ARTICLE IN PRESS

Scientific Data

https://doi.org/10.1038/s41597-026-06865-5

Article in Press

A century long ensemble streamflow dataset in the Pacific Northwest to support water security assessments

Received: 29 September 2025

Accepted: 9 February 2026

Cite this article as: Mizukami, N., Gutmann, E.D., Wood, A.W. *et al.* A century long ensemble streamflow dataset in the Pacific Northwest to support water security assessments. *Sci Data* (2026). https://doi.org/10.1038/s41597-026-06865-5

Naoki Mizukami, Ethan D. Gutmann, Andrew W. Wood, Bart Nijssen, Jane M. Harrell, Christopher D. Frans, Michael D. Warner & Chanel Mueller

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

A century long ensemble streamflow dataset in the Pacific Northwest to support water security assessments

Naoki Mizukami[1], Ethan D. Gutmann[1], Andrew W. Wood[1], Bart Nijssen[2], Jane M. Harrell[3], Christopher D. Frans[4], Michael D. Warner[3], Chanel Mueller[5]

1 NSF National Center for Atmospheric Research, Boulder, CO, USA
2 University of Washington, Civil and Environmental Engineering, Seattle, WA, USA
3 U.S. Army Corps of Engineers, Seattle District, USA
5 Bureau of Reclamation, Research and Development Office, CO, USA
4 U.S. Army Corps of Engineers, Headquarters, Washington, DC, USA

Corresponding author: N. Mizukami (mizukami@ucar.edu)

ARTICLE IN PRESS

ARTICLE IN PRESS

Abstract

In the Pacific Northwest, ensemble projected streamflow datasets are valuable for assessing vulnerabilities and the resilience of reservoir systems across the region. These datasets are typically generated using a climate-hydrologic modeling chain that begins with coarse-resolution outputs from selected Earth System Model (ESMs) and socioeconomic scenarios, which are spatially downscaled, followed by hydrologic simulations driven by the downscaled meteorological forcing. In this work, a calibrated catchment-based hydrology and river model is forced by ESM outputs for several socioeconomic scenarios from the archives of the Coupled Model Intercomparison Project phases 5 and 6, which are downscaled using a computationally efficient weather model. The dataset comprises twenty-nine daily hydrologic traces from 1950 to 2099 for 18,000 river reaches. The dataset also includes a retrospective hydrologic simulation forced by an observation-based meteorological dataset used for hydrologic model calibration. Naturalized flows for 221sites are used to assess historical simulation fidelity. This dataset supports a variety of applications including reservoir modeling, ecological impact assessments, and hydrologic analyses under historical to projected climate conditions.

Background & Summary

Long-term observational meteorological datasets indicate a significant warming trend in the Pacific Northwest during the latter part of the 20th century[1]. Their impacts on snowpack and runoff, such as an earlier timing shift of the annual peak, have been documented accordingly[2–6]. These trends have been projected to continue throughout the 21st century[7,8].

Under such changing hydrologic conditions, long-term water resources planning is critical. In the Pacific Northwest, more than 250 reservoirs serve multiple purposes including power generation, irrigation, recreation, and flood risk management. Traditional approaches to long-term water resources planning rely primarily on historical hydrologic observations, supplemented with statistical scaling applied to historical events where synthetic events are derived from statistical extrapolation of the historical record. However, such approaches do not fully incorporate the process understanding of hydrological variability driven by climate dynamics and potential future change.

ARTICLE IN PRESS

With new information on the meteorological projections through the recent Coupled Model Intercomparison Projects (CMIPs), and improved modeling capability, recent studies[9–12] generated hydrologic projection datasets for the Pacific Northwest using a climate-hydrology modeling chain[13]. The climate-hydrology modeling chain begins with the selection of ESMs and socioeconomic scenarios, followed by spatial downscaling of each of the ESM outputs, before running the selected hydrologic models. Bias-corrections are often performed at multiple stages in this modeling chain, e.g. before and/or after spatial downscaling of the ESM outputs. Hydrologic model outputs are often bias-corrected where streamflow observations are available. The outcomes of the modeling chain are applied by end users for a wide range of purposes, for example reservoir modeling to assess vulnerability under future projected conditions. Individual elements in the modeling chain introduce uncertainties associated with the "methodological choices", such as selection of ESMs, scenarios, and model methodological choices for downscaling methods and hydrologic models[14]. For example, Chegwidden et al.[12] selected two emission scenarios, ten CMIP-Phase 5 (CMIP5) ESMs, two statistical downscaling methods, and four process-based hydrologic model configurations to generate a 160-member ensemble of streamflow projections over the Pacific Northwest. Their study provided insights into the sources of the projection uncertainty. In their study, the choice of emission scenario and ESM had the largest impact on regional hydrologic projections, though the hydrological model choice strongly affected low flow projections.

Although previous applications of climate-hydrology modeling chains have contributed to an improved process-understanding of future hydrologic change in the Pacific Northwest, most of these modeling chains implement statistical downscaling methods to spatially downscale ESM outputs. Statistical downscaling methods are computationally inexpensive sufficiently to apply multiple ESM outputs (e.g., 20 outputs of ESM-scenario combination over the Pacific Northwest by Chegwidden et al.[12]). Statistical-downscaling methods popular for hydrologic projection studies are 1) Bias-Correction and Spatial Downscaling (BCSD)[15], 2) Constructed Analogues (CA)[16,17] and 3) Localized Constructed Analogues (LOCA)[18]. These statistical methods share several well-documented limitations[17,19]. First, the statistical method assumes stationarity. That is, statistical relationships identified during a historical period remain the same under future climates. This is unlikely for century-long projections of variables such as temperature. Second, the above statistical methods downscale meteorological variables separately; therefore may not

ARTICLE IN PRESS

preserve the climate model's projected coherence among meteorological fields. Though multi-variate downscaling methods[20,21] have been recently developed to mitigate this caveat, these methods have not been used practically for the hydrologic projection studies. Finally, some statistical downscaling methods applied at monthly resolution prefer to resample and perturb sequences from the historical record rather than downscaling daily meteorological sequences that are based on the atmospheric outputs of the very coarse-scale ESMs. This practice may introduce artifacts such as an unrealistic frequencies or sequence of precipitation events, which can adversely affect hydrologic simulations[22–24]. Sequences of meteorological events affecting hydrological processes are as critical to water resources infrastructure vulnerability as the impacts of individual events. These limitations are not universal, however; for instance, Gutmann et al.[22] shows that the monthly BCSD method can produces realistic sequences that scale well spatially.

Our study contrasts a few elements of the climate-hydrologic modeling chain with the previous CMIP based hydrologic projection study[12] to generate a new ensemble of streamflow projections in the Pacific Northwest. First, we leverage a computationally efficient quasi-dynamical downscaling capability instead of statistically methods, specifically using meteorological forcing data downscaled by the Intermediate Complexity Atmospheric Research model (ICAR)[25]. This ICAR-downscaled data preserves the weather event sequences from the ESMs and provides process-based atmospheric fields and meteorological variables at finer resolution, offering a novel forcing for our projected hydrologic simulations. Second, we use catchment-based hydrologic modeling, which is calibrated with hundreds of sites that provide naturalized flow data to adjust the model parameters over 84% of the domain. The river routing model is also catchment-based, representing the river reaches and catchments more accurately than traditional gridded river models, including small basins. Finally, the dataset of hydrologic projections includes six ESMs from CMIP5 with two emission scenarios for each ESM and six ESMs from CMIP-Phase 6 (CMIP6) with three emission scenarios for each ESM. Including both CMIPs for the hydrologic projection simulations allow a user to carefully evaluate difference between two CMIPs in hydrologic changes.

The uncertainty of the simulated projection is based solely on multiple ESMs and scenarios. Though a use of multiple hydrologic models may reveal hydrologic model uncertainty[14], well-calibrated models may reduce hydrologic model uncertainty[26], therefore the effort of this study is

ARTICLE IN PRESS

made toward the hydrologic model calibration. We also do not consider meteorological forcing uncertainty based on ESM's initial conditions (i.e., ensemble members). Recently, downscaled large ensemble ESMs datasets[27,28] that incorporate uncertainty due to ESM, scenarios, and ensemble members have been developed based on statistical downscaling methods for the continental US domain. The use of these datasets for hydrologic projection studies is on-going effort. We also do not include any statistical downscaling methods; therefore, direct comparison between dynamical and statistical methods are not possible.

As in all the past studies, our hydrologic modeling aims to simulate the natural hydrologic conditions without human influences such as reservoir operations, irrigations, and inter-basin transfer. By minimizing confounding anthropogenic effects, such dataset enables clearer attribution of hydrologic variability and change to climate forcing and facilitate more robust model evaluation across regions. From a water resources perspective, the dataset serves as a counterfactual baseline for quantifying the human alterations, including streamflow changes, ecosystem impacts, and long-term water availability. For example, the current practice of water resources uses the naturalized streamflow simulations as an input into their impact models such as reservoir models for their planning, where future management strategies and infrastructure are uncertain. Hence, the dataset is valuable both for climate change impact assessments and planning applications.

ARTICLE IN PRESS

Methods

Climate-hydrologic modeling chains

Overall climate-hydrologic modeling chains are illustrated in Fig. 1. Our climate-hydrology modeling chain consists of 26 different ESM and scenario combinations, one downscaling method and one hydrologic model configuration, therefore, uncertainty quantification can be made based on ESM and scenario choices. The simulation period is from 1950 to 2099, split into historical periods (1950-2004 for CMIP5 and 1950-2014 for CMIP6) and future scenario periods (2005-2099 for CMIP5 and 2015-2099 for CMIP6). As a final post-processing step, our study also added streamflow bias correction using an open-source software package called bmorph[29].

ARTICLE IN PRESS

Our specific choices used for each modeling chain component in Fig. 1 are described in the following subsections. Because this paper focuses on streamflow simulations, we discuss the ICAR downscaling only briefly and focus on the hydrologic model description.

Figure 1 inserted

ESM selection

ESMs simulate global weather patterns, with the model's own evolution of natural variability based on simulated surface and atmospheric states. Thus, each model simulation has its own sequence of natural variability for historical and future periods. The CMIP5 and CMIP6 each host climate outputs simulated using more than 30 ESMs. For most of these ESMs, multiple simulations are available representing different combinations of initial conditions and future CO_2 concentrations from different possible global emissions scenarios. Table 1 shows all the ESMs and socioeconomic scenarios included in this study. There are 11 and 15 historical through future climate traces from CMIP5 and CMIP6 ESMs, respectively. The CMIP6 ESM selections are based on ranking based on skill scores evaluated against multiple observed dataset[30], as well as availability of all the vertical profile dataset necessary for ICAR simulations. The CMIP5 ESM selections were based on a combination of skill scores, literature review, and past usage assessments[31]. For each ESM, one ensemble member is selected, therefore uncertainty due to internal variability is not accounted for.

Table 1. List of the 26 ESM-scenario pairs selected from the CMIP5 and CMIP6 archives including the specific ensemble member and model resolution. ESMs from CMIP6 (CMIP5) archive are blue (red) highlighted.

ESMs	Scenarios	ESM Resolution
CanESM5	SSP245, 370, 585	1.875° x 1.25°
CMCC-CM2-SR5	SSP245, 370, 585	0.942° x 1.25°
MIROC-ES2L	SSP245, 370, 585	1.865° x 1.875°
MPI-ESM1-2-LR	SSP245, 370, 585	2.791° x 2.813°
NorESM2-MM	SSP245, 370, 585	1.25° x 0.9375°
CanESM2	RCP45, 85	2.023° x 2.5°
CCSM4	RCP85	0.94° x 1.25°
CMCC-CM	RCP45, 85	1.5° x 2°
CNRM-CM5	RCP45, 85	1.4° x 1.4°
MIROC5	RCP45, 85	1.401° x 1.406°

ARTICLE IN PRESS

MRI-CGCM3	RCP45, 85	1.121° x 1.125°

Quasi-dynamical downscaling method

While the physical process representations of ESMs have become more detailed, CMIP ESM simulations use spatial resolutions of 50-200 km which do not adequately capture finer-scale processes such as local interactions of the atmosphere with the land surface and its topography. Using coarse ESM outputs as boundary conditions for complex, process-based, finer-scale numerical weather models, termed dynamical downscaling, is generally accepted as the gold standard to generate consistent, finer-scale solutions of atmospheric fields in regions with complex terrain and is ideal for hydrological applications, particularly for the representation of extreme events.

However, computational demands of dynamical downscaling still prohibit thousands of model years simulation, even for regional domains like the Pacific Northwest. ICAR simulates atmospheric physical processes similar to those in the widely used Weather Research and Forecasting model (WRF)[32] for mountainous areas, but with greater computational efficiency[25]. ICAR is over 100 times faster than comparable WRF simulations for the same domain; however, to archive this efficiency, ICAR simplifies the solution for the dynamics by relying on linear mountain wave theory. The model has been validated against simulations using WRF and against PRISM[33] estimates of surface meteorology over the US: ICAR explains 85-95% of the variability represented in WRF in complex terrain, and ICAR's precipitation fields are highly correlated with observations. The simplifications in ICAR mean that it does not represent non-linear atmospheric dynamics such as blocking or explicit convection; ICAR relies on the Tiedtke cumulus parameterization to simulate convective precipitation as in coarser resolution WRF simulations. In these simulations, ICAR used the Thompson-Eidhammer microphysics parameterization[34], the Noah-MP land model[35], a simplified planetary boundary layer scheme, the RRTMG radiation code[36] for longwave radiation, and an empirical formulation for shortwave radiation[37]. As in Currier et al[31], the ESM three-dimensional fields are first interpolated to the ERA-Interim[38] three-dimensional grid and then performed a bias correction to match the ERA-Interim data on that grid, then bias-corrected ESMs are downscaled to 6-km spacing grid. This harmonizes the boundary conditions across ESMs and mitigates some of the disparities in the underlying ESM grids and biases.

ARTICLE IN PRESS

Although ICAR simulates all the atmospheric variables at the surface needed for the hydrologic model: precipitation, temperature, specific humidity, short- and long-wave radiations, wind, pressure at an hourly time step, we use the MetSim program[39] to estimate all the meteorological forcing data needed for the hydrologic model from ICAR's daily aggregated precipitation and temperature data. MetSim uses the MTClim v3.4[40] algorithm to disaggregate daily precipitation and temperature (maximum and minimum) time series to a sub-daily time step and also estimates humidity, shortwave, and longwave radiations at the same sub-daily time step. MetSim is run at the hydrologic model's spatial discretization, using ICAR daily precipitation and temperature remapped from the 6km grid data. This approach ensures that the final forcing for the hydrologic model is consistent with the observational-based retrospective meteorological dataset used for hydrologic model calibration and for the retrospective hydrologic model simulations. Before performing the daily disaggregation, we compare historical ICAR downscaled daily temperature and precipitation to the observational dataset, then adjust ICAR temperature and precipitation for the entire 150-year simulation period to reduce biases between the downscaled data and observational data using quantile mapping.

Hydrologic modeling

We use land surface hydrology and river routing models that run in sequence to produce streamflow across the Pacific Northwest domain. The Structure for Unifying Multiple Modeling Alternatives (SUMMA)[41,42] simulates land-surface water and energy fluxes. MizuRoute[43] routes the total runoff simulated with SUMMA to produce streamflow estimates at many locations in the river network. Hereafter in this paper, the modeling framework is called SUMMA-mizuRoute. The initial SUMMA-mizuRoute model configuration and parameters used in this study were developed in prior SUMMA and mizuRoute applications projects[44–46], for which a western US domain SUMMA-mizuRoute model, forcing dataset, and parameter estimation (i.e., calibration) approach for streamflow had been created. The details of this default implementation (which had not been specifically calibrated for the Pacific Northwest) are described further below.

SUMMA enables a hierarchical hydrologic simulation for spatial units in which grouped response units (GRUs) of arbitrary shape (grid or polygon) may enclose smaller hydrologic response units (HRUs), also of arbitrary shape and size. The west-wide SUMMA modeling from

ARTICLE IN PRESS

which the Pacific Northwest subset model adopts a configuration in which each GRU contains only one HRU, and the majority of these are specified using US Geological Survey Hydrologic Unit Code Level 12 (HUC12; median area: 59 $km^2 \pm 16$ km^2 standard deviation). Canadian portions of the domain that extend beyond the HUC12 dataset are drawn from the global vectorized river hydrography dataset (MERIT-basin)[47]. This meso-scale representation of regional hydrology has approximately the spatial resolution of previous gridded hydroclimate datasets that used 1/16th-1/8th degrees, or 36-144 km^2, and was chosen to balance the computational demand (of century-scale ensemble projection and prediction applications) with a need to represent the spatial heterogeneity of the region.

SUMMA computes the water and energy states and fluxes for each HRU at a 3-hour time step, based on meteorological forcings and physiographic characteristics including slope, elevation, soil depth, vegetation and soil type. SUMMA offers several water and energy flux parameterization schemes (i.e., process algorithms). In this study as in the earlier ones, we use the Ball-Berry parameterization[48] for simulating stomatal resistance, one of the main physiological factors controlling transpiration, similar to the Noah-MP land surface model[35]. We also adopt a logarithmic wind profile below the vegetation canopy, described in Mahat et al.[49], and implement the parameterization for vegetation roughness length and displacement height by Raupach[50]. We use Beer's law[51], similar to the Variable Infiltration Capacity (VIC) model[52] to compute the radiative transfer through vegetation. For the vertical moisture transmission in the soil column, the mixed form of the Richards equation[53] is used with a vertically constant hydraulic conductivity. For snow, we use a constant albedo decay rate, and the thermal conductivity is parameterized using the approach by Jordan[54]. Further, each HRU incorporates an unconfined aquifer at the bottom of the soil column, which contributes to baseflow generation. Soil column depth varies by HRU based on depth to bedrock data[55], which was an update for this study (prior applications had used a fixed depth of 1.5 m). No lateral water movement occurs between HRUs in SUMMA. However, SUMMA computes "delayed runoff" or hillslope routed runoff, by accounting for travel time of total runoff (surface and subsurface runoffs) to the river channel using the gamma distribution-based unit hydrograph. The gamma distribution provides flexibility of unit-hydrograph distribution that allows us to calibrate travel time of runoff from several hours to multiple days. However, one limitation is that this does not capture much slower groundwater recharge to river. Water fluxes and states, including evapotranspiration (ET), snow

ARTICLE IN PRESS

water equivalent (SWE) and volumetric soil moisture, are output at a daily time step, except for routed runoff which is output at a 3-hour time step. This sub-daily runoff is input into mizuRoute for river routing, which is performed at a 3-hour time step.

MizuRoute[43], as applied in this study, uses a vector-river network based on the MERIT-basin river network dataset. The Pacific Northwest domain subset of the larger western US model includes 17,957 river reaches (median length 6 km $\pm$ 7 km standard deviation) and their associated contributing catchments or HRUs (median area: 37 km^2 $\pm$ 36 km^2 standard deviation). Since SUMMA's HRUs and mizuRoute's MERIT-basin HRUs differ in size, runoff is remapped within mizuRoute using a mass conserving averaging approach. MizuRoute routes runoff through each river reach in the order defined by the river network topology or downstream-upstream reach connectivity. Among several reach routing schemes offered by mizuRoute, we use unit hydrograph-based routing as this is the most computationally efficient method. This scheme uses an unique unit hydrograph (or impulse response function) for each reach, derived from diffusive wave equation with two physical parameters (slope and length) of reaches as well as two hydraulic parameters (wave celerity and diffusivity), to route the inflow from upstream reach(es) to the downstream end of reach. Outflow of each reach, which is an output of mizuRoute as river discharge at each reach, is computed as a sum of local lateral flow and routed inflow from upstream reach(es).

Hydrologic model parameter estimations

While hydrologic model calibration is straightforward for smaller basin studies where river gauge data is available[56], it is difficult for large domains with limited observations. We have previously used techniques such as multi-scale parameter regionalization[57], which uses transfer functions to relate geophysical characteristics to model parameters. In that case, we calibrate the parameters of the transfer functions, which results in parameter fields that mimic the spatial organization of the geophysical characteristics. However, for most model parameters, the form of the transfer function and their geophysical predictors are largely unknown, resulting in ad-hoc transfer function formulations and sub-optimal performance in the calibrated model[58]. Here we regionalize individual basin calibrations, an approach that improves the retrospective streamflow simulations at many target locations, but which also creates a patchwork pattern of hydrologic states and fluxes (i.e., SUMMA outputs). Calibration is performed for the period 1991-10-01 and

ARTICLE IN PRESS

2001-09-30, which includes the driest and wettest years within the 1970-2014 period for which we have retrospective meteorological forcing data used for the calibration.

The SUMMA calibration is conducted in two stages: first, the selected reference flow sites within a region are calibrated simultaneously (combined basin calibration) to provide improved *a-priori* parameter fields. For this calibration, the Pacific Northwest domain is split into four regions that are hydro-climatologically similar (Cascade, eastern Cascade, Snake, and Upper Columbia) and several sites are selected for each region (Fig. 2a). For the combined basin calibration, regional parameter multipliers, which are applied to adjust spatially distributed SUMMA parameters, are calibrated so that the simulations improve at the selected naturalized flow sites. The goal of this stage of calibration is not to obtain optimal parameters at each site, but to provide improved parameters that could serve as prior estimates for the second stage. Following the earlier applications of SUMMA-mizuRoute[44–46], we use a calibration workflow that applies the Dynamical Dimensional Search algorithm (DDS)[59] via the multi-method general parameter optimization program called Ostrich[60], using 100 iterations. The calibrated SUMMA parameters are shown in Table 2. The selection of the calibration parameters is based one-at-a-time method that analyzes the effect of one parameter on the error metric at a time, keeping the other parameters fixed, at 91 CAMELS (Catchment Attributes and Meteorology for Large-sample Studies)[61] sites within the Pacific Northwest. Calibrated regional parameter multipliers are applied to all the HUC12 HRUs inside each region to update the parameter fields. Compared to the prior applications, this study adds parameters of *tempCritRain*, *vGN_n*, and *windReductionParam* as calibration parameters, based on sensitivity analyses.

The second stage of the calibration focuses on individual basin calibrations, following a nested basin approach that proceeds from upstream to downstream[45,46]. For this step, we first calibrate the headwater upstream areas of 65 naturalized flow sites, denoted by red circles in Fig. 2b. We then continue calibrating the areas between each headwater site and the next downstream site while keeping the upstream parameters fixed. This downstream basin calibration continues till the most downstream flow site, which is near the mouth of the Columbia River. We used 33 naturalized flow sites for the downstream calibrations, denoted by the orange triangles in Fig. 2b. Ultimately, we used a total of 98 flow sites for individual basin calibration, covering 84% of the entire domain (red and orange areas in Fig. 2b). Areas upstream of each flow site consist of many HUC12 HRUs and we adjusted the same parameters as the combined basin calibration

ARTICLE IN PRESS

(i.e., first stage calibration) for each HUC12 HRU by applying parameter multipliers determined using DDS with 300 iterations. Due to the stochastic nature of DDS algorithms[59], some basins show limited improvement of the objective function. For basins with less than 10% improvement of the objective function, additional calibrations are performed by re-running DDS with a different random seed. An objective function for the calibration uses modified version of Kling-Gupta efficiency (KGE′)[62], which is scaled or normalized as Eq. 1 so that the value ranges from 0 to 1:

$$n\mathrm{KGE}' = \frac{1}{2-KGE'}$$

$$\mathrm{KGE}' = 1 - \sqrt{(1-\alpha)^2 + (1-\beta)^2 + (1-r)^2} \qquad (1)$$

$$\alpha = \frac{CV_{obs}}{CV_{sum}}, \beta = \frac{\mu_{obs}}{\mu_{sum}}$$

where μ_{obs} and μ_{sim} are the means of the reference and simulated flows, respectively; CV_{obs} and CV_{sim} are the coefficients of variation (ratio of standard-deviation to mean) of reference and simulated flows, respectively; and *r* is the Pearson correlation coefficient between the reference and simulated flows. The objective function (i.e., all the KGE′ components) is computed with daily aggregated streamflow as the reference flow is given at daily step. Note that modified KGE, which uses the coefficient of variance as a measure of flow variability, instead of the original KGE[63] which uses the standard-deviation. nKGE′ is used for only calibration, and all the evaluations of streamflow simulations with the calibrated models use KGE′.

Figure 2 inserted

Table 2. List of the SUMMA parameters included in the calibration.

Parameter names	Process	Descriptions	min, max
Fcapil	snow	capillary retention as a fraction of the total pore volume (-)	0.009, 0.110
tempCritRain	meteorology	critical temperature above which precipitation is rain (K)	269, 275
frozenPrecipMultip	meteorology	frozen precipitation multiplier (-)	0.5, 1.5
routingGammaScale	Hillslope routing	scale parameter in Gamma distribution used for sub-grid routing (s)	360, 72000
routingGammaShape	Hillslope routing	shape parameter in Gamma distribution used for sub-grid routing (-)	1.0, 5.0
k_soil	soil	saturated hydraulic conductivity (m s^{-1})	10^{-7}, 10^{-4}

ARTICLE IN PRESS

theta_sat	soil	porosity (-)	0.3, 0.6
critSoilTranspire	soil	critical volumetric liquid water content below which transpiration is limited (-)	0.1, 1.0
qSurfScale	soil	scaling factor in the surface runoff parameterization (-)	1, 100
aquiferBaseflowExp	soil	baseflow exponent (-)	1, 10
aquiferBaseflowRate	soil	baseflow flux when aquifer storage is equal to aquiferScaleFactor (m s^{-1})	10^{-7}, 0.1
vGn_n	soil	van Genuchten "n" parameter (-)	1.0, 3.0
summerLAI	vegetation	maximum leaf area index at peak of growing season (m^2 m^{-2})	0.01, 10.0
heightCanopyTop	vegetation	height of top of the vegetation canopy above ground surface (m)	0.05, 100
windReductionParam	vegetation	canopy wind reduction parameter (-)	0.0, 1.0

Streamflow bias corrections in river network

The calibration of SUMMA parameters results in reducing bias (β component in KGE′) in historical streamflow simulations over the Pacific Northwest. However, the calibrated SUMMA-mizuRoute simulations still show errors because of shortcomings in the model conceptualizations and parameter settings. Additional errors in simulated streamflow appear when the calibrated SUMMA-mizuRoute models are forced with the downscaled ESM outputs. One source of the error is due to differences between observation-based forcing and downscaled ESMs forcing sequences that lead to different hydrologic states such as seasonal snow accumulation and soil moisture state.

To reduce the streamflow bias in SUMMA-mizuRoute simulation forced by the downscaled ESMs, we use an open-source software package called bmorph[29]. The bmorph bias correction used 200 sites for the CDF matching training during the period between 1981-01-01 and 2000-12-31, which provide complete reference flow data without any missing values. Also, this period includes the driest and wettest year for 45 years (1970-2014) based on GMET data.

The bmorph package first constructs two cumulative distribution functions (CDF) of long-term observed and simulated streamflow, which are used to correct the simulated flows by quantile-mapping from the simulated to observed flows at individual flow sites. This is a traditional streamflow bias correction, which poses one shortcoming. This independent bias correction

ARTICLE IN PRESS

(IBC) at flow sites fails to maintain the spatiotemporal consistency of the streamflow simulations at multiple sites across a river network. This can result in inconsistent and nonsensical incremental flows between gauge locations.

As a separate option, bmorph can also adjust flows at non-gauged reaches in the entire river network to maintain spatial consistency. For this bias correction—Spatially Consistent Bias Correction (SCBC), lateral flow, instead of routed flow, is adjusted based on interpolated CDFs of observed flow at reaches between flow sites. Then, adjusted lateral flow is re-routed with mizuRoute over the entire river network to generate adjusted or bias corrected routed flow at each reach. The daily flow from SCBC is a final simulated streamflow dataset in this study.

Reference datasets

Reference streamflow and retrospective meteorological datasets are needed for retrospective hydrologic simulations, hydrologic model calibration, hydrologic model evaluation, and bias correction of the simulated streamflow.

Retrospective meteorological data

We use observation based, 6-km meteorological data generated using the Gridded Meteorological Ensemble Tool (GMET)[64–66] as the retrospective forcing dataset for SUMMA. The GMET forcings are drawn from prior SUMMA-mizuRoute applications[44–46], and are available from 1970 through 2023. The detailed GMET methodology is omitted here. The gridded dataset, including daily maximum and minimum temperature and precipitation, are remapped to each SUMMA HRU using a mass conserving mapping. The remapped GMET daily meteorological data is then processed using the MetSim program in the same way as ICAR downscaled meteorological data to estimate all the forcing data for SUMMA.

For this study, we post-process the GMET forcing dataset before using it as input to MetSim. A comparison of the GMET forcings with PRISM[33] reveals that GMET annual total precipitation was biased up to 15-20% low in some areas relative to PRISM, though such biases varied throughout the domain. To ameliorate this bias and bring the precipitation climatology closer to the widely-used PRISM, we compute monthly GMET and PRISM climatology and then scaled GMET to match the monthly PRISM precipitation climatology at each HRU, with a linearly damped scaling extrapolated to the north of the US border because PRISM does not cover

ARTICLE IN PRESS

Canada. As a result, precipitation in the Canadian portion of the basin is largely unmodified. However, the calibration of SUMMA with the adjusted GMET produces reasonable model behaviors at a few sites in Canada (KGE'=0.89, NSE=0.79 at DCD, and KGE'=0.94, NSE=0.88 at MCD for calibrated flow; See the site locations in Figure 2b). Moreover, even using original GMET, calibration results in quite reasonable calibration results (KGE'=0.86, NSE=0.77 at DCD, and KGE'=0.93, NSE=0.89 at MCD for calibrated flow). This adjusted GMET dataset was used in the calibration of the SUMMA model and was used as the reference dataset for the bias correction of ICAR datasets (Hereafter, GMET dataset refers to this adjusted GMET dataset unless indicated otherwise).

Naturalized historical streamflow data

We use daily "naturalized" or "de-regulated" streamflow data derived at 221 locations for the period from 1928 to 2018 (Fig. 2a) from Bonneville Power Administration, referred to as No Regulation-No Irrigation[67]. The naturalized streamflow dataset is reconstructed by correcting for river regulation, reservoir evaporation and irrigation diversions. Naturalized flow should be distinguished from "observed" streamflow, which is affected by human intervention such as dams and irrigation diversions. We subset the dataset for the period between October 1950 and September 2018 for the data archive because the simulations start at 1950. The dataset is complete from October 1950 through 2018 at 179 sites out of 221 sites. All the sites that include missing data are in the Cascade region.

ARTICLE IN PRESS

Data Records

SUMMA and mizuRoute output data are available from the NSF NCAR Research Archive (https://doi.org/10.5065/8SAK-HM25)[68]. Table 3 summarizes the datasets for each ESM-scenario combination. All the datasets are stored in a single NetCDF file January 1950 through September 2099 (1970-2020 for GMET based output) for each ESM-scenario combination. The streamflow dataset includes lateral flow at MERIT-basin catchment (i.e., SUMMA total runoff remapped to MERIT-basin catchment from HUC12) and river discharge (i.e., routed flow) at each MERIT-basin river reach. An ancillary geospatial dataset including SUMMA HUC12 catchments, MERIT-basin river reaches and catchments, with metadata for naturalized gauge

ARTICLE IN PRESS

locations also available. The daily naturalized flow dataset from October 1951 through September 2018 is provided in NetCDF format.

Table 3. List of the archived data.

Dataset name	Descriptions
<ESM>_<scenario>_mizuRoute_daily.nc	mizuRoute daily mean flow at all the reaches in the river network (daily mean of 3hr raw output from mizuRoute).
<ESM>_<scenario>_mizuRoute_daily_site.nc	mizuRoute daily mean flow at 414 sites[1]
<ESM>_<scenario>_bmorph_site_univariate_daily.nc	bmorph at-site bias corrected flow (IBC, SCBC) and raw simulated flow at 200 sites that were used for bmorph bias correction.
<ESM>_<scenario>_summa_daily.nc	Daily SUMMA outputs including SWE, soil moisture, aquifer storage, canopy water, runoff, and ET at each SUMMA HUC12 catchment.
<ESM>_<scenario>_summa_daily_basin_mean.nc	Daily SUMMA variables averaged over upstream HUC12 catchments for 414 sites
<ESM>_<scenario>_ daily_t_p.nc	Daily max. and min. temperature at each SUMMA HUC12 catchment
<ESM>_<scenario>_ daily_t_p_basin_mean.nc	Daily max. and min. temperature averaged over upstream HUC12 catchments for 414 sites

[1]414 sites include 221 naturalized flow sites used for the model validation and additional 193 sites of US Army Corps of Engineer's interest without Naturalized flow available.

Technical Validation

Historical streamflow simulations

We evaluate daily retrospective streamflow simulated with SUMMA-mizuRoute forced by GMET forcing data during the validation period (2001/10-2014/09) against the naturalized flow data at the 221 sites. Since naturalized flow is daily values, 3-hour streamflow simulations from SUMMA–mizuRoute are aggregated to daily values for the evaluation.

Figure 3 shows the six error metrics for daily streamflow values from SUMMA-mizuRoute at each naturalized flow site. The error metrics are KGE′and its three components; variability ratio (α), mean flow ratio (β), correlation (r) as well as the error metrics of high flow and low flow[69]; percent bias in flow duration curve (FDC) high-segment volume (%biasFHV; high volume is defined as the flow at 98 percentile or above of FDC) and percent bias in FDC low-segment volume (%biasFLV; low volume is defined as flow at 10 percentile or below of FDC). There are noticeable spatial patterns in the error metrics. Overall, the KGE′spatial pattern is influenced by

ARTICLE IN PRESS

that of the correlation coefficient, but sub-regional KGE′ patterns are locally modified by the bias or/and variability error. For example, localized areas with poorer performance include the headwater areas of the Snake River (Upper Snake), Puget Sound basins, and Deschutes (See Fig. 2f). These areas show a larger bias and variability error than other regions, resulting in a lower KGE′. The reasons for these high bias and variability errors are unique to each region. In the Deschutes basin, a deeper aquifer with groundwater generates sustained streamflow during the dry summer season, which is not accounted for in SUMMA. The SUMMA-mizuRoute calibration struggles with reproducing naturalized flow at the headwater sites in the Snake River. Naturalized flows may be more uncertain in regions like the Upper Snake due to challenges estimating the effects of intensive irrigation[70].

Figure 3 inserted

In the Puget Sound area, the large errors are associated with a mismatch in drainage areas between naturalized flow sites and the MERIT-basin network, which is a known challenge in comparing simulated with naturalized or observed flows[71]. Figure 4 illustrates an example of a site located near the upstream end of a reach. MizuRoute outputs discharge at the downstream end of the reach, combining routed inflow from all upstream reaches and local lateral flow from the catchment corresponding to the outlet reach. In contrast, the naturalized flow represents the upstream end of the reach and does not include the contributions from local inflow. In wet and small basins in the Puget Sound area where local lateral inflow contributes significantly to total discharge, mizuRoute can appear to overestimate streamflow at sites located near the upstream end, as shown in the Figs. 4b-c. One possible way to mitigate this impact would be to subtract a proportionate amount of the local lateral HRU inflow from streamflow output at the sites located closer to the upstream end. The proportionate local lateral inflow at the site could be estimated based on a ratio of distance between the upstream end of the reach and the site (L_{site}) to the total reach length (L_R). The flows provided in the dataset do not account for this, and users should be aware of this issue particularly for smaller catchment points.

Figure 4 inserted

Finally, high flow is better simulated than low flow. At 48% of sites, %biasFHV is within 10%, while %biasFLV is within 10% at only 8% of sites. For some sites, daily low flows are underestimated by more than 80% in the SUMMA-mizuRoute model. Low flow simulations are

ARTICLE IN PRESS

more difficult to improve than high flow simulations. For example, Lema et al.[72] calibrated SUMMA using an objective function that balances errors in high-flow and low-flow simulations at several Chilean basins and still struggles with improving low flow simulations though overall hydrograph errors are reduced. Low flow simulations are sensitive to both objective functions[73] and model structures related to subsurface flow generation[74].

Figure 5 shows the downscaled ESM forced simulations preserve similar biases (β), variability errors (α), and high-flow/low-flow error metrics to the errors in the non bias-corrected GMET based streamflow simulation during the validation period (2001/10-2014/09). Correlations (r) between ESM based flow simulations and naturalized flow data should not be evaluated, due to discrepancies in the day-to-day sequences of the ESM forced simulations compared to the observation. This difference arises because dynamical downscaling inherits the unique day-to-day weather sequence from the native ESM model, which is shaped by their unique initial conditions and model physics. Although the streamflow daily sequence derived from ESM forcing is dissimilar to observation, long-term annual cycles computed with ESM forced simulation maintain higher correlation as seen in Fig. 5e. This implies that use of long-term daily scale streamflow seasonality for further analysis or water resources application can be appropriate.

Figure 5 inserted

Bias corrected simulated streamflow

The effectiveness of bmorph bias correction was evaluated by comparing the KGE′ values for the raw and bias-corrected retrospective streamflow series, relative to the naturalized flows. For this evaluation, we used the simulations derived by GMET forcing during the period of 2001-2018. This period was outside the SUMMA calibration period and the period used in bmorph to establish a quantile mapping between the reference and simulated flows.

Figure 6 shows the cumulative distribution functions of KGE′ and its three error components for the raw streamflow simulations and the bias-corrected streamflows across the 221 flow sites. Overall, bmorph bias correction improves the KGE′ values, although the highest KGE′ values do not change much (Fig. 6d). This shift is primarily derived from bias improvements (Fig. 6a) and

ARTICLE IN PRESS

moderate improvement of the variability error. The correlation values are largely similar for the raw and bias corrected streamflow.

Figure 6 inserted

Spatial pattern in SUMMA simulations

Selected daily moisture fluxes and states from the SUMMA simulations are made available for each ESM and scenario. Figure 7 displays annual mean precipitation, evapotranspiration (ET), total runoff and annual maximum SWE during the control period; Water Year (WY)1980-2004 (top panel) as well as changes for two future periods; WY2030-2060 (middle panel) and WY2070-2099 (bottom panel). Ensemble mean annual precipitation increases overall, but there is slight decrease leeward of the Cascade Mountains. In the future, the precipitation type becomes rainfall frequently, and liquid precipitant is less likely to be carried across the mountain ridge by westerly wind than snow, depending on wind speed and precipitation intensity[31]. This is a unique feature in dynamical downscaled precipitation patterns in the mountain regions, not seen in statistical downscaled precipitation fields, and contributing to slight reduction of total runoff in the leeward side of the Cascade Mountains. However, localized areas of decreased runoff in the interior regions such as the Snake River and upper Columbia is likely an artifact of the individual basin calibration, though runoff during the control period is small (less than 20 mm/year). Users should be aware of this patchwork patterns in the hydrologic states and fluxes when analyzing spatial analysis of the SUMMA outputs.

ARTICLE IN PRESS

Figure 7 inserted

Projected streamflow characteristics

Figure 8 shows the mean annual cycle at a daily timestep for the control period (WY1980-2004; black lines) and WY2070-2099 period (red line; RCP8.5 for CMIP5 and SSP5-8.5 for CMIP6 and blue line and RCP4.5 for CMIP5 and SSP2-4.5 for CMIP6) at two sites with distinct climate regimes. Libby Dam on the Kootenai River (top panel) is a snow-dominated river, while winter precipitation is a mix of rain and snow along the Snohomish River (bottom panel). The locations of the sites are shown in Figure 1a. At Libby dam, peak flow during the spring occurs earlier and increases with higher emission scenarios, while runoff during fall becomes more frequent due to more frequent snowmelt. In coastal areas like the Snohomish basin (Fig. 8b), a significant

ARTICLE IN PRESS

increase in streamflow is expected during fall through mid-winter due to a transition from snow to rain under future warming. At the same time, future spring streamflow is expected to decrease due to less snow accumulation. For both basins, future increases in flow magnitude during the high flow season are due to increases in precipitation.

As expected, the peak flow timing (Fig. 8a) is shifted earlier at all sites. The magnitude of this shift is larger in the interior regions where streamflow seasonality is driven by snowmelt than in coastal regions where streamflow is rainfall-dominated, even during winter.

Figure 8 inserted

To characterize the projected streamflow at all the 221 sites, Figure 9 shows the centroid of the annual cycle of daily flow, the mean flow, the annual maximum flow for the control period and their future changes. Future changes in the metrics are evaluated for two future periods: WY2030-2060 and WY2070-2099. and the metrics are mean values of all the ESMs with high emission scenarios (RCP8.5 for CMIP5 forced simulations and SSP5-8.5 for CMIP6 forced simulations). As expected, timing of annual peak flow (Figs. 9a-c) is shifted earlier at all the sites, indicated by earlier days of year for annual hydrograph centroid. The magnitude of this shift is larger in the interior regions where streamflow seasonality is driven by snowmelt rather than the coastal regions where rainfall events occur during winter even on present days. For the change in the flow magnitude, the simulations indicate that the mean flow increases in the future at most of the sites, with greater increases in the interior region (>20% for the late 21st century) than in the western part of the domain such as the Cascade region (~10% for the late 21st century). This pattern is similar to that for precipitation and runoff changes. The Deschutes river basin shows a reduction in annual mean flow, consistent with reductions in precipitation and runoff. However, note that the retrospective simulation with GMET forcing shows large bias and variability errors in the Deschutes River. Compared to the annual mean flow, the annual maximum flow increases to a greater degree across the domain. Particularly, many sites in the Cascade region show increases in annual maximum flow over 60%, compared to an increase in annual mean flow of about 20%.

Figure 9 inserted

One way to visualize the shift in annual extremes is to evaluate changes in the return period for flows of a certain magnitude. Figure 10a shows the change in the return period for a flow that is

ARTICLE IN PRESS

equal to the 20-year flow during the 1954-2004 period at MROW site. At many sites, the return period for the historic 20-year flow decreases to 2-year return period for 8.5 emission scenarios (RCP8.5 and SSP5-8.5; Fig. 10b). The changes in the future return periods are consistent with annual maximum flow changes shown in Figs. 9g-i, though the changes also depend on the shapes of annual peak frequency curves.

Figure 10 inserted

We also compare the projected changes of annual mean and maximum flows from the two previous studies[12,75]. The previous streamflow projections are based on the calibrated VIC forced by BCSD downscaled meteorological data (BCSD-VIC)[12] and uncalibrated VIC forced by LOCA downscaled meteorological data (LOCA-VIC)[75]. Figure 11 shows the maps of %changes in annual mean flow (top panels) and annual maximum flow (bottom panels) for WY2070-2099 under RCP8.5 (5 CMIP5 ESM ensembles: CanESM2, CCSM4, CNRM-CM5, and MIROC5) at 144 sites commonly available for three studies. Our study is more aligned with BCSD-VIC, except that our study overall projects increase in annual maximum flows in the mainstem of the Columbia River while BCSD-VIC shows slight decrease there.

Figure 11 inserted

Usage notes

The dataset is suitable for a variety of applications, including serving as input to water-resource models (e.g., reservoir model) or ecological models (e.g., stream temperature models) or specific studies of long-term streamflow characteristics.

However, based on the evaluation of historical simulations, we note that the dataset shows a tendency to underestimate low flows. As such, applications focusing on low flow conditions such as drought projection or analysis should be approached with caution. Scientists and practical water managers must be aware of another limitation related to ESM derived streamflow simulations and ensure that the dataset is used appropriately for their specific applications. The quasi-dynamical downscaling, like fully dynamical downscaling, preserves day-to-day weather sequences derived from large-scale atmospheric dynamics simulated by ESMs, which evolve their own weather sequences. Therefore, during historical periods, ESMs' weather sequences do

ARTICLE IN PRESS

not match actual observed conditions. This limitation imposes several important considerations for the usage of the streamflow dataset:

- Do not analyze specific historical events. The downscaled weather sequences during the historical period differ from observations, so the dataset should not be used to evaluate specific historical hydrologic events.
- Avoid evaluating daily, monthly, and annual series. Instead, use long-term (decadal or longer) climatological mean.
- Use for long-term hydrologic metrics. The dataset is suitable for computing various hydrologic metrics averaged over long period, including extreme flow metrics such as long-term mean annual maximum flow, frequency of high flow event etc. (See Coxon et al.[76] for various hydrologic metrics).
- Do not use daily streamflow series based on ESM ensemble mean for further analysis or applications. Averaging daily streamflow series across ESMs smooths the data, reducing flow variability and leading to an unreliable evaluation of extreme values.
- Use individual ESM-based simulations for analysis. Hydrologic metrics should be computed individually for each ESM derived simulation before applying statistical summary or ensemble interpretation.

ARTICLE IN PRESS

Data availability

SUMMA and mizuRoute output data are available from the NSF NCAR Research Archive (https://doi.org/10.5065/8SAK-HM25)[68]. This archive includes an ancillary geospatial dataset including SUMMA HUC12 catchments, MERIT-basin river reaches and catchments, naturalized gauge points.

Code availability

All the model codes are available from the following GitHub repositories.

1. SUMMA v3.1.2: https://github.com/CH-Earth/summa/tree/v3.1.2
2. mizuRoute v1.2.3: https://github.com/ESCOMP/mizuRoute/tree/v1.2.3
3. MetSim v2.4.1: https://github.com/UW-Hydro/MetSim/tree/2.4.1
4. Bmorph v1.0.0: https://github.com/UW-Hydro/bmorph/tree/1.0.0

ARTICLE IN PRESS

5. GMET v2.0: https://github.com/NCAR/GMET.
6. ICAR v2.1: https://github.com/NCAR/ICAR/tree/2.1

Examples of python notebooks are available at https://github.com/nmizukami/pnw_hydrology_analysis for majority of the technical validations used in this paper. These notebooks are compatible with the dataset downloaded from https://doi.org/10.5065/8SAK-HM25, and users are free to use and customize the codes as desired.

ARTICLE IN PRESS

ARTICLE IN PRESS

References

1. Abatzoglou, J. T., Rupp, D. E. & Mote, P. W. Seasonal Climate Variability and Change in the Pacific Northwest of the United States. *J. Clim.* **27**, 2125–2142 (2014) https://doi.org/10.1175/JCLI-D-13-00218.1.

2. Mote, P. W., Hamlet, A. F., Clark, M. P. & Lettenmaier, D. P. Declining Mountain Snowpack in Western North America. *Bull. Am. Meteorol. Soc.* **86**, 39–49 (2005) https://doi.org/doi:10.1175/BAMS-86-1-39.

3. Casola, J. H. *et al.* Assessing the impacts of global warming on snowpack in the Washington Cascades. vol. 22 2758–2772 (2009).

4. Hidalgo, H. G. *et al.* Detection and Attribution of Streamflow Timing Changes to Climate Change in the Western United States. *J. Clim.* **22**, 3838–3855 (2009) https://doi.org/10.1175/2009jcli2470.1.

5. Stoelinga, M. T., Albright, M. D. & Mass, C. F. A New Look at Snowpack Trends in the Cascade Mountains. *J. Clim.* **23**, 2473–2491 (2010) https://doi.org/10.1175/2009JCLI2911.1.

6. Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M. & Engel, R. Dramatic declines in snowpack in the western US. *Npj Clim. Atmospheric Sci.* **1**, 2 (2018) https://doi.org/10.1038/s41612-018-0012-1.

7. Payne, J. T., Wood, A. W., Hamlet, A. F., Palmer, R. N. & Lettenmaier, D. P. Mitigating the Effects of Climate Change on the Water Resources of the Columbia River Basin. *Clim. Change* **62**, 233–256 (2004) https://doi.org/10.1023/B:CLIM.0000013694.18154.d6.

8. Rupp, D. E., Abatzoglou, J. T., Hegewisch, K. C. & Mote, P. W. Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA. *J. Geophys. Res. Atmospheres* **118**, 10, 810–884, 906 (2013) https://doi.org/10.1002/jgrd.50843.

9. Elsner, M. *et al.* Implications of 21st century climate change for the hydrology of Washington State. *Clim. Change* **102**, 225–260 (2010) https://doi.org/10.1007/s10584-010-9855-0.

10. Werner, A. T., Schnorbus, M. A., Shrestha, R. R. & Eckstrand, H. D. Spatial and Temporal Change in the Hydro-Climatology of the Canadian Portion of the Columbia River Basin under Multiple Emissions Scenarios. *Atmosphere-Ocean* **51**, 357–379 (2013) https://doi.org/10.1080/07055900.2013.821400.

11. Schnorbus, M. A. & Cannon, A. J. Statistical emulation of streamflow projections from a distributed hydrological model: Application to CMIP3 and CMIP5 climate projections for British Columbia, Canada. *Water Resour. Res.* **50**, 8907–8926 (2014) https://doi.org/10.1002/2014WR015279.

12. Chegwidden, O. S. *et al.* How do modeling decisions affect the spread among hydrologic climate change projections? Exploring a large ensemble of simulations across a diversity of hydroclimates. *Earths Future* **0**, (2019) https://doi.org/10.1029/2018EF001047.

ARTICLE IN PRESS

ARTICLE IN PRESS

13. Bosshard, T. *et al.* Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resour. Res.* **49**, 1523–1536 (2013) https://doi.org/10.1029/2011WR011533.

14. Clark, M. P. *et al.* Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Curr. Clim. Change Rep.* **2**, 55–64 (2016) https://doi.org/10.1007/s40641-016-0034-x.

15. Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. P. Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs. *Clim. Change* **62**, 189–216 (2004) https://doi.org/10.1023/B:CLIM.0000013685.99609.9e.

16. Maurer, E. P., Hidalgo, H. G., Das, T., Dettinger, M. D. & Cayan, D. R. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrol Earth Syst Sci* **14**, 1125–1138 (2010) https://doi.org/10.5194/hess-14-1125-2010.

17. Abatzoglou, J. T. & Brown, T. J. A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.* **32**, 772–780 (2012) https://doi.org/10.1002/joc.2312.

18. Pierce, D. W., Cayan, D. R. & Thrasher, B. L. Statistical Downscaling Using Localized Constructed Analogs (LOCA). *J. Hydrometeorol.* **15**, 2558–2585 (2014) https://doi.org/10.1175/jhm-d-14-0082.1.

19. Fowler H, J., Blenkinsop, S. & Tebaldi, C. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* **27**, 1547–1578 (2007) https://doi.org/10.1002/joc.1556.

20. Li, H., Wang, Y., Huang, G., Tao, W. & Lin, P. Generative Downscaling and Bias Correction of Multivariable Earth System Model Simulations. *Geophys. Res. Lett.* **52**, e2025GL117397 (2025) https://doi.org/10.1029/2025GL117397.

21. Eum, H.-I., Gupta, A. & Dibike, Y. Effects of univariate and multivariate statistical downscaling methods on climatic and hydrologic indicators for Alberta, Canada. *J. Hydrol.* **588**, 125065 (2020) https://doi.org/10.1016/j.jhydrol.2020.125065.

22. Gutmann, E. *et al.* An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.* **50**, 7167–7186 (2014) https://doi.org/10.1002/2014wr015559.

23. Mizukami, N. *et al.* Implications of the Methodological Choices for Hydrologic Portrayals of Climate Change over the Contiguous United States: Statistically Downscaled Forcing Data and Hydrologic Models. *J. Hydrometeorol.* **17**, 73–98 (2016) https://doi.org/10.1175/jhm-d-14-0187.1.

24. Wobus, C. *et al.* Modeled changes in 100 year Flood Risk and Asset Damages within Mapped Floodplains of the Contiguous United States. *Nat. Hazards Earth Syst. Sci.* **2017**, 1–21 (2017) https://doi.org/10.5194/nhess-2017-152.

ARTICLE IN PRESS

25. Gutmann, E., Barstad, I., Clark, M., Arnold, J. & Rasmussen, R. The Intermediate Complexity Atmospheric Research Model (ICAR). *J. Hydrometeorol.* **17**, 957–973 (2016) https://doi.org/10.1175/JHM-D-15-0155.1.

26. Mendoza, P. A. *et al.* Effects of hydrologic model choice and calibration on the portrayal of climate change impacts. *J. Hydrometeorol.* **16**, (2015) https://doi.org/10.1175/JHM-D-14-0104.1.

27. Hartke, S. H. *et al.* GARD-LENS: A downscaled large ensemble dataset for understanding future climate and its uncertainties. *Sci. Data* **11**, 1374 (2024) https://doi.org/10.1038/s41597-024-04205-z.

28. Pierce, D. W., Cayan, D. R., Feldman, D. R. & Risser, M. D. Future Increases in North American Extreme Precipitation in CMIP6 Downscaled with LOCA. *J. Hydrometeorol.* **24**, 951–975 (2023) https://doi.org/10.1175/JHM-D-22-0194.1.

29. Bennett, A., Stein, A., Cheng, Y., Nijssen, B. & McGuire, M. A Process-Conditioned and Spatially Consistent Method for Reducing Systematic Biases in Modeled Streamflow. *J. Hydrometeorol.* **23**, 769–783 (2022) https://doi.org/10.1175/JHM-D-21-0174.1.

30. Lybarger, N. D. *et al.* Improving Earth System Model Selection Methodologies for Projecting Hydroclimatic Change: Case Study in the Pacific Northwest. *J. Geophys. Res. Atmospheres* **129**, e2023JD039774 (2024) https://doi.org/10.1029/2023JD039774.

31. Currier, W. R. *et al.* End-of-Century Changes in Orographic Precipitation with the Intermediate Complexity Atmospheric Research Model over the Western United States. *J. Hydrometeorol.* **26**, 577–595 (2025) https://doi.org/10.1175/JHM-D-24-0071.1.

32. Skamarock, W. C. & Klemp, J. B. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Predict. Weather Clim. Extreme Events* **227**, 3465–3485 (2008) https://doi.org/10.1016/j.jcp.2007.01.037.

33. Daly, C. *et al.* Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int J Clim.* **28**, 2031–2064 (2008) https://doi.org/10.1002/joc.1688.

34. G. Thompson & T. Eidhammer. A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J Atmos Sci* **71**, 3636–3658 (2014).

35. Niu, G.-Y. *et al.* The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res. Atmospheres* **116**, (2011) https://doi.org/10.1029/2010JD015139.

36. M. J. Iacono *et al.* Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J Geophys Res* **113**, D13103 (2008).

37. J. W. Finch & M. J. Best. The accuracy of downward short- and long-wave radiation at the Earth's surface calculated using simple models. *Meteor Appl* **11**, 33–39 (2004).

ARTICLE IN PRESS

38. D. P. Dee. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart J Roy Meteor Soc* **137**, 553–597 (2011).

39. Bennett, A., Hamman, J. & Nijssen, B. MetSim: A Python package for estimation and disaggregation of meteorological data. *J. Open Source Softw.* **5**, 2042 (2020) https://doi.org/10.21105/joss.02042.

40. Thornton, P. E. & Running, S. W. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agric. For. Meteorol.* **93**, 211–228 (1999) https://doi.org/Doi:%252010.1016/s0168-1923(98)00126-9.

41. Clark, M. P. *et al.* A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.* https://doi.org/10.1002/2015wr017198 (2015) doi:10.1002/2015wr017198 https://doi.org/10.1002/2015wr017198.

42. Clark, M. P. *et al.* A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resour. Res.* https://doi.org/10.1002/2015wr017200 (2015) doi:10.1002/2015wr017200 https://doi.org/10.1002/2015wr017200.

43. Mizukami, N. *et al.* MizuRoute version 1: A river network routing tool for a continental domain water resources applications. *Geosci. Model Dev.* **9**, 2223–2238 (2016) https://doi.org/10.5194/gmd-9-2223-2016.

44. Broman, D. P. & Wood, A. W. *Better Representation of Low Elevation Snowpack to Improve Operational Forecasts, Final Report No. ST-2019-178-01 to the Science and Technology Program*. https://www.usbr.gov/research/projects/download_product.cfm?id=3099 (2021).

45. Wood, A. W., Sturtevant, J., Barrett, J. L. & Llwewellyn, D. *Improving the Robustness of Southwestern US Water Supply Forecasting, Final Report NO. ST-2018-8117-01 to the Science and Technology Program,*. https://www.usbr.gov/research/projects/download_product.cfm?id=3029, (2021).

46. Sturtevant, J., Wood, A. W., Llewellyn, D. & Barrett, L. Benchmarking analog and ensemble-based reservoir system inflow prediction strategies for seasonal water management and planning. *J Amer Wat Res Assoc* (2025).

47. Lin, P. *et al.* Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. *Water Resour. Res.* **55**, 6499–6516 (2019) https://doi.org/10.1029/2019WR025287.

48. Ball, J. T., Woodrow, I. E. & Berry, J. A. A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis under Different Environmental Conditions. in *In:Biggins, J. (eds) Progress in Photosynthesis Research* (Springer, Dordrecht, 1987).

49. Mahat, V., Tarboton, D. G. & Molotch, N. P. Testing above- and below-canopy representations of turbulent fluxes in an energy balance snowmelt model. *Water Resour. Res.* **49**, 1107–1122 (2013) https://doi.org/10.1002/wrcr.20073.

ARTICLE IN PRESS

50. Raupach, M. R. Simplified expressions for vegetation roughness length and zero-plane displacement as functions of canopy height and area index. *Bound.-Layer Meteorol.* **71**, 211–216 (1994) https://doi.org/10.1007/BF00709229.

51. Mahat, V. & Tarboton, D. G. Canopy radiation transmission for an energy balance snowmelt model. *Water Resour. Res.* **48**, (2012) https://doi.org/10.1029/2011WR010438.

52. Liang, X., Lettenmaier, D. P., Wood, E. F. & Burges, S. J. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.* **99**, 14415–14428 (1994) https://doi.org/10.1029/94jd00483.

53. Celia, M. A., Bouloutas, E. T. & Zarba, R. L. A general mass-conservative numerical solution for the unsaturated flow equation. *Water Resour. Res.* **26**, 1483–1496 (1990) https://doi.org/10.1029/WR026i007p01483.

54. Jordan, R. A one-dimensional temperature model for a snow cover. *Technical documentation for SNTHERM. 89, Special Report* 16–91 (1991).

55. Shangguan, W., Dai, Y., Duan, Q., Liu, B. & Yuan, H. A global soil data set for earth system modeling. *J. Adv. Model. Earth Syst.* **6**, 249–263 (2014) https://doi.org/10.1002/2013ms000293.

56. Bennett, K. E., Werner, A. T. & Schnorbus, M. Uncertainties in Hydrologic and Climate Change Impact Analyses in Headwater Basins of British Columbia. *J. Clim.* **25**, 5711–5730 (2012) https://doi.org/10.1175/jcli-d-11-00417.1.

57. Samaniego, L., Kumar, R. & Attinger, S. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.* **46**, W05523 (2010) https://doi.org/10.1029/2008wr007327.

58. Mizukami, N. *et al.* Towards seamless large-domain parameter estimation for hydrologic models. *Water Resour. Res.* https://doi.org/10.1002/2017WR020401 (2017) doi:10.1002/2017WR020401 https://doi.org/10.1002/2017WR020401.

59. Tolson, B. & Shoemaker, C. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.* **43**, (2007) https://doi.org/10.1029/2005WR004723.

60. Matott, L. S. OSTRICH: an Optimization Software Tool, Documentation and User's Guide. (2017).

61. Addor, N., Newman, A. J., Mizukami, N. & Clark, M. P. The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrol Earth Syst Sci* **21**, 5293–5313 (2017) https://doi.org/10.5194/hess-2017-169.

62. Kling, H., Fuchs, M. & Paulin, M. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* **424–425**, 264–277 (2012) https://doi.org/10.1016/j.jhydrol.2012.01.011.

ARTICLE IN PRESS

63. Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **377**, 80–91 (2009) http://dx.doi.org/10.1016/j.jhydrol.2009.08.003.

64. Bunn, P. T. W. *et al.* Improving Station-Based Ensemble Surface Meteorological Analyses Using Numerical Weather Prediction: A Case Study of the Oroville Dam Crisis Precipitation Event. *J. Hydrometeorol.* **23**, 1155–1169 (2022) https://doi.org/10.1175/JHM-D-21-0193.1.

65. Newman, A. J. A. J. *et al.* Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States. *J. Hydrometeorol.* **16**, 2481–2500 (2015) https://doi.org/doi:10.1175/JHM-D-15-0026.1.

66. Clark, M. P. & Slater, A. G. Probabilistic Quantitative Precipitation Estimation in Complex Terrain. *J. Hydrometeorol.* **7**, 3–22 (2006) https://doi.org/10.1175/jhm474.1.

67. Bonneville Power Administration. *2020 Level Modified Streamflow: 1928-2018*. 1–278 https://www.bpa.gov/-/media/Aep/power/historical-streamflow-reports/2020-level-modified-streamflow.pdf (2020).

68. Mizukami, N. *et al.* SUMMA Hydrology in the Pacific Northwest for Hydrologic System Vulnerability Assessments. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory https://doi.org/10.5065/8SAK-HM25 (2025) https://doi.org/10.5065/8SAK-HM25.

69. Yilmaz, K. K., Gupta, H. V. & Wagener, T. A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour. Res.* **44**, W09417 (2008) https://doi.org/10.1029/2007wr006716.

70. Pytlak, E. *et al. Climate and Hydrology Datasets for RMJOC Long-Term Planning Studies: Second Edition (RMJOC-II) Part I: Hydroclimate Pro (p. 112). River Management Joint Operating Committee (RMJOC)*. 1–112 (2018).

71. Zhou, X., Yamazaki, D., Revel, M., Zhao, G. & Modi, P. Benchmark Framework for Global River Models. *J. Adv. Model. Earth Syst.* **17**, e2024MS004379 (2025) https://doi.org/10.1029/2024MS004379.

72. Lema, F. *et al.* Technical note: What does the Standardized Streamflow Index actually reflect? Insights and implications for hydrological drought analysis. *Hydrol Earth Syst Sci* **29**, 1981–2002 (2025) https://doi.org/10.5194/hess-29-1981-2025.

73. Yang, X. *et al.* Comparison of the Calibrated Objective Functions for Low Flow Simulation in a Semi-Arid Catchment. *Water* **14**, (2022) https://doi.org/10.3390/w14172591.

74. Staudinger, M., Stahl, K., Seibert, J., Clark, M. P. & Tallaksen, L. M. Comparison of hydrological model structures based on recession and low flow simulations. *Hydrol Earth Syst Sci* **15**, 3447–3459 (2011) https://doi.org/10.5194/hess-15-3447-2011.

ARTICLE IN PRESS

75. Vano, J. *et al.* Comparing Downscaled LOCA and BCSD CMIP5 Climate and Hydrology Projections - Release of Downscaled LOCA CMIP5 Hydrology. 96 (2020).

76. Coxon, G. *et al.* CAMELS-GB: hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. *Earth Syst Sci Data* **12**, 2459–2483 (2020) https://doi.org/10.5194/essd-12-2459-2020.

ARTICLE IN PRESS

ARTICLE IN PRESS

Contributions

NM wrote the original draft; BN, AW, JH, EG contributed text for sections of the manuscript; All the authors reviewed and edited the manuscript. EG contributed the generation of ICAR downscaled forcing. BN contributed the bmorph streamflow bias correction. AW contributed the baseline SUMMA-mizuRoute model implementation, the model calibration routines, and the GMET-based forcings and forcing preparation routines. NM further adapted the calibration routines, ran all model calibrations, adjusted the forcings, and performed all analyses, with input and review from the project team. JH, MW, CF and CM contributed the acquisition of the naturalized flow data. EG, JH, MW, CF and CM also managed the projects.

Competing Interests

The author(s) declare no competing interests.

Acknowledgement

This work was primarily funded by U.S. Army Corps of Engineers Grants (W26HM432051759, W26HM432051757, W26HM411547729 and HQUSACE17IIS/003) and was performed at the NSF National Center for Atmospheric Research (NCAR), which is a research and development center of the U.S. National Science Foundation (under Cooperative Agreement No. 1852977). It leveraged modeling and development work underway with sponsorship from the US Bureau of Reclamation Science and Technology Program, and high-performance computing support from the Derecho system (doi:10.5065/qx9a-pg09) provided by the NSF National Center for Atmospheric Research (NCAR), sponsored by the National Science Foundation and CISL for providing computing support.

Figure legends

Figure 1. Overview of climate-hydrology modeling chain. Each component of the modeling chains is detailed in the following subsections.

Figure 2. a) Combined basin calibration (first calibration stages). 4 subregions (Cascade, east Cascade, upper Columbia, and Snake) and basins used for regional combined basin calibrations. Modeled river network is also shown in blue lines. b) Individual basin calibration (second calibration stages). Pink areas are headwater catchments calibrated using the flow data at the red circle sites first, and then their downstream areas in orange are subsequently calibrated using the

ARTICLE IN PRESS

flow data at orange triangle sites with calibrated parameters of their upstream areas fixed. MROW (Snohomish River near Monroe) and LIB (Libby dam) sites are used for annual cycle time series in Figure 8. SUMMA HUC12 HRU not shown in both figures.

Figure 3. Performance metrics of the simulation at 221 flow sites–variability ratio (α), mean flow ratio (β), correlation (r), KGE', %biasFHV and %biasFLV. The error metrics are computed using daily streamflow aggregated from 3hr simulations. Three locations with darker gray shade are mentioned in text as examples of areas with poor performance.

Figure 4. a) Map of an example of a site that is located near the upstream end of the MERIT-basin reach (HAH: Green River at Howard Hanson). Thin blue lines are the stream network represented in mizuRoute and catchment areas associated with the stream network are grey highlighted. The red dot is the location of HAH site. MizuRoute outputs the outflow (local lateral flow and inflow from upstream) at each reach as river discharge. b) 2001-2010 mean annual cycle of naturalized flow at HAH and simulated inflow and outflow at the dark blue reach. c) a scatter plot of annual flows between at-site naturalized flow and simulated flow (blue; inflow, red; outflow) at the thick blue flowline. d) The same as panel c) except for spring (April through June) mean flow.

Figure 5. Distributions of the five error metrics for daily historical streamflow simulated using GMET and ESMs forcings: a) flow variability error α (i.e., ratio of observed flow standard deviation to simulated flow standard deviation), b) mean flow ratio β (i.e., ratio of mean observed flow to simulated flow), c) %biasFHV and d) %biasFLV and e) correlation coefficient of 1980-2014 mean daily seasonal cycle between observed and simulated flows. Distributions are based on 221 flow sites. The whiskers extend from the edge of the box to the farthest data point lying within 1.5x the inter-quartile range (IQR) from the box.

Figure 6. Cumulative distribution functions of a) variability ratio (α), b) mean flow ratio (β), c) correlation (c) and d) KGE′ at 221 naturalized flow sites for the daily simulated streamflow derived by GMET forcing (black) and the streamflow after bmorph bias correction (red). The error metrics are computed for the period 2001-2014.

Figure 7. Top row: Mean annual precipitation, ET, runoff and annual peak SWE during the control period. Middle and bottom rows: Relative changes during the WY2030-2060 (middle row) and WY2070-2099 (bottom row), based on all ESM ensemble means for control period and ensemble mean of high emission scenarios for future changes.

Figure 8. Long-term mean day-of-year streamflow at LIB (Libby Dam in Kootenai River: top) and MROW (Snohomish River near Monroe: bottom). Black dash denotes naturalized flow while black solid line denotes retrospective streamflow simulation with GMET forcing during the WY1980-2004 period. Shades are based on all the ESM forced simulations with 1) grey for control period, 2) red for high emission scenarios (RCP8.5 for CMIP5 and SSP5-8.5 for CMIP6) and 3) blue for low emission scenarios (RCP4.5 for CMIP5 and SSP2-4.5 for CMIP6). Blue and red shades are mean during the WY2070-2099 period.

ARTICLE IN PRESS

Figure 9. Maps of ESM ensemble mean metrics at 221 sites during the control period (WY1980-2004): a) annual centroid, d) annual mean flow and g) annual mean flow, and their changes in two future periods (WY2030-2060 and WY2070-2099) for high-emission scenarios (RCP85 for CMIP5, SSP585 for CMIP6). Circle marker indicates mean difference is statistically significant at 95% level while triangle marker is not significant at 95%.

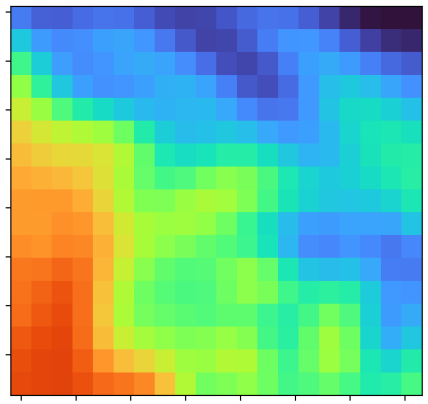
Figure 10. a) change in annual maximum flow frequency curve from WY1955-2004 to WY2050-2099 at MROW site and black arrows show how return period of historical 20-yr annual maximum flow is translated to future return period. Shades indicate ESM uncertainty and point with line is median values of the ESM ensemble. Future flood frequency curve is based on high-emission scenarios (RCP85 for CMIP5, SSP585 for CMIP6). b) Map showing the future return period (during WY2050-2099) for historic flows with a return period of 20 years (during WY1955-2004) at all the sites. Inset figure is a histogram of the future return period.

Figure 11. Percent changes in annual mean flow (top) and annual maximum flow (bottom) for WY2070-2099 period under RCP8.5 at 144 flow sites commonly available for three studies. Marker size is scaled based on values during WY1980-2004. The flow metrics are based on 5 CMIP5 ESMs; CanESM2, CCSM4, CNRM-CM5, and MIROC5.

ARTICLE IN PRESS

Climate-hydrology modeling chain

ESMs & Scenarios



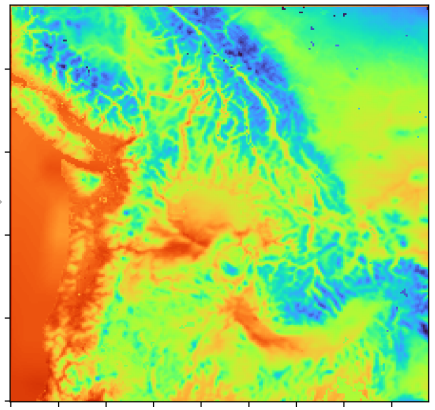
ESM selections

- 6 ESMs (CMIP5)
- 5 ESMs (CMIP6)

Scenario selections

- *RCP4.5, 8.5 (CMIP5)*
- *SSP245, 370, 585 (CMIP6)*

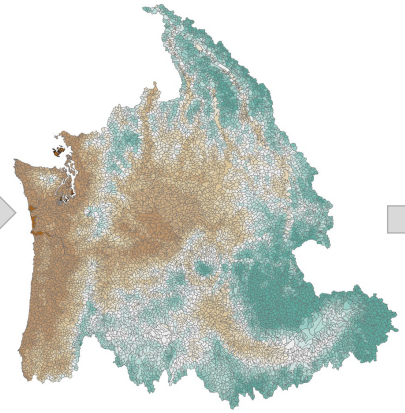
Spatial downscaling



ICAR(Quasi-dynamical downscaling)

- *Extract daily surface temperature and precipitation*
- *Areal weight mapping from 6km grid to HUC12 catchment*
- *Bias correction to GMET retrospective temperature and precipitation*
- *METSIM to produce SUMMA forcing*

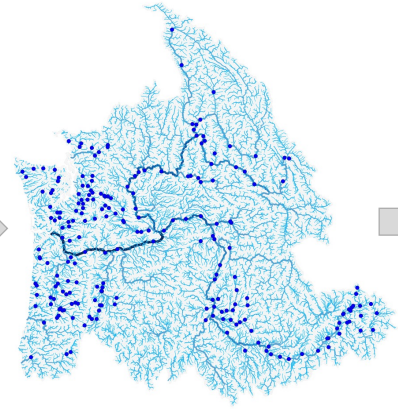
Hydrologic model



SUMMA

- *9,877 HUC12 catchment (~60km^2)*
- *Areal weight mapping of total runoff from HUC12 to MERIT-basin catchment*

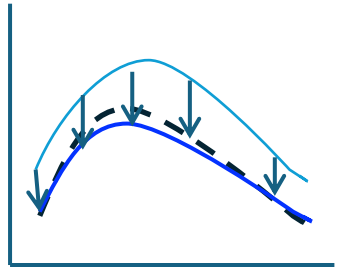
River model



MizuRoute

- *18,164 MERIT-basin catchment (~40km^2) & river reaches (~6km)*

Streamflow bias correction

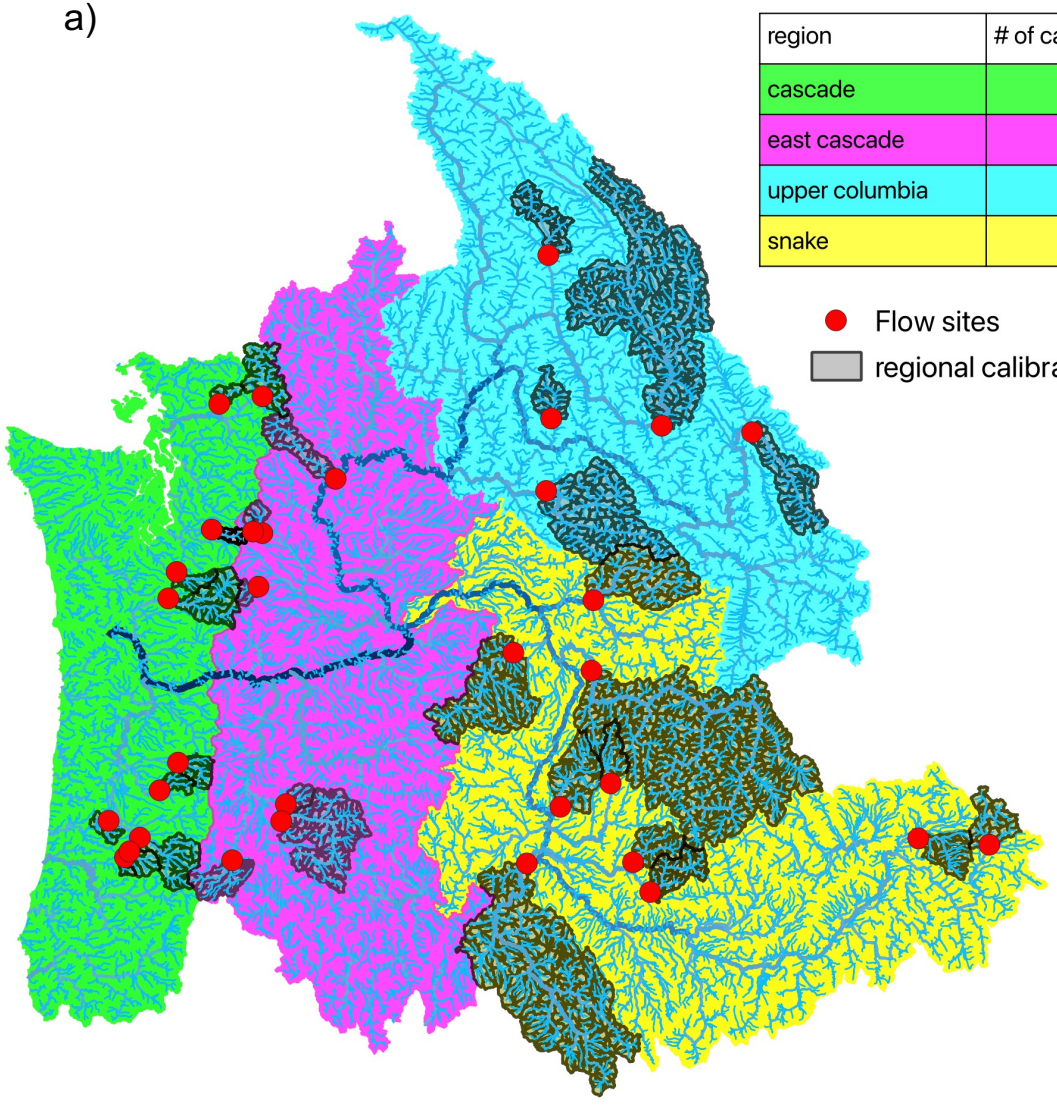


BMORPH

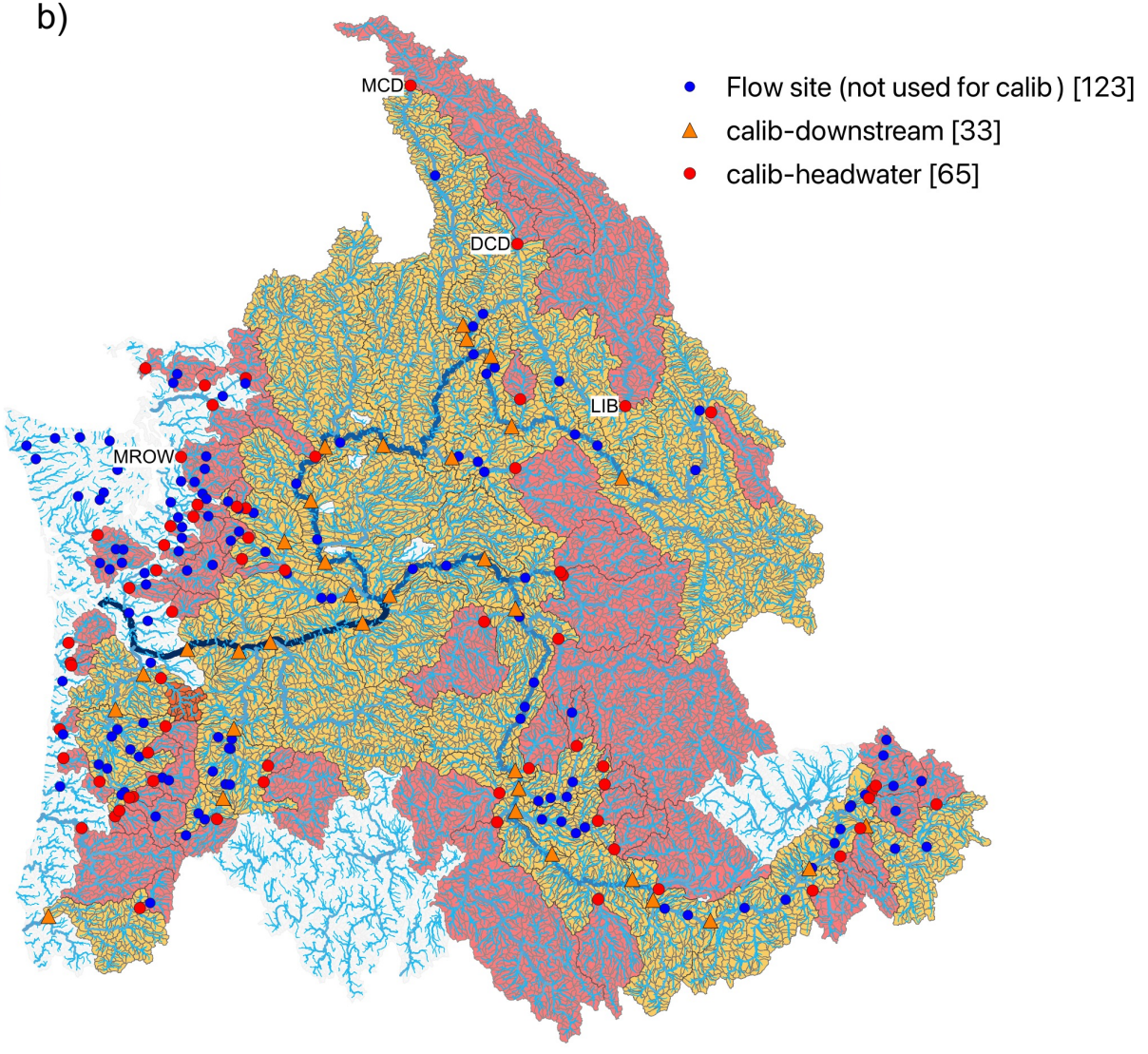
- *Streamflow bias correction at 200 Naturalized site*

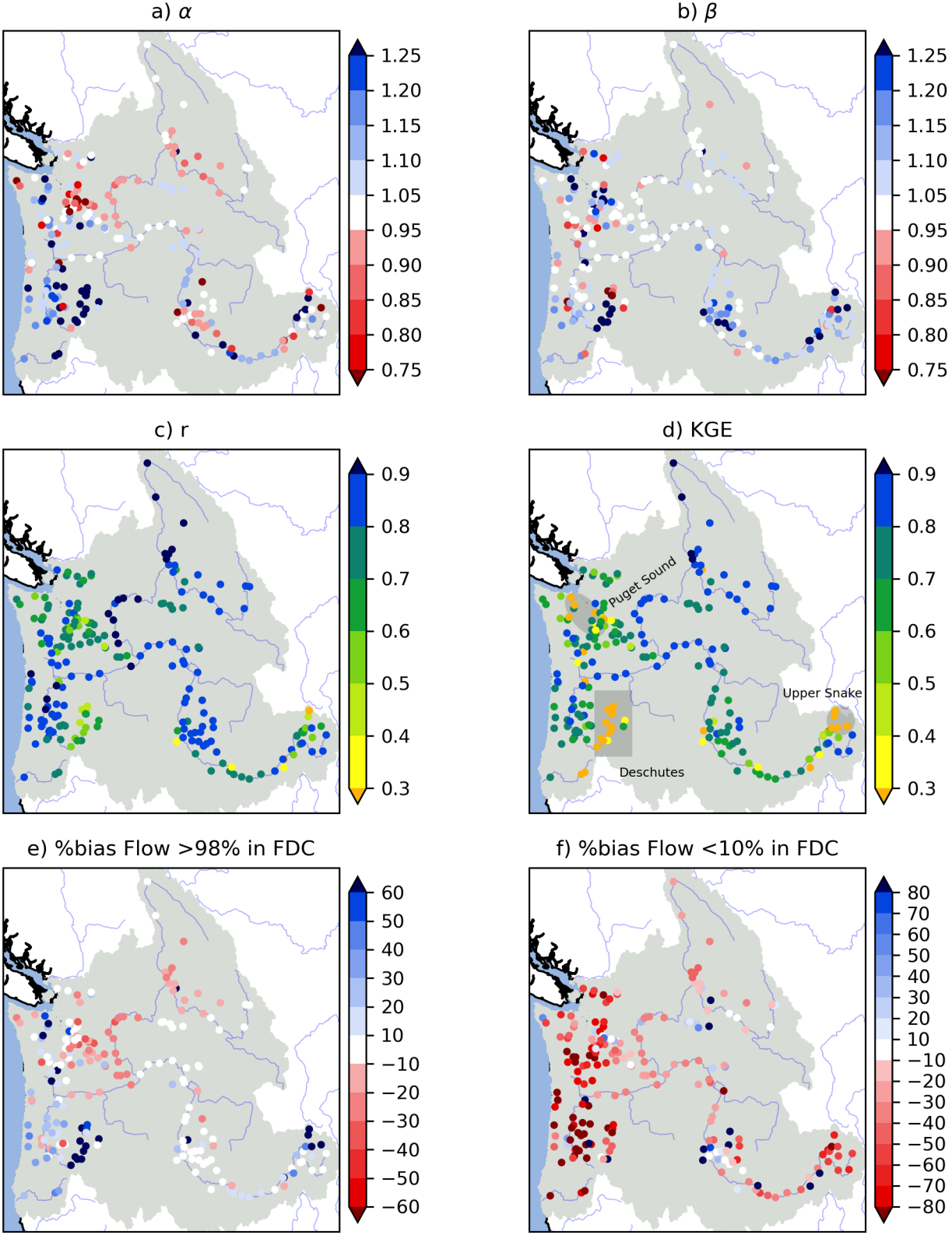
a)

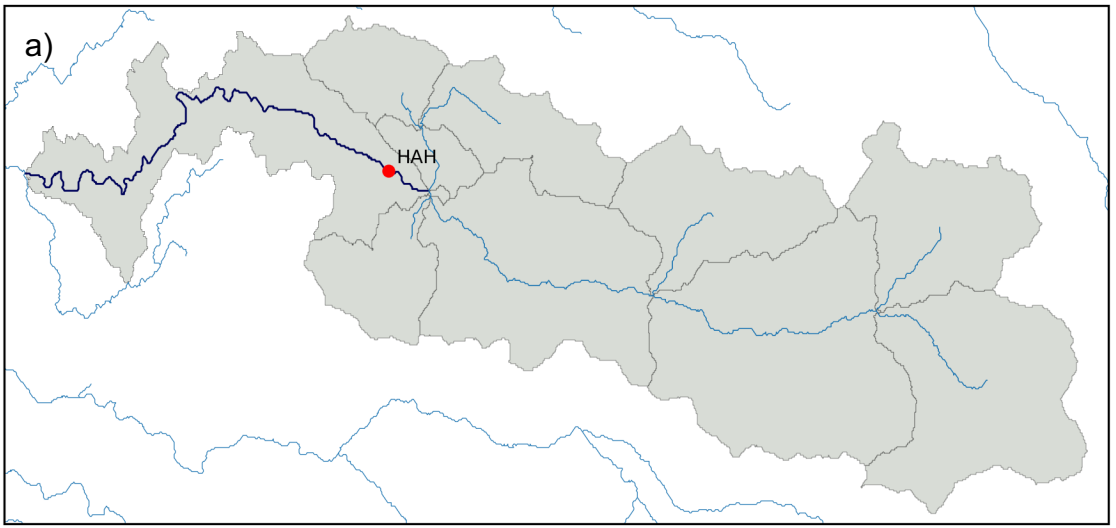
region	# of calibration sites
cascade	11
east cascade	7
upper columbia	5
snake	10

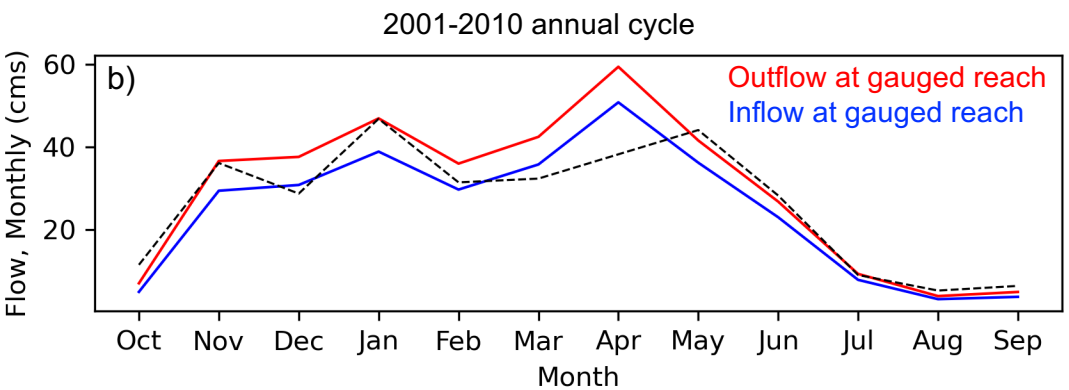


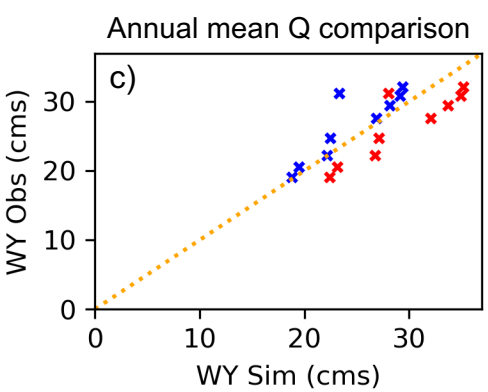
b)

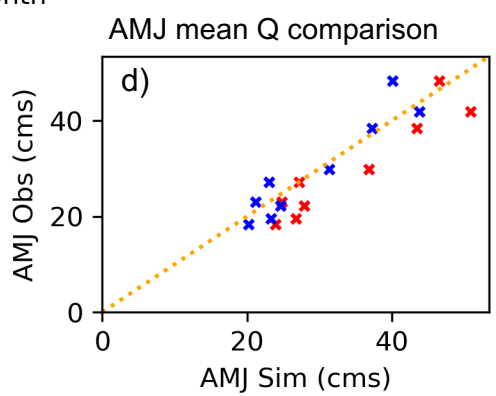


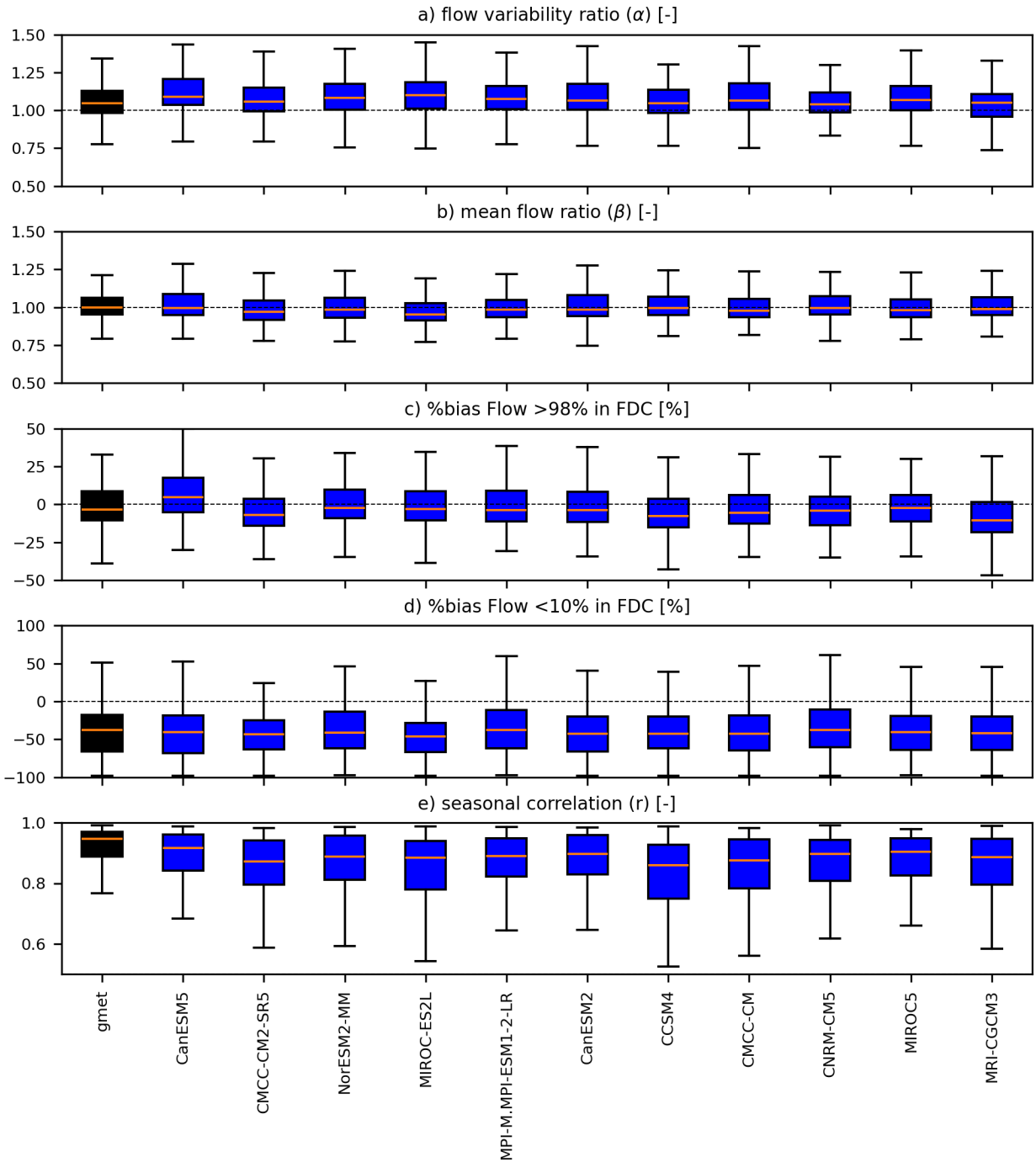

a) α
1.25
1.20
1.15
1.10
1.05
0.95
0.90
0.85
0.80
0.75
b) β
1.25
1.20
1.15
1.10
1.05
0.95
0.90
0.85
0.80
0.75
c) r
0.9
0.8
0.7
0.6
0.5
0.4
0.3
d) KGE
Puget Sound
Upper Snake
Deschutes
0.9
0.8
0.7
0.6
0.5
0.4
0.3
e) %bias Flow >98% in FDC
60
50
40
30
20
10
−10
−20
−30
−40
−50
−60
f) %bias Flow <10% in FDC
80
70
60
50
40
30
20
10
−10
−20
−30
−40
−50
−60
−70
−80

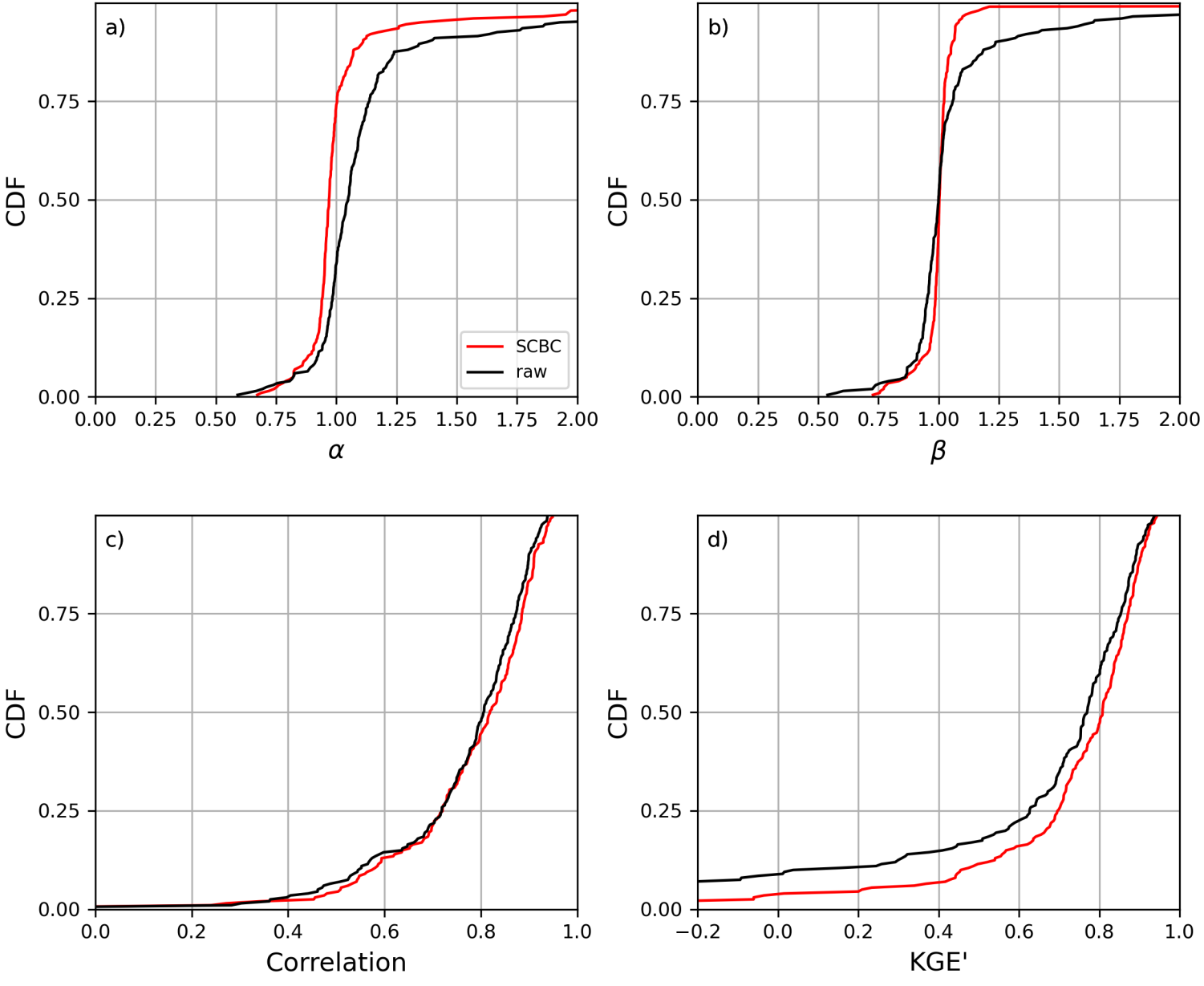

a)
HAH

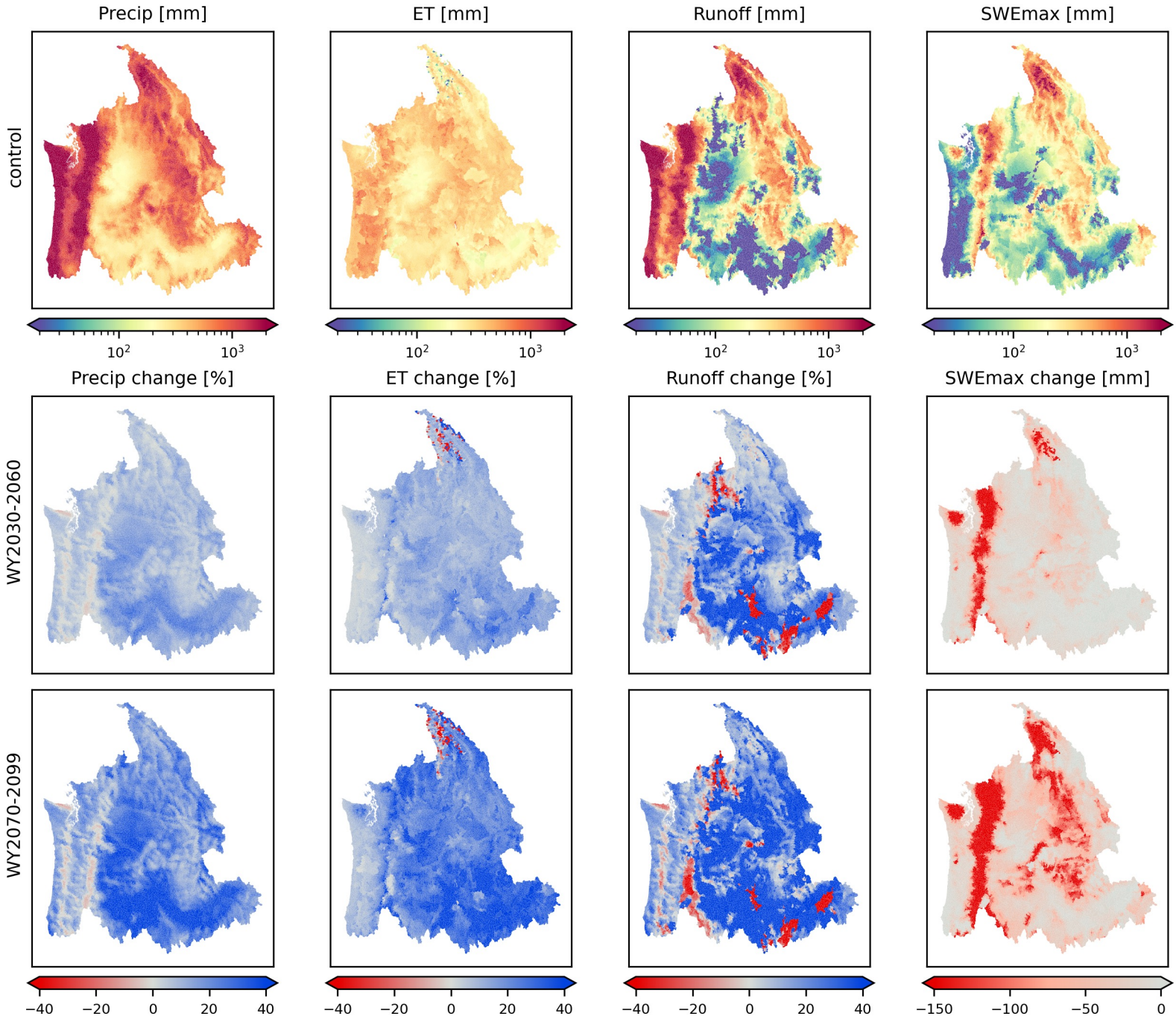

2001-2010 annual cycle
b)
Outflow at gauged reach
Inflow at gauged reach
Flow, Monthly (cms)
60
40
20
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Month

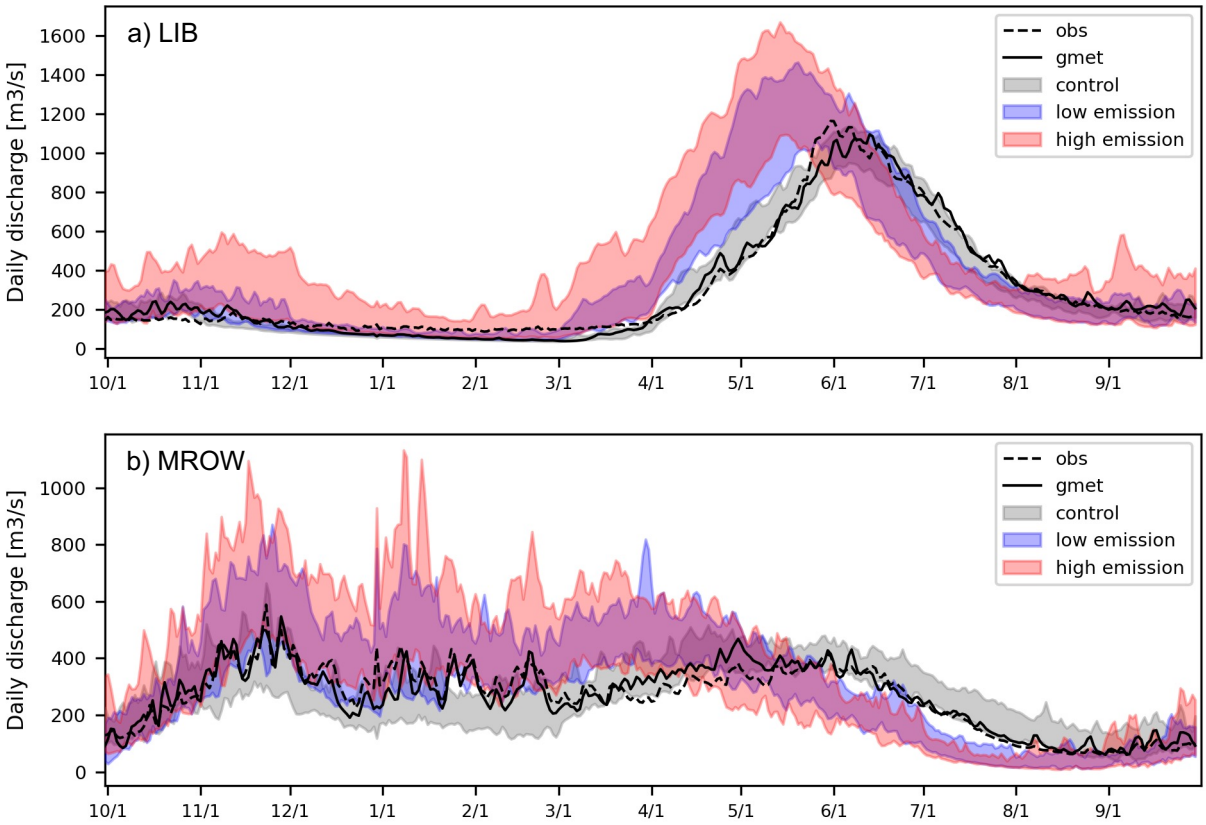

Annual mean Q comparison
c)
WY Obs (cms)
30
20
10
0
0
10
20
30
WY Sim (cms)

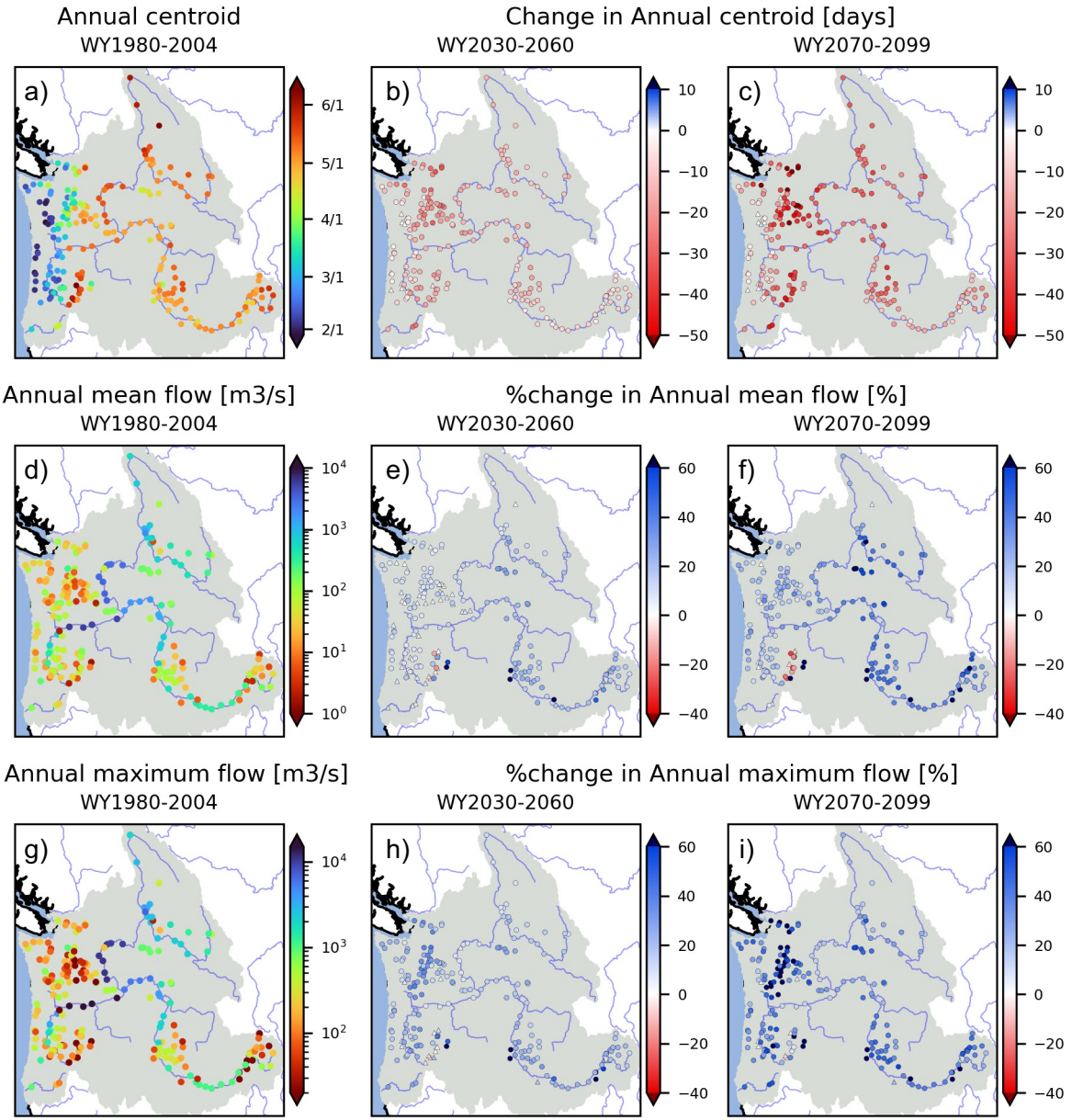

AMJ mean Q comparison
d)
AMJ Obs (cms)
40
20
0
0
20
40
AMJ Sim (cms)


a) flow variability ratio (α) [-]
1.50
1.25
1.00
0.75
0.50
b) mean flow ratio (β) [-]
1.50
1.25
1.00
0.75
0.50
c) %bias Flow >98% in FDC [%]
50
25
0
−25
−50
d) %bias Flow <10% in FDC [%]
100
50
0
−50
−100
e) seasonal correlation (r) [-]
1.0
0.8
0.6
gmet
CanESM5
CMCC-CM2-SR5
NorESM2-MM
MIROC-ES2L
MPI-M.MPI-ESM1-2-LR
CanESM2
CCSM4
CMCC-CM
CNRM-CM5
MIROC5
MRI-CGCM3

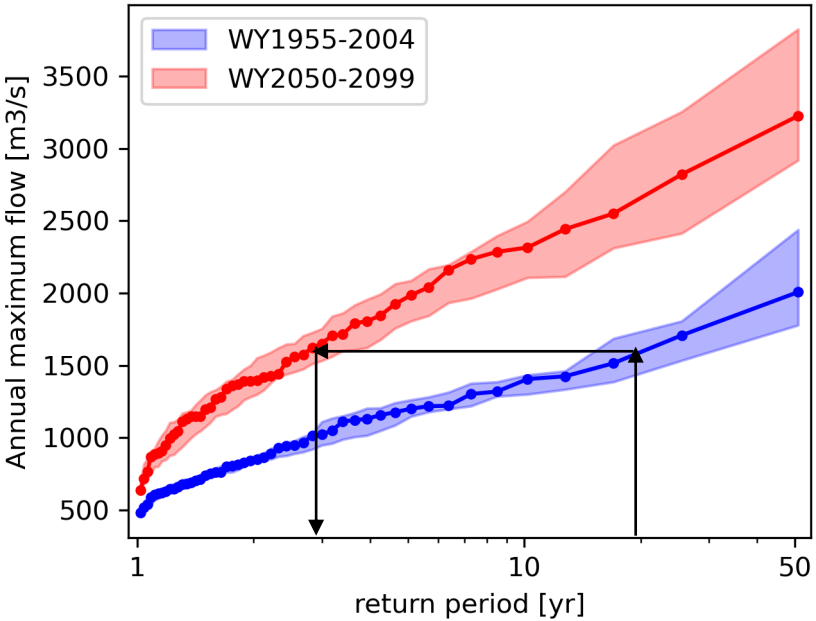

a)
b)
c)
d)
CDF
0.00
0.25
0.50
0.75
SCBC
raw
0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00
α
β
0.0 0.2 0.4 0.6 0.8 1.0
Correlation
−0.2 0.0 0.2 0.4 0.6 0.8 1.0
KGE'


Precip [mm]
ET [mm]
Runoff [mm]
SWEmax [mm]
control
10^2
10^3
Precip change [%]
ET change [%]
Runoff change [%]
SWEmax change [mm]
WY2030-2060
WY2070-2099
−40
−20
0
20
40
−150
−100
−50
0

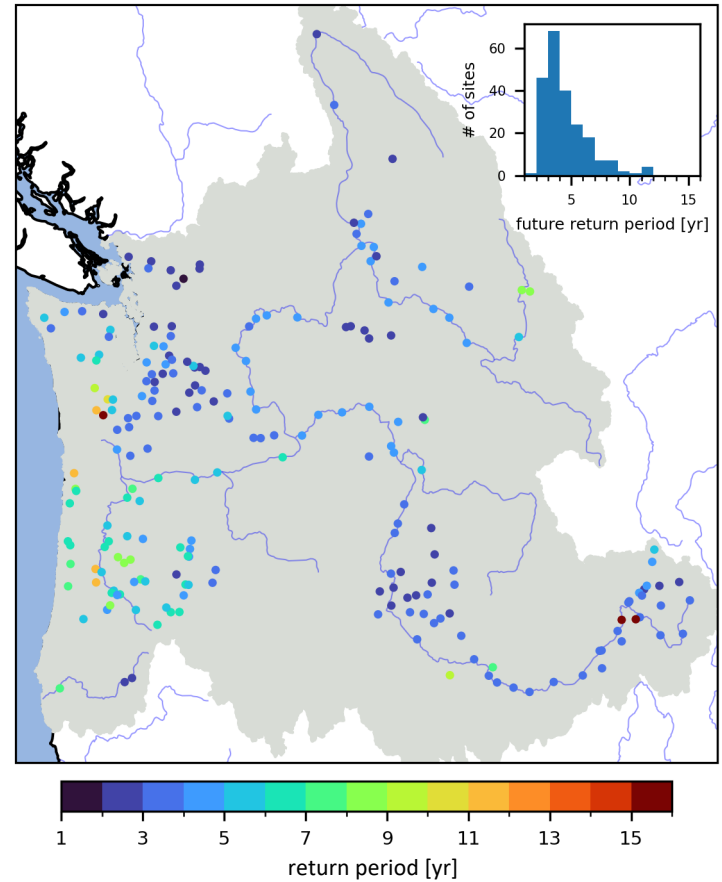

a) LIB
Daily discharge [m3/s]
0
200
400
600
800
1000
1200
1400
1600
10/1
11/1
12/1
1/1
2/1
3/1
4/1
5/1
6/1
7/1
8/1
9/1
obs
gmet
control
low emission
high emission
b) MROW
Daily discharge [m3/s]
0
200
400
600
800
1000
10/1
11/1
12/1
1/1
2/1
3/1
4/1
5/1
6/1
7/1
8/1
9/1
obs
gmet
control
low emission
high emission

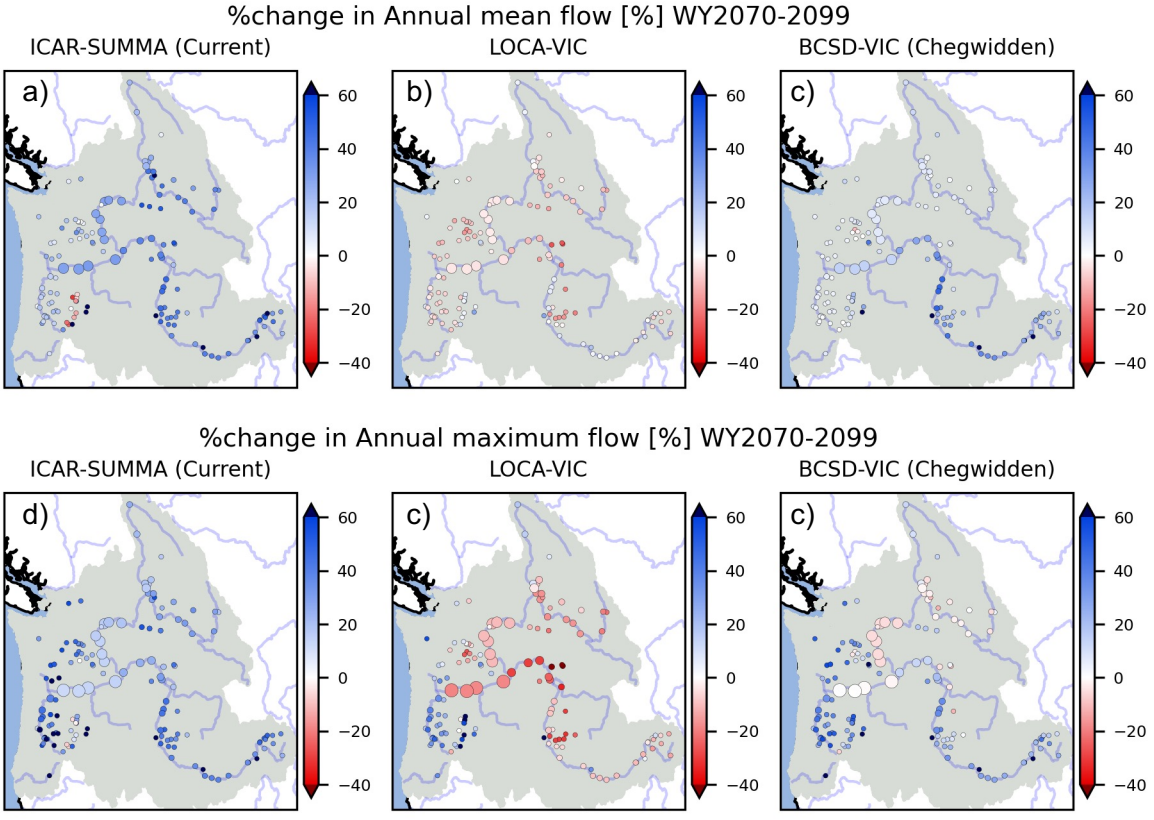

Annual centroid
WY1980-2004
Change in Annual centroid [days]
WY2030-2060
WY2070-2099
a)
b)
c)
6/1
5/1
4/1
3/1
2/1
10
0
−10
−20
−30
−40
−50
Annual mean flow [m3/s]
WY1980-2004
%change in Annual mean flow [%]
WY2030-2060
WY2070-2099
d)
e)
f)
10^4
10^3
10^2
10^1
10^0
60
40
20
0
−20
−40
Annual maximum flow [m3/s]
WY1980-2004
%change in Annual maximum flow [%]
WY2030-2060
WY2070-2099
g)
h)
i)
10^4
10^3
10^2
60
40
20
0
−20
−40

a) Shift of return period from historical to future at MROW



b) Future return period [yr] during WY2050-2099 for historic 20-yr annual max. flow




%change in Annual mean flow [%] WY2070-2099
ICAR-SUMMA (Current)
LOCA-VIC
BCSD-VIC (Chegwidden)
a)
b)
c)
60
40
20
0
−20
−40
%change in Annual maximum flow [%] WY2070-2099
ICAR-SUMMA (Current)
LOCA-VIC
BCSD-VIC (Chegwidden)
d)
c)
c)