

Original research article

A statewide, weather-regime based stochastic weather generator for process-based bottom-up climate risk assessments in California – Part I: Model evaluation

Nasser Najibi ^{a,*}, Alejandro J. Perez ^b, Wyatt Arnold ^b, Andrew Schwarz ^b, Romain Maendly ^b, Scott Steinschneider ^a

^a Department of Biological and Environmental Engineering, Cornell University, 111 Wing Drive, Riley-Robb Hall, Ithaca, NY 14853, USA

^b California Department of Water Resources, 715 P Street, Sacramento, CA 95814, USA

HIGHLIGHTS

- A novel weather regime-based stochastic weather generator simulates daily precipitation and temperature in California.
- The model reproduces a wide range of climate statistics and extremes with high fidelity at various spatiotemporal scales.
- The model's strong performance statewide supports climate impact assessments on water systems across California's watersheds.

ARTICLE INFO

Keywords:

Stochastic weather generator
Weather regimes
Climate change
Bottom-up
Downscaling
Water resources
California

ABSTRACT

This study is the first of a two-part series presenting a novel weather regime-based stochastic weather generator to support bottom-up climate vulnerability assessments of water systems in California. In Part 1 of this series, we present the details of model development and validation. The model is based on the identification and simulation of weather regimes, or large-scale patterns of atmospheric flow, which are then used to condition the simulation of local, daily weather at a 6 km resolution across the state. We conduct a thorough validation of a baseline, 1000-year model simulation to evaluate its ability to accurately simulate daily precipitation and minimum and maximum temperature at various spatial scales (grid cell, river basin) and temporal scales (daily, event-based, monthly, annual, inter-annual to decadal). Results show that the model effectively reproduces a large suite of climate statistics at these scales across the entire state, including moments, spells, dry and wet extremes, and extreme hot and cold periods. Moreover, the model successfully maintains spatial correlations and inter-variable relationships, enabling the use of model simulations in hydrologic and water resources analyses that span multiple watersheds across California. The weather generator can simulate physically plausible extreme events (e.g., multi-day extreme precipitation and severe drought) that extend beyond the worst case conditions observed historically, independent of climate change. Thus, the baseline simulation can be used to understand the impacts of natural climate variability on both flood and drought risk in regional water systems. Scenarios of climate change are discussed in Part 2.

Practical Implications

Water resource planners in California must prepare for the increased stress of climate change. However, there is significant uncertainty about how the climate will evolve over the coming

decades. While most climate projections agree that California's future will be warmer and precipitation will intensify, the rate of these changes is less clear, as are more nuanced climate changes like shifts in average precipitation or changes to atmospheric circulation that can impact regional weather. Furthermore, water systems in California are very vulnerable to natural swings in climate unrelated to climate change, and the range of this natural variability must also be considered in future planning efforts.

* Corresponding author at: 111 Wing Dr, Riley-Robb Hall 325, Ithaca 14853, NY, USA.

E-mail address: nn289@cornell.edu (N. Najibi).

<https://doi.org/10.1016/j.cliser.2024.100489>

Received 26 August 2023; Received in revised form 3 May 2024; Accepted 13 May 2024

Available online 22 May 2024

2405-8807/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

This two-part series comprises a pair of articles documenting the development of a stochastic weather generator for California that can create large ensembles of climate traces to support water resources planning under future climate uncertainty. The weather generator is an efficient tool that can quickly create long (1000-year) traces of statewide weather, and allows water managers to flexibly develop and explore climate scenarios associated with different signals of climate change. These include ‘thermodynamic’ changes (i.e., more intense precipitation linked to an increase in the moisture-holding capacity of warmer air), as well as more nuanced signals related to shifts in atmospheric circulation (i.e., dynamic climate change).

In Part (1) of this series (this article), we present the development and validation of the model. The research in Part (1) demonstrates that baseline simulations from the weather generator (i.e., simulations without climate change) can reproduce historical climate statistics across California very well, based on a large suite of variables that include extreme precipitation events and droughts, average and extreme temperatures, wet and dry spells, and heat and cold waves, among others. Performance was evaluated at various spatial scales, including at the native resolution of the observations (6 km grid cell) and aggregated to HUC-8 and HUC-4 basin scales. Performance metrics were chosen to showcase a range of statistical features inherent in the simulations, with an emphasis on the reproduction of variability across timescales and the representation of extreme events that are important for water resources planning activities. Particular attention was given to evaluating whether the weather generator could simulate extreme precipitation and drought events that exceed those observed in the historical record, but that are also statistically plausible. The analysis in Part (1) also demonstrates that baseline weather generator simulations compare well against the historical simulations from the LOCA v.2 downscaled product, which is based on the most recent generation of global climate model (GCM) simulations from the CMIP6 experiment.

Using the weather generator, we create a publicly available dataset of 30 unique climate change scenarios, each consisting of 1000 years of simulated climate data (precipitation, maximum temperature, minimum temperature) at a 6 km resolution across the entire state of California. The 30 scenarios represent a range of plausible climate changes to temperature, average precipitation, and precipitation extremes. An additional two climate scenarios are also created that reflect changes to the frequency of large-scale weather patterns. The generation of these scenarios is detailed in Part (2) of this series.

The datasets created in this work highlight how even in the absence of climate change, water managers in California should plan for extreme precipitation events and droughts beyond the worst case from the historical record, because such events are quite plausible due to California’s natural climate variability alone. However, when such natural extremes are combined with the effects of anthropogenic climate change, extremes in California become severe and will likely require significant investment in water resources systems to sustain adequate water services across the state. The weather generator and associated datasets should be viewed as a complementary tool to more commonly available downscaled GCM projections. Unlike GCMs, the weather generator is not designed to create scenarios of future climate based on the physical laws of the Earth system and future greenhouse gas emission scenarios. Rather, the model provides a way to help translate various signals of climate change from GCMs into traces of weather that are tailored to support water resource planning efforts. In this way, the California weather generator is envisioned as a tool to help promote collaboration between climate scientists and water resource planners across the state.

Stakeholders across California can utilize the climate scenarios produced by the stochastic weather generator to infer the joint impact of both climate change and natural climate variability on their systems. We recommend that stakeholders use the data products developed in this work in stages to help develop robust

adaptations of water resources infrastructure under climate change. For example, initial adaptation strategies (e.g., new reservoir operational policies, new infrastructure for managed aquifer recharge) could be developed using the historical record to ensure these strategies are able to meet performance requirements under past and recent extreme events. Then, these strategies could be re-evaluated using: 1) a 1000-year baseline weather generator simulation; and 2) climate change scenarios applied to the 1000-year weather generator simulation. These evaluations would provide information on how robust a given adaptation strategy is to natural climate variability, climate change, and a combination of the two.

Such information could be used in different ways to alter the initial adaptation strategy. For instance, if a strategy appears vulnerable to the 1000-year baseline weather generator simulation (i.e., natural climate variability) or scenarios from the model reflective of likely future climate conditions (e.g., a low degree of future warming projected in most GCMs over the next few decades), this might suggest an immediate need for a more robust strategy. Alternatively, in cases where an adaptation strategy is only vulnerable to the most extreme climate change scenarios produced by the model, the current adaptation strategy may be deemed adequate if coupled with plans for continued climate monitoring and retrofits/adjustments that could be implemented later if needed.

1. Introduction

Climate change poses a major threat to the sustainability of water systems in California. Over the last two decades, California has experienced four periods of drought (2001–2004, 2007–2009, 2012–2016, 2020–2022), which when taken together rank as the driest 22-year period in at least 1,200 years (Williams et al., 2022). Each of these drought periods were ended by a string of atmospheric rivers (ARs) (Dettinger, 2013; Zechiel and Chiao, 2021), some of which led to record flooding, threatened major infrastructure projects (Henn et al., 2020), and most recently, even drove the re-emergence of the once-dry Tulare Lake. These extremes are only projected to worsen over the next several decades, highlighting the need for water resource planners across the state to begin preparing their systems for additional climate stress.

Conventionally, water resource planners have used downscaled climate scenarios from global climate models (GCMs) to help plan for necessary system adaptations to mitigate the impending impacts of climate change. GCMs are an invaluable tool for providing internally consistent scenarios that can be used to examine the possible pathways of climate change under anthropogenic forcing. However, the scenarios produced by GCMs contain a mixture of climate change signals, including thermodynamic signals (e.g., the increased moisture holding capacity of a warmer atmosphere) and dynamic signals (e.g., shifts in atmospheric circulation), with the former projected to occur with more confidence than the latter (Emori and Brown, 2005; Seager et al., 2010, 2014). GCM scenarios also exhibit significant variability and biases in hydrologically important variables like precipitation linked to parameterized physics and coarse model resolution. This poses two challenges for water resources planning. First, it is often difficult and time consuming to separate out thermodynamic and dynamic signals of change from natural variability in GCM simulations. However, such separation may be of high interest to water resource planners who want to base their planning efforts on more detectable signals of change linked to thermodynamic mechanisms, but who may be wary of using climate data reflecting dynamic change in which there is less scientific consensus. Second, it is also very challenging to effectively remove biases from GCM simulations. Statistical corrections to certain types of model bias are far from straightforward, since they can be linked to modeled physical processes that could change under global warming and thus change the bias over time (Stephenson et al., 2012; Maraun

et al., 2017).

Given these challenges, water resource planners can benefit from an efficient alternative to climate scenario generation that can complement downscaled GCM simulations and help investigate water system risk under climate stress. *Stochastic weather generators* provide one such alternative. Weather generators are statistical models that are parameterized based on existing meteorological records and used to generate large ensembles of simulated daily or hourly weather records that are statistically similar to the observations but differ in the sequencing, duration, and magnitude of individual events (Richardson, 1981; Wilks and Wilby, 1999; Fowler et al., 2007; Bird et al., 2023). For water system applications, weather generators must often develop sequences of multiple weather variables (e.g., precipitation, maximum and minimum temperature) at multiple locations while maintaining realistic persistence and covariance structures associated with transient, multi-day storm events and over longer (seasonal to interannual) timescales. Once fit to historical data, model parameters can be systematically altered to produce new traces of weather that exhibit a wide range of change in their distributional characteristics that may be experienced under climate change, including the intensity and frequency of average and extreme precipitation, heatwaves, and cold spells (Wilks, 2002, 2010, 2012; Acharya et al., 2017; Mukundan et al., 2019).

Over the past few decades, there has been growing interest in using stochastic weather generators to create large scenario ensembles for bottom-up (or vulnerability based) climate change impact assessments of water and agricultural systems (Semenov and Barrow, 1997; Steinschneider and Brown, 2013; Harris et al., 2014; Ailliot et al., 2015; Yang et al., 2020; Khazaei et al., 2020; Dyreson et al., 2022; LeNoir et al., 2023). The main purpose of this study is to develop a stochastic weather generator and multiple ensembles of future climate scenarios across California to help evaluate the vulnerability of water systems and the robustness of adaptation strategies under climate change across the state. The stochastic weather generator will facilitate the integration of climate change considerations into statewide water policies and long-term planning, ensuring resilience and preparedness in the face of a changing climate. The effort presented in this work builds on a prototype tool (Steinschneider et al., 2019) that was used in a proof-of-concept application for the Tuolumne River basin, which was further refined and expanded to additional basins in California (Najibi et al., 2021; Rahat et al., 2022). This work advances and broadens the scope of the model developed in those previous studies to generate internally consistent climate scenarios across the entire state of California.

The present study (Part 1 of the series) focuses on the development, calibration, and validation of the stochastic weather generator across California. The subsequent study (Part 2) presents the development of publicly available climate change scenarios, which include both thermodynamic and dynamic signals of change and are designed to support bottom-up climate change impact assessments of water systems across the state. In the sections below, we describe the data used in the development of the California stochastic weather generator (Section 2), detail the model itself (Section 3), and present results summarizing baseline model performance (Section 4), before concluding with a brief discussion of key results (Section 5).

2. Meteorological data and atmospheric information

2.1. Precipitation and temperature records in California

We collected observed daily precipitation (P) [mm] between January 1, 1915 and December 31, 2018 (104 years) from the extreme-preserving gridded daily dataset for the conterminous United States developed by Pierce et al. (2021) (also known as the unsplit Livneh precipitation dataset). This extreme-preserving dataset, which has a $0.0625^\circ \times 0.0625^\circ$ (~ 6 km) spatial resolution, follows the same gridding method as employed in Livneh et al. (2013) and Livneh et al. (2015), but omits a time adjustment in the data to avoid excessive smoothing of

extreme events.

We further scrutinized the precipitation time series and noticed that there were 24 dates with extremely high precipitation intensities (>10 in or 254 mm) in the summer (June–August, with a large majority in July), while the closest GHCN-d rain gauges showed much lower magnitudes (often $\sim 1/10$ th of the gridded value). These extreme summertime precipitation intensities were deemed erroneous and were rescaled downward by a factor of 10 to reflect the precipitation intensity of nearby GHCN-d gauges.

We obtained observed daily minimum temperature (Tmin) [$^\circ\text{C}$] and maximum temperature (Tmax) [$^\circ\text{C}$] from the data in Livneh et al. (2013) for the period between January 1, 1915, and December 31, 2015. This dataset was then extended to December 31, 2018 using the PRISM daily dataset (PRISM Climate Group, 2014) to match the timeframe of the precipitation data. Two additional post-processing steps were then employed to prepare the temperature time series before using them in the stochastic weather generator. First, the temperature data in the 1915–2015 timeframe (Livneh et al., 2013) were bias corrected to the monthly PRISM dataset over the entire period. Then, the entire temperature time series (1915–2018) for each grid cell was detrended so that the long-term mean monthly temperatures between 1915–2018 matched those from 1991 to 2020. This ensured that the entire temperature series reflected warming that has already occurred in recent decades due to climate change. This is important for water resources planning activities that are forward-looking and need to accommodate future climate conditions, rather than past conditions that are unlikely to return.

All of these data were collected over and slightly outside of the USGS HUC-2 water resources region that spans the entire state of California and some parts of Nevada and Oregon (Region 18). This led to a final dataset consisting of 13,786 grid cells (Fig. 1). The final time series of both precipitation and temperature were truncated to the period between 1948–2018 to match the timespan of the atmospheric data used for weather regime classification (Section 2.2).

2.2. Atmospheric circulation over the Pacific-North American sector

We obtained daily gridded ($2.5^\circ \times 2.5^\circ$) geopotential heights (GPH) [m] at the 500-hPa level from the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis dataset (NCEP/NCAR Reanalysis 1; Kalnay et al., 1996) between January 1, 1948, and December 31, 2021 (74 years). The gridded data were then extracted for the region between 30°N – 60°N and 180°W – 100°W , covering much of the Pacific-North American sector (see inset in Fig. 1). The GPH data were centered by month to remove their seasonal cycle, producing GPH anomalies (GPHAs).

We also collected the daily occurrences of AR landfalls along the US west coast from the Scripps Institution of Oceanography (SIO)-generated AR catalog (SIO-R1 catalog; Gershunov et al., 2017). This catalog reports individual AR events across western North America from January 1, 1948 to December 31, 2020, detected using integrated vapor transport (IVT) and integrated water vapor (IWV) from NCEP/NCAR Reanalysis 1.

2.3. Annual standardized precipitation index over California

We collected a gridded dataset of a cold-season standardized precipitation index (SPI) ($0.5^\circ \times 0.5^\circ$) across California, taken from the study in Borkotoky et al. (2021). Similar to the work presented by Gupta et al. (2022), these data will be used when fitting the weather generator to ensure it can characterize large-scale weather patterns with appropriate inter-annual variance. We gathered the SPI data between 1948–2021, which is the same period as the GPHA data described above.

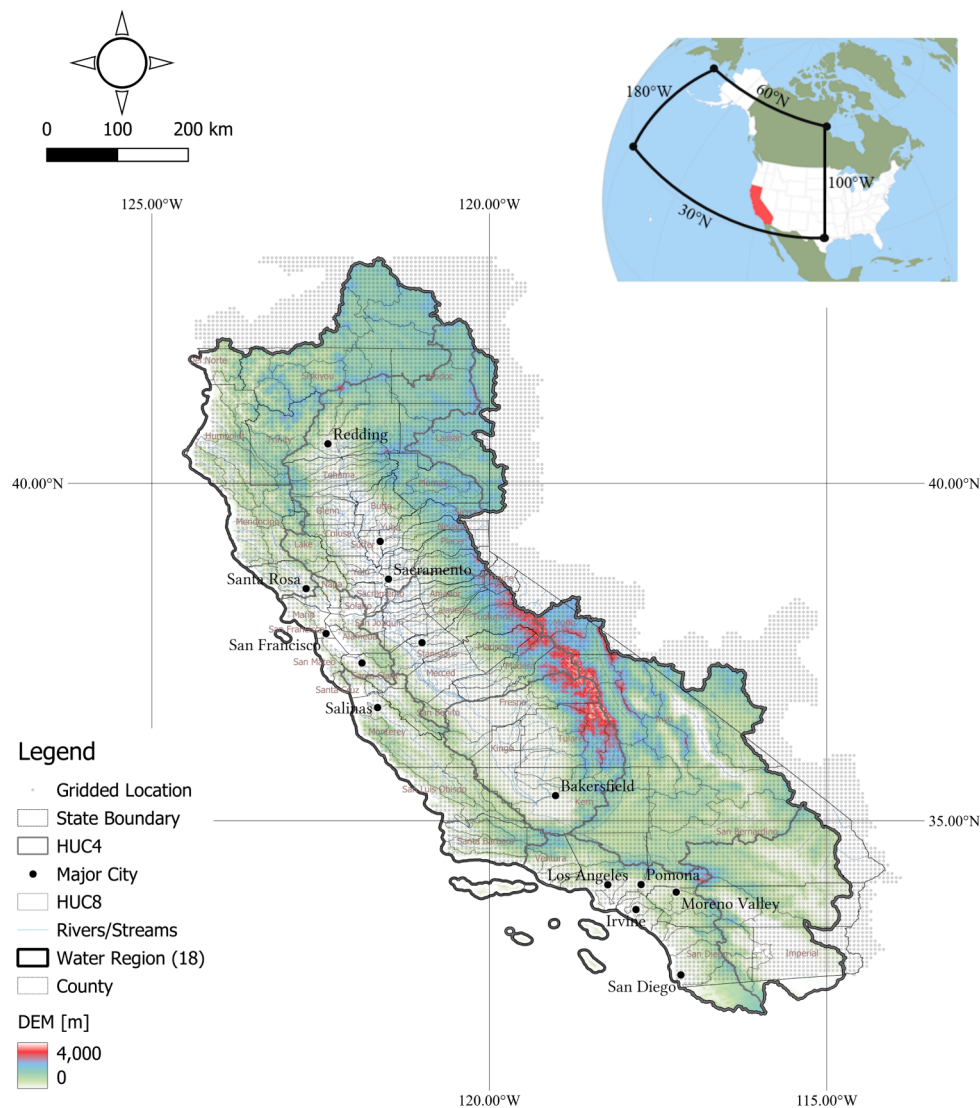


Fig. 1. Domain of the HUC-2 water resources region over California (Region 18), along with HUC-4 and HUC-8 subregions. All climate grid cells modeled within the stochastic weather generator are also shown, and extend slightly beyond the HUC-2 water resources region. Inset shows the atmospheric boundary used for weather regime (WR) identification.

2.4. Downscaled CMIP6 projections (LOCA v.2)

We obtained downscaled precipitation time series at a 3 km resolution from the localized constructed analogs (LOCA) v.2 statistical method (Pierce et al., 2023), based on climate model output from the Coupled Model Intercomparison Project Phase 6 (CMIP6). We collected daily precipitation for the historical period of 1950–2014 from 67 individual CMIP6 climate model simulations. There are 13 separate models in this dataset, each contains some number of initial conditions that varies by model, with a total of 4355 years of data in the historical period across all 67 simulations. These data are used to evaluate the benefit of the weather generator in terms of producing plausible extreme events compared with the baseline simulations provided by downscaled CMIP6 climate models.

3. Weather-regime based stochastic weather generator for California

This work advances a semiparametric, multivariate, and multisite stochastic weather generator that was previously developed for the Tuolumne River basin during the cold season (Steinschneider et al.,

2019) and further refined across several other basins in California (Najibi et al., 2021; Rahat et al., 2022). The weather generator is designed to separately model dynamic and thermodynamic atmospheric mechanisms of climate variability and change through statistical abstractions of these processes. To capture atmospheric dynamics, the weather generator simulates sequences of weather regimes (WRs). WRs are recurring large-scale atmospheric flow patterns (e.g., upper-level, quasi-stationary blocks and troughs) that appear at fixed geographic locations, persist for days-to-weeks within a season, and organize high-frequency weather systems (Robertson and Ghil, 1999; Robertson et al., 2015). They represent intermediary phenomena in the stochastic continuum of atmospheric perturbations that connect local weather to large-scale atmospheric circulation, and provide a parsimonious way of abstracting major patterns of atmospheric circulation into stochastic simulations of weather. To capture thermodynamic mechanisms of climate change, the weather generator post-processes simulated data to reflect patterns of warming and thermodynamic scaling of precipitation rates with that warming (more details in Part 2). These properties of the model are represented in a hierarchical structure composed of three primary modules: 1) identification and simulation of WRs that dictate the large-scale atmospheric flow across the Pacific-North American

sector; 2) simulation of local weather across California conditioned on the WRs; and 3) perturbations to the simulation schemes in (1) and (2) reflective of thermodynamic and dynamic climate change. In the sub-sections below, we focus on the first two of these modules, with the third being the focus of Part (2) in this series.

As part of this work, the stochastic weather generator underwent several refinements to improve the representation of droughts and pluvials, extreme multi-site precipitation events, and other climate statistics. The two largest refinements included changes to how the model simulates new WRs and how it adds noise to resampled large precipitation events. Therefore, we present an overview of the new set of algorithms in the sub-sections below (also see Fig. 2), and then present the mathematical details of these algorithms in Appendix A. Comparisons between the refined and previous model versions are presented in Supplementary Material.

3.1. Weather regime identification and simulation

Following Najibi et al. (2021), we use a Nonhomogeneous Hidden Markov Model (NHMM) to identify WRs. NHMMs are nonlinear statistical models that use latent variables to identify clusters in state-space while simultaneously accounting for the distribution and temporal dynamics of observed data (Rabiner, 1989; Hughes and Guttorp, 1994). In this application, we first divide the 500-hPa GPHAs into two seasons: November–April (cold season) and May–October (warm season). We then project the separated GPHAs onto their first J empirical orthogonal functions (EOFs), where J is chosen using a scree test to ensure that the selected EOFs explain the majority (e.g., $> 90\%$) of the variance in the data (here, $J = 10$). We subsequently evaluate a first-order NHMM on the J PCs of GPHAs to assign each day in the record to one of K separate WRs. This is done separately for the cold and warm season, and K can differ across the two seasons (the selection of K is described further in Section 3.2.2 below). The NHMM is fit using two harmonics as exogenous variables to account for seasonality in the WRs. In addition, the first four PCs of the SPI dataset over California (which explain $\sim 80\%$ of

the variance) are also incorporated as exogenous variables to improve the inference of inter-annual variability of cold-season WRs. Importantly, by using the NHMM, days are classified into WRs in a way that explicitly considers WR persistence, which will lead to persistent weather in the weather generator simulations. The Expectation-Maximization algorithm (Dempster et al., 1977) with the forward-backward algorithm (Baum and Petrie, 1966; Baum et al., 1970) is used to estimate the parameters of the NHMM. The most probable sequence of hidden states is computed using the Viterbi algorithm (Forney, 1973; Rabiner, 1989). More details related to the NHMMs are provided in Appendix A.1.

Future time series simulations of WRs of an arbitrary length could be created through forward simulation of the fitted NHMM, as in Najibi et al. (2021). However, we found that WR simulation using this approach, when coupled with the local weather generation algorithm described below in Section 3.2, underestimated the magnitude of extreme, multi-year droughts and pluvials (i.e., simulations were over-dispersed at inter-annual timescales). Therefore, we developed a novel non-parametric approach to WR simulation that addresses the issue of overdispersion, in which we first cluster the historically identified WRs into four-year segments across the historical record of 1948–2019. We then resample four-year segments with equal probability to develop future time series of WRs of an arbitrary length. This approach (described in more detail in Appendix A.2) ensures that inter-annual WR dynamics are almost completely preserved. The benefits of this approach for model performance as compared to an older model version (Najibi et al., 2021) are shown in Supplementary Material.

3.2. Local weather generation conditioned on weather regimes

3.2.1. Weather generation algorithm

Local weather is simulated by bootstrapping weather data (e.g., daily precipitation, minimum and maximum temperature) based on sequences of simulated WRs. Starting on simulation day t , the algorithm in Section 3.1 will determine a sequence of n days containing the i^{th} WR (i.

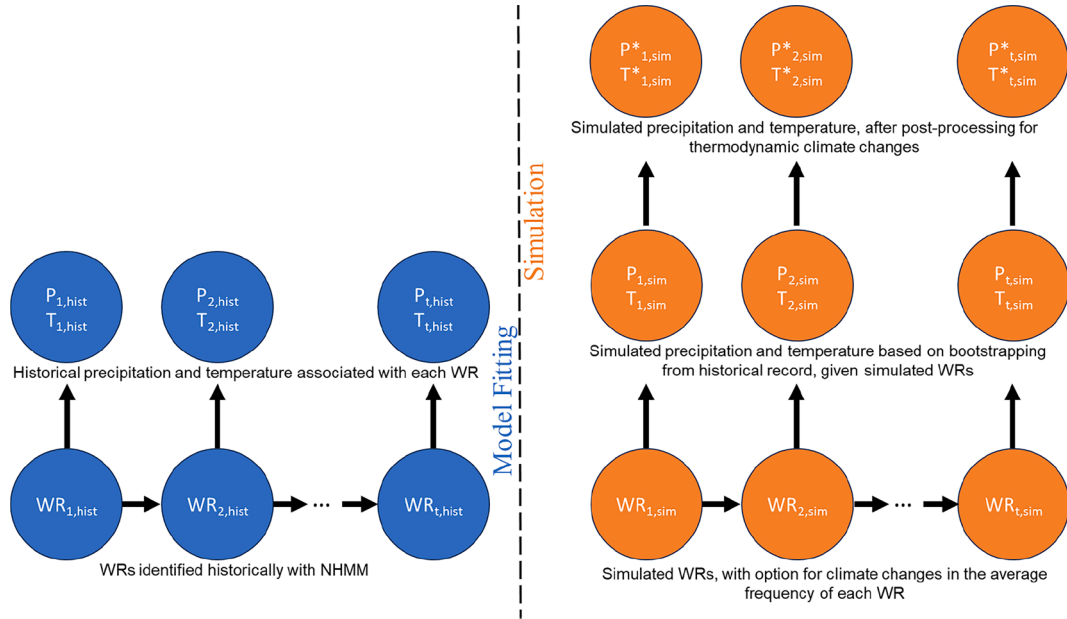


Fig. 2. Overview of the weather regime-based stochastic weather generator algorithm. During model fitting, daily weather regimes (WRs) are identified from historical atmospheric circulation data using a Nonhomogeneous Hidden Markov Model (NHMM). Historical daily precipitation (P_{hist}) and temperature (T_{hist}) data across the state of California are associated with historically identified WRs (WR_{hist}). During simulation, the model creates new sequences of WRs (WR_{sim}), with the option to change their frequency of occurrence as a signal of dynamic climate change. Simulations of daily precipitation and temperature (P_{sim} and T_{sim}) across the state are then generated by bootstrapping values from the historical record, based on the simulated and historical sequences of WRs. Finally, additional thermodynamic climate changes are imposed on the simulations of precipitation and temperature using post-processing methods, creating the final precipitation (P^*_{sim}) and temperature (T^*_{sim}) data.

e., WR_t through WR_{t+n-1} equal i). Here, n usually varies from a single day to a few weeks, although it can extend longer than one month due to the persistence of WRs. To generate weather for those n days, we resample an n^* -day block of historical data that was also classified into the i^{th} WR, based on the absolute difference between the historical and simulated block length (i.e., a historical block with length n^* closer to n will receive a higher probability and will be resampled with a higher likelihood). We also require that any resampled blocks meet two other criteria: 1) the central day of the historical block is within a 3-day window of the day of year for simulation day t ; and 2) the day prior to the historical block has the same state of regionally averaged precipitation (i.e., dry ($p_{avg} < 0.25$ mm) or wet ($p_{avg} > 0.25$ mm), where 0.25 mm is trace precipitation). This ensures that the resampled data will preserve the seasonality of local weather and better maintains precipitation persistence across sites. We define regionally averaged precipitation using daily precipitation averaged over the Calaveras, Stanislaus, Tuolumne, Merced, and Upper San Joaquin watersheds (425 grids in total), which are centrally located in California. This approach helps ensure that inter-daily precipitation dynamics are preserved in the center of the state, which also helps to preserve these dynamics in the northern and southern ends of the state due to regional coherence in storm tracks that pass over California.

If a historical block happens to be resampled with a longer length than the simulated one (i.e., $n^* > n$), we reduce the length of the resampled block by discarding days from that block randomly from one of its two ends until $n^* = n$. If the length $n^* < n$, then the remaining length $n - n^*$ is used as the basis to resample another block for WR i , and this process is continued until data has been resampled for the entire block of n days. At this point, the WR will change states and the resampling procedure begins again. By using this block bootstrap procedure, the resampled data are more likely to capture the entire life cycle of passing storms (and the resulting space–time structure in weather).

The block bootstrap will preserve many of the properties of the marginal and joint distributions of local weather variables, but at the expense of being able to simulate values outside the range of the instrumental record. To address this drawback specifically for heavy precipitation, the weather generator uses a copula-based jittering approach that adds noise to resampled heavy precipitation data as a post-processing step. To do this, we first fit a mixture model to the observed, non-zero precipitation at each site, which uses a gamma distribution for the bulk distribution (by month) up to a threshold and a Generalized Pareto distribution (GPD) to model the tail of the distribution beyond that threshold. For this work, we selected the site-specific 99th percentile of non-zero values as the separating threshold. Then, for heavy precipitation values above the 99th percentile that are resampled in the block bootstrap, we calculate the non-exceedance probability associated with that event based on the fitted gamma-GPD. Spatially correlated random noise is then added to these non-exceedance probabilities across sites, which are then mapped back through the gamma-GPD models to develop new heavy precipitation values across locations. This approach is designed such that final values of simulated, jittered precipitation can extend beyond the range of historical precipitation values, but preserve the marginal distribution of precipitation at each site and the rank correlation structure across sites. More details on this approach are provided in [Appendix A.3](#). The benefits of this approach for model performance as compared to an older model version ([Steinschneider et al., 2019](#)) are shown in [Supplementary Material](#).

3.2.2. Selection of the number of weather regimes

The identification and simulation of the WRs ([Section 3.1](#)) plays a critical role in the performance of the weather generator. A key parameter of the model is K , the number of WRs that should be used in each of the two seasons. To calibrate the value of K , we followed the approach in [Najibi et al. \(2021\)](#) and generated a very long trace (time series) from the weather generator for the entire state and for values of K

ranging from 2 to 10 WRs separately for the cold season and warm season. This simulation is 1008 years long, although for simplicity we hereafter refer to it as the 1000-year simulation (the use of 1008 years is related to a nuance of the simulation strategy described in [Appendix A.2](#)). We then evaluated the distribution of a variety of climate performance measures for a random selection of 100 grid cells across the state of California for different values of K . We quantified performance based on percent bias between simulated and observed statistics of interest, which are listed in [Table 1](#) (for precipitation) and [Table 2](#) (for temperature). These metrics were carefully chosen to showcase a range of statistical features inherent in the simulations, providing valuable insights into their reliability and accuracy for statewide users. We placed an emphasis on the reproduction of variability across timescales (e.g., the distribution of cumulative precipitation across months; standard deviation of annual precipitation totals; distribution of cumulative precipitation across multi-year windows), as well as the representation of extreme events (e.g., extreme 1-day and 5-day annual precipitation maxima; the worst period-of-record drought magnitude for different durations). These features are particularly important to water resources planning activities (e.g., flood risk assessment, drought management, climate risk adaptation, and climate resilience). We select K for each season that results in the best weather generator performance across statistics. No one value of K (i.e., number of WRs) for either season is likely to maximize performance across all these measures for the entire state of California. Therefore, we selected K based on the value that provided the best balance across all statistics over all sampled grid cells.

4. Results

4.1. Weather regime identification

As discussed in [Section 3.2.2](#), we select the number of WRs (i.e., K) for both the cold season and the warm season by evaluating the weather generator's ability to reproduce a variety of weather statistics (see [Tables 1 and 2](#)) for values of K ranging from 2 to 10 WRs, separately in each season. We do not show the model's performance on all statistics above for all combinations of K here, but rather focus on the final selected WRs and their interpretation. In [Section 4.2](#) below, we present the performance of the final selected model across all metrics in [Tables 1 and 2](#).

The calibration procedure resulted in the identification of 7 and 3 WRs in the cold and warm seasons, respectively (10 WRs in total). [Fig. 3](#) presents the composites of 500-hPa GPHAs for days categorized under

Table 1
Statistics of precipitation used for model evaluation.

No	Statistic	Description [unit]
1	Daily mean	Average of daily precipitation [mm]
2	Daily standard deviation	Standard deviation of daily precipitation [mm]
3	Inter-annual standard deviation	Standard deviation of water-year (October 01 – September 30) total precipitation [mm]
4	Seasonal variation	Monthly totals and the distribution of cumulative water-year precipitation across months (mean, median, 10th, and 90th percentiles)
5	Annual maxima of 1-day events	Annual maxima of daily precipitation in each water-year [mm]
6	Daily peaks over threshold	Distribution of daily precipitation greater than or equal to the 99th percentile of site-specific non-zero precipitation [mm]
7	Annual maxima of 5-day events	Annual maxima of 5-day rolling average precipitation in each water-year [mm]
8	Mean and maximum wet spell length	Average and maximum length of a wet spell (consecutive non-zero precipitation days) [days]
9	Inter-annual precipitation totals	Distribution of 1, 2, 3, 5, 10, and 30-year rolling average water-year precipitation totals [mm]
10	Worst multi-year drought intensity	Minimum 1, 2, ..., 10-year water-year precipitation totals across the entire record [mm]
11	Mean and maximum dry spell length	Average and maximum length of a dry spell (consecutive zero precipitation days) [days]

Table 2
Statistics of temperature used for model evaluation.

No	Statistic	Description [unit]
1	Daily mean	Average of daily temperatures [$^{\circ}\text{C}$]
2	Daily standard deviation	Standard deviation of daily temperatures [$^{\circ}\text{C}$]
3	Inter-annual standard deviation	Standard deviation of annual average temperatures [$^{\circ}\text{C}$]
4	Heat wave frequency	Number of instances with three or more consecutive days with temperature over 32.2°C (90°F) [number of events]
5	Heat stress frequency	Number of instances with 3-day rolling mean temperature above 30°C (86°F) [number of events]
6	Mean and maximum of heat wave duration	Average duration and longest duration of heat waves in the record [days]
7	Mean and maximum of heat wave intensity	Average intensity and highest intensity of heat waves in the record [$^{\circ}\text{C}$]
8	Cold wave frequency	Number of instances with five or more consecutive days with temperature below -7°C (20°F) [number of events]
9	Cold stress frequency	Number of instances of 3-day rolling mean temperature below 0°C (32°F) [number of events]
10	Mean and maximum of cold wave duration	Average duration and longest duration of cold waves in the record [days]
11	Mean and maximum of cold wave intensity	Average intensity and highest intensity of cold waves in the record [$^{\circ}\text{C}$]

each WR, along with their average temporal frequency over a calendar year. WRs 1 to 7 occur in the cold season, while WRs 8 to 10 occur in the warm season. WR1 exhibits a widespread trough centered over the eastern Pacific and off the western US coastline, which is located south of a high-pressure anomaly anchored over the Bering Sea. Under WR2, there is a ridge and trough directly over the northwest US and Aleutian Islands, respectively. This pattern is reversed in WR4 and WR7, which differ only in the longitudinal location of the pressure dipole. WR3 exhibits a ridge directly over the northwest US that is slightly more elongated and less intense compared to the ridge in WR2. This ridge is shifted further to the west over the eastern Pacific in WR5. The most notable feature of WR6 is an intense low over the Gulf of Alaska. Finally, the WRs in the summer (WRs 8–10) all exhibit weaker GPHAs, but resemble some of the same spatial patterns as seen in the cold season.

An analysis of California-wide precipitation and temperature anomalies under the WRs in Fig. 3 showed climate varies considerably across the state depending on the prevailing WR. For example, conditions across the Central Valley are wettest in the cold season under WR1, followed by WRs 4, 7, and 6. This is consistent with the deep troughs under each of these WRs that direct storm tracks and moisture over the state. In contrast, conditions across the Central Valley are driest in the cold season under WRs 2, 3, and 5 (in that order), which aligns with the ridging under these WRs that blocks moisture flow over California. Precipitation reaches its absolute lowest under WRs 8–10, because these WRs occur in the warm season when a semi-permanent ridge of high pressure expands further north and pushes storm tracks north of the state (this is not seen in Fig. 3, because seasonality was removed from the GPHAs).

4.2. Validation of simulated weather

Following WR identification, the weather generator was then validated based on the reproduction of weather statistics across the entire state. For the purposes of validation, the precipitation and temperature metrics (Tables 1 and 2) were evaluated at the scale of HUC-4 basins (Fig. 4; Table 3), with some metrics shown based on an average across grid cells within the HUC-4 and other metrics shown for individual grid cells within the HUC-4. Note that some areas at the boundary of the state are located in HUC-4 regions that largely sit outside of California, and for those locations we evaluated the weather generator in HUC-8 basins that are mostly within California (also shown in Fig. 4). For illustration,

Figs. 5–13 below show the validation results for one HUC-4 subregion (HUC-4: 1804 – San Joaquin). All results in these figures are based on a 1000-year baseline simulation of the stochastic weather generator with no climate change. All precipitation and temperature validation statistics for the rest of the HUC-4 subregions in California are available in the [Supplementary Material](#). Results presented for the San Joaquin HUC-4 are very representative of the model's performance for the other HUC-4 regions. Finally, we also present a brief validation of the reproduction of AR landfall frequencies near and along the coast of California.

4.2.1. Precipitation validation

Fig. 5 shows metrics that quantify the characteristics of daily, monthly, and water-year precipitation totals, including: the full distribution of HUC-4 basin-scale water-year precipitation totals (Fig. 5a); inter-annual standard deviation of basin-scale water-year precipitation totals (Fig. 5b); site-specific mean and standard deviation of the daily precipitation time series at each grid cell within the HUC-4 (Fig. 5c,d); basin-scale water-year cumulative precipitation totals (Fig. 5e); and the distribution of monthly precipitation totals (Fig. 5f). The results show that the mean and standard deviation of daily precipitation is extremely well preserved (Fig. 5c,d), and the full distribution and variability of water-year precipitation totals falls well within the range of uncertainty for the observed statistics (Fig. 5a,b). We note though that the inter-annual variance of simulated water-year precipitation falls slightly below that of the observed inter-annual variance, and this pattern is seen for several other basins across the state (see [Supplementary Material](#)). The mean, 10th, and 90th percentiles of basin-scale water-year cumulative precipitation totals also follows the observations very closely (Fig. 5e), with only a slight overestimation of the median cumulative precipitation that emerges during the spring. The median and range of monthly precipitation totals also matches the observations very closely (Fig. 5f). Overall, these results show that the weather generator produces daily, monthly, and water-year precipitation totals that agree very well with observations, both at the HUC-4 scale and at the scale of individual grid cells.

Fig. 6 presents statistics for extreme precipitation attributes. Fig. 6a shows a return period plot for 1-day extreme precipitation at the HUC-4 scale from the observations, based on a GEV distribution fit to annual maxima (red) and a GPD-Poisson model fit to a partial duration series (blue). The black points show annual maxima from the 1000-year stochastic weather generator simulation, sorted and plotted against empirical return periods. The return levels estimated by the weather generator simulation match that of the GEV and GPD-Poisson models fit to the observations very well. Importantly, the weather generator is able to perturb the annual maxima higher than the largest observed value using the copula-based jittering algorithm ([Appendix A.3](#)), and does so in a way that the simulated annual maxima follow the observation-based GEV and GPD-Poisson model estimates for larger return periods (e.g., the 500-year and 1000-year events). We note that this behavior occurs even though the jittering algorithm is conducted at the grid cell (and not HUC-4) scale, and on daily data and not the annual maxima. The reproduction of 500-year and 1000-year extreme precipitation events at the HUC-4 scale is encouraging, as this requires that correlation of extremes be preserved across locations within the basin while still being perturbed upward above the range of the historical data. A similar result is seen for annual maxima of 5-day precipitation events at the HUC-4 scale (Fig. 6b), except that the simulated 5-day annual maxima are somewhat biased below the GEV model fit to the observations for very high return periods. Still, the simulated values are well within the wide GEV-based uncertainty bounds at these high return periods, suggesting that the weather generator adequately preserves the correlation of extreme precipitation across locations within the basin over durations longer than 1 day.

Fig. 6c shows 1-day precipitation extremes that fall over a high threshold (the 99th percentile) in the observations, and also these same events that are bootstrapped in the 1000-year long weather generator

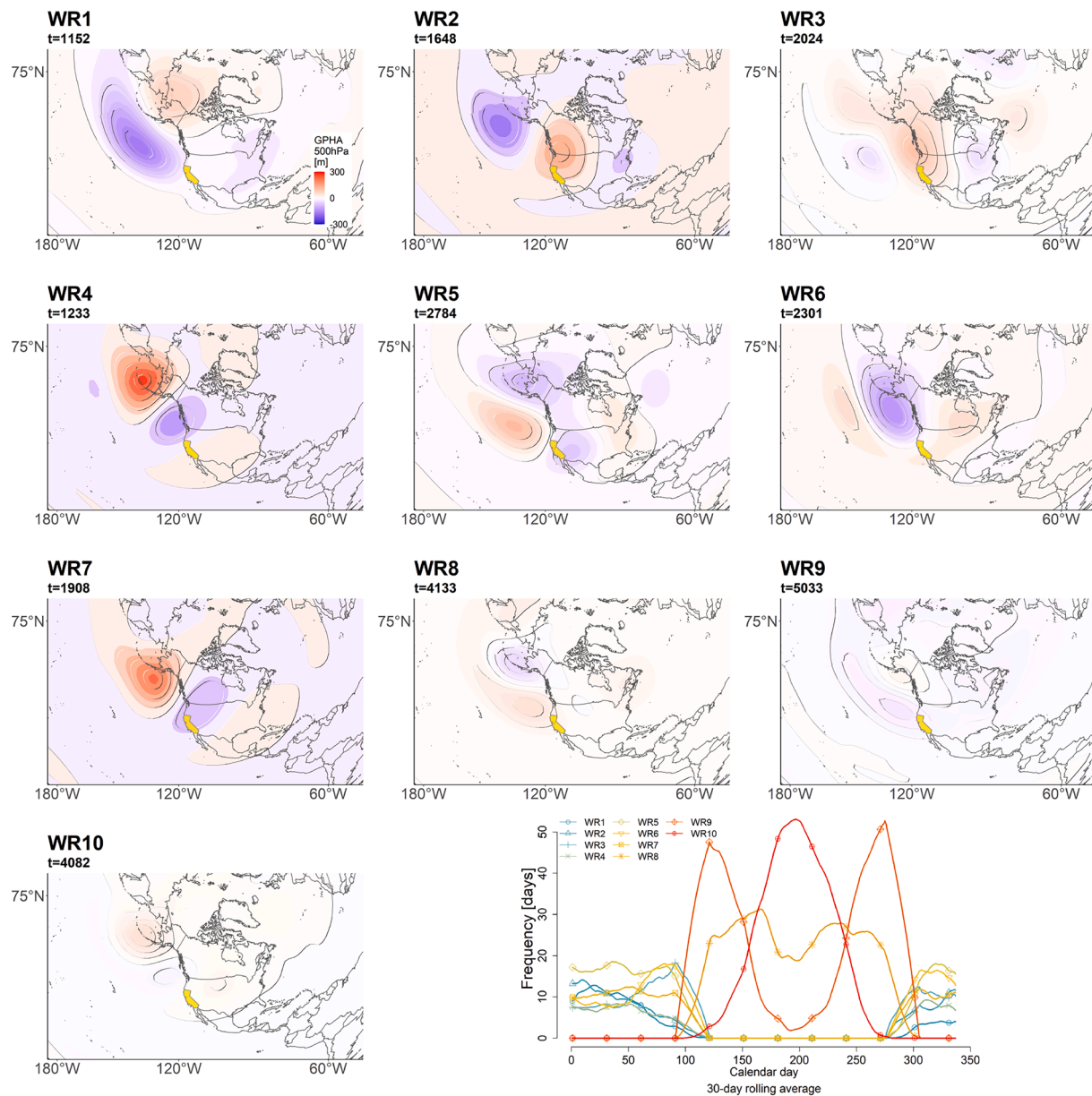


Fig. 3. Composites of 500-hPa GPHAs [m] for days categorized under each WR during the cold season (November–April; WRs 1 to 7) and warm season (May–October; WRs 8 to 10). The number of days (t) classified under each WR is shown above each composite. California is colored with yellow shading on the map. The temporal frequency of WRs per calendar day based on a 30-day rolling average smooth is also shown (1948–2021). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

simulation. Results here are shown for all grid cells within the HUC-4 region. For each event in the observations, the weather generator often resamples that event multiple times in the 1000-year simulation, which accounts for the scatter and horizontal stratified points in Fig. 6c. The results here show that the simulated extremes are centered around and very highly correlated with the historical extremes (Pearson r of 0.99, as expected with the bootstrap), but that the simulation adds noise around these extremes with the copula-based jittering algorithm that allows both smaller and larger values than those seen historically. For instance, the maximum 1-day precipitation in the simulation is 517.2 mm, while it is only 429.8 mm in the observations.

Finally, Fig. 6d shows the distribution of average wet spell lengths across gridded locations within the HUC-4 region. Across grid cells, average wet spell length matches that in the observations well, with only a slight (<1%) downward bias on average.

Fig. 7 focuses on metrics relevant to HUC-4 scale multi-year droughts

and pluvials. Fig. 7a shows the water-year total precipitation during the worst 1-year, 2-year, through 10-year drought in both the observations and the 1000-year simulation. For all durations, the simulation exhibits a more severe drought-of-record than the historical period, as should be expected given the longer period of record. We use independent probability models fit to observed annual precipitation totals to confirm (via order statistics) that the intensity of the most extreme droughts from the 1000-year weather generator simulation are statistically plausible (see [Supplementary Material](#)). We also confirm that the observed worst drought magnitudes for different durations fall well within the simulated distribution if the stochastic weather generator is used to simulate multiple traces of equal length to the historical record (also see [Supplementary Material](#)). The weather generator produces average dry spell lengths across gridded locations within the HUC-4 that align very well with the observations (Fig. 7b), but is also able to produce maximum dry spells that extend beyond the observations (i.e., maximum dry spell of

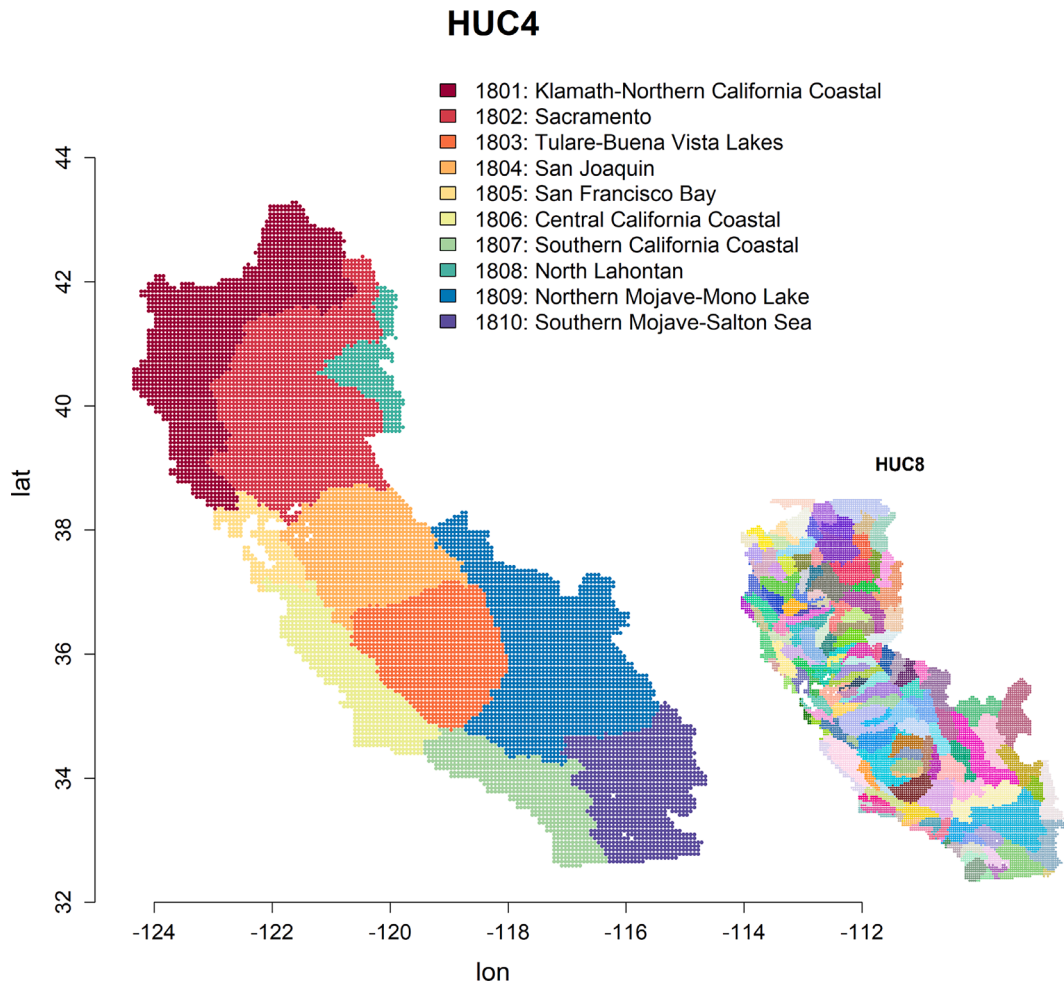


Fig. 4. HUC-4 domains (subregion level) across California, with HUC-8 domain (subbasin level) boundaries shown in the inset.

Table 3
Properties of HUC-4 basins across California.

Basin Number	Basin Name	HUC4	State	Area [km ²]	# Grid-Points
1	Klamath-Northern California Coastal	1801	CA, OR	67,762	1758
2	Sacramento	1802	CA, OR	72,013	1941
3	Tulare-Buena Vista Lakes	1803	CA	42,498	1084
4	San Joaquin	1804	CA	40,986	1072
5	San Francisco Bay	1805	CA	13,910	259
6	Central California Coastal	1806	CA	34,287	728
7	Southern California Coastal	1807	CA, MX	35,863	703
8	North Lahontan	1808	CA, NV	11,791	318
9	Northern Mojave-Mono Lake	1809	CA, NV	73,269	1874
10	Southern Mojave-Salton Sea	1810	CA, MX	44,245	1028

248 days in the simulation vs. 217 days in the observations). Fig. 7c shows the distribution of 1, 2, 3, 5, 10, and 30 water year rolling average precipitation totals, and demonstrates that the frequency of water-year precipitation totals over different multi-year durations is very well preserved in the simulation compared to the observations. However, the simulation is able to generate multi-water year precipitation totals that

extend beyond the observed range for all durations (i.e., more severe worst-case multi-year droughts and pluvials), showing how the model can be used to explore plausible extremes not yet experienced in the observations strictly due to internal climate variability.

We also compared the ability of the weather generator to produce plausible extreme events against the historical simulations derived from the LOCA v.2 dataset. We focus our investigation on basin-average precipitation over the Upper Tuolumne watershed, one of the watersheds in the San-Joaquin basin (HUC4: 1804), and retain a focus on the historical period in the CMIP6 data in order to compare how these model simulations represent extremes due to natural climate variability without climate change. Similar to Fig. 6a, Fig. 8a shows a return period plot of 1-day precipitation extremes at the watershed scale from the observations, based on a GEV distribution fit to annual maxima (red). The black points show annual maxima from the 1000-year stochastic weather generator simulation, sorted and plotted against empirical return periods. The blue triangles show a similar result for the LOCA v.2 downscaled data, where the annual maxima at the watershed scale are concatenated across the 67 model simulations (a total of 4355 years), sorted, and plotted against empirical return periods. The results show that while the 1000-year weather generator simulation of watershed-scale annual maxima follow the observation-based GEV model very well, the LOCA v.2 data are biased low for return periods greater than 10 years. This bias is especially apparent at the higher return periods.

Fig. 8b focuses instead on drought events, and shows the worst 1-, 2-, 3-, 4-, and 5-year drought events in the observed record, the concatenated 4355 historical years from LOCA v.2, and the 1000-year weather generator simulation. Interestingly, both the weather generator and the

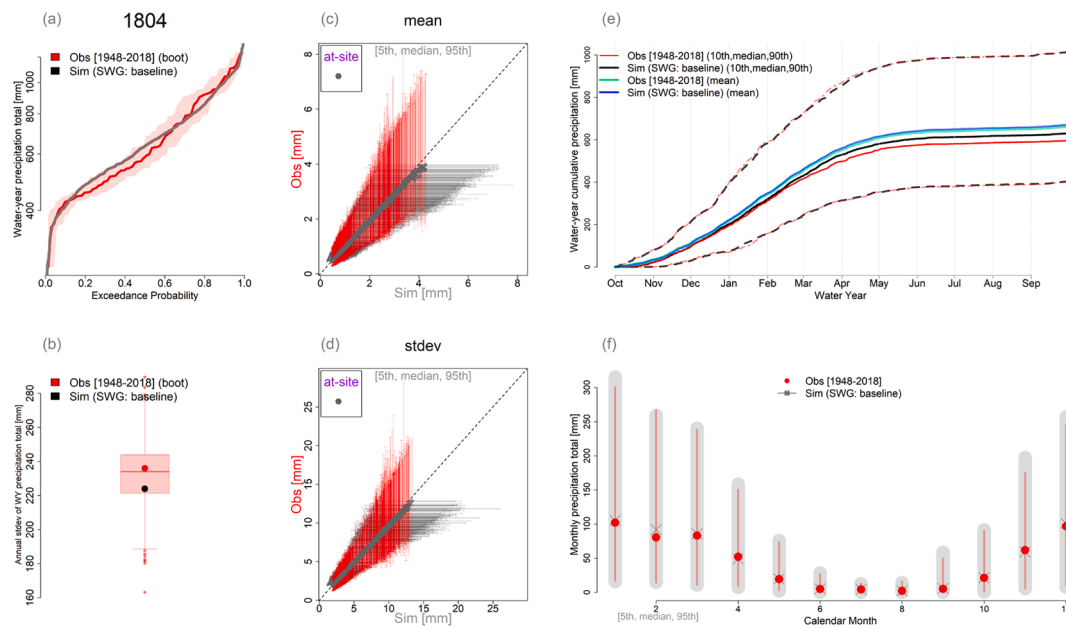


Fig. 5. Observed vs. simulated (a) distribution of basin-scale water-year precipitation totals, with the red shaded area representing the uncertainty in the observed distribution using a 95% confidence interval based on bootstrapping from the historical record; (b) standard deviation of basin-scale water-year precipitation totals, with the boxplot showing the uncertainty in the observed standard deviation based on bootstrapping from the historical record; (c) mean and (d) standard deviation of daily precipitation at individual sites (grid cells) within the HUC-4, with dots and whiskers showing the 50th, 5th, and 95th percentiles across individual water years; (e) basin-scale cumulative precipitation over the water year, including the mean, median, 10th, and 90th percentiles; (f) the distribution of basin-scale monthly total precipitation, including the 5th, 50th, and 95th percentiles. All results are shown for the San Joaquin basin (HUC4: 1804). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

downscaled climate model data produce worst-case droughts that 1) fall well below the worst-case drought in the observational record, and 2) are of very similar magnitude for all durations. This is particularly surprising because these products were developed independently, where the LOCA v.2 data are derived from physically based GCMs, and the weather generator is a statistical model fit to observations. However, it is worthwhile to note that although the most extreme droughts in both the weather generator simulation and the downscaled (historical) climate model data are of similar magnitudes, it took 4355 years of downscaled GCM data to produce these droughts but only 1000 years of the weather generator simulation. This leads to two important implications: 1) the extreme droughts simulated by the weather generator are physically plausible, because an ensemble of GCM simulations produced droughts of similar magnitude; and 2) those droughts may be much more likely to occur due to natural climate variability than suggested by the GCMs, since the weather generator was able to produce them in less than a quarter of the years required to produce them in the GCMs. We note that we repeated this analysis using only 1000 years from the LOCA v.2 dataset and found worst-case droughts to be of the same magnitude as the observations (not shown), further suggesting that the GCMs struggle to produce extreme droughts that are driven by natural climate variability.

We repeated a similar comparison of the LOCA v.2 historical data to the baseline weather generator simulation for additional watersheds (e. g., Upper Feather, Upper American, Lake Millerton), and found that in general the 1000-year baseline weather generator simulation was better able to capture basin-scale extreme precipitation frequencies and produced similar extreme droughts compared to the 4355 years of historical LOCA v.2, although with some variation across watersheds (see [Supplementary Material](#)).

4.2.2. Temperature validation

[Fig. 9](#) is similar to [Fig. 5](#) a-d, but for average temperature. It should be noted that we computed average temperature using simulated daily minimum and maximum temperature at each gridded location. Overall,

the distribution of HUC-4 scale average annual temperatures across water years is well preserved in the simulation compared to the observations ([Fig. 9a](#)), although with a small underestimation of the inter-annual variance ([Fig. 9b](#)). Median daily means and standard deviations across grid cells are very well preserved, as are their range across water years ([Fig. 9c,d](#)).

[Fig. 10](#) highlights characteristics related to temperature extremes across grid cells within the HUC-4 region, including heat waves and cold waves. The average and maximum duration of heat and cold waves are shown in [Fig. 10a,b](#), while the average and maximum intensity are shown in [Fig. 10c,d](#). The model performs well for all these statistics across sites. There is a minor upward bias in the mean duration and mean and maximum intensity of heat waves, but this bias is small. Overall, there is a high level of agreement between the observed and simulated attributes of heat waves and cold waves.

Finally, [Fig. 11](#) shows the spatial distribution of the frequency of heat waves and heat stress events derived from the observations ([Fig. 11a,c](#)) and the 1000-year simulation ([Fig. 11b,d](#)) across the HUC-4 region. A similar result is also shown for cold waves and cold stress events in the observations ([Fig. 11e,g](#)) and simulation ([Fig. 11f,h](#)). The distributions are normalized using the site-specific time-series length, so that results are shown as the average frequency of heat or cold events per year in each gridded location. As shown in [Fig. 11](#), the model is able to reproduce the spatial distribution of these events extremely well as compared to the observations, implying that the stochastic weather generator can fully preserve the spatial organization of multi-day temperature extremes across sites in the basin.

Finally, we note that the validation performance showcased in [Figs. 5-7](#) and [Figs. 9-11](#) for the San Joaquin basin (HUC4: 1804) is consistent across the rest of the HUC-4 subregions throughout California, with a small number of exceptions we highlight here (see [Supplementary Material](#)). There are a few HUC-4 subregions with no cold wave (HUC4: 1805, 1806, 1810) or heat wave (HUC4: 1801, 1808) events, which the stochastic weather generator also replicates in the baseline simulation. The weather generator underestimates interannual

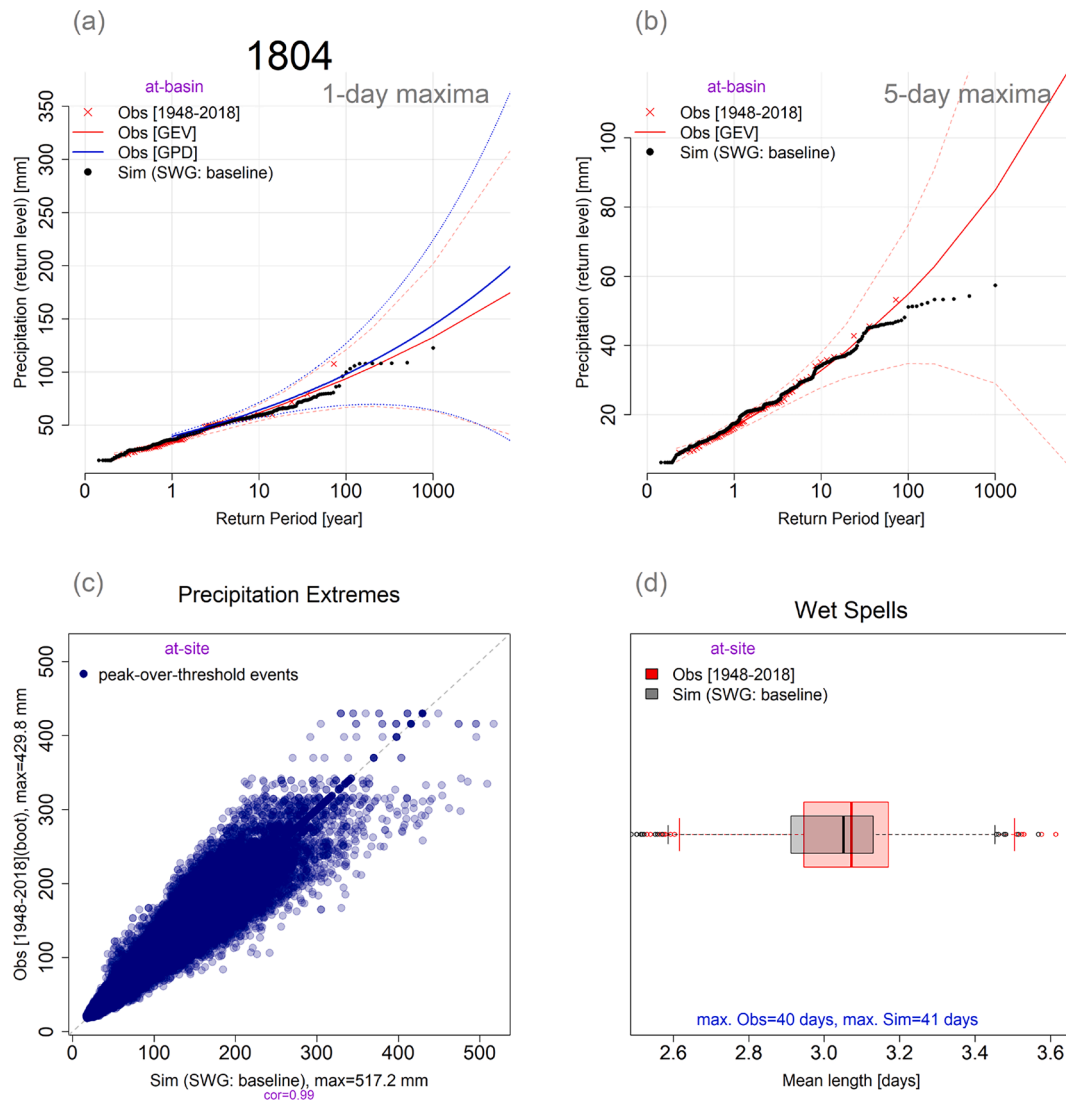


Fig. 6. (a) Observed and simulated empirical distribution of HUC-4 scale water-year 1-day precipitation maxima, along with GEV and GPD-Poisson based return level estimates and their 95% uncertainty bounds fitted to the observed precipitation; (b) similar to (a) but for 5-day precipitation maxima, and without a fitted GPD-Poisson model; (c) the magnitude of precipitation in partial duration series for all grid cells in the HUC-4 region, along with the same events in the weather generator simulation after the values are jittered; and (d) the distribution of average wet spell lengths (in days) across sites. The maximum wet spell length among grid cells for the observations and simulation is shown in blue text. All results are shown for the San Joaquin basin (HUC4: 1804). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

temperature and precipitation standard deviations by less than 5 % in several basins, although still within the sampling uncertainty of this statistic for the observations. Small biases (<10 %) in heat and cold wave intensity and duration are also observed in some grid cells within basins across the rest of the HUC-4 regions. Finally, for one HUC-4 region (HUC4:1809 – Northern Mojave-Mono Lake), the worst 1-year drought is underestimated, even though the maximum dry spells are longer than observed for the gridded locations in this subregion.

4.2.3. AR landfall frequency validation

As a final validation of the 1000-year baseline simulation of the stochastic weather generator with no climate change, we evaluate the frequency of AR landfalls near and along the California coastline (Fig. 12). The observed frequency, expressed as the number of events per year for each calendar month, was taken from the historical AR archive in Gershunov et al. (2017). We focus on five locations and observed that for the most northern location (labeled '1' in Fig. 12), AR landfalls are most frequent in the earlier part of the cold season (November–December) and become less frequent as the cold season progresses.

Conversely, the location furthest south (labeled '5' in Fig. 12) shows an increase frequency of ARs later in the cold season (January–March). This general pattern of more ARs later in the cold season is followed across locations moving north to south. By using the resampled historical dates from the block bootstrap in the stochastic weather generator, we can also derive the monthly frequencies of landfalling ARs in the 1000-year baseline simulation. We find that the weather generator near-perfectly reproduces the monthly frequencies of AR landfalls at all locations.

5. Conclusions and final remarks

This work presented the first of a two-part series of studies introducing the development of a stochastic weather generator suitable for the creation of gridded (~6 km) climate change scenarios across the entire state of California. The stochastic weather generator is novel in comparison with other downscaling techniques, because it is designed to distinguish thermodynamic and dynamic mechanisms of climate change, allowing analysts to separately consider those mechanisms they deem most credible for planning purposes. In this work, the weather

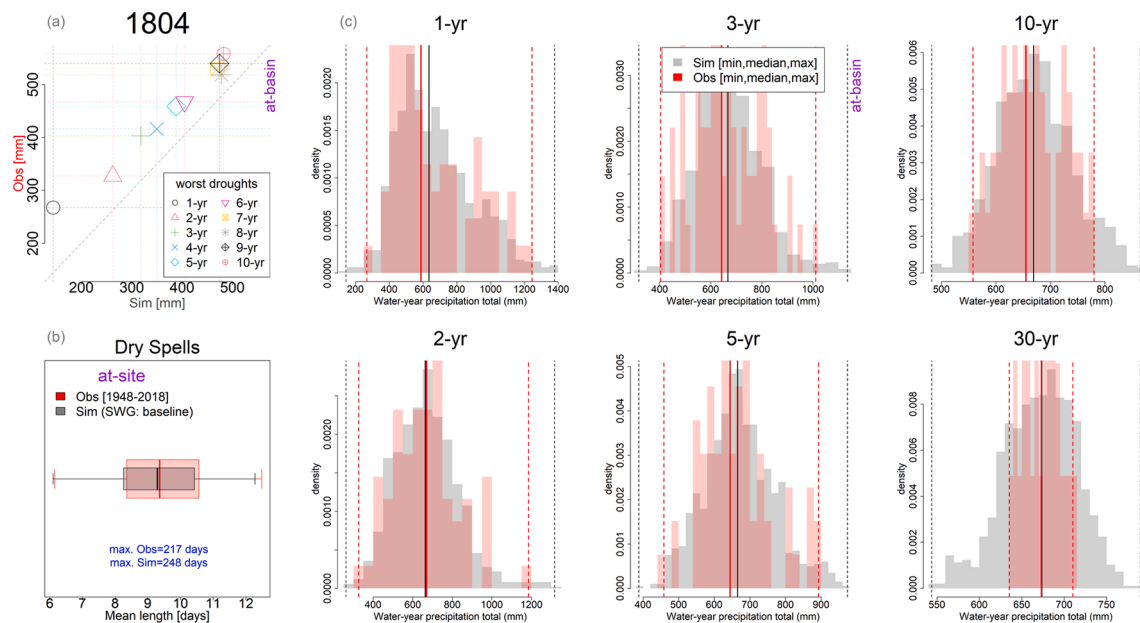


Fig. 7. (a) Observed vs. simulated worst multi-year basin-scale droughts (1 to 10 years) based on water-year precipitation totals; (b) the distribution of average dry spell lengths (in days) across sites. The maximum dry spell length among grid cells for the observations and simulation is shown in blue text. (c) The distribution of basin-scale water-year precipitation totals at 1, 2, 3, 5, 10, and 30-year rolling averages, including the minimum, median, and maximum totals. All results are shown for the San Joaquin basin (HUC4: 1804). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

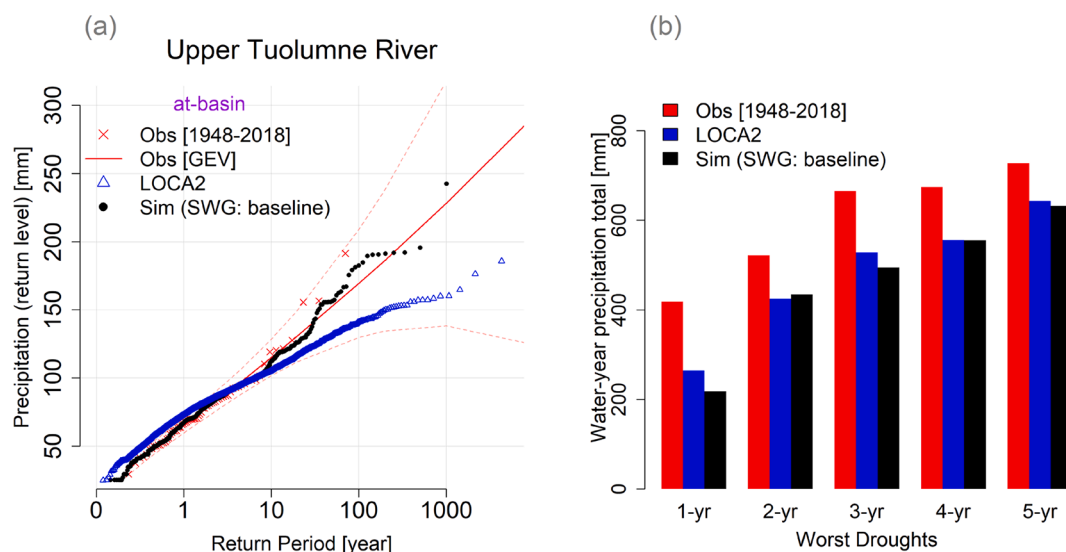


Fig. 8. (a) Observed distribution of watershed-scale annual precipitation maxima, along with GEV-based return level estimates fitted to the observed annual maxima (red). Also shown are annual maxima from the 1000-year weather generator simulation (black) and 4355 years of historical LOCA v.2 data (blue). (b) The worst 1, 2, 3, 4, and 5-year drought events in the observed record, across the historical LOCA v.2 ensemble, and in the 1000-year weather generator simulation. All results are shown for the Upper Tuolumne River watershed within the San Joaquin basin (HUC4: 1804). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

generator was calibrated across the state of California and then used to create a 1000-year baseline weather simulation, which was subsequently validated against observations.

Results showed that the statewide stochastic weather generator is able to simulate long sequences of daily precipitation and minimum and maximum temperature that very accurately mimic the behavior of historical observations at multiple spatial scales (grid cell, basin) and temporal scales (daily, event-based, monthly, annual, inter-annual to decadal). The model reproduces well a large suite of climate statistics at these scales, including moments (averages, variances), spells, both dry and wet extremes, and extreme hot and cold periods. The high degree of

performance is in part driven by refinements to the weather generator algorithm presented in this work, particularly changes to the WR simulation algorithm and the technique used to jitter extreme precipitation events. Moreover, the stochastic weather generator accurately reproduces the monthly frequencies of AR landfalls at locations along the California coastline using resampled historical dates based on the block bootstrap method.

The high level of performance shown for the San Joaquin basin was found to be consistent across other HUC-4 basins throughout the state (shown in [Supplementary Material](#)). However, there were a few notable limitations to model performance that were consistent across HUC-4

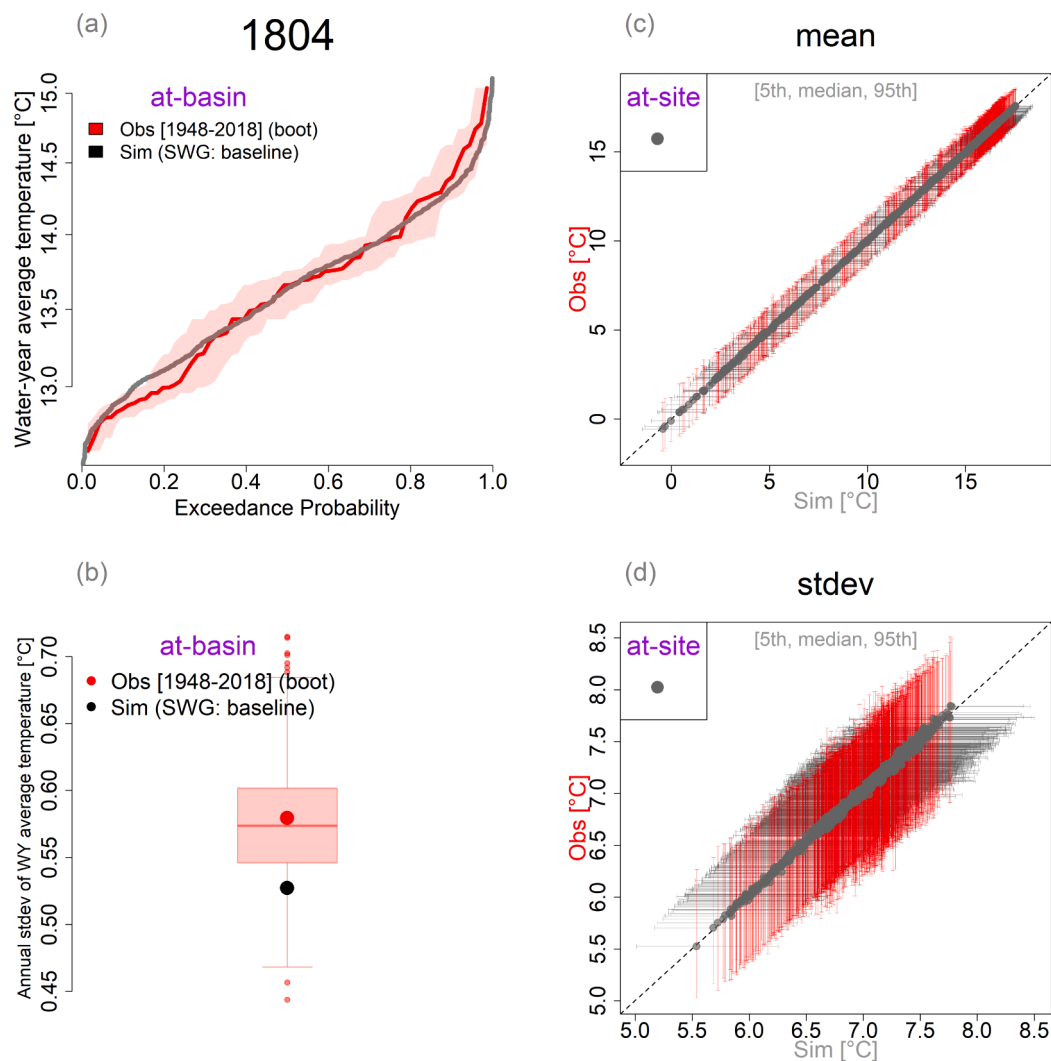


Fig. 9. Observed vs. simulated (a) distribution of basin-scale water-year average temperature (based on the minimum and maximum temperature time series), with the red shaded area representing the uncertainty in the observed distribution using a 95% confidence interval based on bootstrapping from the historical record; (b) standard deviation of basin-scale water-year average temperature, with the boxplot showing the uncertainty in the observed standard deviation based on bootstrapping from the historical record; (c) mean and (d) standard deviation of daily average temperature at individual sites (grid cells), with dots and whiskers showing the 50th, 5th, and 95th percentiles across individual water years. All results are shown for the San Joaquin basin (HUC4: 1804). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

basins and deserve mention here. First, the weather generator tends to underestimate the interannual variance of annual average temperature for almost all HUC-4 basins by approximately 5%. In addition, for several HUC-4 basins, the interannual variance of water-year precipitation totals is also slightly underestimated (by $\sim 5\%$). We also found that for some HUC-4 basins, there were small biases ($<10\%$) in the intensity and duration of heat waves and cold waves for subsets of individual grid cells within each basin. Overall, there were no clear spatial patterns to these biases across the state.

Despite the limitations above, the overall high performance of the model allows the resulting weather generator data products to be used by stakeholders throughout California. Importantly, simulated weather across the state is correctly correlated across space and between different variables (precipitation, temperature), ensuring that weather generator simulations can be used in hydrologic and water resources analyses that span multiple watersheds across California. The stochastic weather generator can also simulate extreme weather conditions, including extreme precipitation across multi-day periods and extreme droughts of varying intensities and durations, that are physically plausible but extend well beyond the range seen in the historical record.

These extremes are simulated even in the absence of climate change, highlighting that water resource managers can use the baseline 1000-year weather generator simulation to better understand how natural climate variability could negatively impact their system.

The ability of the weather generator to generate spatially correlated extremes that extend beyond the magnitude of extreme events in the historical record forms one of the core practical benefits of this model for the water resources community. The 1000-year, ~ 6 km baseline weather generator simulation across the state discussed in this work is publicly available through the California Natural Resources Agency Open Data (<https://data.cnra.ca.gov/dataset/ca-weather-generator-gridded-climate-pr-tmin-tmax-2023>). Practitioners can use these data in conjunction with rainfall-runoff models to create long hydrologic traces for their watershed of interest that contain hydrologic extremes not previously seen in the historical record. Such traces are critical to test the robustness of adaptation strategies (e.g., new reservoir operational policies, new infrastructure for managed aquifer recharge) and to ensure those strategies can meet performance requirements under plausible hydrologic conditions not previously observed. When coupled with different scenarios of climate change, weather generator output can

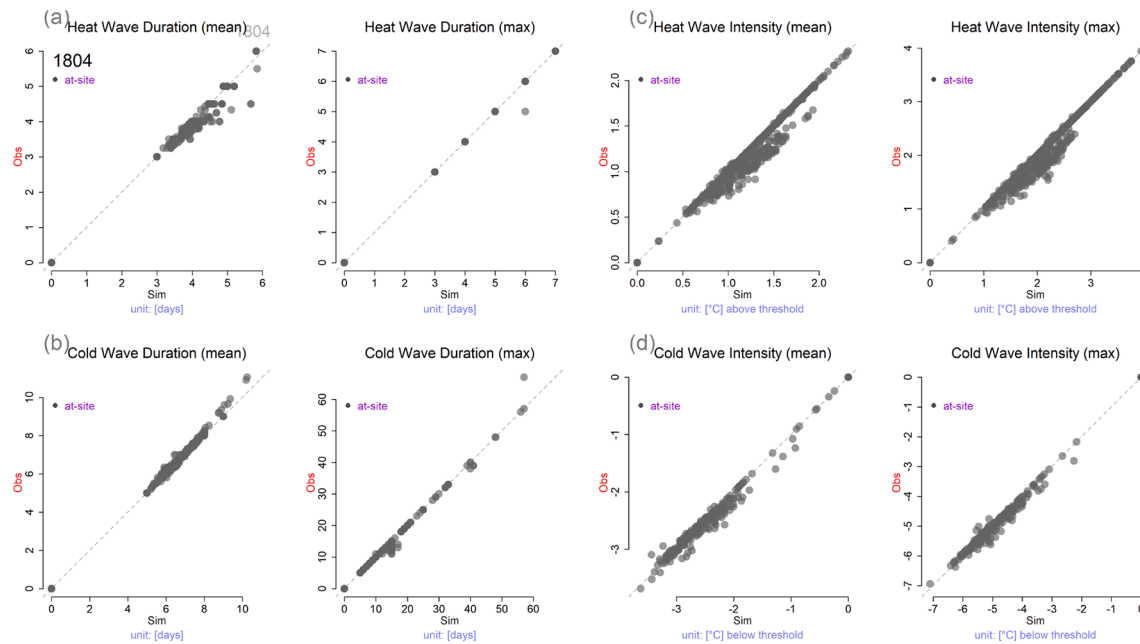


Fig. 10. Observed vs. simulated at-site (a) duration of heat wave (mean, maximum); (b) duration of cold wave (mean, maximum); (c) intensity of heat wave (mean, maximum), and (d) intensity of cold wave (mean, maximum). Each point represents a gridded location in the San Joaquin basin (HUC4: 1804).

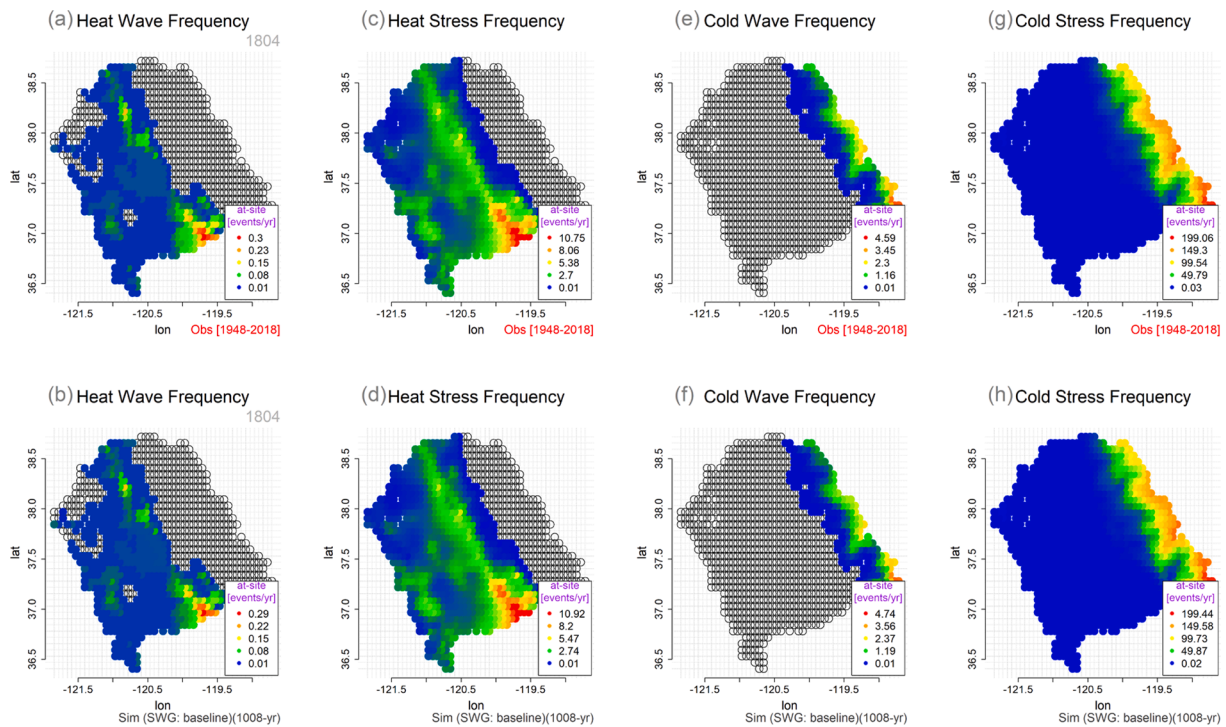


Fig. 11. Frequency of (a) observed and (b) simulated heat waves; frequency of (c) observed and (d) simulated heat stress events; frequency of (e) observed and (f) simulated cold waves; frequency of (g) observed and (h) simulated cold stress events. All results shown for the San Joaquin basin (HUC4: 1804).

then be used to evaluate the combined impacts of natural climate variability and long-term climate change on water system performance. The generation of these climate scenarios is the focus of Part (2).

CRediT authorship contribution statement

Nasser Najibi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alejandro J. Perez:**

Validation, Resources, Project administration, Data curation. **Wyatt Arnold:** Validation, Resources, Project administration, Data curation. **Andrew Schwarz:** Validation, Resources, Data curation. **Romain Maendly:** Validation, Resources, Project administration. **Scott Steinschneider:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

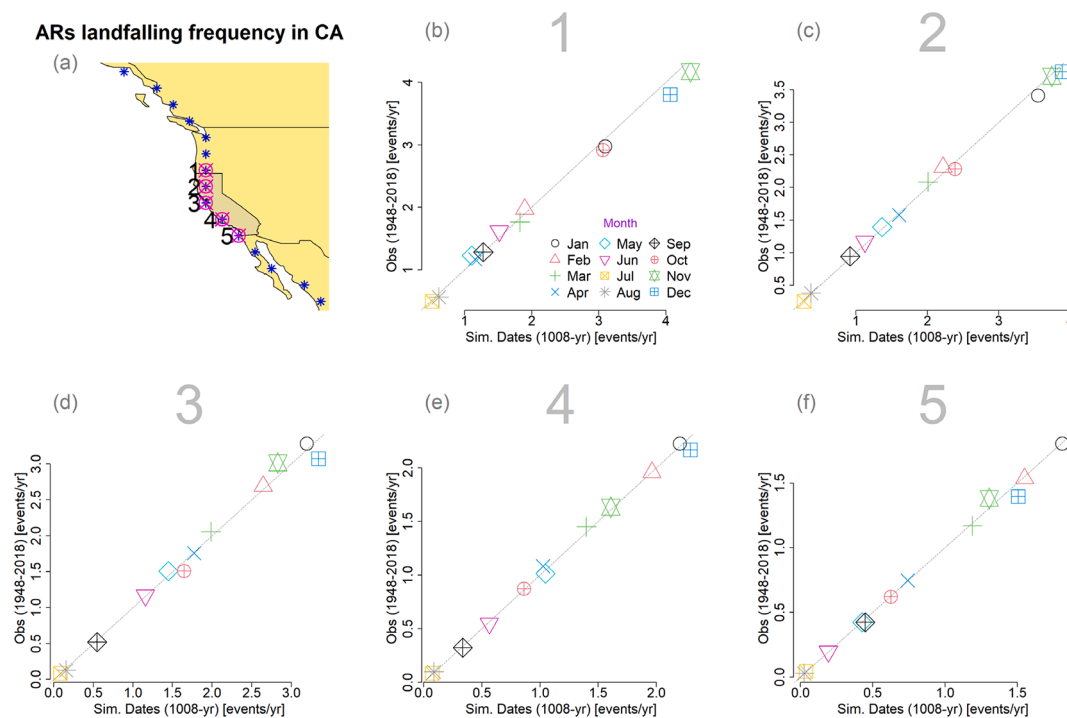


Fig. 12. (a) Location of AR landfalls along Western North America (blue), with 5 locations near and along the California coastline highlighted (pink). (b–f) Observed versus weather generator simulated frequency of AR landfalls, expressed as an average number of days per year for each calendar month, across five locations near and along the California coastline. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Declaration of competing interest

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

Data availability

The climate change scenarios developed under this work are publicly available and can be retrieved from the California Department of Water Resources, who sponsored this work (<https://data.cnra.ca.gov/dataset>,

look up ‘weather generator’, i.e., <https://data.cnra.ca.gov/dataset/ca-weather-generator-gridded-climate-pr-tmin-tmax-2023>).

Acknowledgments

This work was supported by the California Department of Water Resources and Turlock Irrigation District. This project was also partially supported through funds provided under NSF Award No. 2205239. We would like to acknowledge valuable feedback on this work from Alexander Weyant and Alexander Gershunov at the Scripps Institution of Oceanography, UC San Diego.

Appendix A.: Mathematical formulation of the stochastic weather generator

A.1. Non-homogeneous hidden markov models for identifying weather regimes

We utilize a variant of hidden Markov models (HMMs) to identify weather regimes (WRs) from 500-hPa geopotential height anomalies (GPHAs). An HMM involves a finite set of hidden states that transition over time according to the Markov property (Markov, 1954). Each hidden state represents a division of the random field’s state space. This process is akin to cluster analysis, but with the clusters exhibiting Markovian temporal dynamics. At each time step, the spatial field may correspond to a particular hidden state with a certain probability. By leveraging the transition probabilities, each time period can be assigned to a specific state, optimizing the likelihood of the assignment throughout the data’s time span (Rabiner, 1989). Additionally, an external predictor (i.e., a set of exogenous variables, or covariates) can be utilized to influence the transition probabilities over time with a specific period. In such cases, the model is referred to as a nonhomogeneous hidden Markov model (NHMM). NHMMs are an extension of HMMs that allow for time-varying transition probabilities, where the transition probabilities between hidden states change over time according to an external predictor or a set of covariates.

Following Najibi et al. (2021), we use an NHMM to infer WRs from the spatiotemporal evolution of 500 hPa GPHAs, including their persistence, seasonal evolution, and long-term trends. A short description of the mathematical formulation for this approach is provided below.

First, we define the following notations:

K : The number of hidden states (i.e., number of WRs).

T : The number of time steps or observations.

π : The initial state probabilities, a K -dimensional vector where $\pi(i)$ represents the probability of starting in hidden state i .

A : The transition probability matrix, a $K \times K$ matrix that can vary through time, where $A(i,j,t)$ represents the probability of transitioning from hidden state i to hidden state j at time t .

B: The emission probabilities, a $K \times T$ matrix where $B(i,t)$ represents the probability of observing the t -th observation given hidden state i .
X: The external predictors or covariates, a $K \times T$ matrix where each column $X(:,t)$ represents the covariates at time t .

The NHMM workflow can be mathematically described in the following steps:

1. Initialization:
 - o Set the initial state probabilities: $\pi(i)$ for $i = 1$ to K .
2. Time-varying transition probabilities:
 - o Compute the time-varying transition probabilities using the external predictor, or set of covariates: $A(i,j,t) = \Pr(S(t+1) = j \mid S(t) = i, X(:,t))$, where $S(t)$ represents the hidden state at time t . A multinomial regression framework is used here to parameterize the hidden state transitions.
 - o Two types of exogenous variables (i.e., covariates in **X**) are considered in the NHMM: a) two seasonal harmonics $\{1 + \cos(\psi), 1 + \sin(\psi)\}$, where $\psi = \frac{2\pi t}{365}$ represents a periodic signal over the annual cycle; and b) a state-wide wetness index, quantified using the first four principal components from a gridded, water-year standardized precipitation index (SPI) product over California $\{PC1, PC2, PC3, PC4\}$.
3. Emission probabilities:
 - o Compute the emission probabilities using the observations: $B(i,t) = P(O(t) \mid S(t) = i)$, where $O(t)$ represents the vector of observations at time t .
4. Forward-Backward algorithm:
 - o Use the forward-backward algorithm (Baum and Petrie, 1996; Baum et al., 1970) to compute the forward probabilities and backward probabilities (i.e., model coefficients) for each hidden state i and time step t .
5. Baum-Welch algorithm:
 - o Use the Baum-Welch algorithm (also known as the expectation-maximization (EM) algorithm) (Moon, 1996) to estimate the NHMM parameters π , **A**, and **B** based on the forward and backward probabilities.
6. Viterbi algorithm:
 - o Use the Viterbi algorithm (Forney, 1973; Rabiner, 1989) to estimate the most probable sequence of hidden states (i.e., historical sequence of WRs).

The Baum-Welch algorithm iteratively updates the NHMM parameters until convergence, maximizing the likelihood of the observed data (i.e., J PCs of GPHAs in each season). Note that we only used the first four SPI PCs for identifying the WRs in the cold season, but the harmonics in both warm and cold seasons.

We run the EM algorithm 10 times using different random initializations, and utilize the solution with the largest likelihood over all 10 runs to avoid any poor local maxima (Rojo Hernández et al., 2020). We utilized the R-package ‘depmixS4’ (Visser and Speekenbrink, 2010) to fit the NHMM.

A.2. Non-parametric simulation of weather regimes for baseline weather

We developed a novel *non-parametric* approach to WR simulation that addresses the issue of overdispersion in simulated WRs while still allowing for future climate change scenarios with altered WR probabilities. Let $i = 1, \dots, K$ denote the K different WRs, which are available as a daily time series over the historical record (1948–2019). Suppose that the historical time series of WRs are clustered into non-overlapping, consecutive segments, where each segment is D years long and there are N_D segments in total (in this work, $D=4$ and $N_D=18$). In the non-parametric approach for WR simulation, each of the $n = 1, \dots, N_D$ segments is given a sampling probability p_n . To simulate a new sequence of daily WRs for an arbitrary number of years, we simply resample (with replacement) the n^{th} D -year segment of daily WRs with probability p_n , until a sufficient number of years has been generated. The final segment can be truncated to ensure a precise number of years of simulated WRs. In this work, WR simulations are set to 1008 years, which corresponds to 14 times the length of the 72-year historical record. By maintaining the simulation length as a whole multiple of the historical record length, we can compare the historical record directly to 14 separate weather generator traces without any differences in sequence length.

In the baseline scenario for the weather generator with no dynamic climate change, $p_n = \frac{1}{N_D}$ for $n = 1, \dots, N_D$. That is, each segment is considered equally likely. This is the approach taken for those weather generation scenarios that do not incorporate any dynamic climate changes, i.e., no changes to large-scale circulation patterns.

A.3. Copula-based jittering algorithm

We use a copula-based jittering approach that enables bootstrapped values of daily, heavy precipitation to extend beyond the range of the instrumental record. Let $\tilde{p}_t$ be a vector of simulated precipitation values from the bootstrap at time t across all sites. Assume the non-zero, daily precipitation amounts at each site s can be modeled by a distribution with cdf $F(p|\theta_s)$. In this study, we assume precipitation follows an extreme value mixture model (Scarrott and MacDonald, 2012), using a gamma distribution for the bulk density (which varies by month) and a Generalized Pareto distribution (GPD) for values in the tail of the distribution. The cumulative distribution function (cdf) of this model evaluated for precipitation at site s and day t is given by:

$$F(p_{s,t}|\theta_{s,t}) = \begin{cases} \pi_s \frac{F_{\text{gamma}}(p_{s,t}|\alpha_{s,m(t)}, \beta_{s,m(t)})}{F_{\text{gamma}}(u|\alpha_{s,m(t)}, \beta_{s,m(t)})} & p_{s,t} \leq \zeta_s \\ \pi_s + (1 - \pi_s)F_{\text{GPD}}(p_{s,t}|\sigma_s, \xi_s) & p_{s,t} > \zeta_s \end{cases} \quad (\text{A3.1})$$

Here, ζ_s is a threshold that separates heavy from non-heavy precipitation, F_{gamma} is the cdf of a gamma distribution with parameters $\alpha_{s,m(t)}, \beta_{s,m(t)}$ that vary through time based on calendar month $m(t)$, F_{GPD} is the cdf of a GPD with parameters σ_s, ξ_s , and π_s is the probability of precipitation exceeding the threshold ζ_s . The full vector of model parameters is given by $\theta_{s,t}$.

For each simulated day t and site s , let $\tilde{u}_{s,t} = F(\tilde{p}_{s,t}|\theta_{s,t})$ be the non-exceedance probability associated with $\tilde{p}_{s,t}$. We focus specifically on those non-

exceedance probabilities associated with heavy precipitation and utilize the conditional non-exceedance probabilities for the GPD given that $\tilde{p}_{s,t} > \zeta_s$:

$$\tilde{u}_{s,t,GPD} = \frac{\tilde{u}_{s,t} - \pi_s}{(1 - \pi_s)} \quad (A3.2)$$

The $\tilde{u}_{s,t,GPD}$ values are the non-exceedance probabilities for the GPD component of the extreme value mixture model and will range from 0 to 1. At any time t , let $\tilde{u}_{t,GPD}$ denote the vector of values $\tilde{u}_{s,t,GPD}$ only for sites $s \in \psi_t$ where $p_{s,t} > \zeta_s$. That is, ψ_t is the subset of sites with heavy precipitation on day t , so that $|\psi_t| \leq S$.

We perturb the values in the vector $\tilde{u}_{t,GPD}$ to create a new vector of values $\tilde{u}_{t,GPD}^*$ that are centered around but are not equal to $\tilde{u}_{t,GPD}$. The perturbations are simulated using a Gaussian copula. Let Σ be an $S \times S$ Spearman (rank) correlation matrix for the vector of all daily, observed precipitation across sites, and let $\tilde{z}_{s,t,GPD} = \Phi^{-1}(\tilde{u}_{s,t,GPD})$ be a z-score (i.e., Φ is the standard normal cdf) for simulated heavy precipitation at time t and site s in the set ψ_t . Note that for simplicity we drop the notation ψ_t , but emphasize that z-scores at time t ($\tilde{z}_{s,t,GPD}$) are only calculated for sites s with heavy precipitation at time t . We create a covariance matrix to simulate new z-scores $\tilde{z}_{t,GPD}^*$ from a multivariate normal distribution centered around the original scores $\tilde{z}_{t,GPD}$:

$$\tilde{z}_{t,GPD}^* \sim \text{MVN}(\tilde{z}_{t,GPD}, \Omega) \quad (A3.3)$$

With

$$\Lambda = \lambda I$$

$$\Omega = \Lambda \Sigma \Lambda^T \quad (A3.4)$$

Here, Λ is a diagonal matrix of dimension $|\psi_t| \times |\psi_t|$ with constant diagonal term λ , which is a user-defined parameter. If $\lambda = 1$, then $\Omega = \Sigma$ and $\tilde{z}_{t,GPD}^*$ will deviate substantially from the original values $\tilde{z}_{t,GPD}$ (which are based on the bootstrapped precipitation values), but will retain the observed correlation structure across sites. However, as λ is made small, the matrix Ω will have small variances along the diagonal and $\tilde{z}_{t,GPD}^*$ will not vary much from $\tilde{z}_{t,GPD}$. Then, for each site, the perturbed z-score can be back-transformed to a proposed non-exceedance probability for the GPD, $\tilde{u}_{s,t,GPD}^* = \Phi(\tilde{z}_{s,t,GPD}^*)$, and an associated proposed precipitation value $\tilde{p}_{s,t,GPD}^* = F_{GPD}^{-1}(\tilde{u}_{s,t,GPD}^* | \sigma_s, \xi_s)$. The proposed non-exceedance probability will then be selected over the original one based on the following conditional probabilities of observing a different precipitation value given the value that was simulated:

$$\pi = \begin{cases} \Pr(P > \tilde{p}_{s,t,GPD}^* | P > \tilde{p}_{s,t,GPD}) = \frac{\Pr(P > \tilde{p}_{s,t,GPD}^*, P > \tilde{p}_{s,t,GPD})}{\Pr(P > \tilde{p}_{s,t,GPD})} = \frac{\Pr(P > \tilde{p}_{s,t,GPD}^*)}{\Pr(P > \tilde{p}_{s,t,GPD})} = \frac{1 - \tilde{u}_{s,t,GPD}^*}{1 - \tilde{u}_{s,t,GPD}}, & \tilde{p}_{s,t,GPD}^* > \tilde{p}_{s,t,GPD} \\ \Pr(P \leq \tilde{p}_{s,t,GPD}^* | P \leq \tilde{p}_{s,t,GPD}) = \frac{\Pr(P \leq \tilde{p}_{s,t,GPD}^*, P \leq \tilde{p}_{s,t,GPD})}{\Pr(P \leq \tilde{p}_{s,t,GPD})} = \frac{\Pr(P \leq \tilde{p}_{s,t,GPD}^*)}{\Pr(P \leq \tilde{p}_{s,t,GPD})} = \frac{\tilde{u}_{s,t,GPD}^*}{\tilde{u}_{s,t,GPD}}, & \tilde{p}_{s,t,GPD}^* \leq \tilde{p}_{s,t,GPD} \end{cases}$$

$$\tilde{u}_{s,t,GPD}^{final} = \begin{cases} \tilde{u}_{s,t,GPD}^* & \pi \leq r_{s,t} \\ \tilde{u}_{s,t,GPD} & \pi > r_{s,t} \end{cases} \quad (A3.5)$$

where the random variable P is daily precipitation depth and $r_{s,t}$ is a random draw from a uniform distribution between 0 and 1 for site s and time t . The final heavy precipitation value for each site is then set equal to

$$F_{GPD}^{-1}(\tilde{u}_{s,t,GPD}^{final} | \sigma_s, \xi_s) \quad (A3.6)$$

By virtue of the perturbations embedded in $\tilde{z}_{t,GPD}^*$ and thus $\tilde{u}_{t,GPD}^*$, the final values of simulated heavy precipitation can extend beyond the range of historical heavy precipitation values, but they preserve the rank correlation structure across sites and the space-time structure captured by the block bootstrap, as long as λ is small. We calibrate λ to balance the reproduction of extreme events beyond the range of the historical record (e.g., estimates of the 500-year and 1000-year storm) while maintaining the general spatial structure of bootstrapped storms. We find that satisfactory results are generally achieved with $\lambda \in (0.1, 0.5)$. In this work, we set $\lambda = 0.4$.

Appendix B. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cliser.2024.100489>.

References

- Acharya, N., Frei, A., Chen, J., DeCristofaro, L., Owens, E.M., 2017. Evaluating stochastic precipitation generators for climate change impact studies of New York City's primary water supply. *J. Hydrometeorol.* 18 (3), 879–896.
- Ailliot, P., Allard, D., Monbet, V., Naveau, P., 2015. Stochastic weather generators: an overview of weather type models. *Journal De La Société Française De Statistique (J. French Stat. Soc.)* 156 (1), 101–113.
- Baum, L.E., Petrie, T., 1966. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.* 37 (6), 1554–1563.
- Baum, L.E., Petrie, T., Soules, G., Weiss, N., 1970. A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *Ann. Math. Stat.* 41, 164–171. <https://doi.org/10.1214/aoms/1177697196>.
- Bird, L.J., Walker, M.G., Bodeker, G.E., Campbell, I.H., Liu, G., Sam, S.J., Lewis, J., Rosier, S.M., 2023. Deep learning for stochastic precipitation generation – deep SPG v1.0. *Geosci. Model Dev.* 16 (13), 3785–3808.
- Borkotoky, S.S., Williams, A.P., Cook, E.R., Steinschneider, S., 2021. Reconstructing extreme precipitation in the Sacramento River watershed using tree-ring based proxies of cold-season precipitation. *Water Resour. Res.* 57 (4) e2020WR028824.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *J. R. Stat. Soc. Ser. B* 39, 1–22. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.
- Dettinger, M.D., 2013. Atmospheric Rivers as Drought Busters on the U.S. West Coast. *J. Hydrometeorol.* 14, 1721–1732. <https://doi.org/10.1175/JHM-D-13-02.1>.
- Dyreson, A., Devineni, N., Turner, S.W., De Silva M, T., Miara, A., Voisin, N., Cohen, S. and Macknick, J. (2022). The role of regional connections in planning for future power system operations under climate extremes. *Earth's Future*, 10(6), p. e2021EF002554.
- Emori, S., Brown, S.J., 2005. Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophys. Res. Lett.* 32 (17).
- Forney, G.D., 1973. The Viterbi algorithm. *Proc. IEEE* 61 (3), 268–278.
- Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Review: Linking climate change modeling to impact studies: recent advances in downscaling techniques for hydrologic modeling. *Int. J. Climatol.* 27, 1547–1578.
- Gershunov, A., Sulghina, T., Ralph, F.M., Lavers, D.A., Rutz, J.J., 2017. Assessing the climate-scale variability of atmospheric rivers affecting western North America. *Geophys. Res. Lett.* 44 (15), 7900–7908. <https://doi.org/10.1002/2017GL074175>.
- Gupta, R., Steinschneider, S., Reed, P.M., 2022. A Multi-Objective Paleo-Informed Reconstruction of Western U.S. Weather Regimes Over the Past 600 Years. *Clim. Dyn.* <https://doi.org/10.1007/s00382-022-06302-4>.
- Harris, C.N.P., Quinn, A.D., Bridgeman, J., 2014. The use of probabilistic weather generator information for climate change adaptation in the UK water sector. *Meteorol. Appl.* 21 (2), 129–140.
- Henn, B., Musselman, K.N., Lestak, L., Ralph, F.M., Molotch, N.P., 2020. Extreme runoff generation from atmospheric river driven snowmelt during the 2017 Oroville Dam spillways incident. *Geophys. Res. Lett.* 47 (14) e2020GL088189.
- Hughes, J.P., Guttorp, P., 1994. A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water Resour. Res.* 30 (5), 1535–1546.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* 77 (3), 437–472.
- Khazaei, M.R., Zahabiyou, B., Hasirchian, M., 2020. Comparison of IWG and SDSM weather generators for climate change impact assessment. *Theor. Appl. Climatol.* 140, 859–870.
- LeNoir, J.M., Najibi, N., and Steinschneider, S. (2023). Predicted temperature and precipitation values derived from modeled localized weather regimes and climate change in the state of Massachusetts: U.S. Geological Survey data release, <https://doi.org/10.5066/P9KTY3MS>.
- Livneh, B., Rosenberg, E.A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K.M., Maurer, E. P., Lettenmaier, D.P., 2013. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Clim.* 26, 9384–9392. <https://doi.org/10.1175/JCLI-D-12-00508.1>.
- Livneh, B., Bohn, T.J., Pierce, D.S., Munoz-Ariola, F., Nijssen, B., Vose, R., Cayan, D., Brekke, L.D., 2015. A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and southern Canada 1950–2013. *Nature. Sci. Data* 5:150042. <https://doi.org/10.1038/sdata.2015.42>.
- Maraun, D., Shepherd, T.G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J.M., Hagemann, S., Richter, I., Soares, P.M., Hall, A., Mearns, L.O., 2017. Towards process-informed bias correction of climate change simulations. *Nat. Clim. Chang.* 7 (11), 764–773.
- Markov, A. A. (1954). The theory of algorithms. *Trudy Matematicheskogo Instituta Imeni VA Steklova*, 42, 3–375. (Original title: Teoriya algorifmov, Russian Translation of Works of the Mathematical Institute, Academy of Sciences of the USSR, Available at: <http://mi.mathnet.ru/tm1178>).
- Moon, T.K., 1996. The expectation-maximization algorithm. *IEEE Signal Process. Mag.* 13 (6), 47–60.
- Mukundan, R., Archarya, N., Gelda, R.G., Frei, A., Owens, E.M., 2019. Modeling streamflow sensitivity to climate change in New York City water supply streams using a stochastic weather generator. *J. Hydrol.: Reg. Stud.* 21, 147–158.
- Najibi, N., Mukhopadhyay, S., Steinschneider, S., 2021. Identifying weather regimes for regional-scale stochastic weather generators. *Int. J. Climatol.* 41, 2456–2479. <https://doi.org/10.1002/joc.6969>.
- Pierce, D.W., Su, L., Cayan, D.R., Risser, M.D., Livneh, B., Lettenmaier, D.P., 2021. An extreme-preserving long-term gridded daily precipitation dataset for the conterminous United States. *J. Hydrometeorol.* 22 (7), 1883–1895.
- Pierce, D.W., Cayan, D.R., Feldman, D.R., Risser, M.D., 2023. Future Increases in North American Extreme Precipitation in CMIP6 Downscaled with LOCA. *J. Hydrometeorol.* 24, 951–975. <https://doi.org/10.1175/JHM-D-22-0194.1>.
- PRISM Climate Group (2014). Oregon State University, <https://prism.oregonstate.edu>, data created 4 Feb 2014.
- Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 257–286.
- Rahat, S.H., Steinschneider, S., Kucharski, J., Arnold, W., Olzewski, J., Walker, W., Maendly, R., Wasti, A., Ray, P., 2022. Characterizing Hydrologic Vulnerability under Non-Stationary Climate and Antecedent Conditions using a Process-Informed Stochastic Weather Generator. *J. Water Resour. Plan. Manag.* 148 (6) [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001557](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001557).
- Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour. Res.* 17 (1), 182–190.
- Robertson, A.W., Y. Kushnir, U. Lall, and J. Nakamura (2015). Weather and climatic drivers of extreme flooding events over the Midwest of the United States. *Extreme Events: Observations, Modeling, and Economics, Geophys. Monogr.*, 214, Amer. Geophys. Union, 113–124.
- Robertson, A.W., Ghil, M., 1999. Large-scale weather regimes and local climate over the western United States. *J. Clim.* 12, 1796–1813.
- Rojó Hernández, J.D., Mesa, Ó.J., Lall, U., 2020. ENSO dynamics, trends, and prediction using machine learning. *Weather Forecast.* 35 (5), 2061–2081.
- Scarrott, C.J., MacDonald, A., 2012. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT – Stat. J.* 10 (1), 33–59.
- Seager, R., Naik, N., Vecchi, G.A., 2010. Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming. *J. Clim.* 23 (17), 4651–4668.
- Seager, R., Neelin, D., Simpson, I., Liu, H., Henderson, N., Shaw, T., Kushnir, Y., Ting, M., Cook, B., 2014. Dynamical and thermodynamical causes of large-scale changes in the hydrological cycle over North America in response to global warming. *J. Clim.* 27 (20), 7921–7948.
- Semenov, M.A., Barrow, E.M., 1997. Use of a stochastic weather generator in the development of climate change scenarios. *Clim. Change* 35 (4), 397–414.
- Steinschneider, S., Brown, C., 2013. A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments. *Water Resour. Res.* 49 (11), 7205–7220. <https://doi.org/10.1002/wrcr.20528>.
- Steinschneider, S., Ray, P., Rahat, S.H., Kucharski, J., 2019. A weather-regime based stochastic weather generator for climate vulnerability assessments of water systems in the Western United States. *Water Resour. Res.* 55 (8), 6923–6945.
- Stephenson, D.B., Collins, M., Rougier, J.C., Chandler, R.E., 2012. Statistical problems in the probabilistic prediction of climate change. *Environmetrics* 23 (5), 364–372.
- Visser, I., Speekenbrink, M., 2010. depmixS4: An R package for hidden Markov models. *J. Stat. Softw.* 36, 1–21. <https://doi.org/10.18637/jss.v036.i07>.
- Wilks, D.S., 2002. Realizations of daily weather in forecast seasonal climate. *J. Hydrometeorol.* 3, 195–207.
- Wilks, D.S., 2010. Use of stochastic weather generators for precipitation downscaling. *Wiley Interdiscip. Rev. Clim. Chang.* 1 (6), 898–907.
- Wilks, D.S., 2012. Stochastic weather generators for climate-change downscaling, part II: multivariable and spatially coherent multisite downscaling. *Wiley Interdiscip. Rev. Clim. Chang.* 3 (3), 267–278.
- Wilks, D.S., Wilby, R.L., 1999. The weather generation game: A review of stochastic weather models. *Prog. Phys. Geogr.* 23, 329–357.
- Williams, A.P., Cook, B.I., Smerdon, J.E., 2022. Rapid intensification of the emerging southwestern North American megadrought in 2020–2021. *Nat. Clim. Chang.* 12, 232–234.
- Yang, X., He, R., Ye, J., Tan, M.L., Ji, X., Tan, L., Wang, G., 2020. Integrating an hourly weather generator with an hourly rainfall SWAT model for climate change impact assessment in the Ru River Basin, China. *Atmos. Res.* 244, 105062.
- Zechiel, P.R., Chiao, S., 2021. Climate Variability of Atmospheric Rivers and Droughts over the West Coast of the United States from 2006 to 2019. *Atmosphere* 12 (2), 201. <https://doi.org/10.3390/atmos12020201>.