monthly runoff. As shown in Fig. 8c and f, there does not seem to be significant correlations, apart from August in Krinsvatn. We found that in four of the nine catchments (A, E, F and I) no month tends to have correlated changes, whereas for the other catchments, positive correlations may be found typically between August and October (not shown). The effects of correlations in the mean monthly runoff across different simulations are thus somewhat mitigated by considering change in runoff rather than runoff itself.

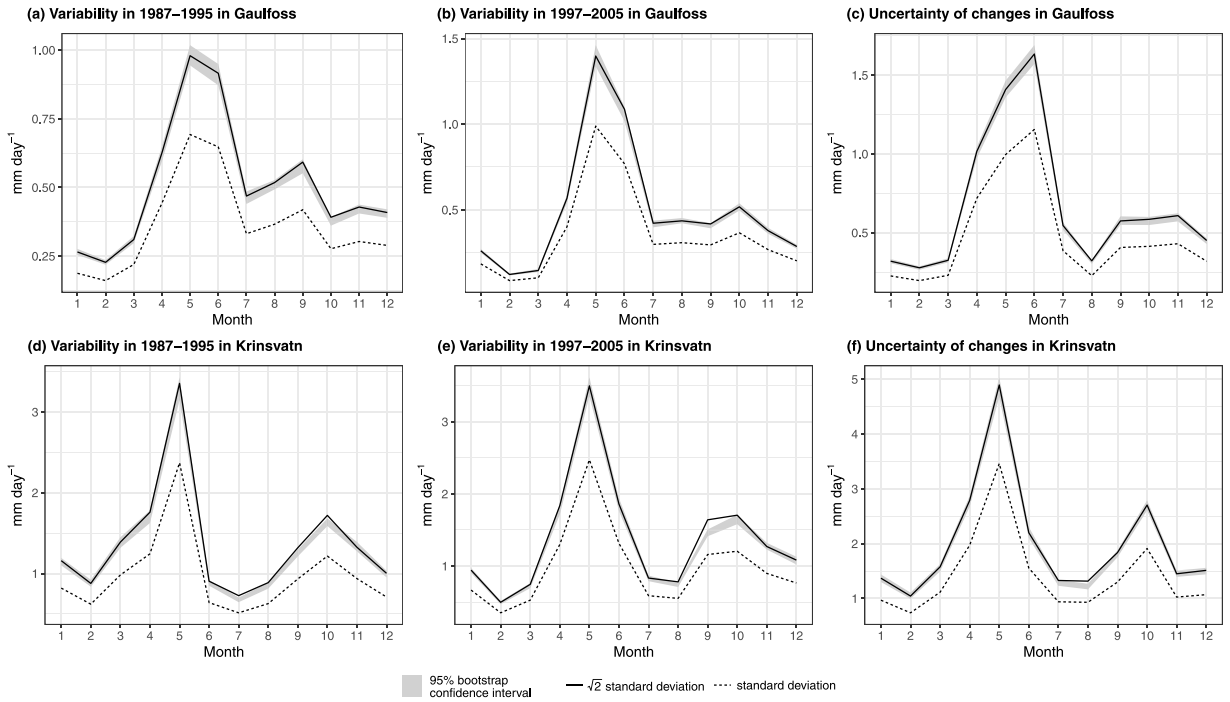

Fig. 8. Bootstrap results to assess model assumptions of independent simulations of runoff: 95% bootstrap confidence interval for the sample standard deviation $\sqrt{V_D}$ of Eq. (9) (gray bars), the sample standard deviation $\sqrt{Vtotal}$ of Eq. (8) (dotted lines) and $\sqrt{2\,Vtotal}$ (solid lines), see Section 3.2.3 for details. Results are shown for the catchments Gaulfoss (top row) and Krinsvatn (bottom row), for the mean monthly runoff in 1987–1995 (left column) and 1997–2005 (middle column), and for the change in mean monthly runoff between the two time periods (right column).

5. Discussion

5.1. Parametrization of hydrological processes in the HBV-model

The simple structure of the parameter parsimonious HBV-model is an advantage in the sense that parameters are more clearly identified, and the problem of parameter equifinality is reduced (Perrin et al., 2001; Abebe et al., 2010; Poulin et al., 2011). However, there are some limitations with simple models. If atmospheric processes such as wind speed, radiation and humidity under climate change has a strong impact on some processes, it may be leading to model parameters not being valid under non-stationary conditions (Beldring et al., 2008; Poulin et al., 2011). As an example, the temperature index approach used for estimating potential evapotranspiration may exaggerate the increase in evaporation in the future (Beldring et al., 2008). A similar argument could be used for snowmelt since the main atmospheric processes determining this process are radiation, wind speed and humidity (Dingman, 2015), whereas the HBV-model uses a temperature index approach in this case as well. Hydrological models with physically based equations could potentially have the possibility to identify changes in the significance of different processes, but on the other hand the lack of ability to identify correct parameter values for the equations describing the hydrological processes limits this possibility, in particular for sub-surface processes like soil moisture and groundwater.

5.2. Significance testing

The results of the significance tests indicated significant changes in mean monthly runoff from 1987–1995 to 1997–2005 for most months and catchments. The changes additionally show a similar seasonal patterns across the different catchments. These are not unexpected since the climate change information is derived from only one RCM, which in particular predicts a negative trend in temperature over the study period. Furthermore, the subsequent hydrological simulations do not introduce additional uncertainty as they are performed using a single optimized set of parameters. However, such detectable and consistent changes may not be common and are not expected when different RCMs, hydrological parameters/models or catchments with more heterogeneity are involved in a case study. Besides, given the large sample size here, a small change may likely have *statistical significance*, yet whether such a change has *practical significance* should be considered in context, a discussion we consider beyond the scope of the current paper.

5.3. The ANOVA framework

In this study, the simulated variability is measured by the sample variance. Applications seeking a variability measure in the same unit as the variable of interest may find a decomposition of the total standard deviation more appropriate (e.g., Addor et al., 2014; Lafaysse et al., 2014). In this case, an equation of the standard deviations may be derived from Eq. (8) using e.g. a scaling method as proposed in Appendix B of Hawkins and Sutton (2011).

The full model chain for assessing hydrological impacts of climate change includes multiple components, each associated with its own sources of uncertainty. Commonly, some components have fewer levels than the other, e.g. there may be two post-processing methods applied to outputs from eight climate models, raising a concern of uneven representation of different sources in the total uncertainty. To address this, Bosshard et al. (2013) propose a subsampling method where components with many levels are downsized to match the size of the component with fewest levels, and iteratively apply ANOVA on each subset of the full sample. This provides technical benefits when sample sizes are restricted, e.g. less than 10 (Bosshard et al., 2013). Here, there are relatively small computational costs associated with each additional precipitation or temperature sample, compared to the costs of increasing the sample size of other components of the full model

chain. Specifically, the number of simulations for temperature and precipitation differing by a factor of 10 was chosen so that each of the stochastic parts in the weather generator is equally represented.

In the current setup, interactions between temperature and precipitation cannot be separated from other unexplained variability, since only one simulation of runoff is obtained for each combination of the inputs. Throughout the study, we found that the unexplained variability is significantly smaller than that directly linked to precipitation and temperature. Nonetheless, interactions between different factors can be important (Bosshard et al., 2013; Meresa et al., 2021), and in order to quantify these interactions, the current framework has to be expanded. One option would be to produce multiple runoff simulations for each combination of inputs using equally suitable parameter sets in the hydrological model (e.g. Lawrence, 2020; Meresa et al., 2021). Accordingly, the ANOVA model Eq. (4) would be extended to include both the main effects and an interaction between temperature and precipitation. As a result, the sample variance would be decomposed into four parts, where the unexplained part in Eq. (8) would be replaced by variability due to the interaction between temperature and precipitation, and the variability due to the hydrological model parameters.

The ANOVA framework applied here is a fixed effect model, in the sense that the effects of individual climate variable simulations are treated as if they were fixed instead of being randomly drawn from a population distribution. Consequently, the analysis is a finite-population inference through a variance decomposition of an existing sample. However, theoretically, the weather generators provide an opportunity of generating more realizations if needed, which corresponds to a superpopulation of effects. Therefore, when there is a need to consider the uncertainty of having a different set of inputs realizations, methods of superpopulation inferences (Gelman, 2005) are more suitable. See also Northrop and Chandler (2014) for an application to climate projections.

The current analysis is performed for the annual runoff cycle, while other statistics such as extreme runoff indicators, or other hydrological fluxes and states e.g. actual evapotranspiration, soil moisture deficits, could be analyzed in a similar manner. While the application here is focused on climate change impacts on regional hydrology, the methods provide a general approach to uncertainty investigation for various time horizons. Other promising applications include seasonal to decadal water resources prediction on the basis of weather and climate predictions. The computationally efficient downscaling methods of temperature and precipitation are particularly useful where large ensembles are required in order to achieve prediction skill and a reliable quantification of sources of variability.

6. Conclusions

This paper carries out a study to simulate and analyze uncertainties in hydrological impact studies that arise from the local-scale internal variability of the climate. To represent that at the catchment scale, we simulate multiple sets of gridded daily precipitation and mean temperature data at 1 km resolution using a weather generator (WG) approach. The WG uses local information from a high-resolution observational data product and climate change information from an RCM. A large set of daily runoff series are then generated based on the temperature and precipitation simulations using a spatially distributed hydrological model. The analysis is focused on the annual runoff cycle of nine catchments in central Norway, with different runoff regimes, climatic types and areas ranging from 144 to 3084 km^2. Significance tests are applied to assess changes in the mean monthly runoff from 1987–1995 to 1997–2005, and sources of variability in the runoff simulations are quantified using analysis of variance (ANOVA) where the total variability is decomposed into contributions from the two input variables as well as an unexplained component. Model assumptions of independence are assessed using a bootstrap procedure.

The change signal in the mean annual runoff cycle is found to be generally significant, indicating that the trends imposed on the WG are not overwhelmed by the simulated internal variability. Variability in the runoff simulations is highest in high flow periods, in the snowmelt season between April and June, and in autumn for some catchments. While the simulated runoff variability in autumn varies substantially between the years, the average level of variability is lower than the snowmelt season. Temperature is the main source of variability in the colder months (November–March), both temperature and precipitation contribute to the variability in the snowmelt season, and variability in July–October is largely due to that in precipitation. The analysis assumes that the individual simulations are independent, even if they may share either a precipitation input or a temperature input. An assessment of this assumption reveals that the simulations are largely independent, with some simulations showing a low degree of correlation between August and October when the variability in runoff is largely dominated by that in precipitation only.

A reliable estimation of the changes in annual runoff cycle and the associated uncertainties can be applied for climate adaptation measures in water management. For example, hydropower companies will need to adapt planning to a changed seasonal pattern of runoff volumes under a changing climate. In Norway, reservoir operation commonly needs to adjust the water levels properly before the snowmelt in order to avoid the spring flood and also meet requirements for environmental low-flow and many other water use purposes during summer. A delayed, more concentrated snowmelt that is accompanied by an increased level of variability, as simulated for some catchments in our case study, may pose challenges such as accommodating to a prolonged low-flow period in early spring and dealing with an increased difficulty in adjusting the water levels such that both the spring flood and possible deficits in summer could be avoided.

This study considers variability in runoff simulations caused by the local-scale internal variability in temperature and precipitation. The results are in line with findings where the large-scale internal climate variability is represented by a multi-member GCM ensemble (Gelfan et al., 2015) or a GCM/RCM chain Seiller and Anctil (2014). This indicates that ignoring the local-scale internal variability may yield an underestimation of the overall variability in projections and projected changes, which may further affect the adaptation strategies not only for the changes in future water resources but also for water-related hazards (e.g. Lawrence, 2020; Zhao et al., 2022).

CRediT authorship contribution statement

Qifen Yuan: Conceptualization, Methodology, Data simulation, Analysis, Visualization, Writing – original draft. **Thordis L. Thorarinsdottir:** Conceptualization, Methodology, Writing – review & editing. **Stein Beldring:** Conceptualization, Methodology, Model calibration, Writing – review & editing. **Wai Kwok Wong:** Model calibration, Visualization, Writing – review & editing. **Chong-Yu Xu:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Research Council of Norway through project no. 255517 “Post-processing Climate Projection Output for Key Users in Norway”. We thank the editor Emmanouil Anagnostou and the two anonymous reviewers for their suggestions for improving the manuscript. We thank Irene Brox Nilsen from Norwegian Water Resources and Energy Directorate (NVE) for her insightful comments on the revision.

References

Abebe, N.A., Ogden, F.L., Pradhan, N.R., 2010. Sensitivity and uncertainty analysis of the conceptual HBV rainfall–runoff model: Implications for parameter estimation. J. Hydrol. 389 (3–4), 301–310. http://dx.doi.org/10.1016/j.jhydrol.2010.06.007.

Abhishek, Kinouchi, T., 2021. Synergetic application of GRACE gravity data, global hydrological model, and in-situ observations to quantify water storage dynamics over Peninsular India during 2002–2017. J. Hydrol. 596, 126069. http://dx.doi.org/10.1016/j.jhydrol.2021.126069.

Abhishek, Kinouchi, T., Sayama, T., 2021. A comprehensive assessment of water storage dynamics and hydroclimatic extremes in the Chao Phraya River Basin during 2002–2020. J. Hydrol. 603, 126868. http://dx.doi.org/10.1016/j.jhydrol.2021.126868.

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., Seibert, J., 2014. Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. Water Resour. Res. 50 (10), 7541–7562. http://dx.doi.org/10.1002/2014wr015549.

Barros, V.R., Field, C.B., Dokken, D.J., Mastrandrea, M.D., Mach, K.J., Bilir, T.E., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., White, L.L. (Eds.), 2014. IPCC, 2014: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Beldring, S., Engeland, K., Roald, L.A., Sælthun, N.R., Voksø, A., 2003. Estimation of parameters in a distributed precipitation-runoff model for Norway. Hydrol. Earth Syst. Sci. 7 (3), 304–316.

Beldring, S., Engen-Skaugen, T., Førland, E.J., Roald, L.A., 2008. Climate change impacts on hydrological processes in Norway based on two methods for transferring regional climate model results to meteorological station sites. Tellus A 60 (3), 439–450. http://dx.doi.org/10.1111/j.1600-0870.2007.00306.x.

Bergström, S., 1976. Development and Application of a Conceptual Runoff Model for Scandinavian Catchments. (RHO 7), Swedish Meteorological and Hydrological Institute, Norrköping Sweden.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., Schär, C., 2013. Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. Water Resour. Res. 49 (3), 1523–1536. http://dx.doi.org/10.1029/2011wr011533.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W.J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A.J., Wehner, M., 2013. Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013-the Physical Science Basis: Contribution of Working Group I To the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1029–1136.

Dingman, S.L., 2015. Physical Hydrology, third ed. Waveland Press, p. 643.

Dobler, C., Hagemann, S., Wilby, R.L., Stötter, J., 2012. Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed. Hydrol. Earth Syst. Sci. 16 (11), 4343–4360. http://dx.doi.org/10.5194/hess-16-4343-2012.

Doherty, J., Brebber, L., Whyte, P., 1994. PEST: Model-independent parameter estimation. In: Watermark Computing, Corinda, Australia, Vol. 122. p. 336.

Fatichi, S., Rimkus, S., Burlando, P., Bordoy, R., 2014. Does internal climate variability overwhelm climate change signals in streamflow? The upper Po and Rhone basin case studies. Sci. Total Environ. 493, 1171–1182. http://dx.doi.org/10.1016/j.scitotenv.2013.12.014.

Finger, D., Heinrich, G., Gobiet, A., Bauder, A., 2012. Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. Water Resour. Res. 48 (2), http://dx.doi.org/10.1029/2011wr010733.

Gao, C., Booij, M.J., Xu, Y.-P., 2020. Assessment of extreme flows and uncertainty under climate change: disentangling the uncertainty contribution of representative concentration pathways, global climate models and internal climate variability. Hydrol. Earth Syst. Sci. 24 (6), 3251–3269. http://dx.doi.org/10.5194/hess-24-3251-2020.

Gelfan, A., Semenov, V.A., Gusev, E., Motovilov, Y., Nasonova, O., Krylenko, I., Kovalev, E., 2015. Large-basin hydrological response to climate model outputs: uncertainty caused by internal atmospheric variability. Hydrol. Earth Syst. Sci. 19 (6), 2737–2754. http://dx.doi.org/10.5194/hess-19-2737-2015.

Gelman, A., 2005. Analysis of variance—why it is more important than ever. Ann. Statist. 33 (1), http://dx.doi.org/10.1214/009053604000001048.

Ghil, M., Lucarini, V., 2020. The physics of climate variability and climate change. Rev. Modern Phys. 92 (3), http://dx.doi.org/10.1103/revmodphys.92.035002.

Giorgetta, M.A., Jungclaus, J., Reick, C.H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K.D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., Stevens, B., 2013. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model intercomparison project phase 5. J. Adv. Model. Earth Syst. 5 (3), 572–597. http://dx.doi.org/10.1002/jame.20038.

Gottschalk, L., Jensen, J.L., Lundquist, D., Solantie, R., Tollan, A., 1979. Hydrologic Regions in the Nordic Countries. Hydrol. Res. 10 (5), 273–286.

Gu, L., Chen, J., Xu, C.-Y., Kim, J.-S., Chen, H., Xia, J., Zhang, L., 2019. The contribution of internal climate variability to climate change impacts on droughts. Sci. Total Environ. 684, 229–246. http://dx.doi.org/10.1016/j.scitotenv.2019.05.345.

Hanssen-Bauer, I., Førland, E.J., Haddeland, I., Hisdal, H., Lawrence, D., Mayer, S., Nesje, A., Nilsen, J.E.Ø., Sandven, S., Sandø, A.B., Sorteberg, A., Ådlandsvik, B., et al., 2017a. Climate in Norway 2100 – a knowledge base for climate adaptation. Report 1, Norwegian Centre for Climate Services (NCCS), https://klimaservicesenter.no/kss/rapporter/kin2100.

Hanssen-Bauer, I., Hygen, H.O., Heiberg, H., Førland, E.J., Nordskog, B., 2017b. User needs for post processed climate data – a survey of the needs for output from the research project PostClim. Report 2, Norwegian Centre for Climate Services (NCCS), https://klimaservicesenter.no.

Hawkins, E., Sutton, R., 2009. The Potential to Narrow Uncertainty in Regional Climate Predictions. Bull. Am. Meteorol. Soc. 90 (8), 1095–1108. http://dx.doi.org/10.1175/2009bams2607.1.

Hawkins, E., Sutton, R., 2011. The potential to narrow uncertainty in projections of regional precipitation change. Clim. Dyn. 37 (1), 407–418.

Hingray, B., Blanchet, J., Evin, G., Vidal, J.-P., 2019. Uncertainty component estimates in transient climate projections. Clim. Dyn. 53 (5–6), 2501–2516. http://dx.doi.org/10.1007/s00382-019-04635-1.

Hingray, B., Saïd, M., 2014. Partitioning Internal Variability and Model Uncertainty Components in a Multimember Multimodel Ensemble of Climate Projections. J. Clim. 27 (17), 6779–6798. http://dx.doi.org/10.1175/jcli-d-13-00629.1.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., Yiou, P., 2014. EURO-CORDEX: new high-resolution climate change projections for European impact research. Reg. Environ. Change 14 (2), 563–578. http://dx.doi.org/10.1007/s10113-013-0499-2.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World Map of the Köppen-Geiger climate classification updated. Meteor. Z. 15 (3), 259–263. http://dx.doi.org/10.1127/0941-2948/2006/0130.

Lafaysse, M., Hingray, B., Mezghani, A., Gailhard, J., Terray, L., 2014. Internal variability and model uncertainty components in future hydrometeorological projections: The Alpine Durance basin. Water Resour. Res. 50 (4), 3317–3341. http://dx.doi.org/10.1002/2013wr014897.

Lawrence, D., 2020. Uncertainty introduced by flood frequency analysis in projections for changes in flood magnitudes under a future climate in Norway. J. Hydrol. Reg. Stud. 28, 100675. http://dx.doi.org/10.1016/j.ejrh.2020.100675.

Lawrence, D., Haddeland, I., Langsholt, E., 2009. Calibration of HBV hydrological models using PEST parameter estimation. Report 1, Norwegian Water Resources and Energy Directorate (NVE).

Lind, P., Belušić, D., Christensen, O.B., Dobler, A., Kjellström, E., Landgren, O., Lindstedt, D., Matte, D., Pedersen, R.A., Toivonen, E., Wang, F., 2020. Benefits and added value of convection-permitting climate modeling over Fenno-Scandinavia. Clim. Dyn. 1893–1912. http://dx.doi.org/10.1007/s00382-020-05359-3.

Lussana, C., Tveito, O.E., Dobler, A., Tunheim, K., 2019. seNorge_2018, daily precipitation and temperature datasets over Norway. Earth Syst. Sci. Data http://dx.doi.org/10.5194/essd-2019-43.

Maraun, D., Widmann, M., 2018. Statistical Downscaling and Bias Correction for Climate Research. Cambridge University Press.

Meresa, H., Murphy, C., Fealy, R., Golian, S., 2021. Uncertainties and their interaction in flood hazard assessment with climate change. Hydrol. Earth Syst. Sci. 25 (9), 5237–5257. http://dx.doi.org/10.5194/hess-25-5237-2021.

Nash, J., Sutcliffe, J., 1970. River flow forecasting through conceptual models part I — a discussion of principles. J. Hydrol. 10 (3), 282–290. http://dx.doi.org/10.1016/0022-1694(70)90255-6.

Northrop, P.J., Chandler, R.E., 2014. Quantifying Sources of Uncertainty in Projections of Future Climate. J. Clim. 27 (23), 8793–8808. http://dx.doi.org/10.1175/jcli-d-14-00265.1.

Peel, M.C., Srikanthan, R., McMahon, T.A., Karoly, D.J., 2015. Approximating uncertainty of annual runoff and reservoir yield using stochastic replicates of global climate model data. Hydrol. Earth Syst. Sci. 19 (4), 1615–1639. http://dx.doi.org/10.5194/hess-19-1615-2015.

Perrin, C., Michel, C., Andréassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. J. Hydrol. 242 (3), 275–301. http://dx.doi.org/10.1016/S0022-1694(00)00393-0.

Poulin, A., Brissette, F., Leconte, R., Arsenault, R., Malo, J.-S., 2011. Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin. J. Hydrol. 409 (3–4), 626–636. http://dx.doi.org/10.1016/j.jhydrol.2011.08.057.

Prein, A.F., Rasmussen, R., Castro, C.L., Dai, A., Minder, J., 2020. Special issue: Advances in convection-permitting climate modeling. Clim. Dyn. 55 (1–2), 1–2. http://dx.doi.org/10.1007/s00382-020-05240-3.

Sælthun, N.R., 1996. The "nordic" HBV model – description and documentation of the model version developed for the project Climate Change and Energy Production. Report 7, Norwegian Water Resources and Energy Directorate (NVE).

Seibert, J., Bergström, S., 2022. A retrospective on hydrological catchment modelling based on half a century with the HBV model. Hydrol. Earth Syst. Sci. 26 (5), 1371–1388. http://dx.doi.org/10.5194/hess-26-1371-2022.

Seiller, G., Anctil, F., 2014. Climate change impacts on the hydrologic regime of a Canadian river: comparing uncertainties arising from climate natural variability and lumped hydrological model structures. Hydrol. Earth Syst. Sci. 18 (6), 2033–2047. http://dx.doi.org/10.5194/hess-18-2033-2014.

Steger, C., Bucchignani, E., 2020. Regional Climate Modelling with COSMO-CLM: History and Perspectives. Atmosphere 11 (11), 1250. http://dx.doi.org/10.3390/atmos11111250.

von Storch, H., Omstedt, A., Pawlak, J., Reckermann, M., 2015. Introduction and summary. In: Second Assessment of Climate Change for the Baltic Sea Basin. Springer, pp. 1–22.

Vidal, J.-P., Hingray, B., Magand, C., Sauquet, E., Ducharne, A., 2016. Hierarchy of climate and hydrological uncertainties in transient low-flow projections. Hydrol. Earth Syst. Sci. 20 (9), 3651–3672. http://dx.doi.org/10.5194/hess-20-3651-2016.

Wang, H.-M., Chen, J., Xu, C.-Y., Zhang, J., Chen, H., 2020. A Framework to Quantify the Uncertainty Contribution of GCMs Over Multiple Sources in Hydrological Impacts of Climate Change. Earth's Future 8 (8), http://dx.doi.org/10.1029/2020ef001602.

Wilby, R.L., Harris, I., 2006. A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. Water Resour. Res. 42 (2), http://dx.doi.org/10.1029/2005wr004065.

Wong, W.K., Beldring, S., Engen-Skaugen, T., Haddeland, I., Hisdal, H., 2011. Climate Change Effects on Spatiotemporal Patterns of Hydroclimatological Summer Droughts in Norway. J. Hydrometeor. 12 (6), 1205–1220. http://dx.doi.org/10.1175/2011jhm1357.1.

Xu, C.-Y., Singh, V.P., 2001. Evaluation and generalization of temperature-based methods for calculating evaporation. Hydrol. Process. 15 (2), 305–319. http://dx.doi.org/10.1002/hyp.119.

Yapo, P.O., Gupta, H.V., Sorooshian, S., 1996. Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. J. Hydrol. 181 (1–4), 23–48. http://dx.doi.org/10.1016/0022-1694(95)02918-4.

Yip, S., Ferro, C.A.T., Stephenson, D.B., Hawkins, E., 2011. A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions. J. Clim. 24 (17), 4634–4643. http://dx.doi.org/10.1175/2011jcli4085.1.

Yuan, Q., Thorarinsdottir, T.L., Beldring, S., Wong, W.K., Huang, S., Xu, C.-Y., 2019. New Approach for Bias Correction and Stochastic Downscaling of Future Projections for Daily Mean Temperatures to a High-Resolution Grid. J. Appl. Meteorol. Climatol. 58 (12), 2617–2632. http://dx.doi.org/10.1175/jamc-d-19-0086.1.

Yuan, Q., Thorarinsdottir, T.L., Beldring, S., Wong, W.K., Xu, C.-Y., 2021. Bridging the scale gap: obtaining high-resolution stochastic simulations of gridded daily precipitation in a future climate. Hydrol. Earth Syst. Sci. 25 (9), 5259–5275. http://dx.doi.org/10.5194/hess-25-5259-2021.

Zhao, W., Abhishek, Kinouchi, T., Ang, R., Zhuang, Q., 2022. A framework for quantifying climate-informed heavy rainfall change: Implications for adaptation strategies. Sci. Total Environ. 835, 155553. http://dx.doi.org/10.1016/j.scitotenv.2022.155553.

Zhuan, M.-J., Chen, J., Shen, M.-X., Xu, C.-Y., Chen, H., Xiong, L.-H., 2018. Timing of human-induced climate change emergence from internal climate variability for hydrological impact studies. Hydrol. Res. 49 (2), 421–437. http://dx.doi.org/10.2166/nh.2018.059.