https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

SWEDE v2: Emulating western US Dynamically Downscaled Snow Water Equivalent from Native CESM2 Large Ensemble

Fanghe Zhao[1], Stefan Rahimi[1], Melissa Bukovsky[1,2], Chad Thackeray[3], Karandeep Singh[3], Jesse Norris[3], Adrienne Marshall[4], Alex Hall[3], Weichen Liu[1], Lei Huang[3]

[1]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA
[2]Haub School of Environment and Natural Resources, University of Wyoming, Laramie, WY, USA
[3]Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, CA, USA
[4]Hydrologic Science and Engineering Program, Colorado School of Mines, Golden, CO, USA

Correspondence to: Fanghe Zhao (fzhao@uwyo.edu)

Abstract.

Snowpack in the western United States serves as a critical seasonal freshwater reservoir, yet Earth system models (ESMs) lack the spatial resolution to resolve snow water equivalent (SWE) at scales relevant to water resource management, and dynamically downscaling a full large ensemble remains computationally prohibitive. Here, we present SWEDE v2, an extension of the Snow Water Equivalent Downscaling Emulator that translates native CESM2 Large Ensemble (CESM2-LE) output to 9 km daily SWE across the western United States. SWEDE v2 retains the dual-encoder Swin-Transformer U-Net of Part 1 but trains two sub-models of identical structure, a baseline model that predicts daily SWE anomalies relative to a fixed training climatology and a recurrent increment model that predicts absolute SWE using the previous day's hybrid field. The emulator was trained on eight CESM2-LE members downscaled with the Weather Research and Forecasting (WRF) model over 1850–2100 and evaluated on a member withheld entirely from training. SWEDE v2 reproduces water-year mean snow water volume, daily SWE in three wet and three dry years, and annual maximum SWE across all six subregions. We then emulate all 100 members of CESM2-LE over the 250 water years from 1850 to 2099, a total of 25000 water years. Across five water management basins that cover the domain, ensemble-median peak volume declines by 37 (Upper Colorado) 67% (Lower Colorado) in 2071–2100 relative to 1981–2010 in the SSP 3-7.0 scenario. Peak timing generally advances, while absolute ensemble variance for April 1 SWE decreases in many snow-covered areas. These results demonstrate the potential of an approach in which an emulator trained on a small dynamically downscaled subset can extend a large ensemble to terrain-resolving resolution and provide a 100-member 9 km SWE dataset for the western United States.

1 Introduction

Snowpack functions as a natural water reservoir across the mountainous western United States, accumulating cold-season precipitation and releasing meltwater during spring and summer when downstream demand is highest (Bales et al., 2006; Li et al., 2017). Snow water equivalent (SWE), the liquid water stored in the snowpack, quantifies this reservoir, and its

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

projected decline under warming is among the most consequential regional hydroclimate changes expected this century (Mote et al., 2018; Barnett et al., 2005). April 1 SWE has declined by roughly 15–30% since the mid-twentieth century, and over 90% of long-record monitoring sites show declines (Mote et al., 2018). Projections show further losses together with earlier and slower melt (Musselman et al., 2017; Siirila-Woodburn et al., 2021). These changes differ among snow metrics. Peak SWE measures the largest seasonal storage, April 1 SWE has long served as the benchmark for seasonal water-supply forecasts, and the timing of accumulation and melt sets when meltwater reaches reservoirs. As the peak moves earlier, April 1 SWE is projected to decline faster than peak SWE, as in the Sierra Nevada (Beltran-Peña et al., 2025), and the interannual variability of peak SWE is projected to shrink where precipitation shifts from snow to rain (Marshall et al., 2019).

Any single realization of the coming decades, however, combines this forced decline with internal variability that can mask or mimic it over the multidecadal horizons relevant to water-resource planning (Deser et al., 2020; Deser et al., 2012; Siler et al., 2019). Initial-condition large ensembles were built to help separate these two sources of uncertainty. The 100-member Community Earth System Model version 2 Large Ensemble (CESM2-LE) samples the internal variability of one ESM under historical and SSP3-7.0 forcing (Rodgers et al., 2021). Yet CESM2-LE is integrated on a grid of approximately 1 degree, which smooths mountain barriers, lowers model terrain, and shifts cold season precipitation phase toward rain (Ikeda et al., 2010; Rasmussen et al., 2011; Lundquist et al., 2019). Native CESM2 consequently underestimates SWE in every subregion of the western United States, most severely in the narrow ranges of the Sierra Nevada and the Northern Rocky Mountains (Zhao et al., 2026). Thus, CESM2-LE samples variability extensively but does not resolve the terrain that controls where snow accumulates.

To address the spatial resolution limitations of the native CESM2-LE, ten members of the CESM2-LE have been dynamically downscaled to 9 km with the Weather Research and Forecasting (WRF) model as part of the Western U.S. Dynamically Downscaled Dataset (WUS-D3). (Rahimi et al., 2024b; Rahimi et al., 2024a). These simulations recover the elevation-dependent SWE that the native grid cannot. (Zhao et al., 2026) Ten members provide estimates of the forced trend and central spread. However, in the extreme tails, where record snow drought and accumulation years are counted, the interval around a risk ratio is set by how many such years ten members contain. The 2026 western United States snow drought shows this. Using these ten members to attribute the event to anthropogenic climate change since the pre-industrial period, the risk ratio in the Upper Colorado basin was 14 with a 95% confidence interval of 0.09 to 4300 (Marshall et al., 2026). The point estimate identifies a forced contribution, and the width of the interval reflects the number of tail years ten members can supply. Machine learning-based downscaling of large ensembles has recently been identified as a key method to reduce uncertainty in this type of attribution event (National Academies of Sciences and Medicine, 2026).

Process-based downscaling is computationally expensive, and two routes around the downscaling cost have been taken. Norris et al. (2025) regressed end-of-century snowfall loss in the downscaled ensemble on the warming and precipitation change of the driving model and applied the coefficients to the full CMIP6 archive. This yielded a period-mean snowfall change for every model but not the daily, member-specific SWE fields that ensemble analysis of timing, extremes, and variance requires. On the other hand, deep learning emulators of regional climate models have reproduced downscaled atmospheric

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

fields at a small fraction of the dynamical cost, including generative emulators trained on WUS-D3 itself (Doury et al., 2023; Hobeichi et al., 2023; Lopez-Gomez et al., 2025), but SWE poses a different problem. Snowpack integrates atmospheric forcing over a season, so supplying daily SWE for members that lack a regional simulation requires an emulator that also carries memory of past accumulation and melt.

Part 1 of this work introduced the Snow Water Equivalent Downscaling Emulator (SWEDE), a dual-encoder Swin-Transformer U-Net that emulates WRF-simulated 9 km daily SWE from coarse daily atmospheric predictors, such as near-surface temperature, humidity, precipitation, and radiation and their lagged means, and from high-resolution terrain, without receiving any coarse SWE as input. We found that the composition of training-data strongly affected projection robustness (Zhao et al., 2026). Trained on WRF downscaling of 15 CMIP6 simulations from 14 ESMs over both the historical and future periods, SWEDE reproduced the forced trend and ensemble spread of the ten WRF-downscaled CESM2-LE members, which were withheld from training, through 2100. Trained on the historical period alone, it degraded under late-century warming (Zhao et al., 2026). SWEDE v1.0 was developed in a perfect-model setting, with WRF fields coarsened to ESM scale as inputs. This setting isolates the downscaling step and produced the training-data result above. Driven instead by an independent coarse simulation, it overestimated snowpack amplitude, a sensitivity to input mismatch that van der Meer et al. (2023) also found when emulators trained on upscaled regional model output were applied to global model output. Native, uncorrected ESM fields at about 1 degree, produced by a different model with its own biases, pose a larger mismatch, and Part 1 did not test them.

Here, we extend SWEDE to operate directly on native CESM2-LE output, with no bias correction applied to the inputs. Because the WRF targets were driven by bias-corrected boundary conditions, the emulator learns the bias correction and the downscaling as one mapping. SWEDE v2 couples a memoryless baseline network, which predicts each day's SWE anomaly from forcing and terrain, with a recurrent increment network, which carries the previous day's snowpack, and blends the two with a fixed seasonal weight. The increment network supplies the snowpack memory that Part 1 carried only through lagged means of the forcing. SWEDE v2 is trained on eight of the ten WRF-downscaled CESM2-LE members, while one is used for validation, and one is used for evaluation. We then emulated all 100 CESM2-LE members over the 250 water years from 1850 to 2099 to create 25000 water years of daily 9 km SWE. Because the application targets additional members of the CESM2-LE, training on a subset of that ensemble is appropriate. The emulator adds no snow physics beyond what WRF provides. This paper presents the change to SWEDE (Section 3), the withheld-member model evaluation (Section 4.1), a descriptive characterization of the resulting ensemble in median, variance, and timing (Section 4.2), and a demonstration of the sampling gain on the 2026 snow-drought risk ratio (Section 4.3). The evaluation includes a multiple linear regression baseline fit on the same predictors, training members, and target climatology, in pooled and per-pixel forms and with and without the CESM2 simulated snowpack as an additional predictor. Beyond that demonstration, scientific analysis of the ensemble, including elevation dependence, emergence timing, and the separation of externally forced signals from internal variability, is reserved for separate work that will use this dataset.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

2 Data Description and Processing Pipeline

2.1 Native and WRF-downscaled CESM2-LE

CESM2-LE is a 100-member initial-condition ensemble of the Community Earth System Model version 2 (Rodgers et al., 2021; Danabasoglu et al., 2020). It is integrated from 1850 to 2100 under CMIP6 historical forcing through 2014 and the SSP3-7.0 pathway thereafter, on a 365-day calendar. The atmosphere and land components run on a nominal 1 degree grid (0.9 degrees latitude by 1.25 degrees longitude). The members differ marginally in their initial conditions, so the ensemble spread measures internal variability under a shared forcing pathway. The ensemble is comprised of two 50-member subsets with differing biomass burning. One is driven by the CMIP6 biomass burning emissions and one by a temporally smoothed version of those emissions (Rodgers et al., 2021). The ten members downscaled in WUS-D3 (Table B1) are the ten macro-initialization members of the smoothed biomass-burning subset of the CESM2-LE (Rodgers et al., 2021). The emulator is trained on and applied to native CESM2-LE fields without regard to subset membership. The native fields used as SWEDE v2 inputs are daily means of the seventeen atmospheric variables listed in Appendix Table A1. The underlying native fields are taken directly from the CESM2-LE archive with no bias correction. CESM2's own SWE is not an input to SWEDE v2. It is used in this study only as an additional predictor in the SWE- augmented multiple linear regression baselines of Section 3.4.

The training targets are the ten dynamically downscaled CESM2-LE members described in Rahimi et al. (2026). Each WUS-D3 simulation employs two one-way nested WRF domains, an outer 45 km domain and an inner 9 km domain covering the western United States. The 9 km daily SWE from the inner domain on the 340 by 270 grid is the target for SWEDE throughout this paper. The ten members span water years 1850 through 2100. Before downscaling, the CESM2 boundary conditions (temperature, moisture, horizontal winds, surface pressure, and sea surface temperature) were bias corrected against the ERA5 mean-state climatology (Hersbach et al., 2020). The adjustment varies by month and removes the mean CESM2 bias while preserving synoptic-scale variability (Rahimi et al., 2024a). This correction is applied to the WRF driving fields only. The native CESM2-LE inputs to SWEDE v2 retain the CESM2 biases. The emulator therefore maps uncorrected atmospheric fields to targets produced downstream of a bias-corrected boundary condition. The WUS-D3 simulations were run as independent, parallelized years, each initialized on 1 August from the driving GCM state, with output retained from 1 September after one month of spin-up. High-elevation snow present in late August is reset at the year boundary rather than carried forward (Rahimi et al., 2024a). SWEDE v2 is trained to reproduce WRF and thus inherits this cold-start discontinuity in September fields.

2.2 Inputs, targets, and member split

SWEDE v2 takes two groups of inputs that enter through separate encoder branches (Zhao et al., 2026). It differs from SWEDE v1 in the source and number of dynamic channels. The baseline network's dynamic branch receives twenty daily channels on a 32 by 25 forcing grid with 90 km spacing. This grid covers the domain of the 9 km target grid, coarsened by a factor of ten, and its spacing is close to that of the native CESM2 grid (0.9 degrees latitude by 1.25 degrees longitude).

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

These channels are comprised of the seventeen CESM2-LE atmospheric variables listed in Appendix Table A1, the terrain height on the forcing grid, and the sine and cosine of the day of year. The atmospheric variables are native CESM2-LE daily fields regridded bilinearly from the nominal 1 degree grid to the forcing grid (90 km), with no bias correction. The static branch receives six fields on the 320 by 250 target grid after removing a 10-grid-point halo from the original dataset. These are a land mask, land use, elevation, subgrid elevation variance, and the two components of the elevation gradient. Longitude is not an input, while latitude enters only through an explicit bias term in the network (Section 3.1). All inputs are standardized to zero mean and unit variance using statistics computed from the eight training members spanning the 1850-2100 period. They are held fixed for validation, testing, and inference on all 100 members.

The target is 9 km daily SWE from the WRF-downscaled members, stored as water-year files spanning 1 September to 31 August. The data loader returns the anomaly of each daily field from a frozen per-pixel, per-day-of-water-year climatology computed over the eight training members from the full training period. Absolute SWE in millimeters is recovered by adding that climatology back and clamping at zero (Section 3.2).

The ten WRF-downscaled members were divided into eight training members, one validation member, and one test member. LE2-1151.008 is the validation member, used for early stopping and hyperparameter selection, and LE2-1191.010 is the test member (Table B1). Every metric reported in Section 4.1 is computed on the test member. The regression baselines of Section 3.4 were fitted to the same eight training members and evaluated on the same test member. This split tests generalization across internal variability within a fixed forced response and a fixed set of model biases. That is the generalization that emulating the remaining ninety members requires.

2.3 Evaluation regions and snow metrics

Emulator evaluation and ensemble characterization in the main text use the six snow climate subregions of Part 1. They follow the regional definitions of (Rahimi et al., 2022) and are shown in the Figure S1. The Pacific Northwest (PNW) and the Sierra Nevada (SN) represent maritime regions with warm temperatures, high precipitation rates, and deep seasonal accumulation. The Northern Central Rockies (NCR) and the Northern Rocky Mountains (NRM) represent continental snowpacks, where cold conditions favor persistence but lower precipitation limits accumulation. The South Central Rockies (SCR) and the Southwest (SW) represent intermountain and arid-mountain environments, where accumulation is episodic and marginal. Regional metrics are computed by spatially averaging or integrating all terrestrial grid cells within each subregion boundary on the 9 km grid. All statistics are reported at daily resolution unless otherwise noted. For water-resource users, evaluation over the California and Upper Colorado basins are also presented. Results are also reported for river basins, and the two aggregation schemes serve different purposes. The subregions group grid cells by snow climate and keep the evaluation comparable with Part 1. Water-resource questions are posed for river basins, so the observational comparison (Section 2.4), the risk ratio (Section 2.5), and the basin-scale projections use the five hydrologic unit code level 2 (HUC2) basins that cover the domain: Upper Colorado (14), Lower Colorado (15), Great Basin (16), Columbia River Basin (17), and California (18) regions.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

All snow metrics are computed from daily SWE. Snow water volume is the sum of SWE multiplied by cell area over all terrestrial 9 km cells of a region, expressed in km³. The water-year mean volume averages the daily volume over the 365 days of a water year, and the peak snow water volume is the annual maximum of the daily volume. Annual peak SWE is the annual maximum of daily SWE at each grid cell, and mean annual maximum SWE averages it over the cells of a region. We also evaluate April 1 SWE. Snow onset and melt-out are the first and last days of the water year on which SWE exceeds 10 mm, regardless of intervening melt, and the peak date is the day of annual peak SWE, with ties assigned to the earliest day. Years in which SWE never exceeds 10 mm are excluded from the timing metrics, which are therefore conditional on snow occurrence. To test the emulator in extreme years, the three wettest and three driest water years of the test member were selected once for the whole domain from the detrended WRF peak snow water volume anomaly, and the same six years are evaluated in every subregion. Detrending removes the forced decline, so the selection samples internal variability across the whole record.

2.4 Observational data

The emulated ensemble is evaluated over the historical period against the University of Arizona 4 km daily SWE product (UA PRISM SWE) (Broxton et al., 2019; Zeng et al., 2018). UA SWE interpolates station snow observations from the SNOTEL and COOP networks using PRISM precipitation and temperature and performs relatively well against lidar-based SWE estimates (Ritchie et al., 2025; Lundquist et al., 2026). Water years 1981 through 2021 are used for this assessment and labelled by the same convention as the emulated fields. UA SWE is integrated over each subregion to give observation-based annual peak snow water volume. That volume is compared with the 100-member envelope of emulated peak volume in Section 4.2. Because CESM2-LE members are free-running, the comparison is statistical. It tests whether the observed median lies near the ensemble median of the same years and whether individual observed years fall within the ensemble spread.

2.5 Snow-drought risk ratio

The 2026 snow drought attribution conducted in Marshall et al. (2026) is repeated on the emulated ensemble with their procedure unchanged. SWE volume on 15 March is standardized to a z-score within each ensemble, generalized extreme value distributions are fitted to the counterfactual (1850 to 1900) and factual (2016 to 2036) member years, and the risk ratio is the factual probability of a SWE z-score at least as low as observed in 2026 divided by the counterfactual one, with 95% confidence intervals from member years. The SWE field is SWEDE v2 daily SWE rather than WRF output, integrated over the same four domains, the California, Columbia, and Upper Colorado HUC2 basins and the western United States as they defined it. The ten-member calculation uses the SWEDE v2 emulation of the ten WUS-D3 members, and the 100-member calculation uses all members. The observed 2026 z-score is taken directly from Marshall et al. (2026), computed from ERA5-Land (Muñoz-Sabater et al., 2021) relative to 1950 to 2026, so that any difference in risk ratio arises from the ensemble and not from the observational anchor. The annual probability of a snow drought at least as severe as 2026 (Figure 8b) is computed by decade as in their Figure 2B.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

3 SWEDE v2 Model Description

SWEDE v2 changes SWEDE v1.0 in four respects. Its inputs change from WRF fields coarsened to ESM scale to native CESM2-LE fields (Section 2.2). Its training data change from 15 WRF-downscaled CMIP6 simulations to eight raw CESM2-LE members (Section 3.3). A recurrent increment network now complements the memoryless network of Part 1 (Section 3.2), and the hidden widths of each network are halved (Section 3.1).

3.1 Architecture

Both networks in SWEDE v2 use the dual-encoder Swin-Transformer U-Net of Part 1 (Zhao et al., 2026). Relative to Part 1, the hidden channel widths of the CNN encoder and decoder and the hidden dimensions of the Swin-Transformer blocks are halved in each network. A high-resolution encoder ingests the six static terrain and land-surface fields on the 320 by 250 target grid. A second encoder ingests the low-resolution daily CESM2-LE forcing on the 32 by 25 forcing grid. The two branches are fused in a shared decoder with windowed self-attention and attention-gated skip connections. The network carries an explicit latitude bias term. The two networks are trained separately and need not share an input channel set. At inference, each reads only the channels it was trained on, addressed by position in a shared data layout. A channel can therefore be added to one network without retraining the other.

3.2 Baseline and increment networks

SWEDE v2 couples two networks that fail in complementary ways. One is memoryless and cannot drift (baseline), while the other carries snowpack state and is more accurate while snowpack persists (increment). The baseline network predicts the same-day SWE anomaly from forcing and terrain alone, with no knowledge of the previous day. The anomaly is converted to absolute SWE (in millimeters) by adding the frozen training climatology of Section 2.2 and clamping at zero. The climatology is indexed by the day of the water year (September-August). Because it has no memory, its error cannot compound across days. It also cannot represent the accumulation and ablation history that governs snowpack on any given day. Part 1 supplied that history only through lagged temporal means of the forcing. Its compression of interannual amplitude is consistent with that limitation. In contrast, the increment network predicts same-day absolute SWE from forcing, terrain, and the previous day's SWEDE v2 field, so it carries snow storage from one day to the next. It was fine-tuned over two-day rollouts so that its own output is a realistic input. It is more accurate on every day from November through August. As a single increment model run without observations, however, it can accumulate residual error over long integrations.

The two networks are trained with different losses. The baseline network uses the composite loss of Part 1, which combines a core loss, a log-space loss, and a temperature-constraint loss and is built for a target that is zero on most pixels, spans four orders of magnitude, and is right-skewed. The core term combines binary cross-entropy for snow detection with a Smooth L1 regression term on snow-covered pixels. The log-space term penalizes relative error so shallow snow is not neglected. The temperature-constraint term penalizes nonzero SWE where the daily minimum temperature exceeds 30 ºC,

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

which suppresses spurious snow under ESM forcing outside the training distribution. The increment network uses a Huber loss with = 100 mm, quadratic for errors below 100 mm and linear above.

The two networks are combined each day as a convex blend, clamped at zero and masked to land. The combined field, hereafter the hybrid field, is fed back as the increment network's input on the following day. The weight w is a fixed function of calendar day alone (Figure S2). It holds at 0.99 from 1 November through 15 July before decreasing by a cosine ramp to 0.1 between 15 July and 15 August, and then holding at 0.1 through 30 September. Lastly, the weight returns by a cosine ramp to 0.99 between 1 October and 1 November. The near-unity winter ceiling was chosen because the increment network run alone outperforms the baseline network on every day from November through August. Mixing in more of the baseline model would pull the recurrent state toward the weaker network when the stronger one is most reliable. The summer floor serves a different purpose. At weight w the state relaxes toward the memoryless baseline at a rate of (1 - w) per day. The floor is therefore an annual reset that clears the residual a free-running increment network accumulates. It is held for several weeks before the new snow year begins.

3.3 Training and inference

The two networks were trained separately on the eight training members (Section 2.2) following the training procedure from Part 1. Part 1 showed that an emulator trained on the historical climate alone degrades under sustained warming, and that training on the combined historical and future periods removes this degradation. The training set here therefore covers 1850 to 2100 of each training member, so the networks see the pre-industrial, historical, and SSP3-7.0 climates they are later asked to emulate. The baseline network was trained on the anomaly target. The increment network was trained on absolute SWE with the previous day's field as an additional input. The validation member was used for early stopping and model selection. After single-step training, the increment network was fine-tuned over two-day rollouts (Lam et al., 2023). In each rollout its own first-day output replaces the target field as the second day's input, so the errors it sees at inference are represented in training.

The projection used for evaluation and for the dataset is a free run, in which the hybrid is integrated forward from its own previous-day output and no snow field enters at any step. Integration begins on 1 September 1850, at the seasonal snow minimum, where the increment weight already sits at its summer floor. It continues through 31 August 2100, covering 250 complete water years. No WRF target fields are supplied during inference. The hybrid is initialized from the baseline network on the first day. The only re-anchoring in the entire integration is the seasonal relaxation toward the memoryless baseline each summer. The same configuration was run for each of the 100 CESM2-LE members, giving 25000-member water years of daily 9 km SWE. The free run for the test member, LE2-1191.010, covers 1850 to 2099 and is evaluated in Section 4.1.

Inference took 4.7 GPU-seconds per simulation year on a single NVIDIA V100 GPU, or 0.33 GPU-hours per member for 250 water years. The full 100-member ensemble took 33 V100-hours and 8.3 h of wall-clock time on one node. The WRF pipeline costs approximately 7500 CPU-hours per simulation year (Zhao et al., 2026). The full daily archive is approximately 3.4 TB. The deposited subset and the access route are given in the code and data availability section.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.4 Regression baseline

A multiple linear regression (MLR) (Baño-Medina et al., 2020; Yang et al., 2022) provides a linear model benchmark for SWEDE v2. The baseline uses the same predictors, training members, standardization, target climatology, and evaluation as SWEDE v2. The comparison isolates the model rather than the data. Each design row is one land pixel on one day. The twenty dynamic channels were bilinearly interpolated from the forcing grid to the target grid and combined with the six static fields and an intercept, giving 27 predictors. The response is absolute SWE (in millimeters). Four variants were fit, crossing two choices. Pooled variants estimate one coefficient vector for the domain, and per-pixel variants estimate one for each of the 57344 land pixels. The SWE-augmented MLR appends CESM2's own coarse SWE at the current and preceding day, giving 29 predictors. All variants were solved by ridge regression (Shao et al., 2022) with a penalty of 10^{-6} times the mean diagonal of the normal matrix. Predictions were clipped at zero. Before clipping, 43.5%, 37.0%, 44.6%, and 36.0% of raw predictions were negative for pooled MLR, per-pixel MLR, pooled SWE-augmented MLR, and per-pixel SWE-augmented MLR respectively. Each variant predicts daily fields for the test member over 1850 to 2100 and is scored with the same evaluation script as SWEDE v2 over 1980 to 2099. The SWE-augmented variants receive the CESM2 simulated snowpack, which SWEDE v2 does not.

4. Results

4.1 Evaluation of the withheld member

The emulator reproduces the fine scale spatial snowpack patterns that the native grid cannot represent (Figure 1). Native CESM2-LE smooths ensemble-mean peak and April 1 SWE into broad lobes over the Cascades and the northern Rockies. It places almost none over the Sierra Nevada, the Wasatch, or the Colorado Rockies. The emulated fields reproduce the WRF range-by-range structure at 9 km, including the Sierra Nevada crest and the individual ranges of Colorado. These maps average the ten WRF-downscaled members, eight of which were used in training, so the remaining evaluation in this section uses the withheld member alone.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

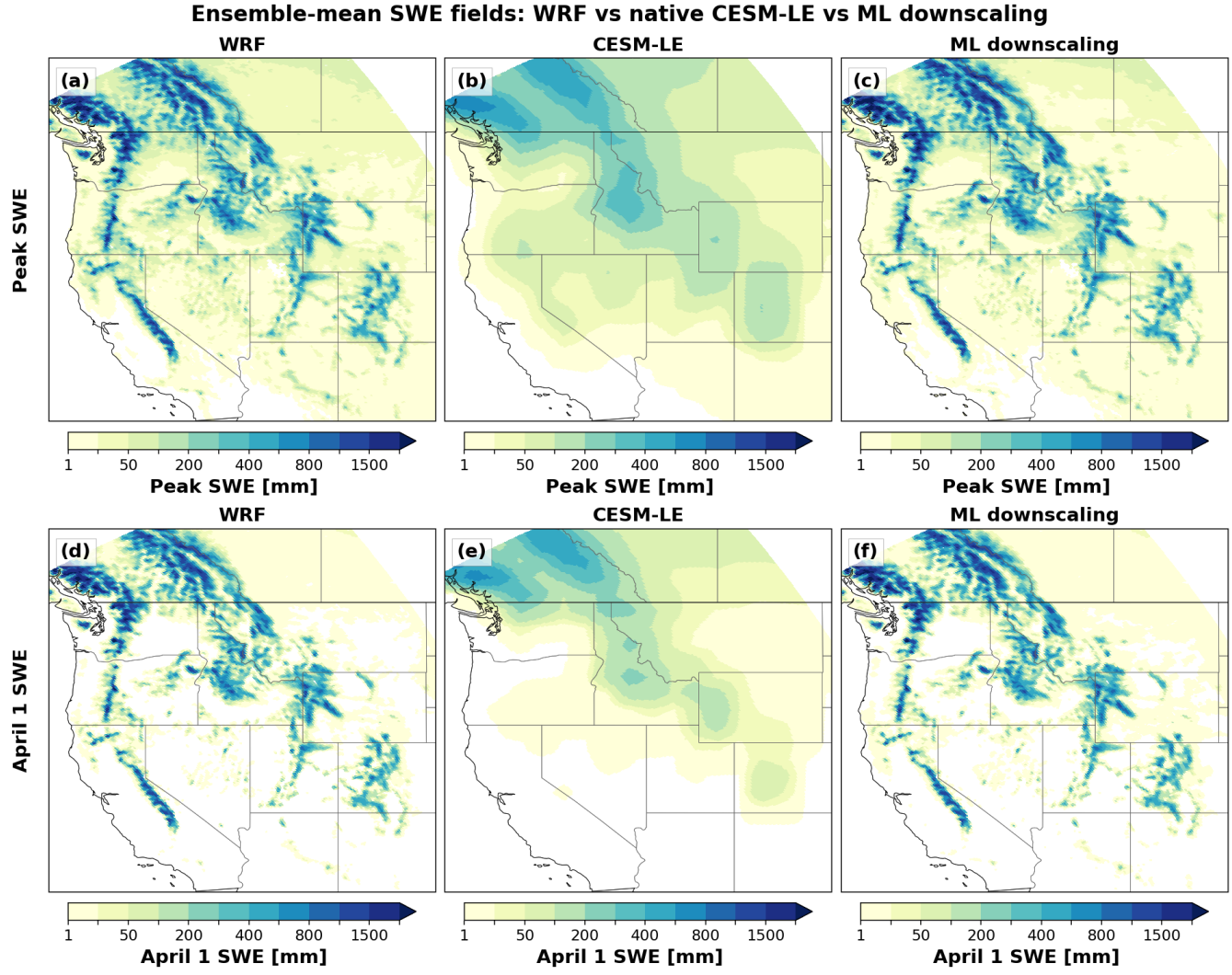

Figure 1. Ensemble-mean (a to c) annual peak SWE and (d to f) April 1 SWE from the ten WRF-downscaled CESM2-LE members (left), the same ten members on the native nominal 1 degree grid (center), and the SWEDE v2 emulation of the same members (right).

We then assess the ability of SWEDE v2 to reproduce regional SWE from the withheld test member. Over 1980 to 2099 the all-pixel daily RMSE against WRF is 12.1 mm. The snow-covered pixel daily RMSE against WRF is 36.7 mm, compared with a WRF mean of 121.6 mm and standard deviation of 238.2 mm. The date of annual peak SWE has a mean absolute error of 3.1 days and the melt-out date 3.0 days (Table 1). Across the six subregions, the correlation between emulated and WRF water-year mean snow water volume ranges from 0.983 (NRM) to 0.996 (PNW) (Figure 2). Mean absolute error of the water-year mean volume ranges from 0.25 km³ in SN to 0.90 km³ in NCR, the largest-volume subregion, against WRF

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

long-term mean volumes of 2.3 km³ in SN and 19.5 km³ in NCR. Emulated and WRF water-year volumes lie on the 1:1 line from the driest years near zero to the wettest years above 50 km³ in NCR (Figure S3). The largest departures are in SN, where the emulator underestimates the three wettest years by roughly 1 km³. The daily regional maximum SWE tracks WRF over 1980-2099 (Figure S4), with daily correlations over 1980 to 2099 of 0.985 (SW, SN) to 0.991 (NCR). Mean annual maximum SWE, defined as the regionally averaged annual maximum SWE at each grid cell, is reproduced to within about 4% in every subregion and every state (Figure S5). The emulator overestimates in PNW and Washington by less than 3% and underestimates elsewhere, most in SW, SN, and Nevada. Pooled over all land cells and days, the emulated SWE distribution matches WRF across four orders of magnitude (Figure S6). The quantile mean absolute error is below 3 mm in every subregion.

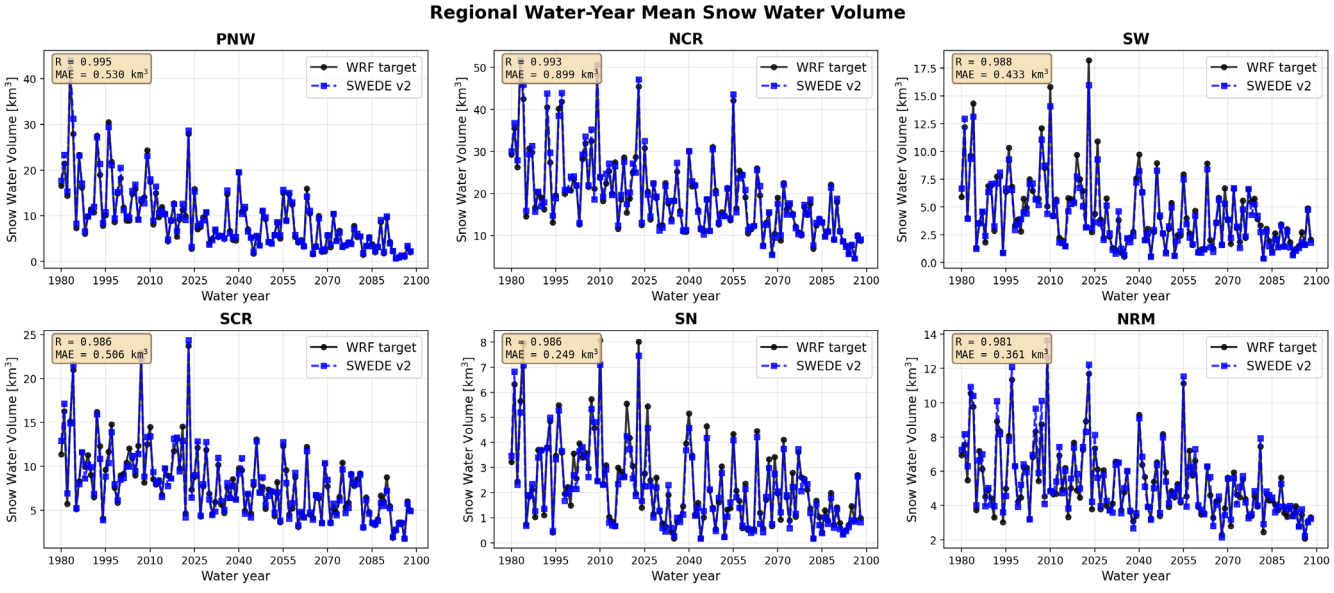

Figure 2. Water-year mean of the daily snow water volume integrated over each subregion for the test member, 1980 to 2099, from the WRF target (black) and SWEDE v2 (blue). Snow water volume is the sum of SWE over all terrestrial 9 km cells in the subregion, expressed in km³.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 1. Regional model performance for 1980–2099. Skill is summarized for the full domain and six snow-climate subregions using spatial, volumetric and seasonal-timing metrics for snow water equivalent (SWE).

Region	All-pixel RMSE (mm)	All-pixel bias (mm)	Interannual volume r	WY-mean volume MAE, km³ (% of mean volume / % of mean annual peak volume)	Peak-date MAE (d)	Melt-out MAE (d)	Mean annual max SWE bias (%)	Daily regional-max r
Domain	12.08	+0.29	0.997	2.72 (3.3%/1.2%)	3.10	3.03	+2.5	0.989
PNW	16.98	+0.57	0.996	0.53 (5.7%/1.8%)	3.19	5.36	+2.8	0.988
NCR	16.66	+0.57	0.994	0.90 (4.6%/1.5%)	5.64	2.53	−1.4	0.991
SW	14.62	−0.62	0.988	0.43 (9.8%/2.7%)	6.33	6.01	−4.1	0.985
SCR	12.64	−0.14	0.986	0.51 (6.2%/1.8%)	6.13	3.13	−3.8	0.988
SN	30.67	−2.75	0.986	0.25 (10.7%/3.0%)	5.77	6.54	−3.8	0.985
NRM	22.06	+1.14	0.983	0.36 (6.8%/2.2%)	3.68	3.40	−1.7	0.989

Daily SWE in the extreme years matches WRF in accumulation timing, peak, and melt-out. The three wettest and three driest water years were selected once for the whole domain, by detrended WRF peak snow water volume anomaly over 1850-2100 in the withheld member, and the same six years are shown for every subregion. Detrending removes the forced decline, so the selection samples internal variability across the whole record. Figures S7 to S9 show the snow water volume time series for the PNW, NCR, and SW, while Figures S10 to S12 show the same for SCR, NRM, and SN. Daily MAE is 10.2 mm or less in every selected year (Table 2). In the three wet years, the PNW bias is positive, +0.7 to +1.9 km³, consistent with the PNW overestimate in Figure S5. In NCR, where wet-year peaks reach 120 to 140 km³, the largest bias is +1.7 km³ in the wettest year, 2024. The SW emulation underestimates the two wettest years, 2024 and 2056, by 2.2 and 0.5 km³, with the error concentrated in the late-winter plateau. Absolute mean volume biases in SCR and NRM are at most 0.62 km³ (Table 2). The SN bias is negative in all three wet years, ranging from 0.2 to 0.6 km³. In the three dry years, the PNW bias remains positive, +0.2 to +0.7 km³. In NCR, the third-driest year, 1995, carries +1.6 km³ on the February to April plateau. Absolute mean volume biases in SCR and NRM are at most 0.57 km³, and the NRM emulation melts out about three weeks early in the second-driest year, 2066. In SN, the bias is positive in the driest year, 2046, where the emulator predicts 1 km³ more than WRF through February.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 2. Regional performance during six extreme water years. Each cell reports regional-mean daily mean absolute error (MAE; mm) / water-year-mean snow-water-volume bias (km³), followed in parentheses by the bias as a percentage of that year's target water-year-mean volume / that year's target annual peak volume. Water years run September–August and are labeled by their ending year. The three wet and three dry years were selected using domain-wide detrended WRF peak-volume anomalies. Bias is prediction minus target.

Region	Wet 1984	Wet 2024	Wet 2056	Dry 1995	Dry 2046	Dry 2066
PNW	4.61 / +1.92 (+4.6% / +1.6%)	3.02 / +0.67 (+2.4% / +0.8%)	1.59 / +0.69 (+4.6% / +1.6%)	1.69 / +0.65 (+8.3% / +2.7%)	0.84 / +0.37 (+20.6% / +2.1%)	0.53 / +0.23 (+14.9% / +3.5%)
NCR	2.24 / +0.33 (+0.6% / +0.2%)	3.50 / +1.73 (+3.8% / +1.4%)	2.69 / +1.41 (+3.3% / +1.1%)	2.99 / +1.62 (+12.3% / +4.2%)	2.03 / −1.03 (−9.1% / −2.5%)	1.08 / −0.01 (−0.1% / <0.1%)
SW	1.50 / −0.28 (−2.9% / −0.8%)	4.95 / −2.22 (−12.2% / −4.1%)	1.44 / −0.45 (−5.7% / −2.0%)	0.36 / −0.04 (−4.6% / −0.8%)	0.71 / +0.18 (+6.5% / +2.0%)	0.73 / −0.32 (−25.3% / −3.7%)
SCR	1.53 / −0.21 (−1.4% / −0.4%)	1.80 / +0.62 (+2.6% / +0.9%)	0.96 / +0.47 (+3.8% / +1.1%)	0.73 / −0.21 (−5.1% / −1.3%)	1.28 / +0.36 (+4.5% / +1.4%)	0.54 / −0.20 (−4.9% / −1.2%)
SN	8.48 / −0.45 (−8.0% / −2.1%)	10.17 / −0.56 (−6.9% / −2.5%)	6.27 / −0.24 (−5.6% / −1.8%)	1.87 / +0.06 (+14.9% / +2.1%)	4.94 / +0.25 (+24.7% / +6.3%)	2.67 / −0.14 (−24.5% / −4.8%)
NRM	4.62 / +0.39 (+3.7% / +1.4%)	5.97 / +0.53 (+4.5% / +1.7%)	5.82 / +0.43 (+3.8% / +1.3%)	5.79 / +0.57 (+18.8% / +6.0%)	4.67 / −0.37 (−6.8% / −2.3%)	5.06 / −0.49 (−15.0% / −4.1%)

The MLR benchmarks, fit on the same predictors, do not reproduce the skill of SWEDE v2 in daily SWE or in the timing of peak SWE and melt-out. All four regression variants (pooled and per-pixel MLR, each with and without CESM2 SWE as a predictor) overestimate SWE through the melt season and the summer minimum (Figure S13). The pooled variants hold 8 to 10 mm of spurious domain-mean snow in September. SWEDE v2 stays within 1.5 mm of the target on every day of the year, and its daily RMSE of domain-mean SWE never exceeds 2 mm. Over all 43,800 days the domain-mean daily SWE

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

from SWEDE v2 lies on the 1:1 line at every SWE level (Figure S14). Its RMSE is 0.88 mm and its correlation 0.9985. Over the same 120 years, SWEDE v2 reaches an all-pixel RMSE of 12.1 mm, in contrast with 26.9 mm for the best regression variant, per-pixel SWE-augmented MLR, and 47.1 mm for the training climatology (Figure S15). The pooled variants reach 65.5 and 68.5 mm RMSE with respect to WRF. In every region, the per-pixel regression variants carry 2.1 to 2.8 times more error than SWEDE v2 while the pooled variants have even larger errors (4.2 to 6.3 times) (Table 3). The performance gap is even larger when it comes to SWE timing. Per-pixel SWE-augmented MLR places the annual peak SWE within 10 days and melt-out within 35 days, against 3 days for SWEDE v2 (Figure S15). The pooled variants miss melt-out by 65 days, twice the error of the climatology. Per-pixel SWE-augmented MLR receives the CESM2 simulated snowpack that SWEDE v2 does not. Even with that input, its peak-day error is three times greater than SWEDE v2 while its melt-out error is ten times larger.

Table 3. Regional all-pixel RMSE (mm) of daily SWE on the withheld member LE2-1191.010 over 1980 to 2099 for SWEDE v2, the four regression variants, and the training climatology. Columns are the six subregions and the California (CA) and Upper Colorado (UCRB) basins. Values in parentheses are the ratio to the SWEDE v2 error in the same region.

Model	CA	NCR	NRM	PNW	SCR	SN	SW	UCRB
SWEDE v2	16.31 (1.0)	16.66 (1.0)	22.06 (1.0)	16.98 (1.0)	12.64 (1.0)	30.67 (1.0)	14.62 (1.0)	15.17 (1.0)
MLR pooled	86.39 (5.3)	95.28 (5.7)	118.33 (5.4)	106.62 (6.3)	55.07 (4.4)	158.35 (5.2)	76.70 (5.2)	69.00 (4.5)
MLR per-pixel	41.74 (2.6)	42.58 (2.6)	55.40 (2.5)	48.09 (2.8)	28.40 (2.2)	78.52 (2.6)	35.55 (2.4)	35.00 (2.3)
SWE-augmented MLR pooled	86.54 (5.3)	90.35 (5.4)	117.01 (5.3)	105.18 (6.2)	53.69 (4.2)	160.37 (5.2)	75.59 (5.2)	67.15 (4.4)
SWE-augmented MLR per-pixel	41.46 (2.5)	39.20 (2.4)	53.40 (2.4)	47.30 (2.8)	27.15 (2.1)	77.95 (2.5)	35.23 (2.4)	33.13 (2.2)
Climatology	62.26 (3.8)	65.22 (3.9)	72.28 (3.3)	86.75 (5.1)	35.16 (2.8)	120.16 (3.9)	53.20 (3.6)	44.96 (3.0)

4.2 Ensemble characterization

Having evaluated the emulator on the withheld member, we now characterize the full emulated ensemble, 100 members over 250 water years under historical and SSP3-7.0 forcing. We describe its median SWE, its variance across members, and its snow-season timing, and then compare its peak snow water volume with UA SWE over the five HUC2

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

basins. As expected, annual peak SWE decreases across the domain between 1981-2010 and 2071-2100 (Figure 3a to d). The largest losses occur along the Cascades crest and the Sierra Nevada, 400 mm or more, and over the lower-elevation margins of every range (exceeding 60%). By end-of-century, the high interior of the Northern Rocky Mountains and the Colorado Rockies retains the largest fraction of its peak snowpack. These high continental ranges have some of the coldest winters in the domain. April 1 SWE decreases more than peak SWE in relative terms (Figure 3e to h). Its losses exceed 400 mm along the Cascades crest and the Sierra Nevada and exceed 80% over the lower-elevation margins of every range. Part of the April 1 loss is a shift of the peak earlier in the season due to warming that exacerbates the loss of winter accumulation. The regional changes in peak SWE indicate declines of about 22% in NRM and 40 to 61% in the other five subregions (Figure 4). In every subregion, the April 1 and annual-mean volume declines exceed the peak decline. The error bars in Figure 4 span the range across all 100 members.

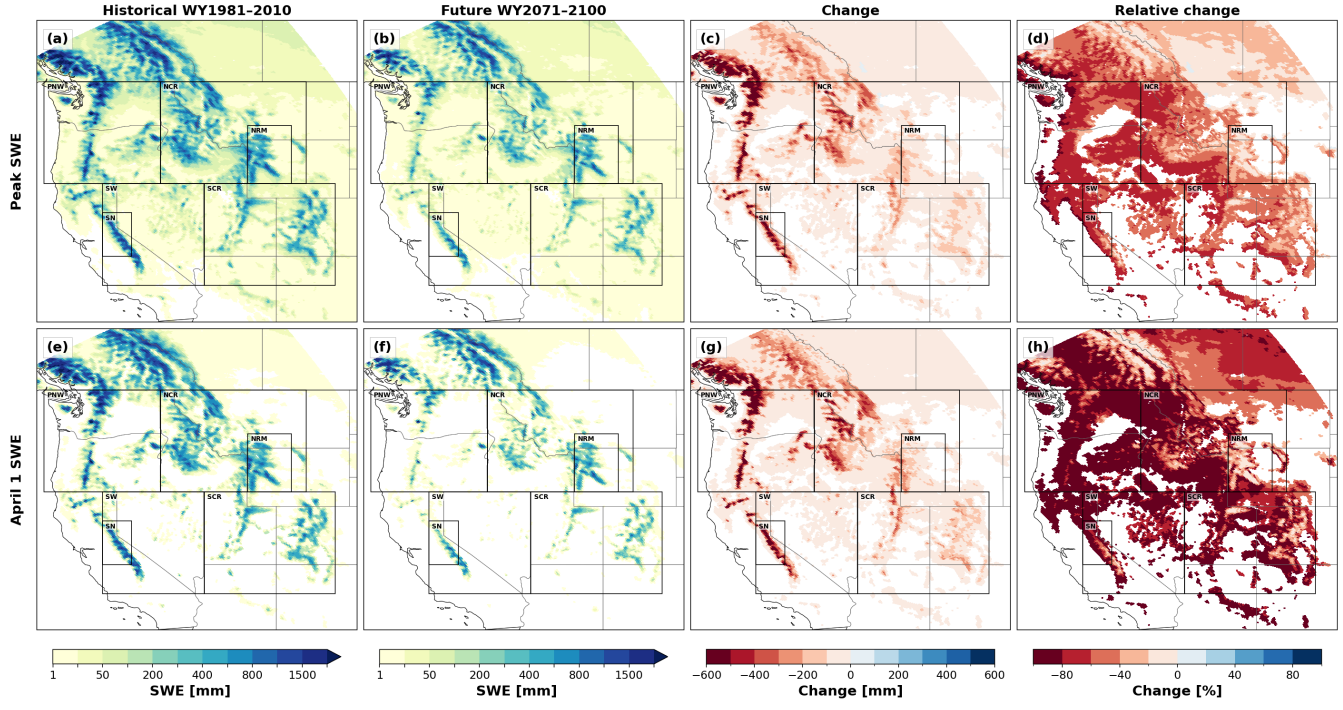

Figure 3. Ensemble-median (a to d) annual peak SWE and (e to h) April 1 SWE for (a, e) 1981 to 2010 and (b, f) 2071 to 2100 under SSP3-7.0, and (c, g) the absolute and (d, h) the relative change between the two periods, from the 100-member SWEDE v2 ensemble. Boxes outline the six subregions.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

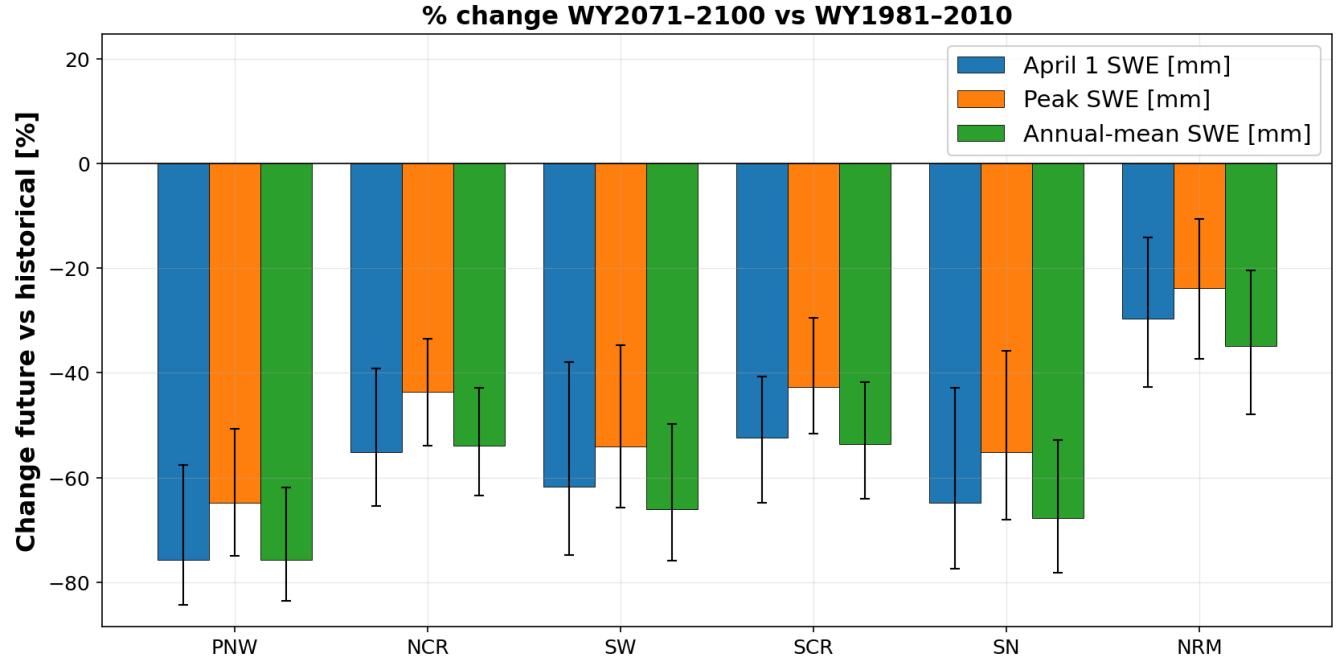

Figure 4. Ensemble-median percentage change between 1981-2010 and 2071-2100 under SSP3-7.0 in April 1 SWE (blue), annual peak SWE (orange), and annual-mean SWE (green) averaged over each subregion, from the 100-member SWEDE v2 ensemble. Within each subregion, the bars appear in this order from left to right. Percentage change is computed separately for each member as the difference between its 2071–2100 and 1981–2010 means, relative to its 1981–2010 mean. Bars show the median of the 100 member changes, and error bars span the full range across members. Annual peak SWE is the annual maximum of daily subregion-mean SWE. Annual-mean SWE is the water-year mean of daily subregion-mean SWE.

The ensemble median of April 1 SWE declines and the ensemble variance contracts with it (Figure 5). The median falls from pre-industrial (1851 to 1880) to historical (1981 to 2010) and again to end of century (2071 to 2100) in every range. Variance is computed across the 100 members after averaging April 1 SWE over the years of each period within each member, so it measures the spread of period means and excludes interannual variability. Cells that carried a variance above 1000 mm² in the pre-industrial Cascades and Sierra Nevada fall below 300 mm² by 2071 to 2100, a drop in ensemble standard deviation from above 31.6 mm to below 17.3 mm. The decrease in variance accompanies the decline in mean April 1 SWE and describes reduced dispersion of period means across members. Timing spread moves the other way, since the ensemble variance of snow onset rises over the marginal snow zones (Figure S16). The dataset therefore narrows in amplitude where snow persists and widens in timing where snow becomes intermittent.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

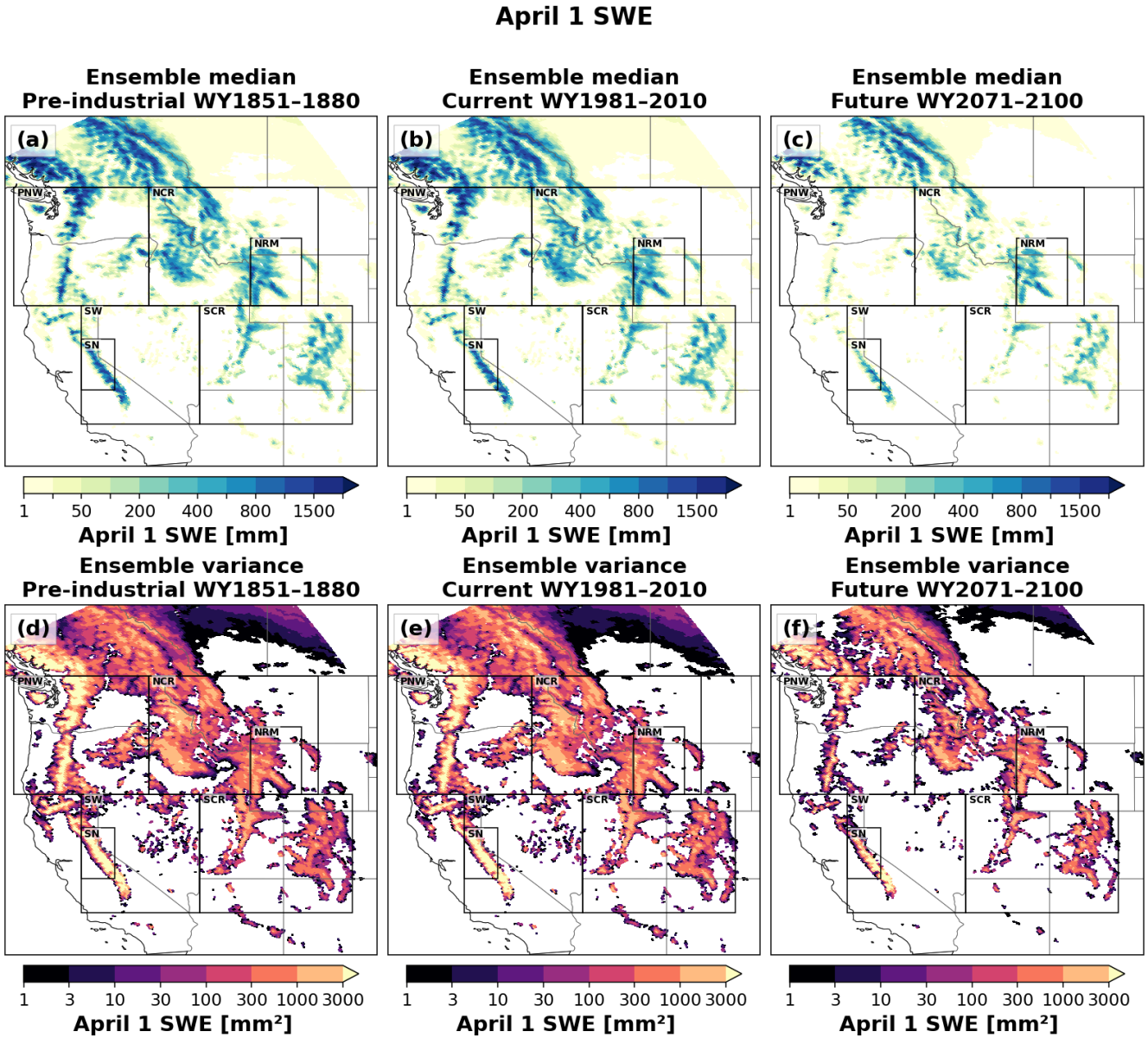

Figure 5. Ensemble median (a to c) and ensemble variance (d to f) of April 1 SWE across the 100 SWEDE v2 members for the pre-industrial period 1851 to 1880, the historical period 1981 to 2010, and the end-of-century period 2071 to 2100 under SSP3-7.0. Variance is on a logarithmic scale in mm^2.

Under future warming, SWEDE v2 also shows large shifts in the snow seasonal cycle. The ensemble-median day of peak SWE advances over most snow-covered grid cells (Figure 6a to c). The advance is 12 to 36 days over the maritime ranges and the mid-elevation interior. It is less than 12 days at the highest elevations of the Northern Rocky Mountains and the

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Colorado Rockies. Only isolated cells in the Sierra Nevada and the Southwest show a later peak. Melt-out advances by 40 to 60 days along the Cascades, the Olympics, and the Sierra Nevada and by 20 to 40 days over the interior ranges (Figure 6d to f), so the interval from peak to melt-out shortens over the maritime ranges. The Northern Rocky Mountains and the highest Colorado ranges retain a June melt-out at end-of-century. Onset shifts later over the lower and marginal snow zones and changes little on the crests, and its ensemble variance rises over the margins (Figure S16). The crests therefore keep their season length while the margins lose it at both ends.

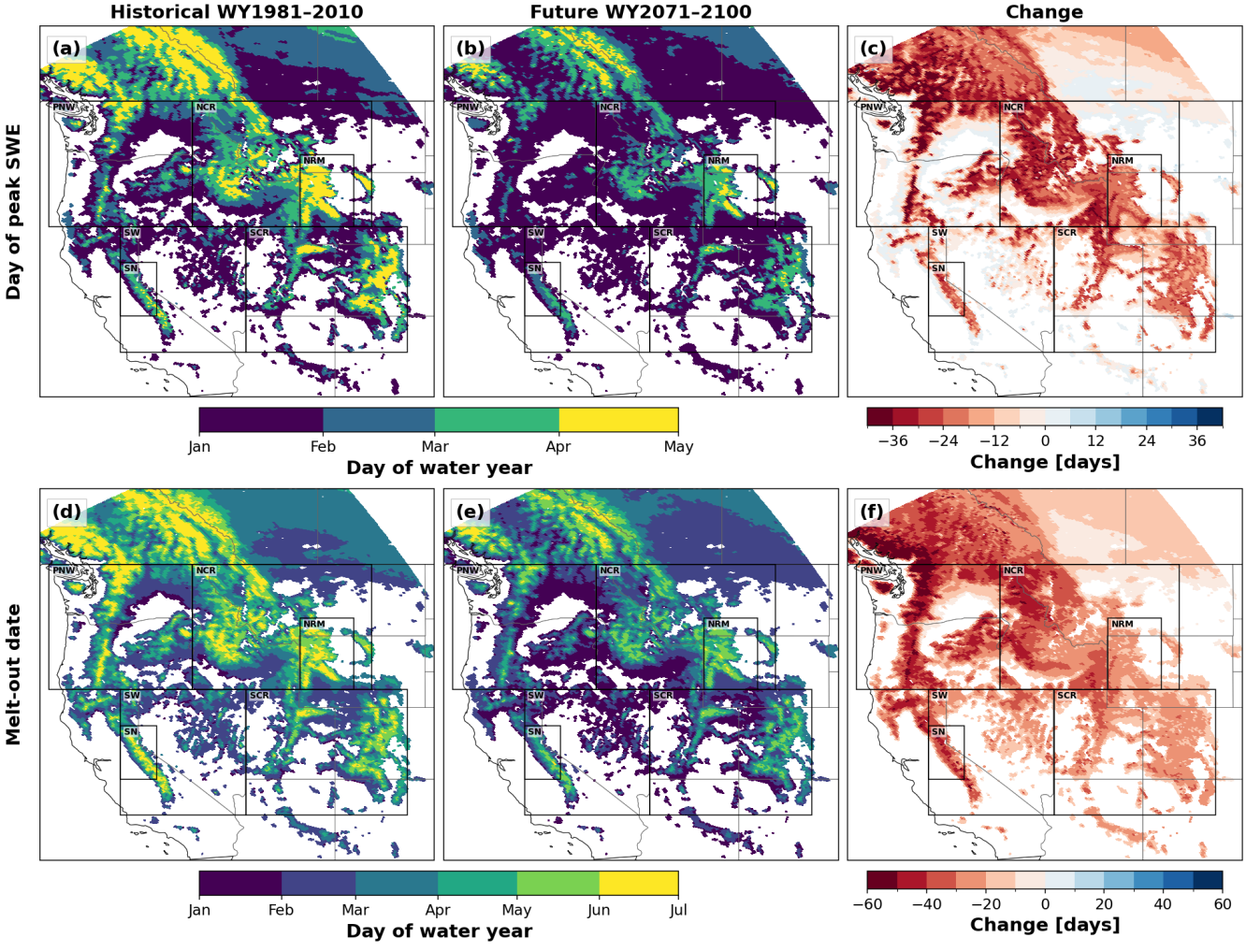

Figure 6. Ensemble-median (a to c) day of annual peak SWE and (d to f) melt-out date, expressed as day of the water year, for (a, d) 1981 to 2010 and (b, e) 2071 to 2100, and (c, f) the change in days. Cells whose ensemble-median peak SWE is below 25 mm are masked white. Melt-out is defined as the last day of the water year on which SWE exceeds 10 mm.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Relative to the ensemble median, observed peak volumes are close over the Upper Colorado and Columbia basins, lower over California, and higher over the Great Basin and the Lower Colorado. Observation-based peak snow water volumes lie within the full 100-member range in every year of the observed record over all five basins (Figure 7 and Figures S17 to S20). Over the Upper Colorado basin (HUC2 region 14), UA SWE peak volume for water years 1981 to 2022 also falls inside the 5th to 95th percentile band of the ensemble in 38 of 41 years and exceeds the 95th percentile in 1993, 2008, and 2019. The observed series sits below the ensemble median in most years only over California (26 of 41 water years, with a 1981 to 2010 median 3.5 km^3 or 14% below the ensemble-median normal), where it falls under the 5th percentile in four years. It tracks the median over the Upper Colorado (22 of 41 years below, 6% below the normal) and over the Columbia footprint (18 of 41, 4% above), and it runs above the ensemble over the Great Basin and the Lower Colorado (15% and 22% above the normal, with six and four years above the 95th percentile). The California deficit matches the wet bias of WRF over the Sierra Nevada (Rahimi et al., 2024b), which SWEDE v2 inherits from its training data. The interior surplus has two candidate sources, an underestimate of peak volume in the WRF training data and biases in the observational product itself (Dixit et al., 2026). The UA SWE and SWEDE v2 differences therefore combine model bias and observational bias, and neither series is an unbiased reference for the other.

Relative to the 1981-2010 ensemble-median annual peak snow volume of 23.3 km^3 in the Upper Colorado River Basin, the median is 1% higher in 1851-1880, 23% lower in 2041-2070, and 37% lower in 2071-2100. The four other HUC2 basins show the same pattern of decline (Figures S17 to S20). Their 1981-2010 normals are 3.6 km^3 (Lower Colorado), 15.2 km^3 (Great Basin), 143.0 km^3 (Columbia), and 24.1 km^3 (California). Their end-of-century declines are 67%, 59%, 52%, and 61%, and their mid-century declines 48%, 39%, 32%, and 42% under the SSP3-7.0 warming trajectory. Across the five basins, the ensemble median of peak volume therefore declines by 23 to 48% by 2041-2070 and by 37 to 67% by 2071-2100 relative to 1981-2010. For the Columbia basin the observed series is compared with the ensemble median over the UA product footprint, which holds 68% of basin volume, because the basin extends into Canada. The Lower Colorado has the smallest normal, the largest relative decline, and a strongly right-skewed distribution of peak volume in every period.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

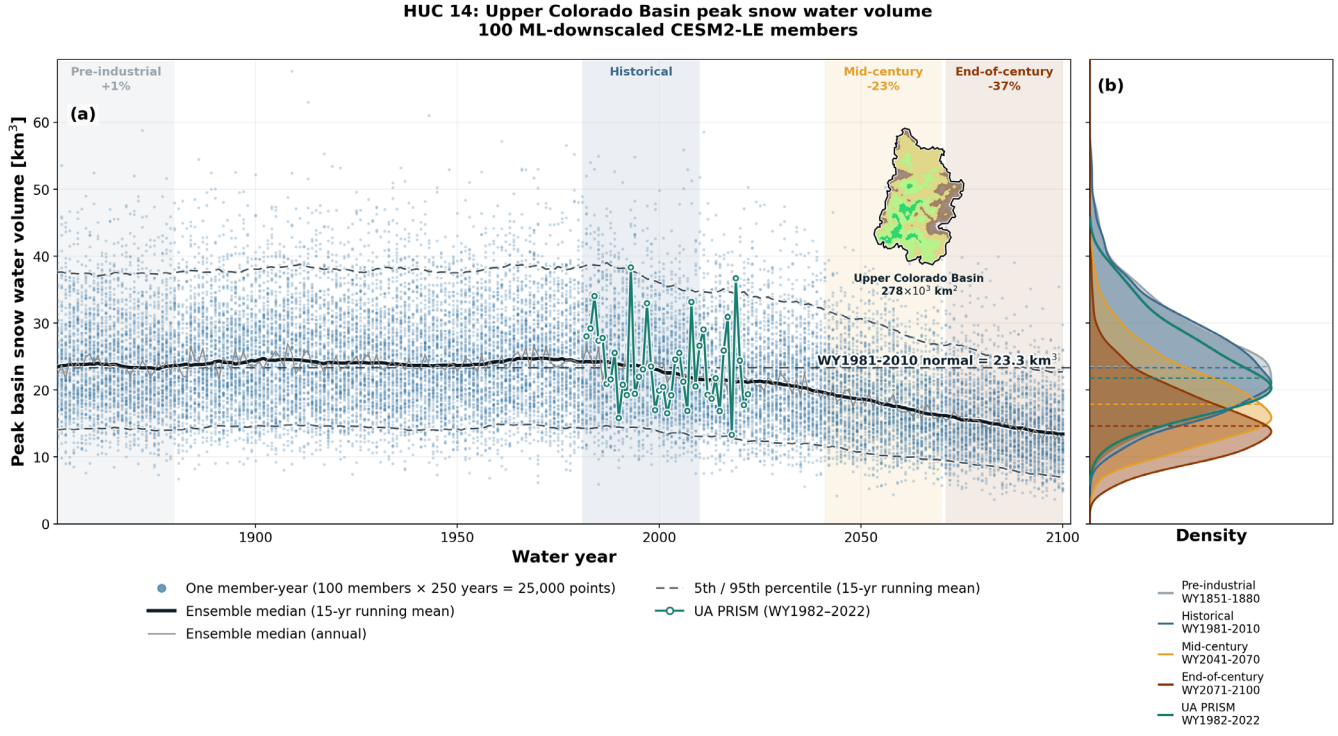

Figure 7. Annual peak snow water volume over the Upper Colorado basin (HUC2 region 14, basin area 278 × 10^3 km²) in the 100-member SWEDE v2 ensemble. (a) Each blue dot is the peak volume of one member in one water year, for 100 members and water years 1851 to 2100. The thick black line is the ensemble median smoothed with a 15-year running mean, the thin grey line is the unsmoothed annual ensemble median, and the dashed grey lines are the 5th and 95th percentiles across members, also smoothed with a 15-year running mean. Green circles joined by a green line are the annual peak volume from the University of Arizona 4 km SWE product integrated over the basin for water years 1981 to 2021. The horizontal dashed line marks the ensemble-median normal for 1981 to 2010, with its value printed beside it. Four shaded bands mark the pre-industrial (1851 to 1880), historical (1981 to 2010), mid-century (2041 to 2070), and end-of-century (2071 to 2100) periods, and each band is labelled with the change of the period ensemble median relative to the historical normal. The inset map shows the basin outline over shaded terrain. (b) Kernel density of peak volume across all member-years in each of the four periods, using the same colors as the band labels, and of the University of Arizona annual peak volume (green), with the median of each distribution marked by a horizontal dashed line. The vertical axis of (b) is shared with (a).

The range of possible peak volumes narrows through the twenty-first century (Figure S21). In every subregion the ensemble mean of peak snow water volume is stationary through 1950, declines slowly to 2000, and declines steadily thereafter. The upper bound of the full member range falls faster than the mean, indicating that the occurrence of extremely high snow years is likely to experience particularly large decreases . The decline is smallest in NRM, where the mean falls

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

from about 19 to 13 km³. It is largest in PNW and SN, where the mean approaches one-quarter of its historical value near the end of simulation.

4.3 Sampling uncertainty of a snow-drought risk ratio

The value of ninety additional members is a narrower confidence interval on a tail statistic, and the 2026 snow drought provides a test with a published benchmark. Marshall et al. (2026) estimated the risk ratio (RR) of the 2026 western United States snow drought from the ten WRF-downscaled CESM2-LE members as the probability of a March 15 SWE volume at least as low as observed in 2026, evaluated in 2016 to 2036 divided by the same probability in 1850 to 1900, using generalized extreme value fits to SWE volume z-scores. We repeated their calculation on the SWEDE v2 emulation of the same ten members and on all 100 members (Figure 8). On the ten emulated members the RR is 3.9 [95% CI 2.5, 6.7] for the western United States, 3.0 [2.0, 5.1] for California, 2.9 [2.2, 3.9] for the Columbia basin, and 13 [0.13, 13000] for the Upper Colorado, against 4.4 [2.6, 9.4], 2.9 for both California and the Columbia basin, and 14 [0.09, 4300] from the WRF members. Thus, the emulator approximately reproduces the WRF-derived risk ratios for the ten WRF-downscaled members (eight training, one validation, and one test).

With 100 members the medians for the western United States, California, and the Columbia change little, to 4.1, 2.8, and 2.8. However, the intervals narrow to [3.5, 4.9], [2.5, 3.2], and [2.6, 3.1]. The interval width in log space falls by a factor of 2.9 to 3.8, close to the factor of 3.2 expected from a tenfold increase in sample size. In the Upper Colorado the 100 members raise the median RR from 13 to 53 and the lower bound from 0.13 to 6.5, so an event that ten members could not distinguish from no change becomes at least sixfold more likely. The median and upper bound remain poorly constrained in both ensembles, with upper bounds above 10^4, because the observed 2026 z-score lies beyond the counterfactual sample. These intervals treat the ninety emulated members as independent realizations. Table 2 shows that the emulator's errors are generally smallest in the dry years, which supports that treatment, but the confidence intervals are conditional on the fitted emulator and the assumed sampling structure; they do not quantify shared emulator error.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

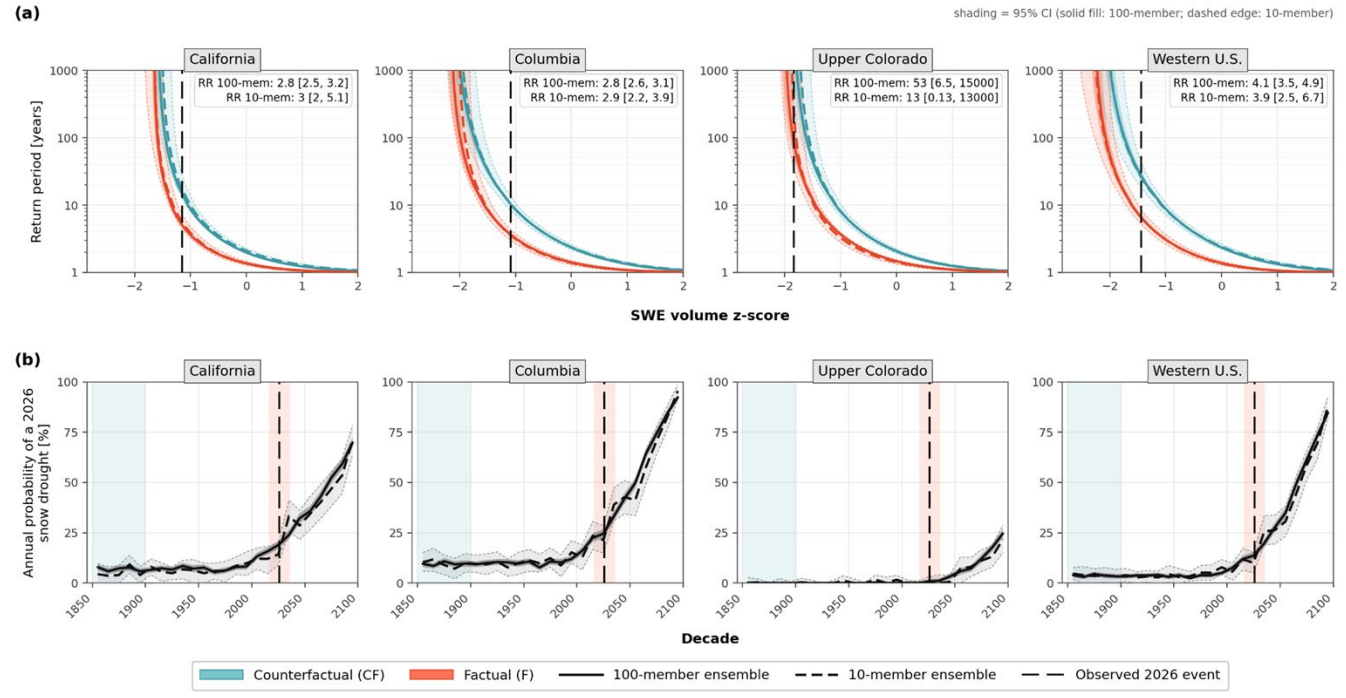

Figure 8. Risk ratio of the 2026 western United States snow drought from the 10-member and 100-member SWEDE v2 ensembles, for the California, Columbia, and Upper Colorado basins and the western United States. (a) Return period of March 15 SWE volume z-score in the counterfactual (1850 to 1900, blue) and factual (2016 to 2036, red) periods from the 10-member (dashed) and 100-member (solid) SWEDE v2 ensembles, with 95% confidence intervals shaded. The vertical line is the observed 2026 z-score from ERA5-Land, and the inset gives the risk ratio with its 95% interval. (b) Annual probability of a snow drought at least as severe as 2026 by decade, 1850 to 2100, median and 95% interval for both ensembles; bands mark the counterfactual and factual periods.

5. Discussion and Conclusions

SWEDE v2 emulates WRF-downscaled daily SWE directly from native, uncorrected CESM2-LE output, which makes WRF-consistent SWE available for every member of the ensemble. The product is the first 100-member, 9 km daily SWE dataset for the western United States, covering 250 water years. With eight WRF-downscaled training members, SWEDE v2 achieves a daily all-pixel RMSE of 12.1 mm and a peak-date MAE of 3 days on the withheld member. The emulated ten members approximately reproduce the 2026 snow-drought risk ratio obtained from WRF, and the hundred members narrow its 95% interval by a factor of about three without moving its median in the three domains where ten members already bound it. The projected changes in the emulated ensemble follow the established response of western snowpack to warming. As in previous studies, relative losses are larger in the maritime ranges than in the interior ranges (Siirila-Woodburn et al., 2021),

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

April 1 SWE declines more than peak SWE as the peak moves earlier (Beltran-Peña et al., 2025), and variability contracts with the declining snowpack (Marshall et al., 2019; Cho et al., 2021). Multi-model assessments constrain the magnitude of this response better than one ESM can, because they also span model uncertainty (Siirila-Woodburn et al., 2021; Norris et al., 2025). The distinct value of the emulated ensemble lies in sampling, with 100 realizations of internal variability at 9 km from one ESM under one scenario.

The dataset also fills the gap that the two existing routes around the downscaling cost leave open. Norris et al. (2025) extended end-of-century snowfall loss in WUS-D3 to the full CMIP6 archive through its regression on driving-model warming and precipitation change. That route broadens sampling of model-dependent snowfall change, but it returns one period-mean change per model and no daily fields. Lopez-Gomez et al. (2025) replaced the 9 km stage of WUS-D3 with a diffusion model conditioned on the intermediate-resolution dynamical fields. That route lowers the cost of the final refinement, but each new member still requires the intermediate dynamical run, and it targets meteorological fields rather than a land state with seasonal memory. SWEDE v2 removes the regional model from the loop for the emulated members and returns daily 9 km SWE for all 100 initial-condition members of one ESM. The Norris et al. (2025) extrapolation spans model uncertainty across the CMIP6 archive, and this dataset complements it by spanning internal variability with one hundred realizations of one model under one scenario. Together, these products expand the sampling of model differences and internal variability, while uncertainty from emissions scenarios, the regional model, and the emulators remains.

SWEDE v2 takes raw CESM2-LE fields and returns SWE consistent with WRF driven by bias-corrected boundary conditions. The network thus learns the bias correction together with the downscaling, as emulators trained on global model inputs learn the inconsistencies between global and regional models (Van Der Meer et al., 2023). A user who wants SWE for a new CESM2-LE member needs only the native daily archive and the trained weights. The regression comparison demonstrates an improvement over the linear benchmarks. A linear map of the same inputs overestimates summer snow. Even when given the CESM2 snowpack, it misplaces melt-out. The added skill of SWEDE v2 therefore comes from its network design. The memoryless baseline anchors the seasonal climatology and the recurrent increment carries the accumulation and ablation history. The seasonal relaxation is designed to limit the accumulation of recurrent error and to reproduce the annual reset in the WRF targets. What remains is the amplitude of the deepest seasons. The emulator underestimates the wettest years in Southwest and in Sierra Nevada, and overestimates mean annual maximum SWE in Pacific Northwest. These are the regimes Part 1 identified as most challenging.

The dataset is built for questions that are difficult to address with ten members. Snow-drought attribution, extreme-value analysis of peak SWE and melt-out, and water-year risk for a fixed basin all need the full distribution across members. The ensemble also supports estimates of the frequency of snow deluges (Marshall et al., 2024) and of consecutive snow-drought years (Marshall et al., 2019). Internal variability has slowed the decline of western snowpack since the 1980s (Siler et al., 2019), and the ensemble can show how often such slowdowns occur over a planning horizon and when the forced loss emerges from internal variability in each range. Daily 9 km fields from many realizations of the same climate can also drive

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

hydrological and water-management models to test reservoir operations against snowpack sequences that ten members rarely sample.

Applications involving rare events require evaluation of the relevant tails and an explicit treatment of sampling and emulator uncertainty. The dataset carries the snow physics of the WRF configuration in WUS-D3, the biases of CESM2 under SSP3-7.0, and the September cold start of the WRF water-year runs. It represents one ESM and one scenario. The ten WRF members are distributed alongside it as the benchmark for any use of the emulated ninety. Extending SWEDE to another ESM requires retraining or fine-tuning, because the network learns the mapping from one model's biased fields to WRF. Elevation dependence, transfer learning applicability, emergence timing, and the separation of externally forced signals from internal variability are the subject of separate work. More broadly, the results suggest that future downscaling campaigns can be designed around emulation, in which an emulator trained on a small set of dynamically downscaled members that spans the forcing range of interest extends them to the full ensemble.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix A: Model Input Features

Table A1. Input features for SWEDE v2

Original variable name	CF-compliant variable name	Full name	Model
QREFHT	specific_humidity_at_2m	Reference-height (2 m) specific humidity	Baseline, Increment
TREFHT	air_temperature_at_2m	Reference-height (2 m) air temperature	Baseline, Increment
PRECT	total_precipitation_rate	Total precipitation rate (convective + large-scale)	Baseline, Increment
PS	surface_air_pressure	Surface pressure	Baseline, Increment
TREFHTMN	daily_minimum_air_temperature_at_2m	Daily minimum reference-height air temperature	Baseline, Increment
TREFHTMX	daily_maximum_air_temperature_at_2m	Daily maximum reference-height air temperature	Baseline, Increment
FLDS	surface_downwelling_longwave_flux	Downwelling longwave flux at the surface	Baseline, Increment
FSDS	surface_downwelling_shortwave_flux	Downwelling shortwave flux at the surface	Baseline, Increment
TREFHT_lag90d_mean	air_temperature_at_2m_90d_trailing_mean	90-day trailing mean of TREFHT	Baseline
PRECT_lag110d_mean	total_precipitation_rate_110d_trailing_mean	110-day trailing mean of PRECT	Baseline
FLDS_lag80d_mean	surface_downwelling_longwave_flux_80d_trailing_mean	80-day trailing mean of FLDS	Baseline
FSNS_lag130d_mean	surface_net_shortwave_flux_130d_trailing_mean	130-day trailing mean of net shortwave flux at the surface	Baseline
QREFHT_lag45d_mean	specific_humidity_at_2m_45d_trailing_mean	45-day trailing mean of QREFHT	Baseline
PRECT_B1C_ACCUM_180d	precipitation_below_1c_180d_accumulation	180-day accumulated precipitation falling on days with TREFHT < 274.15 K (1 °C)	Baseline
PRECT_B1C_lag30d	precipitation_below_1c_30d_trailing_mean	30-day trailing mean of below-1 °C precipitation	Baseline
PRECT_A1C_lag30d	precipitation_at_or_above_1c_30d_trailing_mean	30-day trailing mean of at-or-above-1 °C precipitation	Baseline
PDD_TREFHTMX_lag30d	positive_degree_day_sum_30d	30-day trailing positive degree-days, Σ max(TREFHTMX − 273.15 K, 0), formed on the native CESM grid	Baseline
(lowres terrain)	coarse_grid_surface_altitude	Coarse CESM-grid terrain height	Baseline, Increment
(DOY sin)	day_of_year_sine	sin(2π·DOY/365)	Baseline, Increment
(DOY cos)	day_of_year_cosine	cos(2π·DOY/365)	Baseline, Increment

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Original variable name	CF-compliant variable name	Full name	Model
LANDMASK	land_binary_mask	Land–water mask	Baseline, Increment (static)
LU_INDEX	land_cover_category_index	Land-use / land-cover category index	Baseline, Increment (static)
HGT	surface_altitude	High-resolution terrain height	Baseline, Increment (static)
VAR	subgrid_surface_altitude_variance	Sub-grid terrain height variance	Baseline, Increment (static)
GRADHGT_X	surface_altitude_gradient_x	Terrain height gradient, x	Baseline, Increment (static)
GRADHGT_Y	surface_altitude_gradient_y	Terrain height gradient, y	Baseline, Increment (static)
(previous-day hybrid SWE)	previous_day_hybrid_snow_water_equivalent	Recurrent state H(t−1)	Increment only

Appendix B: CESM2-LE members downscaled in WUS-D3

Table B1. The ten CESM2-LE members downscaled by WRF in WUS-D3 and their role in SWEDE v2. All ten belong to the smoothed biomass-burning subset and start from the listed year of the pre-industrial control run.

Member	Control start year	Role
LE2-1011.001	1011	Training
LE2-1031.002	1031	Training
LE2-1051.003	1051	Training
LE2-1071.004	1071	Training
LE2-1091.005	1091	Training
LE2-1111.006	1111	Training
LE2-1131.007	1131	Training
LE2-1171.009	1171	Training
LE2-1151.008	1151	Validation
LE2-1191.010	1191	Test

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Code and data availability

The trained SWEDE v2 model code, training scripts, evaluation scripts, and trained checkpoint are available on Zenodo at https://doi.org/10.5281/zenodo.24483810. All data used for model training and evaluation in this study were obtained from WUS-D3 and are publicly available in the Amazon S3 bucket s3://wrf-cmip6-noversioning/ at https://registry.opendata.aws/wrf-cmip6/ (Rahimi et al., 2024b), accessible through the AWS Open Data Registry. SWEDE is written in Python 3.12 (https://www.python.org/) with PyTorch 2.5.1 (https://pytorch.org/get-started/previous-versions/).

Author contributions

FZ and SR conceived the study and designed the methodology. FZ developed all model and evaluation code, conducted the experiments, analyzed the results, prepared the figures, and wrote the original manuscript. SR, MB supervised the research. CT and JN contributed substantially to manuscript review and editing. MB, KS, AH, AM, and WL also reviewed and edited the manuscript. MB and SR acquired the funding.

Competing interests

At least one of the (co-)authors is a member of the editorial board of Geoscientific Model Development.

Acknowledgements

We would like to acknowledge the use of the Derecho system (doi:10.5065/qx9a-pg09) supported by the NSF National Center for Atmospheric Research (NCAR) at the NSF NCAR-Wyoming Supercomputing Center, sponsored by the National Science Foundation and the State of Wyoming.

Financial support

FZ, MB, and SR acknowledge funding from the U.S. Department of Energy Office of Biological and Environmental Research, Earth and Environmental Systems Sciences Division, through Lawrence Livermore National Lab for the HyperFACETS project, “A Framework for Improving Analysis and Modeling of Earth System and Intersectoral Dynamics at Regional Scales” (award no. DE-SC0016605). SR acknowledges the financial support from a National Oceanic & Atmospheric Administration MAPP-NIDIS grant, “Towards predicting drought and subsequent water resource challenges at landscape-resolving scales across the western U.S.” (award no. NA230AR4310633). MB and SR also acknowledge the National Science Foundation-supported project, “Wyoming Anticipating the Climate-Water Transition (WY-ACT)” (award no. 2149105). SR acknowledges support from the NSF ASCEND Engine (award no. 24-0767-A0002-SUB01). CT and JN acknowledge funding from the

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.



Regional and Global Model Analysis Program for the Office of Science of the U.S. Department of Energy through the PCMDI project.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., and Dozier, J.: Mountain hydrology of the western United States, Water Resources Research, 42, https://doi.org/10.1029/2005WR004387, 2006.

Baño-Medina, J., Manzanas, R., and Gutiérrez, J. M.: Configuration and intercomparison of deep learning neural models for statistical downscaling, Geoscientific Model Development, 13, 2109-2124, 2020.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303-309, 10.1038/nature04141, 2005.

Beltran-Peña, A., Rhoades, A., Burakowski, E., Girotto, M., Michalak, A. M., Diffenbaugh, N. S., Inda-Diaz, H., and D'Odorico, P.: Future implications of enhanced hydroclimate variability and reduced snowpack on California's water resources, Environmental Research: Water, 1, 025004, 2025.

Broxton, P., Zeng, X., and Dawson, N.: Daily 4 km Gridded SWE and Snow Depth from Assimilated In-Situ and Modeled Data over the Conterminous US, Version 1, NASA National Snow and Ice Data Center Distributed Active Archive Center [dataset], 10.5067/0GGPB220EX6A, 2019.

Cho, E., McCrary, R. R., and Jacobs, J. M.: Future changes in snowpack, snowmelt, and runoff potential extremes over North America, Geophysical Research Letters, 48, e2021GL094985, 2021.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., Emmons, L., Fasullo, J., Garcia, R., and Gettelman, A.: The community earth system model version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001916, 2020.

Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability, Climate Dynamics, 38, 527-546, 10.1007/s00382-010-0977-x, 2012.

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., Fiore, A., Frankignoul, C., Fyfe, J. C., Horton, D. E., Kay, J. E., Knutti, R., Lovenduski, N. S., Marotzke, J., McKinnon, K. A., Minobe, S., Randerson, J., Screen, J. A., Simpson, I. R., and Ting, M.: Insights from Earth system model initial-condition large ensembles and future prospects, Nature Climate Change, 10, 277-286, 10.1038/s41558-020-0731-2, 2020.

Dixit, A., Rahimi, S., Huang, L., Musselman, K. N., Vano, J. A., Addor, N., and Lehner, F.: High-resolution models better simulate historical snowpack declines in the Upper Colorado River Basin, Environmental Research Letters, 21, 054012, 10.1088/1748-9326/ae4113, 2026.

Doury, A., Somot, S., Gadat, S., Ribes, A., and Corre, L.: Regional climate model emulator based on deep learning: concept and first evaluation of a novel hybrid downscaling approach, Climate Dynamics, 60, 1751-1779, 10.1007/s00382-022-06343-9, 2023.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999-2049, https://doi.org/10.1002/qj.3803, 2020.

Hobeichi, S., Nishant, N., Shao, Y., Abramowitz, G., Pitman, A., Sherwood, S., Bishop, C., and Green, S.: Using Machine Learning to Cut the Cost of Dynamical Downscaling, Earth's Future, 11, e2022EF003291, https://doi.org/10.1029/2022EF003291, 2023.

Ikeda, K., Rasmussen, R., Liu, C., Gochis, D., Yates, D., Chen, F., Tewari, M., Barlage, M., Dudhia, J., and Miller, K.: Simulation of seasonal snowfall over Colorado, Atmospheric Research, 97, 462-477, 2010.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., and Hu, W.: Learning skillful medium-range global weather forecasting, Science, 382, 1416-1421, 2023.

Li, D., Wrzesien, M. L., Durand, M., Adam, J., and Lettenmaier, D. P.: How much runoff originates as snow in the western United States, and how will that change in the future?, Geophysical Research Letters, 44, 6163-6172, https://doi.org/10.1002/2017GL073551, 2017.

Lopez-Gomez, I., Wan, Z. Y., Zepeda-Núñez, L., Schneider, T., Anderson, J., and Sha, F.: Dynamical-generative downscaling of climate model ensembles, Proceedings of the National Academy of Sciences, 122, e2420288122, 10.1073/pnas.2420288122, 2025.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Lundquist, J., Abel, M. R., Gutmann, E., and Kapnick, S.: Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks, Bulletin of the American Meteorological Society, 100, 2473-2490, 2019.

Lundquist, J. D., Abel, M. R., Currier, W. R., and Jackson, D. L.: Snow What? Strengths and Limitations of Different Snow Products in Western U.S. Mountains, Water Resources Research, 62, e2026WR043728, https://doi.org/10.1029/2026WR043728, 2026.

Marshall, A. M., Abatzoglou, J. T., Link, T. E., and Tennant, C. J.: Projected changes in interannual variability of peak snowpack amount and timing in the Western United States, Geophysical Research Letters, 46, 8882-8892, 2019.

Marshall, A. M., Cowherd, M., Rahimi, S., and Ye, Y.: The 2026 western US snow drought was about four times more likely due to climate change, Proceedings of the National Academy of Sciences, 123, e2612961123, 10.1073/pnas.2612961123, 2026.

Marshall, A. M., Abatzoglou, J. T., Rahimi, S., Lettenmaier, D. P., and Hall, A.: California's 2023 snow deluge: Contextualizing an extreme snow year against future climate change, Proceedings of the national academy of sciences, 121, e2320600121, 2024.

Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., and Engel, R.: Dramatic declines in snowpack in the western US, Npj Climate and Atmospheric Science, 1, 2, 2018.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J. N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349-4383, 10.5194/essd-13-4349-2021, 2021.

Musselman, K. N., Clark, M. P., Liu, C., Ikeda, K., and Rasmussen, R.: Slower snowmelt in a warmer world, Nature Climate Change, 7, 214-219, 2017.

National Academies of Sciences, E. and Medicine: Attribution of extreme weather and climate events and their impacts, 2026.

Norris, J., Rahimi, S., Huang, L., Bass, B., Thackeray, C. W., and Hall, A.: Uncertainty of 21st Century western U.S. snowfall loss derived from regional climate model large ensemble, npj Climate and Atmospheric Science, 8, 134, 10.1038/s41612-025-01002-2, 2025.

Rahimi, S., Krantz, W., Lin, Y. H., Bass, B., Goldenson, N., Hall, A., Lebo, Z. J., and Norris, J.: Evaluation of a reanalysis-driven configuration of WRF4 over the western United States from 1980 to 2020, Journal of Geophysical Research: Atmospheres, 127, e2021JD035699, 2022.

Rahimi, S., Huang, L., Norris, J., Hall, A., Goldenson, N., Krantz, W., Bass, B., Thackeray, C., Lin, H., and Chen, D.: An overview of the Western United States dynamically downscaled dataset (WUS-D3), Geoscientific Model Development, 17, 2265-2286, 2024a.

Rahimi, S., Huang, L., Norris, J., Hall, A., Goldenson, N., Risser, M., Feldman, D. R., Lebo, Z. J., Dennis, E., and Thackeray, C.: Understanding the Cascade: Removing GCM Biases Improves Dynamically Downscaled Climate Projections, Geophysical Research Letters, 51, e2023GL106264, https://doi.org/10.1029/2023GL106264, 2024b.

Rahimi, S., Huang, L., Thackeray, C. W., Risser, M. D., Tarouilly, E., Norris, J., Rhoades, A., Bass, B., Hall, A., Lebo, Z. J., Liu, W., Lehner, F., Corrie, T., and Dixit, A.: Bias Correction Can Strongly Shape California's Projected Megaflood Risk, Journal of Geophysical Research: Atmospheres, 131, e2026JD046503, https://doi.org/10.1029/2026JD046503, 2026.

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., Tewari, M., Barlage, M., Dudhia, J., and Yu, W.: High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: a process study of current and warmer climate, Journal of Climate, 24, 3015-3048, 2011.

Ritchie, E., Wood, A. W., Johnson, R., Marshall, A., Sturtevant, J., Liljestrand, D., and Golitzin, E.: Benchmarking Catchment-Scale Snow Water Equivalent Datasets and Models in the Western United States, EGUsphere, 2025, 1-38, 10.5194/egusphere-2025-5514, 2025.

Rodgers, K. B., Lee, S. S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., Edwards, J., Kim, J. E., Simpson, I. R., Stein, K., Stuecker, M. F., Yamaguchi, R., Bódai, T., Chung, E. S., Huang, L., Kim, W. M., Lamarque, J. F., Lombardozzi, D. L., Wieder, W. R., and Yeager, S. G.: Ubiquity of human-induced changes in climate variability, Earth Syst. Dynam., 12, 1393-1411, 10.5194/esd-12-1393-2021, 2021.

Shao, D., Li, H., Wang, J., Hao, X., Che, T., and Ji, W.: Reconstruction of a daily gridded snow water equivalent product for the land region above 45° N based on a ridge regression machine learning approach, Earth Syst. Sci. Data, 14, 795-809, 10.5194/essd-14-795-2022, 2022.

https://doi.org/10.5194/egusphere-2026-7362
Preprint. Discussion started: 26 November 2026
© Author(s) 2026. CC BY 4.0 License.

Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., Nico, P. S., Feldman, D. R., Jones, A. D., and Collins, W. D.: A low-to-no snow future and its impacts on water resources in the western United States, Nature Reviews Earth & Environment, 2, 800-819, 2021.

Siler, N., Proistosescu, C., and Po-Chedley, S.: Natural Variability Has Slowed the Decline in Western U.S. Snowpack Since the 1980s, Geophysical Research Letters, 46, 346-355, https://doi.org/10.1029/2018GL081080, 2019.

van der Meer, M., de Roda Husman, S., and Lhermitte, S.: Deep Learning Regional Climate Model Emulators: A Comparison of Two Downscaling Training Frameworks, Journal of Advances in Modeling Earth Systems, 15, e2022MS003593, https://doi.org/10.1029/2022MS003593, 2023.

Yang, K., Musselman, K. N., Rittger, K., Margulis, S. A., Painter, T. H., and Molotch, N. P.: Combining ground-based and remotely sensed snow data in a linear regression model for real-time estimation of snow water equivalent, Advances in Water Resources, 160, 104075, https://doi.org/10.1016/j.advwatres.2021.104075, 2022.

Zeng, X., Broxton, P., and Dawson, N.: Snowpack Change From 1982 to 2016 Over Conterminous United States, Geophysical Research Letters, 45, 12,940-912,947, https://doi.org/10.1029/2018GL079621, 2018.

Zhao, F., Rahimi, S., Bukovsky, M., Singh, K., Thackeray, C., Norris, J., Hall, A., and Liu, W.: SWEDE v1.0: emulating dynamically downscaled snow water equivalent across the western United States depends on training data composition, EGUsphere, 2026, 1-33, 10.5194/egusphere-2026-5377, 2026.