Journal of Hydrology 670 (2026) 135228

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Uncertainty-aware reservoir operation projections using multi-model weighting and adaptive hedging rules

Zeynep Beril Ersoy [a,b], Okan Fistikoglu [a], Umut Okkan [b,*,1]

[a] Department of Civil Engineering, The Graduate School of Natural and Applied Sciences Dokuz Eylul University, Izmir, Turkey
[b] Department of Civil Engineering, Hydraulic Division, Balikesir University, Balikesir, Turkey

ARTICLE INFO

This manuscript was handled by Y. Huang, Editor-in-Chief, with the assistance of Jing-Cheng Han, Associate Editor

Keywords:
Reservoir operation
Adaptive hedging
Vulnerability
Uncertainty reduction
Dynamically weighted ensembles
Climate change
Tahtali Reservoir

ABSTRACT

While reservoirs designed based on historical records are increasingly vulnerable to inflow variability driven by climate change, conventional rules such as the standard operating policy (SOP) often amplify projection uncertainties by failing to anticipate future shortages. To tackle these challenges, this study adopts an integrative, uncertainty-aware reservoir operation framework that jointly addresses upstream inflow projection uncertainty and downstream operational uncertainty, representing a research gap not yet explored. Specifically, a dynamic weighting scheme, Uncertainty Optimizing Multi-Model Ensemble (UO-MME), was used for credible inflow projections, which were then incorporated into a two-dimensional hedging model (HDG-2d) to optimize reservoir releases. The methodology was applied to the Tahtali Reservoir (western Türkiye) considering a modeling chain comprising five GCMs, two emission scenarios, two downscaling methods, and seven hydrological models. The climate inputs needed by hydrological models and UO-MME include bias-corrected CORDEX outputs and statistically downscaled CMIP5 data covering the projection span 2021–2099. The inflow projections from different variants were used to operate both SOP and HDG-2d models. Compared with SOP operated under unweighted data, where projection uncertainties tend to accumulate, HDG-2d emerged as the main uncertainty-reducing mechanism for reservoir outputs, while UO-MME played a supportive role. Overall, the combined strategy achieves ensemble consistency in most variants, with dimensionless vulnerability remaining below 0.25, reduces its related uncertainty by 23% through mitigating single-period shortages, and further narrows projected release and sustainability index uncertainties by 43% and 36%, respectively. Combining hydrological model weighting with climate-informed hedging is shown to provide a more robust basis for reservoir planning.

1. Introduction

Climate change gradually challenges the reservoir sustainability, and reservoirs designed from past hydrological records have become vulnerable under changing climatological regimes and streamflow dynamics (Eum and Simonovic, 2010; Adeloye and Dau, 2019). Especially while the standard operating policy (SOP) providing a simple operating rule may be somewhat convenient under stationary conditions, it may lead to projecting insufficient reservoir storage levels under climate change scenarios since it releases water up to the target demand without any reserving for future use (Gorguner and Kavvas, 2020; Nguyen et al., 2020). Hence, the fact that the ability of reservoirs to achieve

Abbreviations: ANOVA, Analysis of Variance; Awbm, Australian Water Balance Model; BMA, Bayesian Model Averaging; CMIP5, Coupled Model Inter-comparison Project Phase 5; CORDEX, Coordinated Regional Climate Downscaling Experiment; DE, Differential Evolution; DM, Downscaling Method; Dynwbm, Dynamic Water Balance Model; ECMWF, European Centre for Medium-Range Weather Forecasts; En-BMA, Entropy-Based BMA; ES, Emission Scenario; GCM, General Circulation Model; HDG-2d, Two-Dimensional Hedging Rule; HIST, Reference Scenario; HM, Hydrological Model; KMRB, Kucuk Menderes River Basin; LNSE, Logarithmic Nash–Sutcliffe Efficiency; MENA, Middle East and North Africa; mTUV, Long-Term Mean Total Projection Uncertainty; NSE, Nash-Sutcliffe Efficiency; NSGA-II, Non-Dominated Sorted Genetic Algorithm II; PM, Penman–Monteith; QDM, Quantile Delta Mapping; RBFN, Radial Basis Function Network; RCP, Representative Concentration Pathway; RLS, Release; RPM, Reservoir Performance Measure; RT, Time-Based Reliability Index; RV, Volumetric Reliability Index; SI, Sustainability Index; SOP, Standard Operating Policy; TINT, Total Interaction; Twbm, Thornthwaite Water Balance Model; UO-MME, Uncertainty Optimizing Multi-Model Ensemble; VUL, Vulnerability Index.

* Corresponding author.
E-mail address: umutokkan@balikesir.edu.tr (U. Okkan).
[1] http://orcid.org/0000-0003-1284-3825.

https://doi.org/10.1016/j.jhydrol.2026.135228
Received 19 August 2025; Received in revised form 5 December 2025; Accepted 27 February 2026
Available online 2 March 2026

operational objectives with SOP under changing conditions may markedly deteriorate entail a shift toward climate change adaptation-oriented reservoir operations. This need is particularly important in cases where reservoir performances are widely sensitive to hydroclimatic changes (Mateus and Tullos, 2017) and in regions where operational objectives must be carefully managed under uncertain future scenarios (Thomas et al., 2021).

One of the ways to decrease infrastructure construction or rehabilitation costs in water resources management under adaptation to climate change is to update reservoir operating curves through optimization models (Elgendy et al., 2024). In this context, certain reservoir operation policies, known as hedging rules, have been formulated to restrict immediate water releases, guided by a parameterized rule curve, in order to mitigate the risk of severe deficits during future periods (Tu et al., 2008; Celeste and Billib, 2009). It has already been shown that hedging models, which are optimized with observed or stochastically generated inflow data, can clearly provide a more dynamic approach to reservoir operation compared to the SOP (e.g., Bayesteh and Azari, 2021). It has also been revealed that these adaptive rules can effectively mitigate drought impacts by preventing complete reservoir depletion during multiyear critical periods (e.g., Luo et al., 2023; Anvari et al., 2023). However, only a limited number of studies have focused on developing hedging strategies based on global climate model (GCM)-driven projections, particularly with the goal of maintaining reliability and tempering system vulnerability in the long term (e.g., Adeloye et al., 2016; Adeloye and Dau, 2019; Ahmadianfar and Zamani, 2020; Beshavard et al., 2022; Moghaddasi et al., 2022; Okkan et al., 2023; Thiha et al., 2023; Erfanian et al., 2023).

In addition, as highlighted in recent studies, reservoir operation models are rather responsive to uncertainties in inflow predictions. It has been shown that disregarding these uncertainties may affect key performance indices and complicate informed decision-making (Soleimani et al., 2016; Nourani et al., 2025). Moreover, Liu et al. (2019) emphasized that the uncertainty associated with predictive inflow data on reservoir water level forecasts can outweigh that of rule curve parameters derived through a Bayesian deep learning method. They also showed that employing probabilistic inflow forecasts for model training can reduce uncertainty related to reservoir operation decisions. However, given that the aforementioned studies were not conducted under GCM-driven simulations, it can be argued that the combined uncertainties originating from different sources used in the modeling chain, such as multiple GCMs, emission scenarios (ESs), downscaling methods (DMs), and hydrological models (HMs), tend to propagate differently within the ensemble of hydrological projections. More specifically, the uncertainty level rises progressively at each stage of the impact modeling chain, since each component builds upon the outputs of the preceding one. In particular, studies that decompose uncertainty in streamflow projections displayed that GCMs, along with their interactions with ESs, appeared as important sources of the uncertainty (Vetter et al., 2015; Wang et al., 2020). These hydrological projection uncertainties undoubtedly influence hedging rule parameterization and also bring forth the need to measure how different sources of uncertainty contribute to the uncertainty in projected reservoir operations.

Considering that improving the credibility of hydrological projections can play a pivotal role in reducing uncertainty associated with future operational strategies based on hedging policies, the use of an ensemble of weighted models may represent one of the effective options. The virtue of employing such ensembles is that they can compensate for the flaws of individual models, including structural limitations and potential biases arising from parameterization choices or climate inputs (Pastén-Zapata et al., 2022). From the perspective of climate modeling, assigning different weights to GCMs according to certain performance metrics can build a more uncertainty-aware framework for climate change impact assessments (e.g., Wanders and Wood, 2016; Knutti et al., 2017). However, such approaches are not yet fully adopted by the climate modeling community because all GCMs designed to generate large-scale climate processes are considered to have the same probability of occurrence (Collins, 2017). In contrast, HMs are often applied following automatic calibration stages to better simulate hydrological processes at a specific catchment scale. Besides, the uncertainties inherited from downscaled GCM simulations accumulate nonlinearly in hydrological response by means of HMs. Therefore, weighting HMs may offer a more appropriate means of boosting the credibility of projected outputs (Pastén-Zapata et al., 2022; Castaneda-Gonzalez et al., 2023).

As seen in existing studies, weights are generally assigned to HMs according to their historical performance only, and these weights are then applied to projections in a static manner, assuming that they remain valid under future conditions. However, it should be noted that HMs possessing adequate simulation performance under past climate conditions may not guarantee robust projections under future emission scenarios (Coron et al., 2014). In this regard, the use of static weights, even when obtained by using iterative methods such as Bayesian Model Averaging (BMA), does not necessarily provide a viable way to reduce the uncertainty variance propagated throughout an impact modeling chain. This limitation underscores the requirement for alternative weighting schemes that can dynamically adapt to evolving future conditions. Regarding this issue, Ersoy et al. (2025a) demonstrated that uncertainty in runoff projections can be significantly reduced using a framework that balances the trade-off between past simulation performance and projection uncertainties, unlike studies based on statically weighting HMs for hydrological projections such as Pastén-Zapata et al. (2022) and Castaneda-Gonzalez et al. (2023).

Another notable gap in the literature is that only a scant number of studies have addressed reservoir operation strategies under uncertainty in the hydroclimatic impacts of climate change, but these have employed ensemble streamflow projections without explicitly attempting to manage the associated uncertainty. For example, Raje and Mujumdar (2010) examined the impacts of climate projection uncertainty on stochastic dynamic programming-based operations, without undertaking a decomposition or variance reduction related to source-specific uncertainties. On the other hand, Okkan et al. (2023) quantified the contribution of GCMs and ESs to the overall uncertainty in reservoir performance measures (RPMs), but they reduced the uncertainty in projected reservoir releases to some extent by choosing a subset of GCMs. That is, all these few papers did not adopt a strategy that allows dynamic ensemble generation tailored to both historical simulation accuracy and to narrowing projection uncertainty.

Therefore, this study aims to bridge that gap by integrating a framework that dynamically weights the multi-member hydrological ensemble—developed in a companion study (Ersoy et al., 2025a)—with a two-dimensional hedging rule (HDG-2d) to provide a more robust decision-support basis under climate change. To the best of our knowledge, no previous work has investigated an integrative approach that enables the reduction of upstream modeling uncertainties via weighting HMs and also yields more credible reservoir operation strategies under climate change scenarios. The potential of this dual-level uncertainty management has been scrutinized through the case of the Tahtali Reservoir in western Türkiye, which is operated primarily for domestic water supply. Despite increasing evidence of climate change-derived shifts in hydroclimatic regimes within the Kucuk Menderes River Basin, where the Tahtali Reservoir is located (e.g., Yagbasan, 2016; Rotbeei et al., 2025), no research has examined how such changes may influence reservoir operation uncertainty for any existing reservoir in this drought-prone basin. Accordingly, this study links projected hydroclimatic changes over this region to reservoir-operation decision-making, and it seeks to address the following objectives. These include: (i) examining the advantages offered by the HDG-2d rule over SOP under uncertainty propagated from a complex modeling chain with multiple sources; (ii) evaluating how decomposed projection uncertainties related to long-term RPMs differ across reservoir operation models; (iii) investigating the efficacy of multi-hydrological model

weighting schemes for reducing uncertainties in reservoir operation projections; and (iv) quantifying the temporal evolution of the influence of uncertainty-driving factors on optimized reservoir releases under unweighted and weighted ensembles over the projection period 2021–2099.

2. Data and projection framework

2.1. Study area and in situ observations

This study considered the Tahtali watershed, which has a drainage area of 554 km^2 within the Kucuk Menderes River Basin (KMRB) in western Türkiye (Fig. 1a). The Tahtali Reservoir was selected as the case study because it constitutes one of the most critical water supply systems in the KMRB and is characterized by a typical Mediterranean climate regime. The reservoir has a maximum operational storage capacity of 306.65 Mm^3 and a dead volume of 15.5 Mm^3, and its area–storage relationship (Fig. 1b) shows nonlinear characteristics, with surface area increasing sharply at lower storage levels but expanding only gradually at higher volumes. In this context, the representation of sector-specific water demand is a critical component of reservoir operation modelling and shapes the system responses under drought conditions (Anvari et al., 2023). This study focuses only on domestic (urban) water supply, as the Tahtali Reservoir is operated primarily to meet an annual demand of 128 Mm^3 for İzmir, the third largest city in Türkiye, serving a population of more than four million. In recent years, noticeable changes in hydrological regime have made optimal reservoir management more essential and entailed reliable reservoir operation projections to fortify long-term water resources planning in KMRB (Ersoy et al., 2025b). All these factors make the Tahtali Reservoir a highly relevant case for analyzing climate-driven impacts in reservoir operation optimization.

The study utilized streamflow observations from Derebogazi station (D06A007) for the water year period October 1970–September 1988, a period during which no reservoir existed. After the construction of the Tahtali Dam, this station was permanently closed, and no upstream gauges have been operated since then. Consequently, the period 1970–1988 constitutes the only reliable and natural hydrological window for calibrating and validating hydrological models used in this study. Although the data range of the Derebogazi station does not fully overlap with the 1981–2005 historical period of the GCM used and meteorological data sets, it is assumed that, consistent with standard practice, hydrological models calibrated on a homogeneous natural

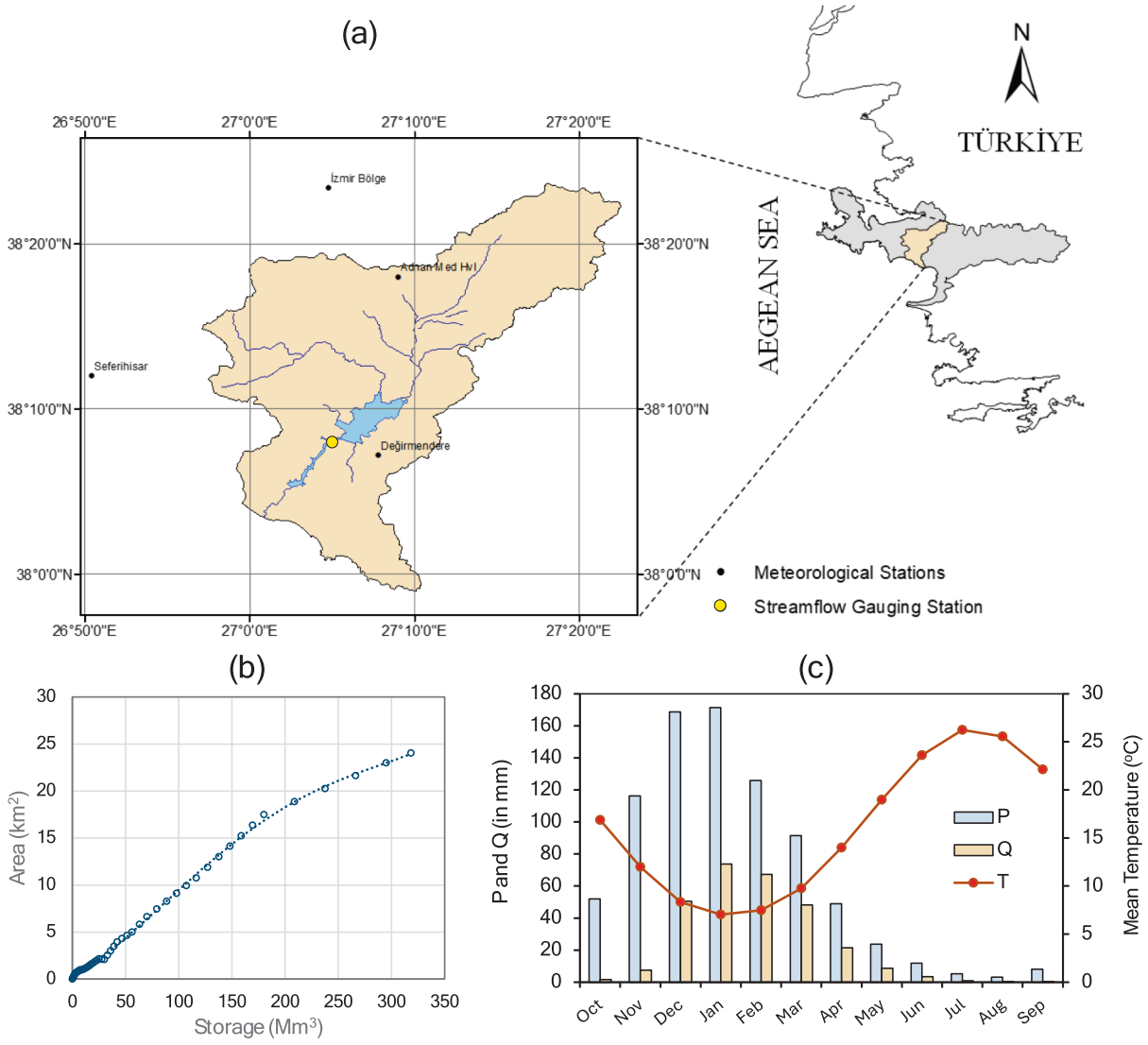


Fig. 1. Location and key characteristics of the Tahtali watershed in western Türkiye. (a) Map of the case study, including watershed boundary, river network, and Tahtali Dam reservoir, the upstream streamflow gauging station (D06A007) and representative meteorological stations used. (b) Area–storage relationship of the reservoir. (c) Monthly mean precipitation (P), natural streamflow-derived runoff (Q), and temperature (T) regimes for streamflow observation period.

period (i.e., the first 10 years of the available 1970–1988 record) can robustly capture the rainfall-runoff relationship and be applied for future projections as well. The streamflow records, the area–storage function for the dam site and the operational characteristics mentioned above were obtained from the General Directorate of State Hydraulic Works of Türkiye, together with data pertaining to representative climatological stations in the region. During the natural streamflow observation period, the annual precipitation and mean temperature were 825 mm and 16.0 °C, respectively, and the watershed belonged to the dry subhumid climate zone. Moreover, the annual mean runoff was observed to be 285 mm, which is almost equivalent to one-third of the annual mean precipitation. Both precipitation and runoff reach their maximum values in January, and nearly 84% of the annual runoff is obtained between December and March, highlighting the strong winter dominance of the hydrological regime (Fig. 1c). While precipitation and runoff show non-monotonic trends during the common observation period, increasing temperature trends are likely to intensify in the future, and the resulting net evaporative losses may place further operational stress on the reservoir. Although the study was carried out only in the Tahtali watershed, it is considered to be sufficient to introduce the developed integrative strategy and is expected to inspire similar applications in other regions facing uncertain climate change impacts.

2.2. Climate projections: Precipitation and temperature

To assess future hydroclimatic changes in the Tahtali watershed, data derived from five GCMs were employed under two representative concentration pathway (RCP) scenarios: RCP4.5, which aims to stabilize radiative forcing at 4.5 W/m^2 by the end of the century, and RCP8.5, which is based on continuously rising emissions and projects a radiative forcing of 8.5 W/m^2 by 2100. An overview of the RCPs is available in van Vuuren et al. (2011), to which readers may refer for further details. In addition to two RCPs, the historical (HIST) outputs from the GCM simulations were also used. In this context, the 1981–2005 interval represents the common overlap period between all GCMs and the available observational meteorological datasets. It was adopted as the reference window for correcting biases in precipitation and temperature projections as well as calculating future anomalies. Moreover, by incorporating two DMs (i.e., dynamically and statistically downscaled data) and seven lumped HMs listed in Table 1 into the modeling chain, the study allowed for the analysis of an ensemble of 140 streamflow projection combinations. It should be noted that the hydroclimatic projections obtained from that chain were previously compiled from Ersoy et al. (2025a) to serve as input data for the reservoir operation optimization modeling conducted in this study.

2.2.1. Dynamically downscaled data

Dynamically downscaled data for the reference scenario period of 1981–2005 (HIST) and four future horizons of 2021–2039, 2040–2059, 2060–2079, and 2080–2099 were extracted from the Coordinated Regional Climate Downscaling Experiment (CORDEX) database. Even though this platform contains different regional climate model simulations for the Middle East and North Africa (MENA) domain, the outputs from only five GCMs, namely CNRM-CM5, GFDL-ESM2M, EC-EARTH, HadGEM2-ES, and MPI-ESM-MR, are available under both RCP4.5 and RCP8.5 (Table 1). The regional climate models mentioned in Table 1 were driven by boundary conditions from these five GCMs as part of the Coupled Model Intercomparison Project Phase 5 (CMIP5). Following the weighting of multiple representative grids for each GCM, the quantile delta mapping (QDM) algorithm was applied in order to correct systematic biases in the outputs. This method was also found superior to standard quantile mapping in terms of preserving relative changes in quantiles (Tong et al., 2021).

In the study, the mean value anomalies for each future period, relative to the HIST, are given in Fig. 2. In various parts of the text, the notation 'GCM-CRX' is used to denote dynamically downscaled data for a given GCM within the CORDEX-MENA framework, while 'GCM-SD' refers to the statistically downscaled counterparts. Under RCP4.5, projected precipitation changes for the Tahtali watershed appear to remain statistically insignificant almost until the end of the century. As for annual precipitation time series, no persistent long-term trend is evident during 1981–2005 or under RCP4.5 (Fig. S1). According to findings under RCP8.5, only the GFDL-CRX and EARTH-CRX variants tend to display significant decreases in precipitation during certain periods, whereas the trend signals remain inconsistent across the GCMs (Fig. 2a). Moreover, all CORDEX models showed an unequivocal signal in terms of temperature projections, with ensemble-mean temperatures rising steadily across both RCP scenarios (Fig. S1). Under RCP4.5, temperature anomalies ranged from 1.0 to 3.0 °C across the GCMs during 2040–2099. As for RCP8.5, the warming intensified, with anomalies reaching between 3.2 and 6.0 °C during 2080–2099 (Fig. 2b). Also, HadGEM-CRX stands out with a remarkable warming trend (Fig. S1), which is consistent with earlier findings obtained for the MENA domain (Ozturk et al., 2018).

2.2.2. Statistically downscaled data

Statistical downscaling was performed using radial basis function networks (RBFNs) that establish empirical transfer functions between ERA5 reanalysis data, provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), and local surface observations. Then, these large-scale climate predictors were re-gridded to closely match the spatial resolution of the CMIP5 models, and the re-gridded

Table 1
Overview of the modeling chain consisting of five GCMs, two downscaling methods (DMs), and seven hydrological models (HMs), all operated under RCP4.5 and RCP8.5 scenarios. CMIP5 models are expressed in latitude × longitude format, while CORDEX regional simulations are provided at a uniform 0.44° rotated grid spacing. Regional climate models specified in the CORDEX framework are given in parentheses (i.e., RCA4 and RegCM4.4). For detailed descriptions of the HMs used, governing equations, and calibration procedures, which are identical to those used in this study, see Ersoy et al. (2025b).

RCPs	GCMs			DMs	HMs
	GCM Name	Institution/Origin	Resolution		
RCP4.5RCP8.5	CNRM-CM5 (RCA4)	SMHI-Sweden	0.44° (~50 km)	Dynamical downscaling*	abcde
	GFDL-ESM2M (RCA4)	SMHI-Sweden	0.44° (~50 km)		Awbm
	EC-EARTH (RCA4)	SMHI-Sweden	0.44° (~50 km)		Dynwbm
	HadGEM2-ES (RegCM4.4)	iklimBU-Türkiye	0.44° (~50 km)		Gr2m
	MPI-ESM-MR (RegCM4.4)	iklimBU-Türkiye	0.44° (~50 km)		Guo
	CNRM-CM5	Centre National de Recherches Météorologiques – France	1.4° × 1.4°	Statistical downscaling**	TemezTwbm
	GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory – USA	2.0° × 2.5°		
	EC-EARTH	Irish Centre for High-End Computing – Ireland	1.125° × 1.125°		
	HadGEM2-ES	Met Office Hadley Centre – United Kingdom	1.25° × 1.875°		
	MPI-ESM-MR	Max Planck Institute – Germany	1.875° × 1.875°		

* Dynamically downscaled datasets were obtained from the CORDEX-MENA framework (referred to as CRX throughout the text).
** Statistically downscaled (SD) climate series were derived via radial basis function networks as transfer functions.

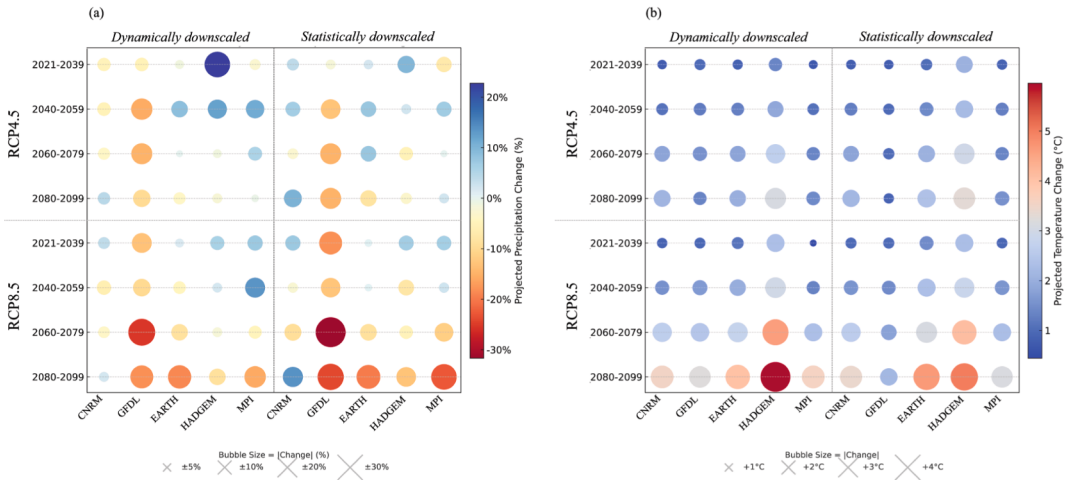


Fig. 2. Projected changes in annual mean precipitation (a) and temperature (b) over the Tahtali watershed under RCP4.5 and RCP8.5 for four future horizons. Each bubble represents a GCM–DM combination, with bubble size indicating the absolute value of the projected anomaly, and bubble color denoting the magnitude and direction of these changes.

areal data corresponding to each GCM underwent a predictor selection procedure based upon the least absolute shrinkage and selection operator method (Ersoy et al., 2025a). Once the explanatory predictors were identified, they were employed in the training and validation of RBFNs. The way in which the data is standardized before training, as well as the general flowchart followed, are mostly similar to Okkan and Inan (2015). Consequently, the Nash-Sutcliffe efficiency (NSE) values exceed 0.75 in both stages, indicating that the transfer functions are robust enough to statistically downscale GCM outputs to catchment-scale variables.

The QDM approach was also used to correct biases in statistically downscaled data, and mostly marginal discrepancies ($<\pm10\%$) were observed between the precipitation anomalies derived from statistical downscaling and those from the CORDEX models (Fig. 2a). In addition, as can be inferred from Fig. S2, RBFN-based annual precipitation time series exhibit similar precipitation variability to their CORDEX-based counterparts because both variants are driven by the same GCM forcings. Besides, temperature projections from the statistical approach also reflected the general warming trend, consistent with the CORDEX variants, with both downscaling approaches showing similar warming trajectories (Fig. S1 and S2). Yet, under RCP8.5, GFDL-SD and HadGEM-SD produced about 1.0 °C lower anomalies than their CORDEX counterparts during the period 2080–2099 (Fig. 2b). This may be attributed to the fact that some regional climate models may amplify the raw signals due to their parameterization schemes (Thomas et al., 2021).

2.3. Multi-model ensemble hydrological projections

The ensemble of hydrological projections employed in this study is derived from seven lumped HMs, which were previously calibrated and validated for the same watershed (see Ersoy et al., 2025b). These models are the abcde model, the Australian water balance model (Awbm), the dynamic water balance model (Dynwbm), Gr2m, the Guo model, the Témez model, and the Thornthwaite water balance model (Twbm), each featuring different runoff generation mechanisms despite sharing similar conceptual frameworks. In this study, both HMs and reservoir operation models need not only monthly total precipitation but also potential evaporation (in mm), estimated from the Kharrufa equation (Kharrufa, 1985), which was calibrated for the study region based on monthly mean temperature. The consistently high NSE scores (i.e., NSE values exceeding 0.80) obtained during the calibration and validation periods indicate that the rainfall-runoff relationship in the watershed can be adequately captured by these HMs (Fig. S3). Nevertheless, the Dynwbm, Gr2m, and Guo models produced rather weaker results in terms of low-flow simulation quality, as evidenced by the insufficient values of LNSE based on NSE calculations from log-transformed flows. In contrast, Twbm stood out slightly in that it was able to better represent the overall hydrograph without compromising the simulation of low-flow conditions.

While all these employed HMs have a certain simulation ability under past conditions, they may respond more sensitively to the projected climate signals. This sensitivity is probably due to their structural limitations or internal assumptions and can lead to inflated projection uncertainty. In this context, weighting HMs for the studied region was adopted by Ersoy et al. (2025a) as a way to reduce uncertainty in hydrological projections. In the unweighted case, a total of 140 runoff projections, generated by coupling seven HMs with five GCMs and two DMs under RCP4.5 and RCP8.5 scenarios, were evaluated. In the weighted ensemble, a subset of 20 projections was compiled using two different weighting schemes: one based solely on historical simulation performance, following the BMA approach, and the other using the Uncertainty Optimizing Multi-Model Ensemble (UO-MME), which addresses the trade-off between simulation accuracy (i.e., measured by NSE) and projection uncertainty (Fig. S4).

Using simulated runoff data (i.e., calibration data) from seven HMs, the constrained weighting approach BMA, where the sum of assigned weights is equal to one, was evaluated considering Darbandsari and Coulibaly (2019). The resulting weights were then applied to both the validation data and future runoff projections (see Table S1a). Accordingly, the BMA method predicted a higher weight for Twbm, which performed more consistently in terms of NSE–LNSE balance, while the weights of the remaining HMs showed only moderate correlation with their individual performance. It was also seen that BMA tended to penalize models with slightly lower NSE scores and, in some cases, assigned relatively high weights to HMs with poor LNSE performance (e.g., Gr2m), suggesting that the approach could only partially reflect multi-metric performance. Darbandsari and Coulibaly (2019) drew similar conclusions for BMA, albeit with other metrics. In terms of predictive performance, the BMA-based ensemble gave the highest NSE for the calibration data by a small margin. Interestingly, it ranked behind all individual HMs in the same metric for the validation data, while LNSE scores indicated that it retained low-flow simulation capability.

The impact of different factors on projection uncertainty (i.e., projected runoff changes in annual means relative to the baseline scenario HIST) was quantified by means of the analysis of variance (ANOVA) framework. For detailed description of multi-factor ANOVA, readers are referred to Vetter et al. (2015) and Chawla and Mujumdar (2018). It

should be noted that the interaction terms in the employed version of ANOVA may have implicitly embedded the effect of internal variability. However, as Lemaitre-Basset et al. (2022) pointed out, the contribution of natural climatic variability to total uncertainty is likely to become less influential when projected anomalies are computed over long time slices. A four-way ANOVA decomposed the contribution of the full set of sources and their interactions to the overall projection uncertainty (see Fig. S5). This analysis showed that the uncertainty contribution caused by the HM selection was relatively small (<1%), likely due to similar NSE scores of HMs. Yet, the overwhelming variance from multiple GCMs, despite bias correction, is increasingly propagated through HMs that have inherent structural uncertainties as well. Therefore, the potential to alleviate this inflated projection variance by means of weighted HMs, as stated previously, was discussed. A three-way ANOVA was applied to the weighted data so as to examine the extent to which weighting lumped HMs could narrow the uncertainty associated with GCMs used, their interactions with emission scenarios (GCM*ES), and the total projection uncertainty. Changes in the decomposed uncertainty variances of the weighted projections, compared to those obtained from the four-way ANOVA, are given in Table S1b. It is clear from this table that BMA has only a marginal impact on reduction in projection variance across future time horizons. As can be inferred from Pastén-Zapata et al. (2022), when HMs exhibit similar simulation skill, weighting schemes that rely on historical performance may be less effectual in reducing projection uncertainty. Hence, the limited capability of BMA for the Tahtali watershed may not be surprising, given that the NSE performances of the individual HMs were observed to be relatively close.

The fact that statically determined weights that generate balanced simulations for calibration and validation data do not guarantee a reduction in runoff projection uncertainty has motivated the usage of the UO-MME approach, which estimates the weights of HMs under a bi-objective structure (Ersoy et al., 2025a). Considering that assigning negative weights to HMs may enable offsetting adjustments between them (Castaneda-Gonzalez et al., 2023), UO-MME was also formulated in an unconstrained form that allows negative weights and includes a constant term for bias correction. The UO-MME framework driven by the non-dominated sorted genetic algorithm II (NSGA-II, Deb et al., 2002) produced Pareto frontiers in Fig. S4. This revealed that the trade-off between NSE-based simulation accuracy and long-term mean total projection uncertainty (mTUV) during 2021–2099 is clearly evident. Within the identified trade-off space, certain solutions offered a significant reduction in mTUV in exchange for tolerating up to a 5% decrease in NSE value compared to that of the best-performing model, Twbm. Especially, one candidate solution (Table S1a) was observed to provide a satisfactory NSE–LNSE balance. As shown in Table S1b, the total uncertainty was reduced by about 31–35% through that decision point of UO-MME. Besides, this framework brought about marked uncertainty reductions in the GCM and GCM*ES components, which together accounted for nearly two-thirds of the total uncertainty, in contrast to the minor changes introduced by the BMA approach.

Based on the weighted combinations derived from the UO-MME, the projected changes in annual mean runoff relative to the HIST baseline are presented in Fig. 3. As is immediately apparent from the patterns in this figure, the disparities in the runoff changes are mostly attributable to higher uncertainty in the GCMs. For instance, while GFDL-ESM2M, which shows small sensitivity to choice of DM, stands out with a stark decrease in runoff, with an average of 21–25% across existing periods and ESs. In contrast, other variants such as CNRM-CM5 and MPI-ESM-MR produced even anomalies nearly one-fifth or one-sixth of those values. As for the ESs, of which interacting with GCMs contributed significantly to overall uncertainty, early- and mid-term runoff changes vary considerably. However, the RCP4.5 scenario during the 2080–2099 period has relatively consistent declines, with an average decrease of 10% across DMs. Under RCP8.5, the runoff declines are more

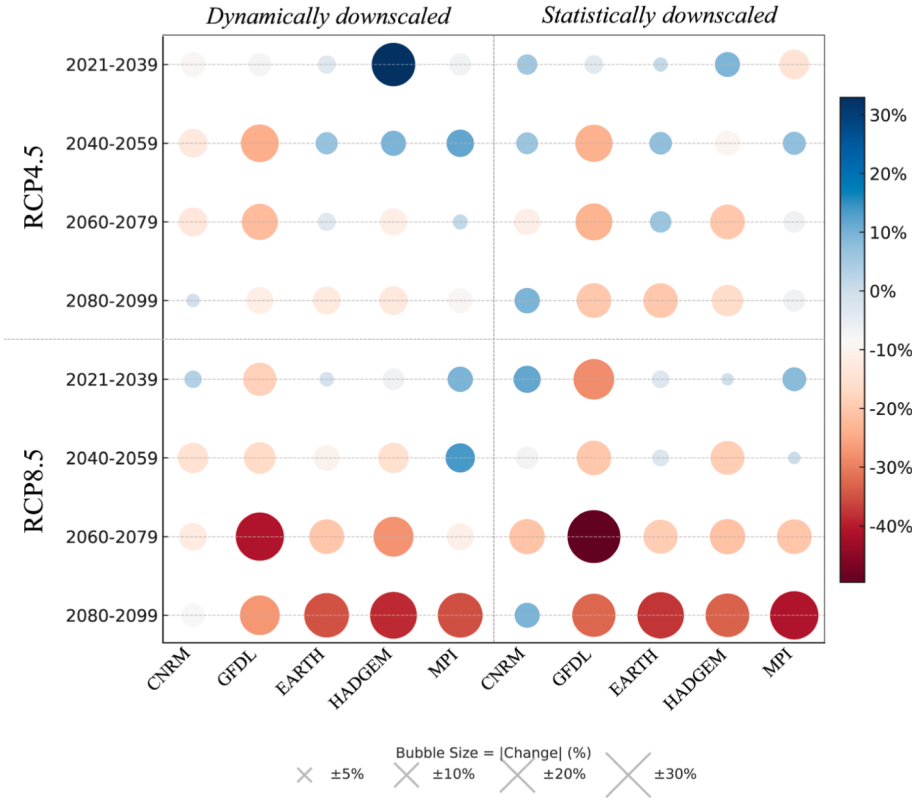


Fig. 3. Projected changes in annual mean runoff generated through the UO-MME weighting scheme, presented for each GCM–DM combination under the two RCP scenarios over the four future periods. Bubble sizes show the absolute values of the projected changes, while the colors vary based on the magnitude and direction of these changes.

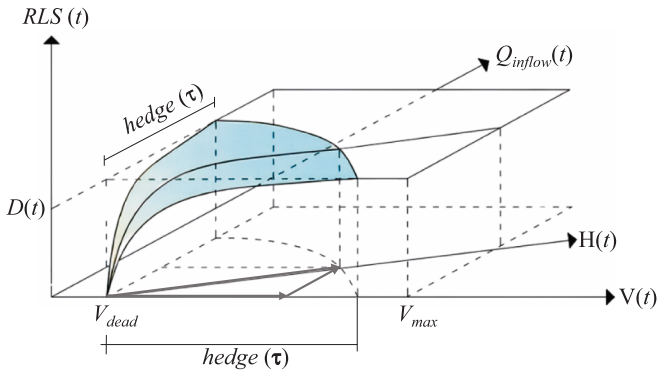


Fig. 4. Conceptual representation of the two-dimensional HDG-2d model used for adaptive reservoir operation, re-adapted from Celeste and Billib (2009).

pronounced across GCMs, reaching up to nearly 30% on average by the end of the century.

Fig. 3 also suggests that if a single GCM were to be relied upon, the impact of DM choice on hydrological projections would be relatively important. For example, CNRM-CM5 exhibited the largest differences between DMs, averaging 9%, with its CRX variant consistently producing lower runoff anomalies across almost all future periods. Unlike that case, CRX indicated an opposite direction change for HadGEM2-ES under RCP4.5 and led to a markedly wetter tendency since its related anomalies exceeded those of SD by +24% in 2021–2039 and by +19% in 2040–2059. But notwithstanding, given that diversity from multiple GCMs exerted a dominant effect on runoff projection spread, UO-MME's strength in narrowing total projection uncertainty lies in its ability to manage the climate change signals from GCMs, whose variance is further propagated together with other sources of uncertainty (Ersoy et al., 2025a).

3. Climate change-informed reservoir operation

3.1. Adaptive hedging rule

In the study, the projected runoff data obtained from both the unweighted and weighted ensembles were converted into inflows and subsequently used as inputs to two reservoir operation models: one based on SOP, which does not retain water for future use, and the other a hedging policy that employs a two-dimensional correlation rule between the release and the reservoir storage value for any month index t (Fig. 4). The hedging model, termed HDG-2d, restricts reservoir releases through a parameterized rule curve. This design helps mitigate probable water shortages in the future, as opposed to SOP, which may lead to severe short-term shortages during drought conditions (Celeste and Billib, 2009).

When implementing HDG-2d, the main objective of reservoir operation is to determine releases that best satisfy the corresponding demands. Accordingly, minimizing the sum of squared deficits is typically adopted as the objective function:

$$minimize \sum_{t=1}^{L} \left(\frac{D(t) - RLS(t)}{D(t)} \right)^2 ; t \in L \tag{1}$$

The constraints are as follows:

$$V(t+1) = V(t) + Q_{inflow}(t) - RLS(t) - E_{net}(t); \forall t \tag{2}$$

$$V_{dead} \le V(t+1) \le V_{max}; \forall t \tag{3}$$

where L is the length of the operating period (in months); $D(t)$ and $RLS(t)$ denote the demand and release in month t, respectively; $V(t)$ and $V(t+1)$ represent the reservoir storage volumes at the beginning and the end of month t; $Q_{inflow}(t)$ and $E_{net}(t)$ are the inflow and net evaporation volume during month t, respectively; V_{dead} and V_{max} correspond to the reservoir dead volume and maximum storage volume, respectively.

In the continuity equation (Eq. (2), $RLS(t)$ is governed by the hedging rule parameters, while $E_{net}(t)$ representing evaporation loss associated with the reservoir's average surface area is determined iteratively using the reservoir area-storage curve. All variables in this equation are typically expressed in Mm^3. The operating rule applies hedging within the limits defined by $hedge(\tau)$, based upon the combined value of active storage and inflow during month t, as shown in Eq. (4):

$$RLS(t) = \begin{cases} 0, & if\ V(t+1) \le V_{dead} \\ D(t) \times \left(\dfrac{H(t)}{hedge(\tau)} \right)^{m(\tau)}, & if\ H(t) \le hedge(\tau) \\ D(t), & if\ H(t) > hedge(\tau)\ or\ V(t+1) \ge V_{max} \end{cases} \tag{4}$$

where the combined indicator $H(t) = \sqrt{[V(t) - V_{dead} - E_{net}(t)]^2 + Q_{inflow}(t)^2}$; $m(\tau)$ is the parameter that controls the nonlinearity of the hedging function.

In accordance with Celeste and Billib (2009), the HDG-2d model parameters $hedge(\tau)$ and $m(\tau)$ are defined for each month ($\tau = 1, 2, \cdots, 12$), resulting in a total of 24 parameters to be calibrated. The parameter $m(\tau)$ is constrained to lie within the closed interval [0, 1], while $hedge(\tau)$ must be positive and less than or equal to the active storage capacity. According to Eq. (4), the reservoir cannot meet even a certain part of the downstream water demand (i.e., $RLS(t)$ is set to zero), when the end-of-month storage falls below the dead storage level. Additionally, demand is fully met without any restrictions on water allocation when $H(t)$ exceeds the calibrated threshold parameter $hedge(\tau)$ or when the reservoir reaches its maximum capacity, which results in excess water being spilled (Bayesteh and Azari, 2021).

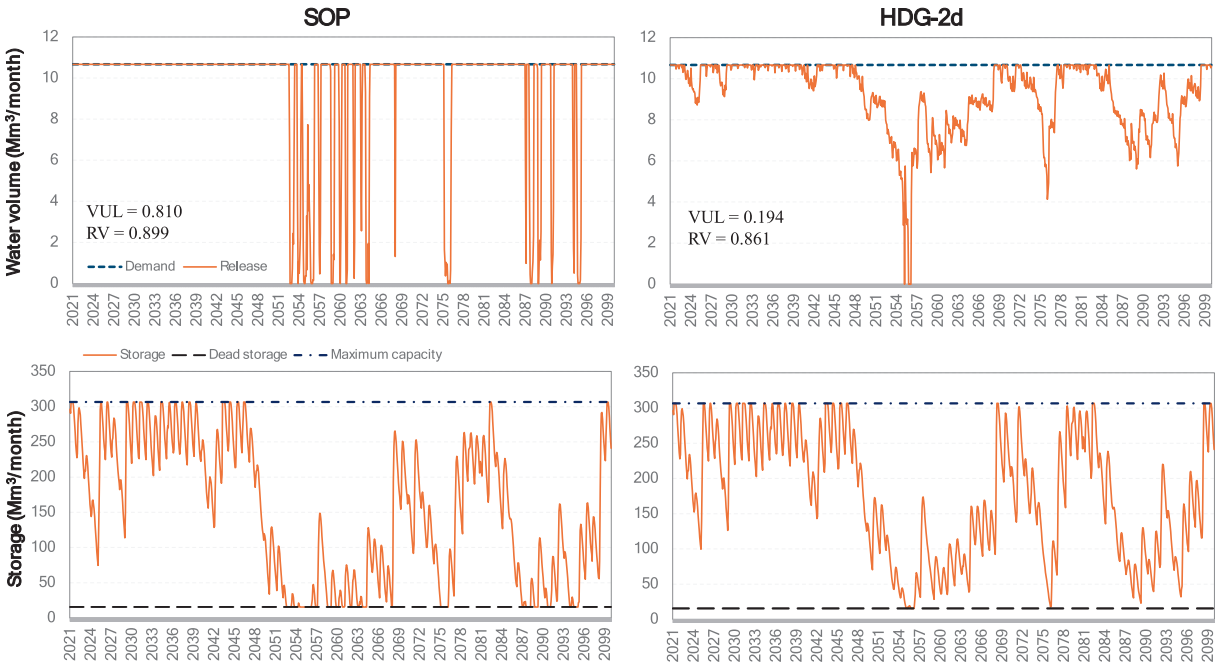
The optimization of the HDG-2d model is initiated by the calculation of the objective function value by means of a predefined set of hedging rule parameters. These parameters are then iteratively updated to minimize this objective function. Given both the nonlinear and multi-parameter nature of reservoir operation models, the use of evolutionary algorithms (e.g., frequently genetic algorithms) has become a standard practice in the literature (Labadie, 2004; Moghaddasi et al., 2022). Therefore, this study employs a population-based differential evolution (DE) algorithm, which is equipped with a modified mutation scheme. The details about this DE variant were provided by Okkan et al. (2023), and the algorithm settings in that study were adopted without any modification. In total, 180 inflow projections were evaluated while optimizing the HDG-2d model: 140 from the unweighted ensemble and 20 each from the BMA- and UO-MME-based weighted ensembles. Despite operating with a large number of projections, executing the DE algorithm with a population size of 50 and a maximum of 300 iterations was found sufficient due to its fast-converging and robust nature. The RPMs used to interpret the effects of establishing adaptive hedging rules on reservoir performance under multiple projection variants are introduced in the next section.

3.2. Reservoir operation performance criteria

Using reservoir releases and storage volumes projected under various conditions, the long-term performance of the Tahtali Reservoir was evaluated based on three key RPMs: vulnerability, volumetric reliability, and time-based reliability.

The vulnerability index (VUL), which quantifies the expected value of deficits, can be formulated as follows (Sandoval-Solis et al., 2011):

$$VUL = \frac{\sum_{t=1}^{L} [D(t) - RLS(t)]/N_F)}{D_{mean}} \tag{5}$$



Pars.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.
hedge (τ)	160.04	196.40	218.27	291.15	262.00	245.59	235.39	240.20	268.57	233.92	262.55	241.47
m (τ)	0.339	0.244	0.230	0.170	0.236	0.321	0.334	0.304	0.300	0.280	0.249	0.253

Fig. 5. Comparison of reservoir operation behavior under the SOP and HDG-2d models for an example variant. The results are based on the CNRM–CRX configuration under the RCP8.5 emission scenario, with projected inflows derived using the Twbm hydrological model. The upper panels illustrate monthly release and demand volumes over the 2021–2099 horizon, while reservoir storage volumes are shown in the lower panels. Calibrated monthly HDG-2d parameters are also listed at the bottom of the figure.

where N_F is the number of months in which $RLS(t) < D(t)$.

This dimensionless index can be defined as the ratio of the mean deficit for the months in which the water demand is not fully met to the mean demand over the same months (D_{mean}).

In addition, the volumetric reliability index (RV) represents the portion of the total water demand that is supplied over the entire operation period and can be defined as follows:

$$RV = \sum_{t=1}^{L} RLS(t) / \sum_{t=1}^{L} D(t) \tag{6}$$

Assuming that the total number of months in which the end-of-month storage volume is less than or equal to V_{dead} (i.e., the case of no water release) represents the critical failure duration (F_D), time-based reliability index (RT) can be expressed as (Okkan et al., 2023):

$$RT = 1 - \frac{F_D}{L}; F_D \leq L \tag{7}$$

Moreover, there have also been efforts to convert the above-mentioned indices into a single metric termed the sustainability index (SI), but the relative contributions for each component might be elusive when calculating such a metric. To deal with this limitation, Sandoval-Solis et al. (2011) recommended using the geometric mean of the individual performance indices, which can be formulated as:

$$SI = \sqrt[3]{(1 - VUL) \times RV \times RT} \tag{8}$$

The next section shows how the combined use of multiple hydrological model weighting schemes and adaptive hedging policies influenced the projected uncertainty in system reliability, vulnerability, and sustainability.

4. Results and discussion

4.1. Reservoir performance implications under full-chain ensemble projections

This section presents the evaluation of the reservoir operation models under a set of 140 inflow projections. Before detailing the following analyses, it should be noted that a uniform monthly water demand of 10.67 Mm^3 (i.e., 128 Mm^3/year) from Tahtali Reservoir in line with the related planning situation provides a more operationally viable option for this study. The reservoir simulation–optimization process was conducted based on these values, since using higher demands (or a gradually increasing trend) led to unrealistic regulation ratios (not shown). While the SOP does not require an iterative procedure except for the estimation of net evaporation loss, the HDG-2d-based simulation–optimization framework was executed 10 times per variant using DE algorithm to identify the optimal parameter sets that minimize the objective function, and the stopping condition in each was set to 300 iterations. Fig. 5 exemplifies the behavioral differences between the SOP and HDG-2d models, based on Twbm-generated inflows derived from the CNRM-CRX outputs under the RCP8.5 scenario. Accordingly, while the SOP fully meets demand up to a certain period, the frequency of single periods of severe shortages increases as the intensity of hydrological drought becomes more pronounced after the 2050s.

In contrast, the HDG-2d model was characterized by a higher hedging threshold and a lower rule curve shape parameter for wetter months (i.e., December to March), while it still allows for allocating partial amounts even during dry periods. These calibrated parameters gradually restrict releases in the earlier phase of the 2021–2099 projection horizon in a controlled manner, leading to a slight reduction in

volumetric reliability. This parametrization is consistent with a system behavior that prepares the reservoir for potential drought conditions, and it has kept the VUL value at nearly one-fourth of that obtained under the SOP. However, as reported by Bayesteh and Azari (2021), the hedging process may predominantly shift toward the dry periods, being sensitive to hydroclimatic effects. In our study, it was observed that some variants exhibiting limited and uncertain inflow conditions during some winter months in the future horizon led to a similar delayed hedging behavior. Therefore, not only the SOP but also the HDG-2d model could give varying degrees of uncertainty responses to different sources.

Although Fig. 5 involves a detailed time-series assessment for a single variant, plotting the rest of the projections would be redundant. We therefore present a compact yet informative visualization of the variability in projected RPMs across each uncertainty source using box plots (i.e., box-and-whisker plots). Given each uncertainty source, box plots were made based on the combination of the remaining components. For example, as shown in Fig. 6, each HM-specific box plot represents the distribution of VUL values obtained from 5 GCMs × 2 ESs × 2 DMs (i.e., 20 projections), while GCM-specific plots were generated with 28 projections according to the modeling chain structure. Box plots for RV, RT, and SI were also produced and are given in Figs. S7, S8, and S9, respectively. It is seen from Fig. 6 that HDG-2d significantly curbs the severe vulnerabilities introduced by the SOP, regardless of the source of uncertainty (with an overall mean reduction of 83%). This improvement is generally consistent with the findings reported by Moghaddasi et al. (2022) and Adeloye et al. (2016), who also demonstrated substantial VUL decreases through other hedging rules. Given that the tolerable threshold value for VUL is 0.25, as noted by Adeloye and Dau (2019), the use of SOP resulted in only few variants being below this threshold, while 86% of HDG-2d-based combinations did so. Another finding is that the influence of HM selection on VUL values is rather apparent under SOP, with notable median shifts observed across HM variants (Fig. 6a). This is fortified by the four-way ANOVA (Fig. 7a), which indicates that variations among HMs account for 36% of the total variance in projected VUL values. What is more interesting is that when projecting system vulnerability with SOP, the second- or third-order nonlinear interactions of HMs with other sources contribute even more to the total variance. In contrast, HDG-2d adequately dampened the HM-induced variances for VUL and hence the related total uncertainty when compared to SOP (see Fig. 7b). These results suggest that the degree to which different sources contribute to VUL uncertainty depends on the reservoir operation model used. Several underlying factors may be responsible for these differences:

Although the projected streamflow changes in annual means according to the HIST scenario imply that the HM-driven uncertainty is marginal (Fig. S5), the HMs used differ in their representation of key hydrological processes. Therefore, the flows generated through such monthly HMs under alternative climates can exhibit disparities in both quantity and interannual variability, as pointed out by Panagoulia and Dimou (1997). Notably, LNSE results show that some HMs have a limited capacity to represent low-flow conditions (Fig. S3), while the corresponding baseflow index values and even low-flow durations across models were not compatible with each other (not shown). A reservoir operation model like SOP cannot ensure an adaptive response to these differences, instead routing the projected inflows directly to the target demand. This often leads to a lack of release sustainability, thereby resulting in higher and also more uncertain VUL levels. Considering that the uncertainty of HMs plays a more dominant role in the low-flow projections than in annual means (see Wang et al., 2020), SOP can be incapable of regulating HM-based low-flow variability, which in turn amplifies overall uncertainty. On the other hand, the HDG-2d model can preemptively govern the releases through its parametric rules even during low-flow periods. This is made possible by its buffer mechanism that prevents the HM-induced inflow disparities from directly reflecting on the projected VUL uncertainty. It should be also noted that we used a fixed demand assumption as in Celeste and Billib (2009), rather than incorporating time-varying demand as done in other studies (e.g., Zamani et al., 2017), which may lead reservoir operation models to show greater uncertainty responses.

Even though HDG-2d reduced total uncertainty in VUL by 65% by narrowing the substantially greater variances associated with HMs and their interactions, GCMs and RCPs together contributed 62% of the uncertainty for this key index, indicating that it remains responsive to

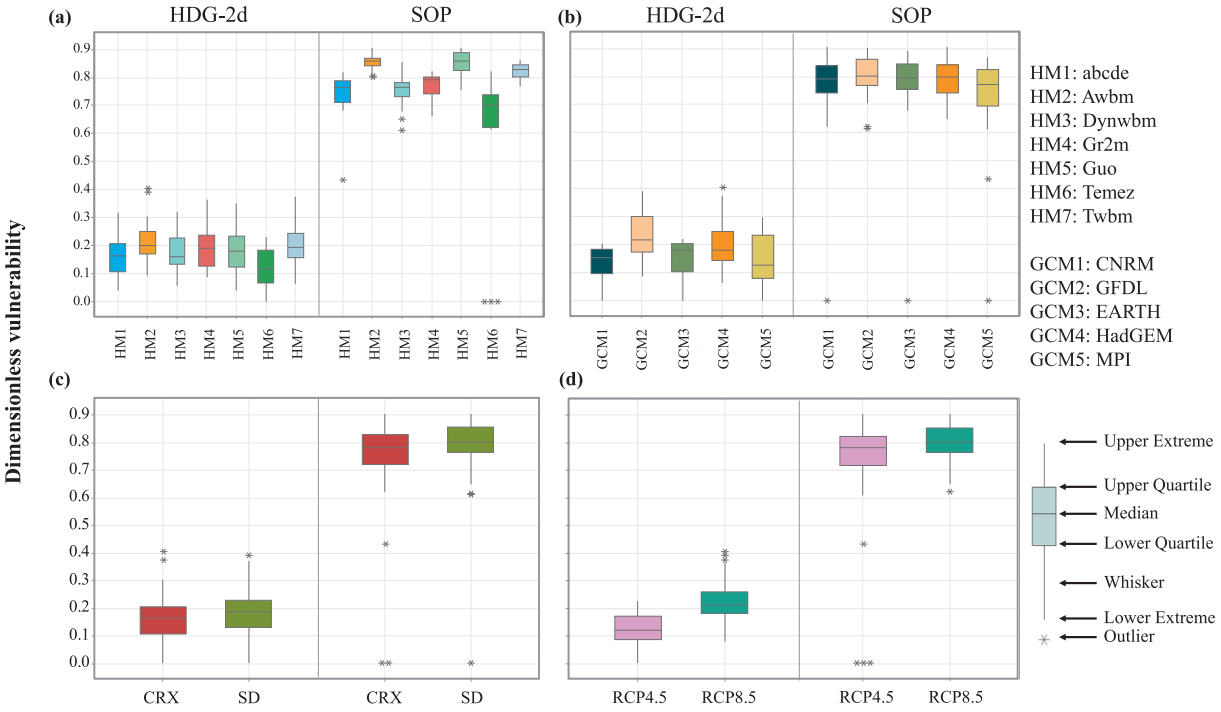


Fig. 6. Vulnerability projections under the SOP and HDG-2d reservoir operation models, stratified by four key sources of uncertainty. Each panel shows the distribution of vulnerability values associated with a single factor: (a) HMs, (b) GCMs, (c) DMs, and (d) RCP scenarios.

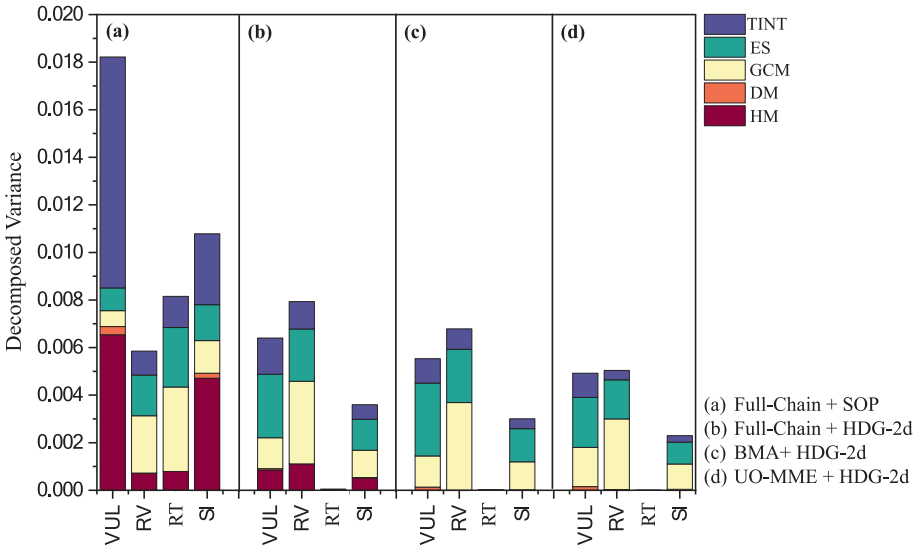


Fig. 7. Decomposed variances for projected reservoir performance metrics, attributed to major uncertainty sources, together with their combined interactions (TINT), which represent the total contribution of all interactions among these sources. Results for full-chain ensembles under the SOP and HDG-2d models are given in panels (a) and (b), respectively. Panels (c) and (d) display the results for weighted ensemble configurations using BMA and UO-MME, both coupled with the HDG-2d rule.

uncertain climate signals embedded in the ensemble projections (Fig. 7b; Fig. S6). Throughout the broad modeling chain, since the HDG-2d framework prioritizes mitigating the impact of flow deficits on unmet demand depending on HM selection, the dominance of GCM and ES in VUL variance is likely to become more evident when the hedging model attempts to reflect the changing climate signal in operational decisions. This reverse pattern compared to SOP highlights that relatively distinct climate forcings can make both the hedging responses and VUL projections more scattered. It can be inferred from Fig. 6b that the optimization of HDG-2d displays increased sensitivity to GCMs with strong drying signals, such as GFDL, which may have heightened the share of GCM uncertainty in the VUL projections. According to Fig. 7b, the GCM-induced variance in VUL under HDG-2d was nearly twice that decomposed for SOP, with the corresponding variance fraction increasing from 4% to 20%.

However, it cannot be claimed that the downscaling of GCM simulations is always directly responsible for the above-mentioned findings. Also, it has been emphasized that bias correction procedures applied to downscaled data can have implications for reservoir performance characteristics, and the difficulty of verifying their stationarity assumption may introduce further uncertainty into storage projections (Nguyen et al., 2020). In this respect, the use of QDM, a stationary bias correction algorithm, together with GCMs having strong climate signals may have caused a parameterization–simulation–optimization model such as HDG-2d to produce somewhat more dispersed VUL distributions. Of course, the fact that both statistically downscaled data and CORDEX simulations were subjected to the same QDM procedure made the DM uncertainty marginal for all performance indices (Fig. 6c, Figs. S7–S9). Nevertheless, the propagation of the biases remaining in the hindcast ensembles into future data may have been perceived differently by reservoir operation models and may have led to varied variance implications.

It is also worth noting that HDG-2d recalibrates its hedging behavior for each RCP scenario, and the model parameters are observed to be shaped more strongly by emission scenario pathways rather than inter-GCM variability. In this regard, Fig. 6d provides a preliminary understanding of the relative importance of scenario uncertainty. Under RCP8.5, which is characterized by a pronounced radiative forcing trend, HDG-2d model tends to apply stricter release restrictions, which can widen the interquartile range of ensembles and distinctly emerge the ES-induced uncertainty for system vulnerability (see Fig. 7b). Given that RCP selection has become temporally more influential on projected changes in temperature and evaporation rather than those of annual mean runoff (Ersoy et al., 2025a), the greater RCP uncertainty under HDG-2d essentially revealed that the calibrated model indeed responds to these diverging climate signals. Therefore, the combined evaluation of high VUL values with the lower scenario uncertainties under SOP, which cannot suppress HM-based volatility under low-flow conditions and is insensitive to emission-driven shifts due to its simple heuristic structure, cannot be interpreted as a sign of robustness.

Since VUL offers a direct indication of reservoir stress under climate change impacts, it was placed at the center of the uncertainty analysis, as shown above. However, the reliability metrics (RV and RT) also exhibited notable patterns depending on the uncertainty sources and were therefore briefly discussed herein. The HDG-2d model deliberately allows for a slight compromise in RV to temper vulnerabilities, as reflected in the reduction of the overall mean from 0.91 to 0.875 (Fig. S7). As a result of this trade-off, the total uncertainty in projected RV was observed to be 35% higher compared to SOP across the full-chain ensemble. Besides, two-thirds of this uncertainty already arises from variability in climate drivers (Fig. 7a, b), and it cannot be fully moderated by any operational strategy. As for projected RT, which reflects the frequency of maintaining storage above the dead storage level, an obvious distinction between SOP and HDG-2d was seen (Fig. S8). For the set of 140 combinations, HDG-2d achieved near-perfect time-based reliability (0.995 on average, range [0.96 to 1.00]), while SOP gave relatively lower performance and wider variability (0.885 on average, range [0.61 to 1.00]). As can be inferred from Fig. 7b, HDG-2d resulted in a negligible uncertainty response for RT, regardless of any source in the modeling chain. This is expectable as the model limits releases for stress periods and helps prevent storage depletion. Besides, when examining the contribution of adaptive policies to system sustainability under climate change, trade-offs among the criteria are available. Therefore, similar analyses were conducted for SI, which synthesizes the three measures into a single factor (see Fig. S9 and Fig. 7). Accordingly, the HDG-2d narrowed the large HM-driven uncertainty in SI projected under SOP by nearly 90%, and it did so by almost the same proportion for VUL (Fig. 7b). In fact, variance reductions were detected across all uncertainty sources and their interaction terms for SI, which can also be attributed to the enhanced RT performance under HDG-2d.

Furthermore, the HDG-2d ensured SI values steadily above 0.70 across all combinations and yielded an overall mean of 0.894, which was 0.34 greater than that of SOP (Fig. S9). All these findings confirm the robustness of HDG-2d in managing uncertainty in reservoir operation projections and lay the groundwork for further exploration of its performance together with the weighted lumped HMs in the next section.

4.2. Behavioral responses of reservoir performance to weighted ensembles

To explore how weighted HMs affect the Tahtali Reservoir behavior under climate change, we adopted two different weighting schemes: BMA and UO-MME. As mentioned in Section 2.3, BMA assigned weights to HMs based on historical simulation skill only, while UO-MME also considered the reduction of uncertainty in the ensemble of hydrological projections to make forward-looking robustness. While the impact of UO-MME to narrow runoff projection uncertainty relative to other multi-model averaging methods was discussed in the companion study (Ersoy et al., 2025a), this presented study intends to explore its operational implications when combined with a reservoir hedging rule. Accordingly, Fig. 8 compares a subset of 20 projected RPMs derived from two unequally weighted hydrological model ensembles with the full-chain ensemble comprising 140 projected values, separately under SOP and HDG-2d. For each variant, the difference between the mean performance measure obtained from the weighted ensemble and that of the full-chain ensemble is also annotated on Fig. 8.

It is also clear from Fig. 8 that RPMs did not respond positively to BMA method, even though it provided a skill-informed weighting for runoff simulation. In contrast, under SOP, UO-MME presents a clear advantage, especially in terms of projected VUL and SI, with median values for both measures shifting by 0.09, relative to the BMA-based results. As for the results under HDG-2d, performance gains of UO-MME, although less pronounced, were nonetheless found to be statistically significant by Wilcoxon signed-rank test. Since HDG-2d, by design, provides adaptive releases and makes the reservoir system less susceptible to projected inflow variability, the slightly limited impact of UO-MME on this setting would not be unsurprising. Nevertheless, the combined use of UO-MME and HDG-2d achieved an average increase of 0.03 in SI, along with a tighter interquartile range, pointing to slightly more stable sustainability outcomes (Fig. 8d).

All the findings above suggest that the improvements for long-term RPMs provided by UO-MME compared to the unweighted case are indicative of a good agreement between the reduced inflow uncertainty and reservoir responsiveness. This raises the question of whether HDG-2d or a weighting scheme contributes more to reducing uncertainty in RPMs. As shown in Fig. 7, the total uncertainty variance (TUV) provided by the SOP operated with full-chain data for the VUL measure was 0.0182, while the TUV provided by the HDG-2d under the same conditions decreased to 0.0064, corresponding to a 65% narrowing. When HDG-2d is coupled with UO-MME, the TUV further decreases to 0.0049, and this combination results in a 73% and 23% reduction in TUV relative to those of SOP and HDG-2d, respectively, both of which are operated with unweighted full-chain inflows. Hence, the incremental contribution of UO-MME beyond the dominant reduction already achieved by HDG-2d is about 8% for VUL. A similar pattern is observed for SI, where the additional reduction attributable to UO-MME is 12%. Therefore, moving from panel (a) to (d) in Fig. 7, it can be deduced that HDG-2d serves as the main driver of total uncertainty reduction, while UO-MME acts as a secondary but supportive factor that enhances this effect. To facilitate the readability of Fig. 7b to 7d, the relative changes in decomposed variances, resulting from applying HDG-2d with weighted ensembles, compared to their full-chain ensemble counterparts under the same hedging model, are also summarized in Table 2. Accordingly, filtering out the HMs based only on runoff observations by BMA contributed to balancing uncertainties in reservoir behavior, especially when used in conjunction with an adaptive hedging rule. This combination led to reductions in total uncertainty of up to 16% for all RPMs except for RT, which showed much greater sensitivity. Table 2 also affirms that the further narrowing of GCM- and ES-related uncertainties in the projected runoff data, which is effectively achieved by UO-MME-based weighting scheme, could be even better aligned with

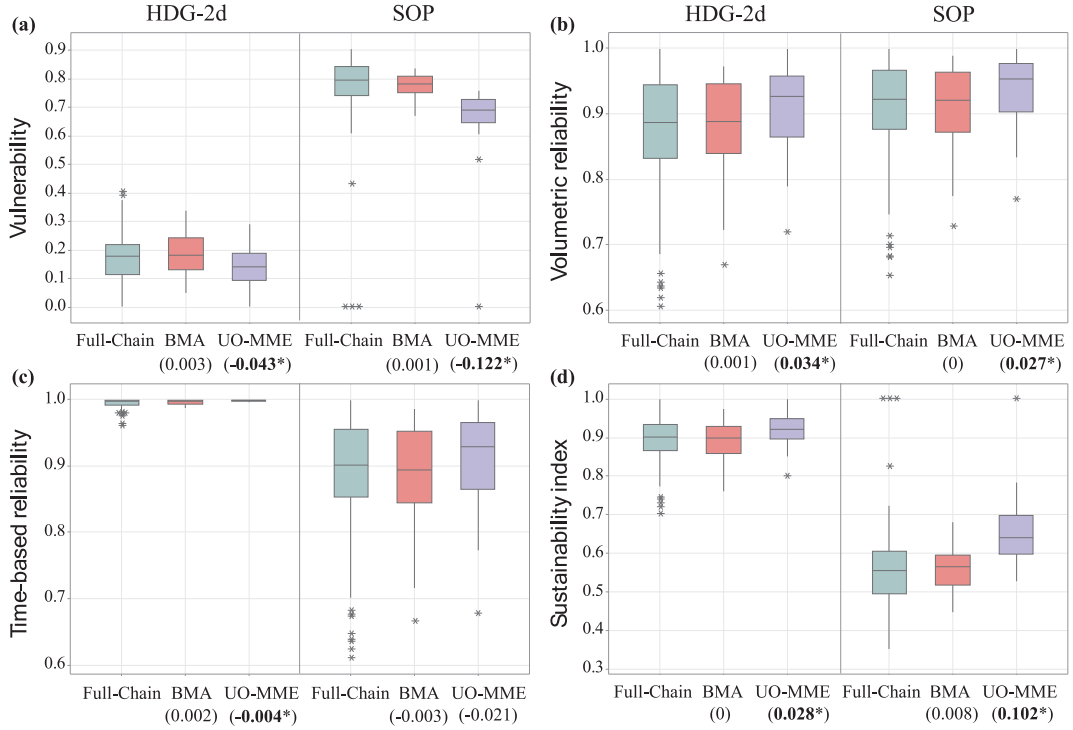


Fig. 8. Performance comparison of reservoir operation models under different ensemble configurations for the RPMs considered. Paired differences in the mean values of each performance metric between weighted and full-chain ensembles are given in parentheses; those marked with the symbol * are statistically significant at the 90% confidence level based on the Wilcoxon signed-rank test.

Table 2
Relative changes in the decomposed variances of major uncertainty sources for the projected RPMs under different weighted ensembles. Columns Δ_{GCM}, Δ_{ES}, and Δ_{TUV} indicate the percentage change in GCM uncertainty, ES uncertainty, and total uncertainty variance, respectively. All these changes were computed relative to the full-chain ensemble baseline, which was also operated with the HDG-2d model.

RPMs	Weighting schemes	Δ_{GCM}	Δ_{ES}	Δ_{TUV}
VUL	BMA	0.4%	14.7%	−13.6%
	UO-MME	26.8%	−21.2%	−23.1%
RV	BMA	6.5%	2.0%	−14.2%
	UO-MME	−14.4%	−25.4%	−36.4%
RT	BMA	−66.6%	−39.5%	−75.2%
	UO-MME	−95.6%	−94.3%	−96.4%
SI	BMA	2.7%	7.0%	−16.5%
	UO-MME	−7.9%	−29.0%	−36.1%

the inherent capacity of HDG-2d to dampen projected inflow variability. From the standpoint of the overall spread under this compound effect, variances in projected VUL, RV, RT, and SI were reduced by 23.1%, 36.4%, 96.4%, and 36.1%, respectively. Compared to BMA, the UO-MME scheme resulted in an additional reduction in overall uncertainty of roughly 20% for measures other than VUL, while the gain for VUL remained at about 10%. These findings partly align with Liu et al. (2019), who handled inflow uncertainty within a historical-data-driven Bayesian deep-learning framework, though we addressed projected inflow uncertainty more straightforwardly via multi-HM weighting.

Another finding is that although the overall variance in projected VUL decreased under UO-MME-based weighted ensembles, this was accompanied by a relative increase in the contribution of GCM-related uncertainty (Table 2). As previously specified, uncertainty stemming from HM selection strongly affects low-flow extremes and, consequently, VUL outcomes. Since HM-induced variances were already suppressed through the UO-MME approach, the residual variance contribution was redistributed to other sources, i.e., mostly GCMs and ESs (Fig. 7d). Moreover, several studies have emphasized that relying solely on the severity of failure may lead to misleading quantification of VUL (e.g., Mateus and Tullos, 2017). Hence, it is not unexpected that efforts to reduce upstream modeling uncertainty have a relatively limited contribution to managing the projected uncertainty associated with the VUL, which in this study reflects the likely magnitude of operation deficits. In this regard, the influence of UO-MME on alternative vulnerability indices, particularly those that account for the system's capacity to respond to varying hydrological conditions, should also be explicitly discussed. Still, all findings derived from this study do not necessarily imply that weighted streamflow projections always bring about reduced uncertainty in projected RPMs, irrespective of reservoir operation policy employed. Such limitations were particularly detected under SOP, where the same weighting schemes failed to provide significant reduction in projection uncertainty despite some performance gains.

4.3. Temporal evolution of optimized reservoir releases across different sources

As in Fig. 8, Fig. S10 presents comparative boxplots of projected release anomalies, this time based solely on the projections under HDG-2d. The figure displays the relative changes in annual mean releases with respect to the historical baseline (HIST) counterparts across four future time windows for three ensemble configurations, including full-chain, BMA, and UO-MME. According to Fig. S10, all ensemble variations suggest no substantial release anomaly during the period of 2021–2039, while their median lines in the following periods show a gradual downward trend. From 2060 onward, the severity of decreases in release anomalies intensifies across all variants, accompanied by an evident widening of the interquartile ranges, which indicate a growing projection uncertainty over time windows. Moreover, the BMA-based ensemble mirrors its limited impact observed for RPMs on release anomalies as well. That is, it does not exhibit a marked tendency to change these anomalies compared to those in full-chain projections. More notably, UO-MME-based ensembles consistently yielded higher medians (i.e., greater than −10%) and, most visibly, mitigated the anomaly decrease by nearly 5% relative to the full-chain case during 2080–2099. Just as the integration of UO-MME with two-dimensional hedging model made the projected RPMs more narrowly spread, its benefit was likewise obvious in the temporal patterns of projected release anomalies, which exhibited more compact ensemble distributions than those of other configurations.

Although operating adaptive rule curves with dynamically weighted flow projections has mitigated the reductions in reservoir releases to a certain degree, it is also clear that most anomalies across GCMs are still negative and become stark temporally (Fig. 9). Additionally, EARTH and MPI under RCP4.5 give neutral values during 2021–2079, standing out as benign cases. Similarly, CNRM shows stable release anomalies under both scenarios, consistent with its optimistic runoff anomaly patterns (Fig. 3). Under RCP8.5, the increasing temporal prominence of negative anomalies is evident for GFDL and HadGEM, with drastic decreases of roughly 30% and 25% by 2060–2079, respectively. These two GCMs also reveal a more pessimistic trend under RCP4.5, albeit with milder anomalies. Interestingly, MPI under RCP8.5 shifts from minor early-century anomalies to a dramatic 26% decline by 2080–2099, reflecting a similar decreasing pattern in its projected runoff. These results indicate that, despite the application of adaptive policies, the substantial influence of GCMs and ESs (i.e., key sources of runoff uncertainty) on reservoir releases persists, in line with Raje and Mujumdar (2010), who drew alike conclusion using stochastic dynamic programming.

To unravel how the uncertainty change in optimized reservoir releases propagates over time, Fig. 10 presents a year-by-year decomposition of uncertainty contributions from each modeling chain component for projected release anomalies under different ensemble simulations. As illustrated in Fig. 10a, the fraction of variance attributable to HM selection remained limited, averaging around 5% over the whole projection period, since release anomalies were evaluated on an annual basis. According to Fig. S6, however, indices such as VUL are understandably more affected by HM selection owing to their sensitivity to baseflow simulations of these models, which are obtained on a monthly basis. As seen in Fig. 10b, eliminating the HM-based variance component—constituting only a small fraction of total uncertainty—through BMA led to a largely neutral shift in the variance structure.

When evaluated within each ensemble configuration, the total variance in these anomalies during the late window (2061–2099) was found to be nearly 3.5 times that of the preceding period (Fig. 10). But this proportional observation cannot be directly attributable to the choice of ensemble weighting scheme. Similarly, the variance ratio for the late window to the earlier part was in the range of 2.5–2.9 for GCM-related uncertainty, and 12–13 for ES-related uncertainty, with minimal variation across ensemble types. Therefore, these escalations are likely due to primary climate modeling drivers (i.e., ESs and GCMs), whose contributions to projected inflow dynamics and the net evaporation regime have temporally intensified (Figs. 2-3). In this context, it has also been reported that models parameterizing potential evaporation as a function of temperature may exaggerate hydrological responsivity to climate change (Clark et al., 2016). Alongside this, discrepancies in evaporative loss signals across two RCPs may have amplified the role of radiative forcing–driven variability in projected release anomalies, independent of weighting of HMs.

Moreover, the extent of the influence of UO-MME-based ensembles on temporally decomposed variances in Fig. 10 appeared somewhat different from that observed for the projected RPMs, which is plausible given that these are compound performance indicators representing the entire future horizon of 2021–2099. Since the variances in release anomalies are typically more sensitive to interannual fluctuations in

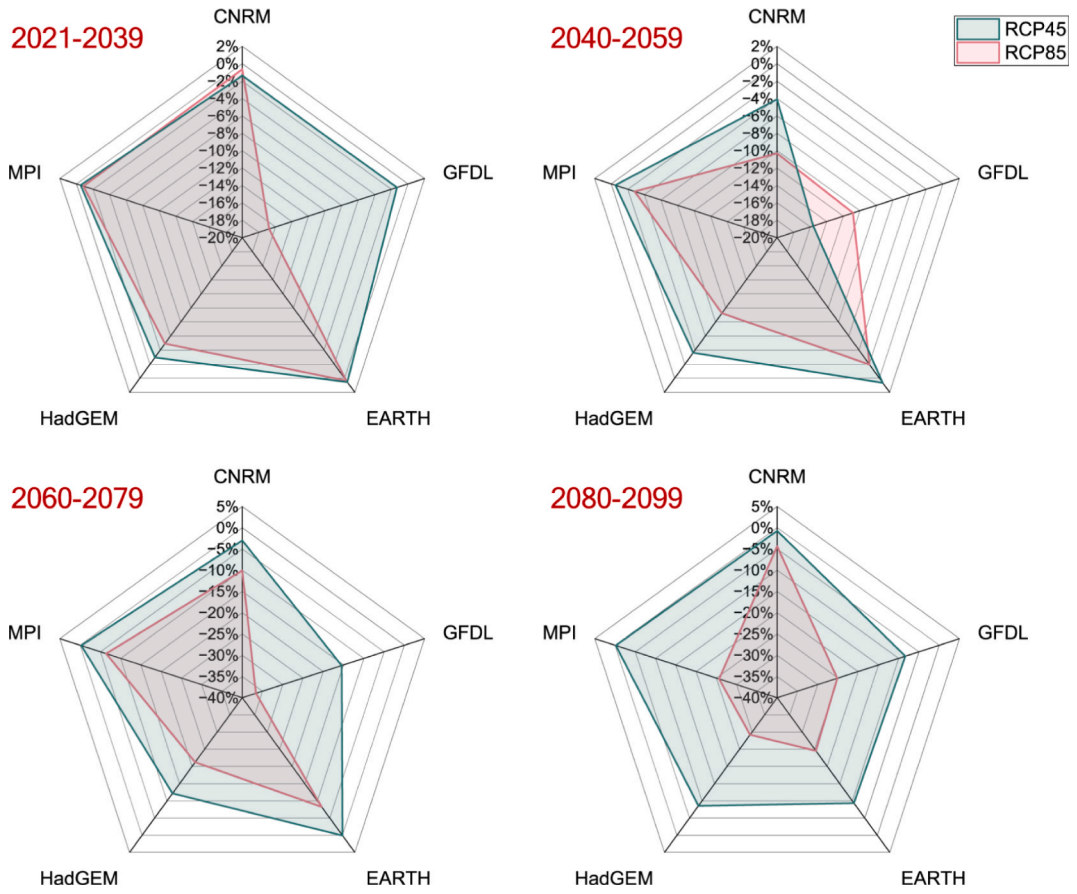


Fig. 9. Projected anomalies in optimized annual mean reservoir releases relative to the historical baseline, considering only hedging-based outputs under dynamically weighted inflow ensembles. The given values represent average anomalies derived from statistically and dynamically downscaled data, presented for each GCM under two RCP scenarios. Each radar chart illustrates the results for the five GCMs within the corresponding future horizon.

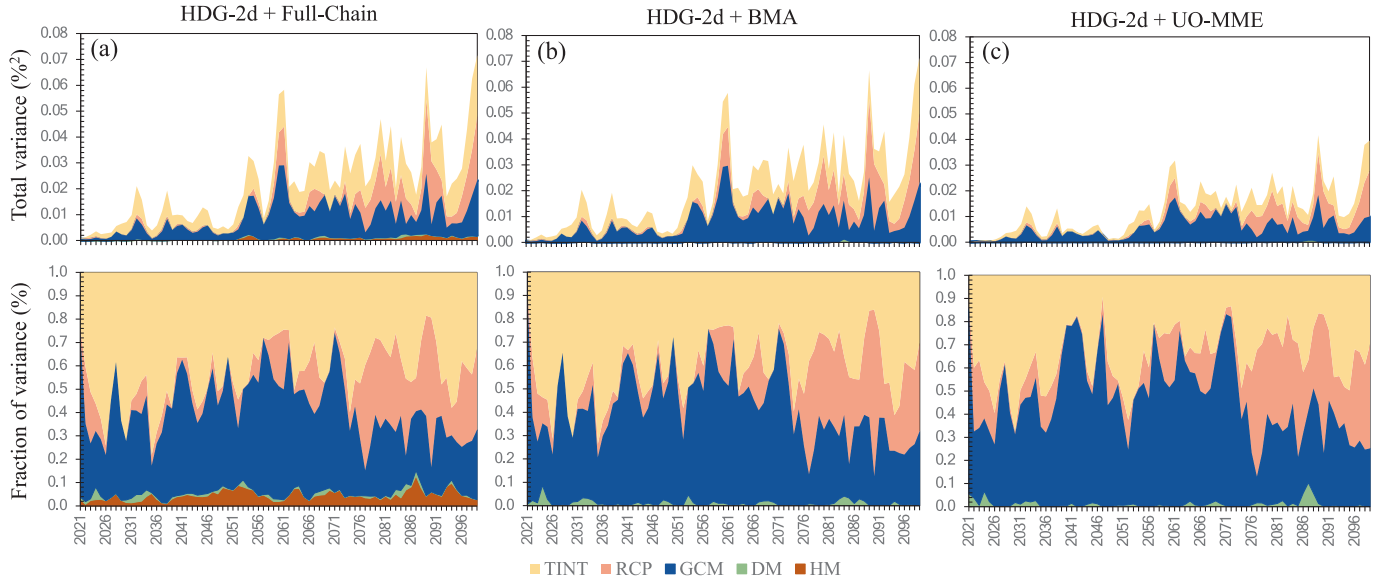


Fig. 10. Temporal evolution of decomposed variances (top panels) and corresponding variance fractions (bottom panels) for projected reservoir release anomalies derived under the HDG-2d operation rule. Variance contributions are compiled for three ensemble configurations, where panels (a), (b), and (c) represent the results for full-chain, BMA-weighted, and UO-MME–weighted ensembles, respectively. Uncertainty decompositions are shown with respect to the main uncertainty sources and their total interaction (TINT).

projected inflows (e.g., Nourani et al., 2025), and the corresponding uncertainties in those projections have already been narrowed by UO-MME in a significant manner, the overall uncertainty reduction for VUL and SI remained less pronounced compared to that computed for the release anomalies. In turn, the implemented strategy is much more influential for temporally evolving outputs like annual release anomalies, unlike RPMs that do not directly capture intra-annual variability. Fig. 10c demonstrates that using UO-MME-based ensembles leads to a prominent reduction in all projection uncertainty components within the full modeling chain, albeit to varying degrees. Accordingly, the results indicate that upstream uncertainty treatment has translated into a 23–42% and 11–36% narrowing of GCM-related and ES-related uncertainties, respectively, in projected reservoir releases. Furthermore, the meaningful tempering of interaction terms-induced uncertainty has collectively contributed to an overall contraction of total variance within the range of 40–52%. To summarize, the incorporation of dynamically weighting the multi-member hydrological ensemble, by which narrowing uncertainty in projected runoff was effectively made, into hedging rule optimization led to more tangible enhancements in the RPMs as well as tighter uncertainties in both aggregate performance indices and time-variant reservoir outputs.

It can also be noted that the relative percentage contributions of the uncertainty sources shown in Fig. 10c differ from those obtained for the performance indicators. While RCP-related uncertainty accounts for only about one-fifth of the total variance in RLS for long-term, its contribution nearly doubles in the case of VUL. This contrast shows that vulnerability responds more directly to emissions-driven differences, whereas long-term releases are influenced more markedly by the cumulative interactions among GCM, RCP, and DM factors, with total interaction contributing 35% of the variance. Within the developed framework, although the dominance of the uncertainty source changes according to RPMs, the significant alleviation of overall variance remains a promising outcome.

4.4. Opportunities and challenges for reducing uncertainty in reservoir operation projections

In characterizing the hydrological impacts of climate change, Clark et al. (2016) underscored not only the quantification of uncertainties but also their integration into the decision-making process for uncertainty-aware operational strategies. In line with that perspective, our study combines uncertainty-reduced streamflow projections with an adaptive hedging model, thereby aiming to reflect the approach they advocate. Although the HDG-2d model is observed to be more reliable in avoiding severe release failures when coupled with appropriately weighted inflow ensembles, its own parameter optimization was not evaluated as a contributing source to projection uncertainty. For instance, the HDG-2d model, when utilized in multi-objective settings, can give varying parameter sets (Bayesteh and Azari, 2021). Besides, implicit-deterministic reservoir operation policies under different inflow ensembles may respond differently to climate change signals (Celeste and Billib, 2009). Therefore, an in-depth characterization of reservoir operation projections along with uncertainty decomposition of alternative reservoir operation optimization models, which have different parameterizations, would address a critical gap for institutional applications.

The study emphasizes that the hedging procedure, based on UO-MME-weighted ensembles rather than using 140 inflow projections, is a more adaptive measure, and its clear contribution to uncertainty reduction compared to standard BMA (Figs. 7–10). Yet, it should be noted that entropy-based BMA (En-BMA), which leads to better representation of predictive uncertainty relative to the standard one, has been implemented for postprocessing precipitation forecasting (Darbandsari and Coulibaly, 2022). Hence, operating the HDG-2d model under En-BMA-based hydrological projections may result in further reductions in uncertainty in RPMs and is worth evaluating against UO-MME in a future study.

In this study, HDG-2d turns out to be the primary contributor to uncertainty reduction in reservoir operation projections, with UO-MME playing a complementary role (Fig. 7). However, the relative effectiveness of these two processes may differ depending on the reservoir characteristics and data sources. Furthermore, no matter how refined the modeling chain is, an irreducible level of uncertainty originating from GCMs persists in projected reservoir releases. This can be partially related to interannual variability effect, whereas our study considered the traditional ANOVA methodology, in which interaction term involve internal variability (e.g., Vetter et al., 2015; Chawla and Mujumdar, 2018; Wang et al., 2020). Another assumption is that GCMs are generally assumed to be equiprobable in multi-model ensembles (Collins, 2017). Hence, applying a nested weighting strategy to both HMs and GCMs (especially those from the CMIP6 archive) may yield more interpretable variance structures when integrated with adaptive hedging models. Finally, the implications for reservoir operation projections of the differences between ANOVA type that seeks to isolate interannual variability from other sources of uncertainty (e.g., Aitken et al., 2023) and those following the latter tradition are planned to be addressed in a future study.

The selection of an adapted potential evaporation equation as input for HMs is also a challenge. While the Penman–Monteith (PM) equation is often referred to as a physically based formulation for estimating reference evapotranspiration, the biases of the estimations from different empirical formulas compared to those of PM can be compensated through HM calibration, adjusting model parameters in response to the inputs (Seiller and Anctil, 2016). In line with this, Oudin et al. (2005) also concluded that various potential evaporation formulations lead to similar runoff simulations. Nevertheless, these choices can meaningfully contribute to runoff projection uncertainty, especially when GCM-driven temperature change ranges are large (Bae et al., 2011). For example, PM and radiation-based formulations may exhibit different responses to future hydroclimatic conditions compared to the temperature-based approach adopted in this study, indirectly affecting reservoir performance metrics such as VUL. Therefore, incorporating multiple empirical potential evaporation methods into the modelling chain that drives reservoir operation optimization appears worthwhile to explore.

Another source of limitation is the assumption of constant water demand used in the reservoir operation models. Although a fixed monthly demand choice (10.67 Mm^3) is operationally consistent with the planning strategy of the Tahtali Reservoir and aligns with previous studies using non-varying demand trajectories (e.g., Celeste and Billib, 2009), this simplification does not account for potential demographic growth or climatic factors that may change water demand in the future. For reservoirs with irrigation purposes, such issues become more critical because demand is driven by projected crop evapotranspiration data (Zamani et al., 2017; Saedi et al., 2021). Since these factors may alter the way uncertainties propagate through reservoir performance metrics, future work should incorporate demand-side uncertainty into the total modelling chain, including climate-induced irrigation requirements under crop patterns and demographic projections as well. Such extensions would allow the integrated framework combining hydrological model weighting and climate-informed hedging rules to be evaluated under additional sources of uncertainty and enable a more comprehensive assessment of system robustness.

5. Concluding remarks

This study explores the potential of combining weighted streamflow projections from a dynamically constructed UO-MME framework with an adaptive hedging model termed HDG-2d to reduce projection uncertainties in climate-informed reservoir outputs. In this respect, the integration is consistent with recent state-of-the-art reviews about climate change adaptation (e.g., Elgendy et al., 2024). Although the

paper employed the Tahtali Reservoir in Türkiye as a case study, the implemented strategy constitutes a transferable procedure for projecting both credible and optimal reservoir releases under climate change impacts in other dams worldwide. The key implications of our study are as follows:

- Upstream uncertainty reduction followed by integration of downstream adaptive hedging, a previously unaddressed topic to the best of our knowledge, offers a powerful alternative to applications based on the use of several traditional approaches such as SOP and BMA. While BMA is a post-processing approach that relies on a statically weighted combination of predictive distribution functions obtained from different competing HMs, the UO-MME dynamically updates HM weights by solving a bi-objective optimization problem that simultaneously evaluates hydrological simulation performance trajectories and narrows runoff projection uncertainty propagated through the modeling chain, particularly from GCMs. Likewise, the HDG-2d provides climate-informed operational flexibility, unlike SOP-type methods that cannot adapt to evolving hydroclimatic conditions. Therefore, this novel, dual-structured strategy allows for consistent treatment of modeling chain uncertainties that exert cumulative pressure on surface water reservoirs.
- Relative to SOP-based operation results, HDG-2d model proves to be the primary contributor to uncertainty reduction in reservoir operation projections, while UO-MME method contributes meaningfully as a secondary factor. Compared to hedging runs using unweighted inflow ensembles, UO-MME-based weighted inflow ensembles led to a consistent reduction in the long-term uncertainty variance of projected release anomalies by about 43% across the whole period. Uncertainty reduction was also noticeable but slightly less pronounced in the RPMs projected under the same weighted data, while performance improvements also remained visible, particularly for VUL (95% of the variants remained below 0.25). All these outcomes correspond to a 36% reduction in SI uncertainty, as well as maintaining SI values nearly 0.3 higher than those obtained under SOP-based variants. By contrast, no substantial performance gains were obtained from the BMA-weighted inflow ensembles, though a mild uncertainty reduction was achieved for projected reservoir outputs.
- The study emphasizes the importance of HM weighting in generating refined, decision-relevant information, in line with recent studies (e.g., Pastén-Zapata et al., 2022). Even though such approaches have not yet been directly applied to real-world reservoir management, the resulting reduction in runoff projection uncertainty clearly has tangible operational implications. In particular, narrowing variances enables more reliable operational decisions for both seasonal and annual release planning, which may help improve drought preparedness. In this regard, to temper long-term water shortages and provide low-cost adaptation and mitigation strategies under uncertain climate conditions, it would be beneficial for institutions or decision-makers to avoid relying solely on traditional, non-hedging operational rules and instead adopt uncertainty-aware methodological frameworks, as suggested in this study.

Given the key implications outlined above, we plan to further develop this study by applying entropy-based weighting to CMIP6-focused GCM ensembles to better address model-induced irreducible uncertainties. In addition to such an extension, a potential cascade framework would also involve multiple reservoir operation optimization models (see Celeste and Billib, 2009), thereby enabling a more elaborate decomposition of projection uncertainties throughout the reservoir operation process. However, such a cascade framework may present several challenges. Implementing multiple reservoir operation models needs consistent calibration procedures, which can lead to higher optimization costs and longer run times. Furthermore, integrating entropy-based weighting into a multi-model operation cascade would significantly increase the dimensionality of the modeling chain, thus complicating uncertainty decomposition. From a methodological perspective, these interactions may also introduce substantial structural variability arising from differences among hedging models, which could exceed parameterization-related variability and further inflate certain variance components.

Furthermore, our study, carried out on a single-reservoir case, has relatively similar motivations to the few studies that evaluate climate-tailored policies in multi-reservoir systems using a large ensemble of hydrological projections (e.g., Espanmanesh et al., 2024). In fact, the quantitative indicators obtained in this study strongly suggest that the proposed framework is suitable for direct integration into the operational management of Tahtali Reservoir. Accordingly, such integrated frameworks are expected to support the authorities (e.g., General Directorate of State Hydraulic Works of Türkiye) that face difficulties due to competition in water allocation and seek to update their operational rule curves. Yet, how the proposed uncertainty-aware framework would respond in complex multi-reservoir systems or in single reservoirs operated for different purposes, including environmental-flow supply (e.g., Dash et al., 2023), where RPMs exhibit varied trade-offs, is a topic worthy of consideration for future studies.

CRediT authorship contribution statement

Zeynep Beril Ersoy: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Okan Fistikoglu:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Umut Okkan:** Writing – review & editing, Writing – original draft, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Funding

This study was funded by the Scientific and Technological Research Council of Türkiye under Grant No. 122Y083.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Umut Okkan reports financial support was provided by Scientific and Technological Research Council of Türkiye.

Acknowledgments

The authors would like to thank the Scientific and Technological Research Council of Türkiye (TÜBİTAK) for providing financial support for this research. The authors also express their gratitude to the editors and all reviewers for their valuable comments that helped significantly improve the quality of this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2026.135228.

Data availability

Data will be made available on request.

References

Adeloye, A.J., Soundharajan, B.S., Ojha, C.S., Remesan, R., 2016. Effect of hedging-integrated rule curves on the performance of the pong reservoir (India) during scenario-neutral climate change perturbations. Water Resour. Manag. 30 (2), 445–470. https://doi.org/10.1007/s11269-015-1171-z.

Adeloye, A.J., Dau, Q.V., 2019. Hedging as an adaptive measure for climate change induced water shortage at the Pong reservoir in the Indus Basin Beas River, India. Sci. Total Environ. 687, 554–566. https://doi.org/10.1016/j.scitotenv.2019.06.021.

Aitken, G., Beevers, L., Parry, S., Facer-Childs, K., 2023. Partitioning model uncertainty in multi-model ensemble river flow projections. Clim. Change 176 (11), 153. https://doi.org/10.1007/s10584-023-03621-1.

Ahmadianfar, I., Zamani, R., 2020. Assessment of the hedging policy on reservoir operation for future drought conditions under climate change. Clim. Change 159 (2), 253–268. https://doi.org/10.1007/s10584-020-02672-y.

Anvari, S., Moghaddasi, M., Bagheri, M.H., 2023. Drought mitigation through a hedging-based model of reservoir-farm systems considering climate and streamflow variations. Theor. Appl. Climatol. 152, 723–737. https://doi.org/10.1007/s00704-023-04402-7.

Bae, D.H., Jung, I.W., Lettenmaier, D.P., 2011. Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin, Korea. J. Hydrol. 401 (1–2), 90–105. https://doi.org/10.1016/j.jhydrol.2011.02.012.

Bayesteh, M., Azari, A., 2021. Stochastic optimization of reservoir operation by applying hedging rules. J. Water Resour. Plan. Manage. 147 (2), 04020099. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001312.

Beshavard, M., Adib, A., Ashrafi, S.M., Kisi, O., 2022. Establishing effective warning storage to derive optimal reservoir operation policy based on the drought condition. Agric Water Manag 274, 107948. https://doi.org/10.1016/j.agwat.2022.107948.

Castaneda-Gonzalez, M., Poulin, A., Romero-Lopez, R., Turcotte, R., 2023. Hydrological models weighting for hydrological projections: the impacts on future peak flows. J. Hydrol. 625, 130098. https://doi.org/10.1016/j.jhydrol.2023.130098.

Celeste, A.B., Billib, M., 2009. Evaluation of stochastic reservoir operation optimization models. Adv. Water Resour. 32 (9), 1429–1443. https://doi.org/10.1016/j.advwatres.2009.06.008.

Chawla, I., Mujumdar, P.P., 2018. Partitioning uncertainty in streamflow projections under nonstationary model conditions. Adv. Water Resour. 112, 266–282. https://doi.org/10.1016/j.advwatres.2017.10.013.

Clark, M.P., Wilby, R.L., Gutmann, E.D., Vano, J.A., Gangopadhyay, S., Wood, A.W., et al., 2016. Characterizing uncertainty of the hydrologic impacts of climate change. Curr. Clim. Change Rep. 2, 55–64. https://doi.org/10.1007/s40641-016-0034-x.

Collins, M., 2017. Still weighting to break the model democracy. Geophys. Res. Lett. 44 (7), 3328–3329. https://doi.org/10.1002/2017GL073370.

Coron, L., Andréassian, V., Perrin, C., Bourqui, M., Hendrickx, F., 2014. On the lack of robustness of hydrologic models regarding water balance simulation: a diagnostic approach applied to three models of increasing complexity on 20 mountainous catchments. Hydrol. Earth Syst. Sci. 18 (2), 727–746. https://doi.org/10.5194/hess-18-727-2014.

Darbandsari, P., Coulibaly, P., 2019. Inter-comparison of different Bayesian model averaging modifications in streamflow simulation. Water 11 (8), 1707. https://doi.org/10.3390/w11081707.

Darbandsari, P., Coulibaly, P., 2022. Assessing entropy-based Bayesian model averaging method for probabilistic precipitation forecasting. J. Hydrometeorol. 23 (3), 421–440. https://doi.org/10.1175/JHM-D-21-0086.1.

Dash, S.S., Sahoo, B., Raghuwanshi, N.S., 2023. An integrated reservoir operation framework for enhanced water resources planning. Sci. Rep. 13 (1), 21720. https://doi.org/10.1038/s41598-023-49107-z.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comp. 6 (2), 182–197. https://doi.org/10.1109/4235.996017.

Elgendy, M., Hassini, S., Coulibaly, P., 2024. Review of climate change adaptation strategies in water management. J. Hydrol. Eng. 29 (1), 03123001. https://doi.org/10.1061/JHYEFF.HEENG-6014.

Erfanian, A., Azizipour, M., Jalali, M.R., Afshar, A., Soleimanian, E., 2023. Operational policy development for dynamic restoration of lakes in a changing climate; application of innovative hedging rules in a system dynamics platform. Appl. Water Sci. 13 (9), 174. https://doi.org/10.1007/s13201-023-01983-6.

Ersoy, Z.B., Fistikoglu, O., Okkan, U., 2025a. Exploring the potential of multi-hydrological model weighting schemes to reduce uncertainty in runoff projections. Water 17 (20), 2919. https://doi.org/10.3390/w17202919.

Ersoy, Z.B., Fistikoglu, O., Okkan, U., Derin, B., 2025b. Convergence and final performances of optimization algorithms for rainfall-runoff model calibration based on the number of function calls. Earth Sci. Inform. 18 (2), 382. https://doi.org/10.1007/s12145-025-01885-y.

Espanmanesh, V., Guilpart, E., Bourgault, M.A., Tilmant, A., 2024. Adapting reservoir operation to climate change in regions with long-term hydrologic persistence. Clim. Risk Manag. 45, 100623. https://doi.org/10.1016/j.crm.2024.100623.

Eum, H.I., Simonovic, S.P., 2010. Integrated reservoir management system for adaptation to climate change: the Nakdong River Basin in Korea. Water Resour. Manag. 24, 3397–3417. https://doi.org/10.1007/s11269-010-9612-1.

Gorguner, M., Kavvas, M.L., 2020. Modeling impacts of future climate change on reservoir storages and irrigation water demands in a Mediterranean basin. Sci. Total Environ. 748, 141246. https://doi.org/10.1016/j.scitotenv.2020.141246.

Kharrufa, N.S., 1985. Simplified equation for evapotranspiration in arid regions. Beiträge Zur Hydrologie 5 (1), 39–47.

Knutti, R., Sedláček, J., Sanderson, B.M., Lorenz, R., Fischer, E.M., Eyring, V., 2017. A climate model projection weighting scheme accounting for performance and interdependence. Geophys. Res. Lett. 44 (4), 1909–1918. https://doi.org/10.1002/2016GL072012.

Labadie, J.W., 2004. Optimal operation of multireservoir systems: State-of-the-art review. J. Water Resour. Plann. Manag. 130 (2), 93–111. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:2(93).

Lemaitre-Basset, T., Oudin, L., Thirel, G., Collet, L., 2022. Unraveling the contribution of potential evaporation formulation to uncertainty under climate change. Hydrol. Earth Syst. Sci. 26 (8), 2147–2159. https://doi.org/10.5194/hess-26-2147-2022.

Liu, Y., Qin, H., Zhang, Z., Yao, L., Wang, Y., Li, J., Liu, G., Zhou, J., 2019. Deriving reservoir operation rule based on Bayesian deep learning method considering multiple uncertainties. J. Hydrol. 579, 124207. https://doi.org/10.1016/j.jhydrol.2019.124207.

Luo, C., Xu, B., Ding, W., Zhang, C., Du, E., 2023. Characteristics of reservoirs to mitigate drought effects with a hedging rule triggered by drought limited water level. J. Hydrol. 617, 129106. https://doi.org/10.1016/j.jhydrol.2023.129106.

Mateus, M.C., Tullos, D., 2017. Reliability, sensitivity, and vulnerability of reservoir operations under climate change. J. Water Resour. Plan. Manage. 143 (4), 04016085. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000742.

Moghaddasi, M., Anvari, S., Akhondi, N., 2022. A trade-off analysis of adaptive and non-adaptive future optimized rule curves based on simulation algorithm and hedging rules. Theor. Appl. Climatol. 148 (1), 65–78. https://doi.org/10.1007/s00704-022-03930-y.

Nguyen, H., Mehrotra, R., Sharma, A., 2020. Assessment of climate change impacts on reservoir storage reliability, resilience, and vulnerability using a multivariate frequency bias correction approach. Water Resour. Res. 56 (2), e2019WR026022. https://doi.org/10.1029/2019WR026022.

Nourani, V., Najafi, H., Nikofar, B., Huang, J.J., 2025. Optimizing reservoir operations with Z-numbers: addressing uncertainty and reliability. J. Hydrol. 656, 132903. https://doi.org/10.1016/j.jhydrol.2025.132903.

Okkan, U., Fistikoglu, O., Ersoy, Z.B., Noori, A.T., 2023. Investigating adaptive hedging policies for reservoir operation under climate change impacts. J. Hydrol. 619, 129286. https://doi.org/10.1016/j.jhydrol.2023.129286.

Okkan, U., Inan, G., 2015. Bayesian learning and relevance vector machines approach for downscaling of monthly precipitation. J. Hydrol. Eng. 20 (4), 04014051. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001024.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C., 2005. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. J. Hydrol. 303 (1–4), 290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026.

Ozturk, T., Turp, M.T., Türkeş, M., Kurnaz, M.L., 2018. Future projections of temperature and precipitation climatology for CORDEX-MENA domain using RegCM4.4. Atmos. Res. 206, 87–107. https://doi.org/10.1016/j.atmosres.2018.02.009.

Panagoulia, D., Dimou, G., 1997. Linking space-time scale in hydrological modelling with respect to global climate change. Part 2. Hydrological response for alternative climates. J. Hydrol. 194, 38–63. https://doi.org/10.1016/S0022-1694(96)03221-0.

Pastén-Zapata, E., Pimentel, R., Royer-Gaspard, P., Sonnenborg, T.O., Aparicio-Ibañez, J., Lemoine, A., Pérez-Palazón, M.J., Schneider, R., Photiadou, C., Thirel, G., Refsgaard, J.C., 2022. The effect of weighting hydrological projections based on the robustness of hydrological models under a changing climate. J. Hydrol. Reg. Stud. 41, 101113. https://doi.org/10.1016/j.ejrh.2022.101113.

Raje, D., Mujumdar, P.P., 2010. Reservoir performance under uncertainty in hydrologic impacts of climate change. Adv. Water Resour. 33 (3), 312–326. https://doi.org/10.1016/j.advwatres.2009.12.008.

Rotbeei, F., Nuri Balov, M., Safari, M.J.S., Vaheddoost, B., 2025. A CMIP6-based drought assessment over Küçük Menderes Basin. Türkiye. Theor. Appl. Climatol. 156 (6), 331. https://doi.org/10.1007/s00704-025-05552-6.

Saedi, F., Ahmadi, A., Abbaspour, K.C., 2021. Optimal water allocation of the Zayandeh-Roud Reservoir in Iran based on inflow projection under climate change scenarios. J. Water Clim. Change 12 (5), 2068–2081. https://doi.org/10.2166/wcc.2021.219.

Sandoval-Solis, S., McKinney, D.C., Loucks, D.P., 2011. Sustainability index for water resources planning and management. J. Water Resour. Plan. Manage. 137 (5), 381–390. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000134.

Seiller, G., Anctil, F., 2016. How do potential evapotranspiration formulas influence hydrological projections? Hydrol. Sci. J. 61 (12), 2249–2266. https://doi.org/10.1080/02626667.2015.1100302.

Soleimani, S., Bozorg-Haddad, O., Loáiciga, H.A., 2016. Reservoir operation rules with uncertainties in reservoir inflow and agricultural demand derived with stochastic dynamic programming. J. Irrig. Drain. Eng. 142 (11), 04016046. https://doi.org/10.1061/(ASCE)IR.1943-4774.0001065.

Thiha, S., Shamseldin, A.Y., Melville, B.W., 2023. Optimal reservoir operation using the improved multi-step-ahead time-varying hedging rule under climate and land-use changes. Hydrol. Sci. J. 68 (8), 1139–1161. https://doi.org/10.1080/02626667.2023.2196427.

Thomas, T., Ghosh, N.C., Sudheer, K.P., 2021. Optimal reservoir operation–a climate change adaptation strategy for Narmada basin in central India. J. Hydrol. 598, 126238. https://doi.org/10.1016/j.jhydrol.2021.126238.

Tong, Y., Gao, X., Han, Z., Xu, Y., Xu, Y., Giorgi, F., 2021. Bias correction of temperature and precipitation over China for RCM simulations using the QM and QDM methods. Clim. Dyn. 57, 1425–1443. https://doi.org/10.1007/s00382-020-05447-4.

Tu, M.Y., Hsu, N.S., Tsai, F.T.C., Yeh, W.W.G., 2008. Optimization of hedging rules for reservoir operations. J. Water Resour. Plan. Manage. 134 (1), 3–13. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:1(3).

Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., et al., 2011. The representative concentration pathways: an overview. Clim. Change 109 (1), 5. https://doi.org/10.1007/s10584-011-0148-z.

Vetter, T., Huang, S., Aich, V., Yang, T., Wang, X., Krysanova, V., Hattermann, F., 2015. Multi-model climate impact assessment and intercomparison for three large-scale river basins on three continents. Earth Sys. Dyn. 6 (1), 17–43. https://doi.org/10.5194/esd-6-17-2015.

Wanders, N., Wood, E.F., 2016. Improved sub-seasonal meteorological forecast skill using weighted multi-model ensemble simulations. Environ. Res. Lett. 11 (9), 094007. https://doi.org/10.1088/1748-9326/11/9/094007.

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), e2020EF001602. https://doi.org/10.1029/2020EF001602.

Yagbasan, O., 2016. Impacts of climate change on groundwater recharge in Küçük Menderes River Basin in Western Turkey. Geodin. Acta 28 (3), 209–222. https://doi.org/10.1080/09853111.2015.1121802.

Zamani, R., Akhond-Ali, A.M., Ahmadianfar, I., Elagib, N.A., 2017. Optimal reservoir operation under climate change based on a probabilistic approach. J. Hydrol. Eng. 22 (10), 05017019. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001559.