https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

Daily Precipitation-Frequency Estimates under Climate Oscillations across the Conterminous United States

Ali Takallou[1], Nibedita Samal[1], and Hamid Moradkhani[1,2]

[1]Department of Civil, Construction, and Environmental Engineering, The University of Alabama, Tuscaloosa, AL, USA

[2]Department of Computer Science, The University of Alabama, Tuscaloosa, AL, USA

Correspondence to: Hamid Moradkhani (hmoradkhani@ua.edu)

Abstract: Precipitation-frequency data are essential for design storms, flood risk assessment, infrastructure design, and erosion and sediment control plans, but remain uncertain due to regional heterogeneity and climate variability. Existing precipitation-frequency products are generally based on stationary estimates and provide limited information on how extreme-precipitation frequency varies with large-scale climate conditions. Accounting for these spatial and climate variability effects in models over large domains is computationally intensive, which has limited the generation of such precipitation-frequency products. In this study, we developed a 0.1° gridded precipitation-frequency dataset, named PreXFOCUS (Precipitation eXtreme Frequency under Oscillations in Climate across the Conterminous United States; Takallou et al., 2026), using Bayesian hierarchical modeling of annual maximum daily precipitation. The dataset is available on Zenodo at https://doi.org/10.5281/zenodo.21979711. Four model formulations were compared, ranging from independent site-specific estimation to spatially varying extreme-value parameter fields with Gaussian-copula dependence among annual maxima. Spatial pooling across all GEV parameters combined with data-level spatial dependence provided the best goodness of fit. The selected model was then used for

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

conditional spatial simulation at held-out sites, where the resulting return levels were validated against estimates obtained from direct model fitting. These return-level estimates were subsequently incorporated into spatially varying Bernoulli occurrence models to quantify climate-conditioned changes in their exceedance probabilities. Adding MEI, NAO, and AMO to these occurrence models improved model fit relative to the time-invariant baseline and shifted exceedance probabilities by up to 40% in some regions. The resulting products provide fine-resolution precipitation-frequency design estimates and climate-informed return levels across the CONUS, supporting hydrologic modeling, infrastructure planning, and regional flood-risk assessment.

1 Introduction

Reliable precipitation-frequency estimates are fundamental to hydraulic infrastructure design, stormwater management, and flood-risk assessment. However, short observational records, irregular gauge networks, and uncertainty in upper-tail behavior make these estimates difficult to obtain (Hu et al., 2020; Patel and Schneider, 2026; Scarrott and MacDonald, 2012). The challenges in estimating extreme precipitation frequency are compounded by regional differences in topography, diverse precipitation-generating mechanisms, and large-scale climate variability (Lu et al., 2025; Obeysekera and Salas, 2016; Sun et al., 2015; Sun and Lall, 2015). Therefore, a large-domain precipitation-frequency dataset is needed to account for regional heterogeneity and temporal climate variability while also quantifying uncertainty in areas with sparse observations.

Bayesian hierarchical models (BHMs) have been widely used in regional hydrologic-extreme analysis to combine site-specific extreme-value distributions with spatial latent processes and dependence structures (Aryal et al., 2009; Atyeo and Walshaw, 2012; Cooley et al., 2007; Davison et al., 2012). This structure enables spatial pooling, uncertainty propagation, and prediction at locations with limited or unavailable observations. However, applying spatial Bayesian extreme-value models to continental-scale gauge networks is computationally demanding, and previous large-domain studies have often

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

relied on likelihood approximations, regional subdivision, or simplified spatial representations (Banerjee et al., 2008; Bracken et al., 2016; Castruccio et al., 2016; Ghosh and Mallick, 2011; Reich and Shaby, 2012; Ribatet et al., 2012; Schliep et al., 2010). Recent advances in JAX-based probabilistic programming have improved the computational feasibility of large-scale Bayesian inference (Phan et al., 2019). This study uses GPU-accelerated Bayesian inference implemented in JAX to enable posterior estimation for a continental-scale gauge network and generate gridded precipitation-frequency products.

BHMs must represent two distinct forms of spatial structure. The latent process layer describes regional variation and spatial pooling in the parameters of the GEV distribution, whereas the data layer may represent dependence among annual maxima after accounting for their marginal distributions. Previous studies have represented data-layer dependence using conditional independence assumptions (Boumis et al., 2023; Cooley et al., 2007; García et al., 2018; Lu et al., 2025), copulas (Beck et al., 2020; Bracken et al., 2016; Renard, 2011), or max-stable processes (Davison et al., 2012; Ribatet et al., 2012; Shang et al., 2011; Stephenson et al., 2016). Conditional independence cannot represent the joint occurrence of precipitation extremes across gauges, potentially underestimating their joint probability and spatial coherence. Copula and max-stable models capture cross-gauge dependence but rely on selected dependence structures and can be computationally demanding for large networks. This study evaluates model formulations that spatially pool different combinations of the GEV parameters and either assume conditional independence or represent inter-gauge dependence at the data layer.

Large-scale climate variability may also alter the probability of extreme-precipitation occurrence. BHMs represent climate associations by incorporating standard climate indices (SCIs) into their latent process structures (Aryal et al., 2009; Lima et al., 2015). Previous studies have represented these effects by allowing GEV parameters to vary with climate indices (Bracken et al., 2018; García et al., 2018; Lu et al., 2025) or by modeling threshold-exceedance probabilities as functions of climate covariates using Bernoulli regression (Renard and Thyer, 2019). While the former can decrease identifiability and increase posterior sampling uncertainty when multiple GEV parameters vary simultaneously (Montanari and Koutsoyiannis, 2014; Prosdocimi and Kjeldsen, 2021; Serinaldi and Kilsby, 2015), the latter provides climate-conditioned exceedance probabilities without assigning temporal variability to predefined GEV parameters. However, precipitation-frequency models that account for climate

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

variability across large spatial domains have been limited because of computational challenges (Bracken et al., 2016). Hence, this study produces precipitation-frequency products that incorporate large-scale climate indices across CONUS.

Despite advances in regional precipitation-frequency modeling, existing design estimates remain difficult to reuse for two reasons. First, return levels are typically reported at individual gauges, and their estimation uncertainty is rarely quantified in a spatially consistent form. This limits their use in ungauged and sparsely monitored areas and requires users to repeat site-specific analyses. Second, return levels are conventionally treated as time-invariant, whereas the frequency of extreme precipitation across the CONUS varies with large-scale climate indices. Therefore, a stationary return level cannot reflect how the probability of exceeding a design threshold changes with large-scale climate variability. A dataset that provides spatial return levels together with their posterior uncertainty addresses the first gap, while climate-conditioned probabilities of exceeding these return-level thresholds associated with MEI, NAO, and AMO address the second gap.

This study presents PreXFOCUS, a 0.1° gridded precipitation-frequency dataset across the CONUS generated by applying established spatial extreme-value and climate-informed occurrence models within a Bayesian hierarchical framework. The data was derived from observations at 936 gauges during 1951–2020 accounting for regional variation in GEV parameters and inter-gauge dependence. The dataset provides reference and climate-conditioned return-level estimates, together with their corresponding credible-interval bounds. These products complement established precipitation-frequency resources, such as NOAA Atlas 14 (National Oceanic and Atmospheric Administration), by providing spatially distributed uncertainty estimates and climate-conditioned precipitation-frequency information. They may be reused for infrastructure and flood-risk analyses, hydrologic and hydraulic model evaluation, and benchmarking of spatial extreme-value methods.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

2 Methods

2.1 Input datasets

Daily precipitation observations were obtained from the Global Historical Climatology Network (GHCN-Daily) database through the National Centers for Environmental Information at (https://www.ncei.noaa.gov/data/global-historical-climatology-network-daily/). The 1951–2020 period was selected for dataset generation to ensure temporal consistency between the GHCN-D precipitation and climate variability indices. At each gauge, years with more than 20% missing daily precipitation observations were considered incomplete and excluded from annual-maximum extraction. For years satisfying this criterion, annual maxima were computed from the available daily observations without imputing the remaining missing values. Gauges were retained only if all 70 years met the completeness criterion and yielded a valid annual maximum. This screening procedure resulted in 936 gauges with complete annual-maxima series for 1951–2020. To evaluate predictive performance before generating spatially gridded products, the sites were randomly divided into calibration sites (used for posterior inference) and held-out sites (excluded from inference and used only for validation), consisting of 327 and 609 sites, respectively. A relatively large held-out set was used to ensure a robust assessment of spatial prediction performance across diverse hydroclimatic regions. The geographical distribution of both calibration and held-out sites across CONUS is shown in Fig. 1, overlaid on an elevation map.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

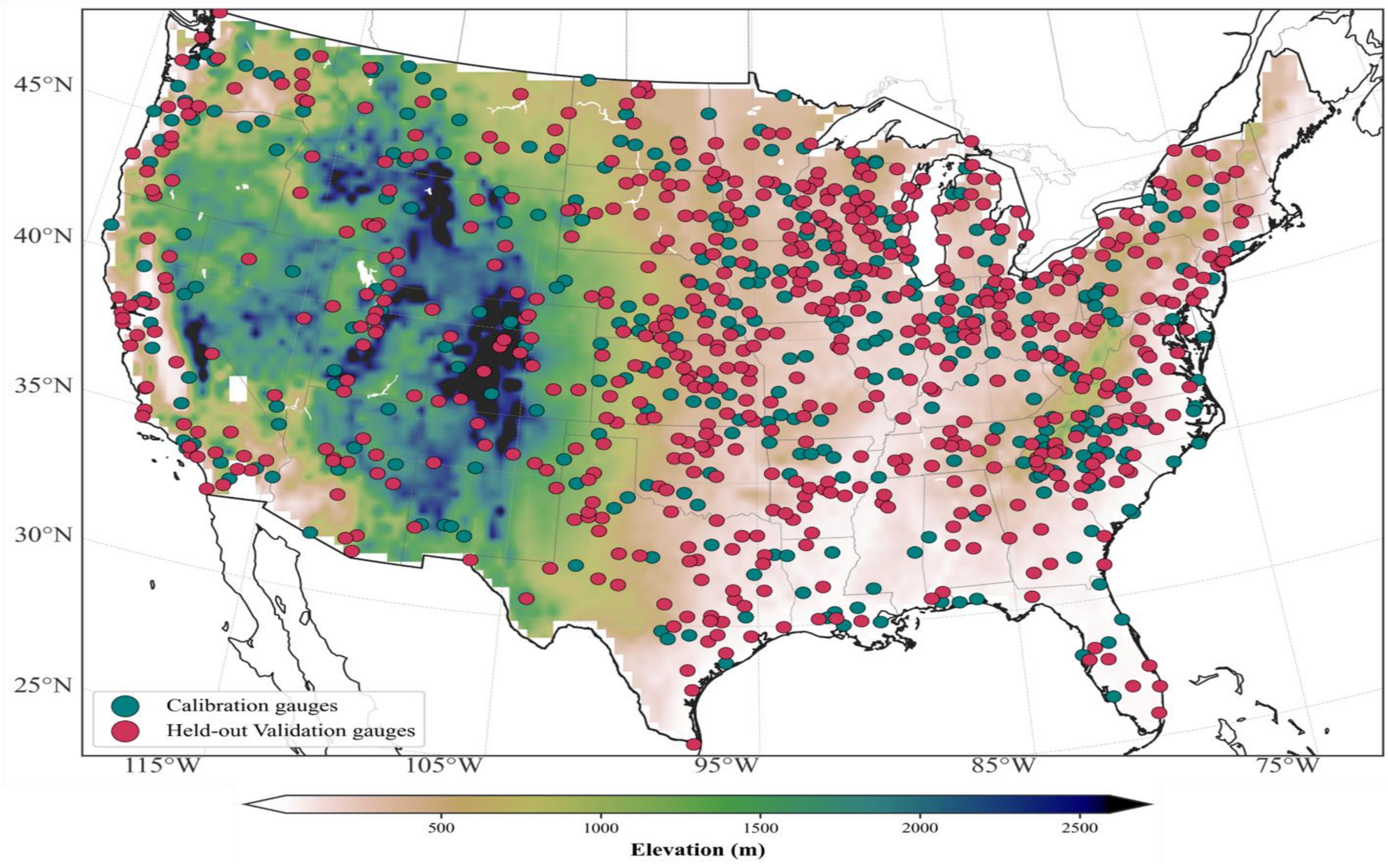


Figure 1. Elevation map and spatial distribution of calibration and held-out precipitation sites across CONUS.

To examine climate dependence in exceedance occurrence probabilities, four SCIs were considered. The Multivariate El Niño–Southern Oscillation Index (MEI) and Pacific Decadal Oscillation (PDO) index were obtained from NOAA (Newman et al., 2016; Wolter and Timlin, 2011). The North Atlantic Oscillation (NAO) index was obtained from the NCAR station-based dataset, and the Atlantic Multidecadal Oscillation (AMO) index was obtained from the NCAR Climate Analysis Section (Hurrell and Deser, 2010; Schneider et al., 2013). Monthly climate indices were averaged to annual values to ensure consistency with the annual maximum precipitation series.

2.2 Statistical models

This study employs established Bayesian hierarchical models to generate precipitation-frequency estimates across the CONUS. Annual maxima are modeled following regional Bayesian hierarchical

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

formulations for spatial extreme-value analysis (Boumis et al., 2023; Bracken et al., 2016; Najafi and Moradkhani, 2015; Najafi and Moradkhani, 2013; Yan and Moradkhani, 2015, 2016), while the probability of occurrence of precipitation extremes is modeled as a function of climate indices using logistic regression (Kwon et al., 2008; Lima and Lall, 2009; Renard and Thyer, 2019). Detailed hierarchical specifications, and covariance formulations are provided in Appendix A1–A4.

The annual maxima, denoted by $Y(s,t)$ at location s and year t, are modeled as block maxima using the GEV distribution, with the parameter vector at location s defined as:

$$\boldsymbol{\theta}(s) = [\mu\,(s), \sigma\,(s), \xi\,(s)], \tag{1}$$

where μ (s), σ (s) > 0, and ξ (s) denote location, scale, and shape parameters, respectively. Once $\boldsymbol{\theta}(s)$ is estimated, the precipitation return level, z_p, corresponding to a specified exceedance probability, p, can then be calculated as:

$$z_p = \mu - \frac{\sigma}{\xi}\left(1 - (-\log(1 - p))^{-\xi}\right), \tag{2}$$

This formulation follows the standard GEV parameterization for block maxima (Coles et al., 2001; Mishra and Singh, 2010). Because the scale parameter must be positive, log[σ(s)] was modeled; for notational simplicity, it is referred to as σ(s) hereafter.

Four model formulations (M1–M4), adapted from established spatial extreme-value modeling approaches, were implemented and comparatively evaluated to determine the representation most suitable for generating the precipitation-frequency dataset. M1 estimates site-specific GEV parameters without spatial pooling (Coles et al., 2001). M2 introduces spatial pooling of the location and log-scale parameters while retaining a spatially homogeneous shape parameter, consistent with previous regional Bayesian extreme-value applications (Boumis et al., 2023; Lu et al., 2025; Najafi and Moradkhani, 2013, 2015; Yan and Moradkhani, 2015, 2016). In the spatially pooled formulations, spatial variation in the GEV parameters was represented using latent Gaussian processes (GPs), with longitude, latitude, and elevation included in the mean structure and an exponential covariance function used to represent distance-based spatial dependence. The complete GP formulation is provided in Appendix A1. M3 additionally allows the GEV shape parameter to vary spatially, as considered in previous hierarchical extreme-value studies (Cooley and Sain, 2010; Lima et al., 2016; Wu et al., 2018, 2019). M4 retains the spatial structure of M3 and additionally represents dependence among annual maxima at the data layer

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

using a Gaussian copula, following previous applications of copula-based spatial extreme-value models (Bracken et al., 2016; Renard, 2011). Such dependence can arise when large storm systems produce concurrent precipitation extremes across multiple locations (Sharkey and Winter, 2019). Complete mathematical specifications of M1–M4 are provided in Appendix A2–A3, and the criteria used to select among them are described in Sect. 2.4.

Models M1–M4 represent the annual maximum precipitation using time-invariant GEV parameters and do not explicitly include climate-related temporal variation. This study applies a Bernoulli model to represent the occurrence of extreme precipitation events, following previous applications to hydrologic extremes (Lima and Lall, 2009; Renard et al., 2022; Renard and Thyer, 2019), and incorporates large-scale SCIs into the latent occurrence process, consistent with previous climate-conditioned extreme-value studies (Aryal et al., 2009; Bracken et al., 2018; Lima et al., 2015). The complete Bernoulli occurrence-model formulation is provided in Appendix A4.

The Bernoulli occurrence model requires a site-specific threshold to define the binary exceedance response. Therefore, the model selected from M1–M4 was used to estimate the 2-, 5-, 10-, 25-, 50-, and 100-year return levels at each site, which were then used as site-specific exceedance thresholds to define binary indicators from the annual maximum precipitation records. Six separate Bernoulli occurrence models were then fitted, one for each return-period threshold. For each threshold, the exceedance probability was modeled as a function of large-scale climate conditions through spatially varying regression coefficients. Additional details on the occurrence-model specification are provided in Appendix A4. Weakly informative priors were assigned to the model parameters and hyperparameters to regularize estimation while allowing the observations to dominate posterior inference.

To assess whether incorporating large-scale climate variability improves the representation of exceedance occurrence relative to a time-invariant baseline, multiple candidate formulations were evaluated. The occurrence-model candidates included an intercept-only baseline model, denoted N1, and candidate formulations incorporating individual and combined climate indices from AMO, MEI, NAO, and PDO. The climate-conditioned formulation selected through the model-selection procedure described in Sect. 2.4.2 was denoted N2. These models are referred to as climate-conditioned because the exceedance probabilities vary annually with the observed climate-index values. This formulation does not imply a deterministic temporal trend or temporal variation in the underlying GEV parameters.

2.3 Posterior estimation and spatial prediction

Posterior distributions were sampled using the No-U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014), an extension of Hamiltonian Monte Carlo (HMC) (Neal, 2011), implemented in NumPyro version 0.18.0 with JAX version 0.4.28 and Python version 3.10.17. NumPyro enables efficient Bayesian inference through JIT compilation, automatic vectorization, and GPU/TPU acceleration (Phan et al., 2019). For each model, three parallel chains were run for 3,000 iterations, with the first 2,000 iterations discarded as warm-up, yielding 3,000 total post-warm-up samples for inference. Convergence was assessed using the split-$\hat{R}$ statistic and effective sample size (N_{eff})(Vehtari et al., 2021). Chains were considered converged when (a) split-$\hat{R} \approx 1.0$, and (b) $\frac{N_{eff}}{N} > 0.1$, where N represents the total number of post-warm-up iterations.

Spatial prediction was performed to estimate latent GEV parameter surfaces at held-out sites and gridded locations. For each posterior draw, the latent GP process was conditionally simulated at target locations using the fitted GP mean function, covariance structure, observed gauge locations, and site covariates, following the conditional GP formulation (Davison et al., 2012; Renard, 2011; Yan and Moradkhani, 2015). This procedure propagates posterior uncertainty in both the regression coefficients and GP covariance parameters to target locations. The conditionally simulated GEV parameters were then used to compute return levels and posterior predictive annual maxima at held-out sites. For M4, posterior prediction additionally conditioned the Gaussian copula field on calibration-site observations to preserve spatial dependence among annual maxima at held-out sites.

2.4 Model selection and evaluation

Although statistical formulations considered here have been previously applied and evaluated in different regions, their relative goodness of fit and predictive performance can depend on the spatial extent, hydroclimatic heterogeneity, and observational network of the application domain. Therefore, before generating the CONUS precipitation-frequency products, the candidate annual-maximum models

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

(M1–M4) and occurrence models (N1–N2) were comparatively evaluated to identify the formulations most suitable for representing the observations across the study domain.

2.4.1 Annual-maximum models

The four model formulations, M1–M4, were compared using the deviance information criterion (DIC) and the widely applicable information criterion (WAIC). DIC combines a measure of goodness of fit, based on the posterior mean deviance, with a penalty for model complexity represented by the effective number of parameters (Spiegelhalter et al., 2014). WAIC is based on the Bayesian predictive distribution and estimates the Bayes generalization loss while accounting for posterior variability; it is asymptotically equivalent to Bayesian leave-one-out cross-validation (Watanabe and Opper, 2010). Lower values of both criteria indicate better-supported models.

After model comparison, predictive performance was evaluated at both calibration and held-out sites using reliability, precision, and probabilistic accuracy metrics. Reliability was assessed using probability integral transform (PIT) values obtained by evaluating the posterior predictive distribution at each observed annual maximum. For a calibrated predictive distribution, the PIT values are expected to follow a uniform distribution. Deviations from uniformity were summarized using the α-index (Renard et al., 2010). Precision was assessed using two complementary metrics: the differential entropy of the fitted GEV distributions (reported in natural units, or nats) and the posterior coefficient of variation (CV) of the 100-year return level ($R_{0.99}$) (Cover, 1999). Conditional GEV entropy was calculated for each posterior parameter draw and averaged across draws to quantify the spread of the fitted at-site GEV distributions. As differential entropy measures the absolute spread of the fitted GEV predictive distribution, the CV of posterior return-level samples were introduced to measure relative uncertainty in high-return-period tail estimates. Utilizing both metrics allows precision to be evaluated in terms of both distributional spread and relative uncertainty in design-relevant return levels. Overall probabilistic accuracy was evaluated using the continuous ranked probability skill score (CRPSS), which evaluates the CRPS of each model against a reference forecast (Matheson and Winkler, 1976) defined using a leave-one-out empirical climatology at each site. The leave-one-out empirical climatology is an

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

observation-based reference distribution: for each target year it is formed from the observed annual maxima at that site with the target year withheld, and therefore represents the predictive skill obtainable from the historical record alone rather than from any fitted model.

2.4.2 Occurrence models

The time-invariant baseline model, N1, was compared with candidate climate-conditioned formulations containing different combinations of AMO, MEI, NAO, and PDO. Candidate models were evaluated using the stepwise DIC-based procedure described in Sect. 2.4.1 (Ntzoufras, 2011). The number of SCIs was increased incrementally, and all candidate combinations were evaluated at each model size. The combination with the lowest DIC was retained, and the selection procedure was terminated when adding another SCI did not reduce DIC relative to the best model from the preceding step. The climate-conditioned model selected through this procedure was denoted N2. Because the responses in the occurrence models are binary, N1 and N2 were evaluated using metrics appropriate for probabilistic binary prediction. The Brier score was used to assess the accuracy of predicted exceedance probabilities, with lower values indicating better probabilistic performance (Brier, 1950). In addition, the area under the receiver operating characteristic curve (AUC) was used to evaluate how well each model ranked exceedance events above non-exceedance events. Together, these metrics were used to demonstrate the added value of incorporating SCIs in N2 for representing the probabilistic occurrence of extreme events relative to the time-invariant N1 baseline.

2.5 Dataset generation

The final objective of this study is to generate PreXFOCUS, a spatially gridded precipitation-frequency dataset across CONUS, using the established model formulations that showed the best overall performance in representing annual maximum precipitation across the study domain. To assess whether return levels obtained through conditional spatial simulation were consistent with those estimated by directly fitting the model to local observations, the held-out sites were randomly divided into two approximately equal validation subsets. For each set, 10- and 100-year return levels were obtained in two ways: (i) conditional spatial simulation from the calibration-site posterior, and (ii) direct refitting of the selected formulation to the validation-set observations. Statistical agreement between the resulting

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

posterior mean return-level estimates was quantified using NSE and percentage bias metrics (Nash and Sutcliffe, 1970). To evaluate the selected annual-maximum model against an existing precipitation-frequency product, its posterior-mean 2-, 5-, and 10-year return levels were compared with NOAA Atlas 14 estimates using empirical quantiles of the observed annual maxima as a reference, with performance quantified using mean absolute error (MAE). Because the observational record spans 70 years, the comparison was limited to these return periods. Following this evaluation, posterior draws of the selected model's GEV parameter fields were conditionally simulated on a 0.1° grid across the CONUS, and the resulting gridded parameter estimates were used to calculate GEV-based return levels for the 2-, 5-, 10-, 25-, 50-, and 100-year return periods at each grid cell.

To generate products accounting for climate variability, the N2 model for each return-period threshold was used to estimate annual exceedance probabilities at each grid cell and year. For an r-year threshold $q_r^{ref}(s)$, $P_r(s,t)$ denotes the probability that annual maximum precipitation at location s exceeds the threshold under the climate conditions of year t. The effective return period of the threshold was calculated as

$$T_{\mathrm{eff},r}(s,t) = \frac{1}{P_r(s,t)}, \quad (3)$$

When $P_r(s,t) > 1/r$, the threshold is exceeded more frequently than its nominal rate and $T_{eff}(s,t) < r$; therefore, a larger precipitation magnitude is required to maintain an r-year return period. Conversely, when $P_r(s,t) < 1/r$, the threshold is exceeded less frequently, $T_{eff}(s,t) > r$, and the corresponding r-year return level is lower than the M4 reference estimate. Climate-conditioned return levels were then obtained by inverting the exceedance-probability curve.

At each grid cell and year, the effective return periods were paired with their corresponding reference return-level magnitudes to construct a local return-level curve. Hydrologic frequency curves are commonly represented with return period on a logarithmic scale (England Jr et al., 2018; Gumbel, 1941; Miniussi and Marra, 2021). Accordingly, monotonic PCHIP interpolation was performed in log-return-period space. As an exploratory analysis, the suitability of this interpolation was additionally assessed using empirical frequency curves derived directly from the observed annual maxima, where interpolation between selected empirical return-period points was evaluated against the remaining

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

empirical quantiles within the interpolation range. This procedure identifies the magnitude whose exceedance probability corresponds to the target return period under the climate conditions of year t.

3 Results and Discussion

The reliability and reusability of the PreXFOCUS dataset were evaluated through comprehensive validation and performance analyses. First, four annual-maximum model formulations were assessed using their held-out observed precipitation extremes and associated predictive uncertainty. Second, the spatial predictive performance of the models was evaluated at held-out gauges using conditional posterior simulation. Third, the Bernoulli occurrence model was used to evaluate whether incorporating standard climate indices in the model improved the return-level exceedance probabilities. Finally, the selected annual-maximum and occurrence models were combined to generate and evaluate the resulting reference and climate-conditioned gridded return-level estimates across the CONUS. The results of these evaluations are presented and discussed in the following sections, along with the dataset's operational strengths and inherent technical limitations.

3.1 Statistical evaluation of spatial model formulations

The four annual-maximum model formulations (M1–M4) were fitted to the calibration-gauge records and evaluated using goodness-of-fit criteria. MCMC sampling converged for all four formulations, and the convergence diagnostics defined in Sect. 2.3 were satisfied in every case, so the posterior samples were considered reliable for the model comparison presented below. To address the risk of unnecessary model complexity introduced by highly parameterized formulations, DIC was used to evaluate the optimal trade-off between fit and effective complexity, where lower values indicate better performance (Spiegelhalter et al., 2002). WAIC was also computed as a supplementary diagnostic. The DIC and WAIC values for M1–M4 are reported in Table 1.

Table 1. DIC and WAIC values for models (M1–M4) fitted to the annual-maximum precipitation records at the calibration sites.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

Model	DIC	WAIC
M1	306151.69	306201.16
M2	305963.78	305983.84
M3	305902.06	305950.12
M4	304087.09	304104.72

M1, which estimates site-specific GEV parameters without spatial pooling, produced the highest DIC and therefore the weakest fit among the four models. Introducing spatial pooling reduced DIC in M2 and M3, with M3 showing better performance by allowing the GEV shape parameter to vary spatially. M4 achieved the minimum DIC among all formulations, indicating that representing data-layer spatial dependence through a Gaussian copula improved the model fit.

While M4 provided the best overall goodness-of-fit, predictive uncertainty and performance of all formulations were further evaluated using several diagnostic metrics. Reliability was assessed using the α-index, which quantifies deviation from the 1:1 line in the predictive Q–Q plot. At calibration sites, posterior predictive samples from each model were compared with the observed annual maxima. At held-out sites, posterior predictive samples from M2–M4 were generated through conditional spatial simulation and compared with the corresponding observations. As M1 lacks a spatial process layer and cannot predict at unobserved sites, it was fitted directly to the held-out records to provide an at-site benchmark. For M4, the held-out posterior predictive samples were additionally conditioned on the calibration-site annual maxima through the Gaussian copula structure described in Sect. 2.3.

To evaluate distributional reliability, α-index values were calculated by pooling predictive probabilities across all site–year combinations, separately for the calibration and held-out sites. Values closer to 1 indicate greater reliability. All configurations exhibited strong predictive reliability, yielding α-index values consistently near the ideal value of 1. At the calibration sites, M3 produced the highest value (0.995), followed closely by M2 and M4, while M1 showed the lowest value (0.982). At the held-out

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

sites, M2 achieved the highest value (0.982), followed by M3 (0.981) and M4 (0.976). Cumulatively, α-index demonstrated that all four formulations produced reliable posterior predictive quantiles.

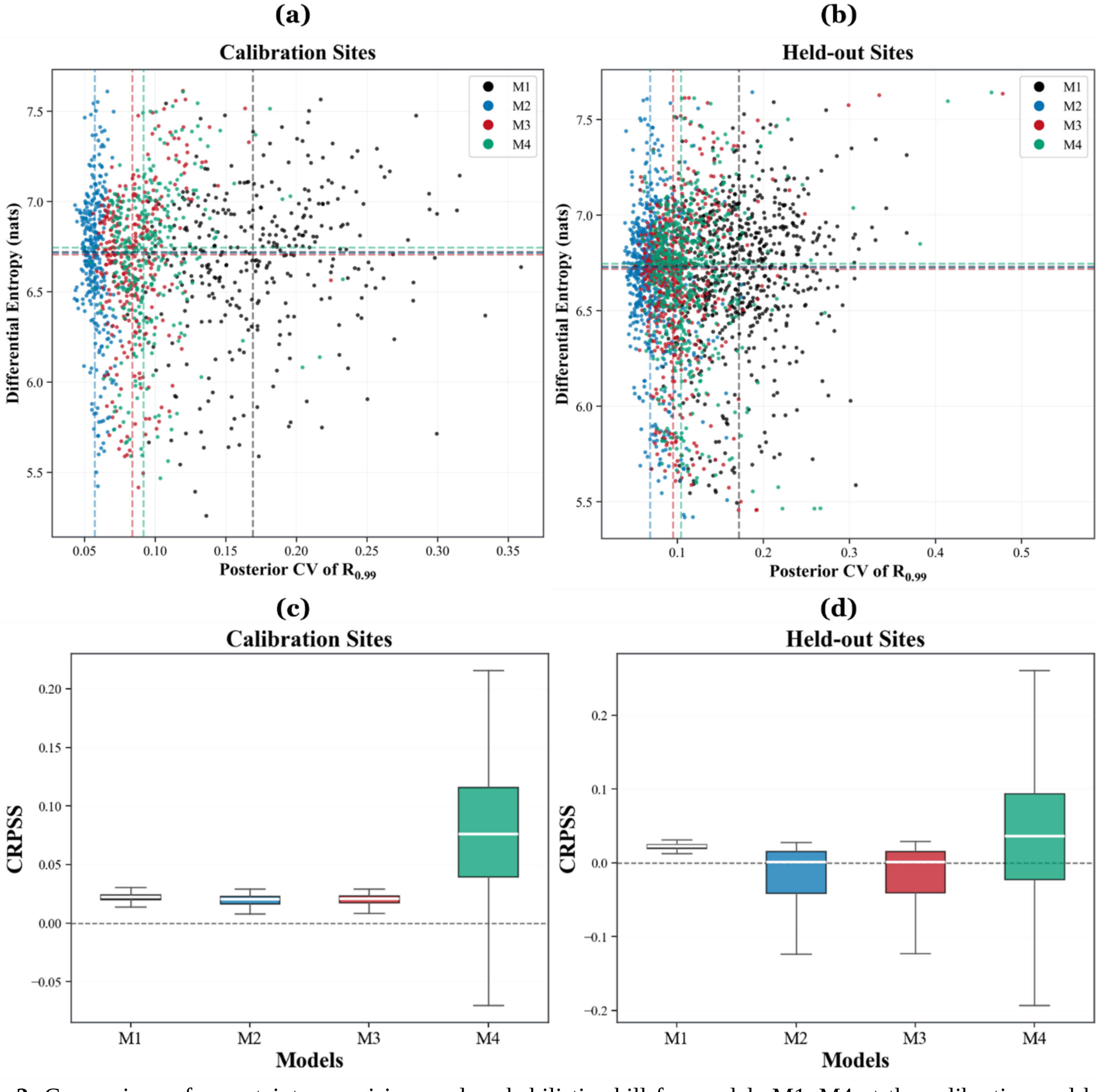


Figure 2. Comparison of uncertainty, precision, and probabilistic skill for models M1–M4 at the calibration and held-out sites. Panels (a) and (b) show differential entropy versus the posterior coefficient of variation of $\boldsymbol{R_{0.99}}$, for the calibration and held-out sites, respectively; dashed lines indicate the median of each metric across sites. Panels (c) and (d) show the distributions of CRPSS values relative to the leave-one-out empirical climatology for the calibration and held-out sites, respectively.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Further, the model formulations were compared in terms of precision and overall probabilistic accuracy using differential entropy, the posterior CV of $R_{0.99}$, and CRPSS. Figure 2 summarizes these results for both calibration and held-out sites. Differential entropy was similar across M1-M4 at both calibration and held-out sites, although M4 exhibited marginally higher values overall. In contrast, the posterior CV of $R_{0.99}$ showed more pronounced differences among the models. M1 produced the largest CV values, indicating higher relative uncertainty in high-return-period estimates when no spatial pooling is used. Models M2 and M3 produced narrower posterior CVs than M1, reflecting the effect of spatial pooling in the latent process of the GEV parameters on return-level estimation uncertainty. Overall, the results suggest that omitting spatial pooling among GEV parameters tends to overestimate predictive uncertainty, whereas ignoring data-layer spatial dependence among annual maxima may result in underestimation.

Figure 2 (c) and (d) show the CRPSS results across calibration and held-out sites evaluating predictive skill of models relative to the leave-one-out empirical climatology at each site. Across calibration sites, M4 produced the highest median CRPSS, indicating improved probabilistic skill relative to the other formulations. M4 outperformed M1 at held-out sites, even though M4's posterior predictive distributions at these sites were generated through conditional simulation rather than direct site-level fitting. However, the wider spread of CRPSS values for M4 at held-out sites suggests that the magnitude of improvement varies spatially, with some locations showing limited or negative skill relative to the empirical climatology. Moreover, sharper uncertainty does not necessarily imply better predictive performance. Although M2 produced the sharpest predictive distributions, its higher DIC and lower CRPSS indicate weaker overall performance compared with M3 and M4.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

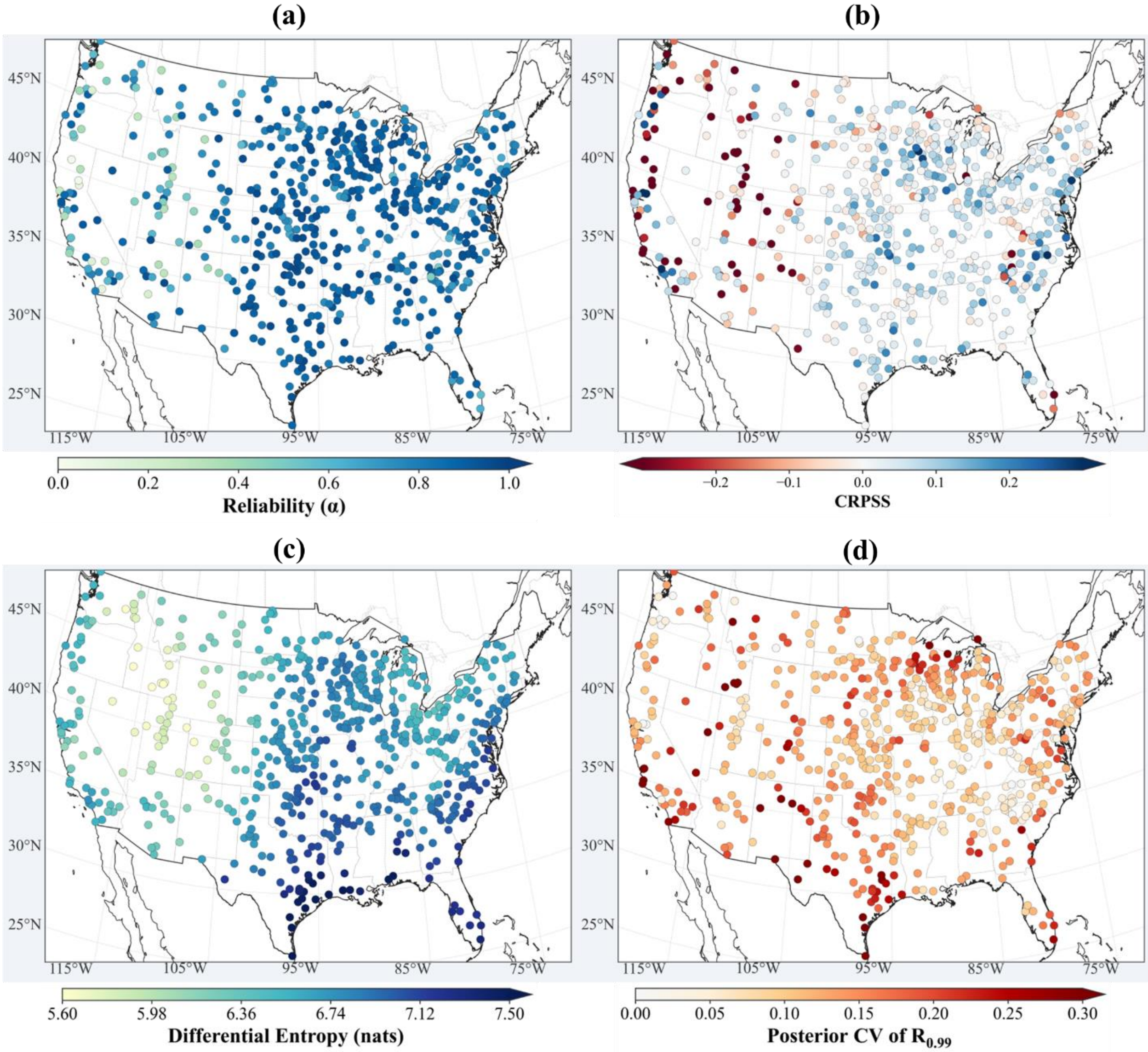


Figure 3. Spatial distribution of M4 probabilistic performance and uncertainty metrics at held-out sites across CONUS: (a) reliability α-index, (b) CRPSS relative to leave-one-out empirical climatology, (c) differential entropy, and (d) posterior coefficient of variation for $\mathbf{R_{0.99}}$. Metrics were computed from posterior predictive samples generated through conditional spatial simulation.

Figure 3 shows the spatial variation of probabilistic metrics for model M4 across held-out sites over CONUS. The α-index values in Fig. 3a were computed separately at each held-out site using its full record of predictive probabilities and therefore show greater variability than the domain-wide values reported earlier. Reliability was generally higher across the eastern U.S. and lower across parts of the

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Rocky Mountains and the western U.S. For example, the median α-index was 0.88 for held-out sites east of 100°W, compared with 0.77 for sites west of 100°W.

A similar spatial pattern was observed for CRPSS, with values as low as -0.2 at some western sites and exceeding 0.2 at several eastern sites relative to the leave-one-out empirical climatology. Figure 3c shows a spatially coherent pattern in differential entropy, with values near 5.6 nats across higher-elevation regions and increasing toward the Gulf Coast, where they reach approximately 7.5 nats. In contrast, the posterior CV of the 100-year return level in Fig. 3d is more spatially heterogeneous and exhibits no clear relationship with elevation. This difference arises because return-level CV reflects the joint posterior uncertainty in all three GEV parameters, whereas GEV differential entropy depends only on the scale and shape parameters.

The calibration network is denser in the eastern CONUS (Fig. 1), which provides more nearby information for conditional spatial simulation than in sparsely sampled western regions. In contrast, the sparser gauge network across the western CONUS, particularly in mountainous regions, limits the information available for estimating held-out-site posterior distributions. The lower differential entropy observed in parts of the Rocky Mountains, together with reduced reliability and CRPSS, suggests that predictive uncertainty may be underestimated in some of these regions. In contrast, higher differential entropy along the Gulf Coast indicates wider predictive distributions, consistent with greater uncertainty in regions affected by intense storm systems (Knight and Davis, 2009). The posterior CV of $R_{0.99}$ was also higher in areas with sparse nearby calibration sites and in parts of the Gulf Coast, indicating larger relative uncertainty in high-return-period estimates. These results show that although M4 improves overall spatial prediction skill, its performance varies regionally.

Although M4 provided the strongest overall performance among the evaluated formulations, its spatial predictions are optimized for specific operational scopes. The Gaussian copula used in M4 represents dependence in the main body of the distribution but has no upper-tail dependence, implying asymptotic independence among extreme precipitation values across sites (Beck et al., 2020; Bracken et al., 2016). In addition, the stationary exponential dependence structure imposes a common spatial decay function across the CONUS, despite documented spatial variation in the local dependence and tail-dependence structures of U.S. precipitation extremes(Castro-Camilo and Huser, 2020). The resulting products

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

should therefore be interpreted with caution in applications that depend specifically on joint extremes at multiple locations. A major limitation of M4 is the restriction to strictly positive GEV shape parameters. In regions where the true shape parameter is near zero or negative, this constraint can impose an excessively heavy upper tail and may overestimate high-return-period precipitation (Papalexiou et al., 2018; Papalexiou and Koutsoyiannis, 2013). Future work should investigate alternative copula families and max-stable processes to better characterize spatial and tail dependence among extremes (Reich and Shaby, 2012; Renard, 2011).

3.2 Out-of-sample validation and comparison with NOAA Atlas 14

The conditional spatial simulations were evaluated using the two non-overlapping subsets of held-out gauges described in Sect. 2.5. The conditional spatial simulations for held-out sites were compared with the direct-fit return-level estimates obtained from M4 using observations at those sites. The direct-fit results were treated as comparison estimates to evaluate spatial transferability rather than treating them as error-free reference values to validate the GEV model. Figure 4 shows the percentage error between the conditional-simulation and direct-fit reference estimates for the mean 10- and 100-year return levels, $R_{0.90}$ and $R_{0.99}$, across the two validation subsets. The percentage error was calculated by normalizing each site-level estimation error by the mean direct-fit reference estimate across all sites. This common denominator facilitated consistent comparisons among regions and prevented small site-level reference values from producing disproportionately large percentage errors. Overall, the conditional simulation estimates showed strong alignment with the direct-fit estimates, with NSE ranging from 0.74 to 0.85 across the four validation cases. Higher NSE values indicate agreement between the two estimations, suggesting that the conditional spatial simulation procedure can reproduce held-out-site return-level estimates.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

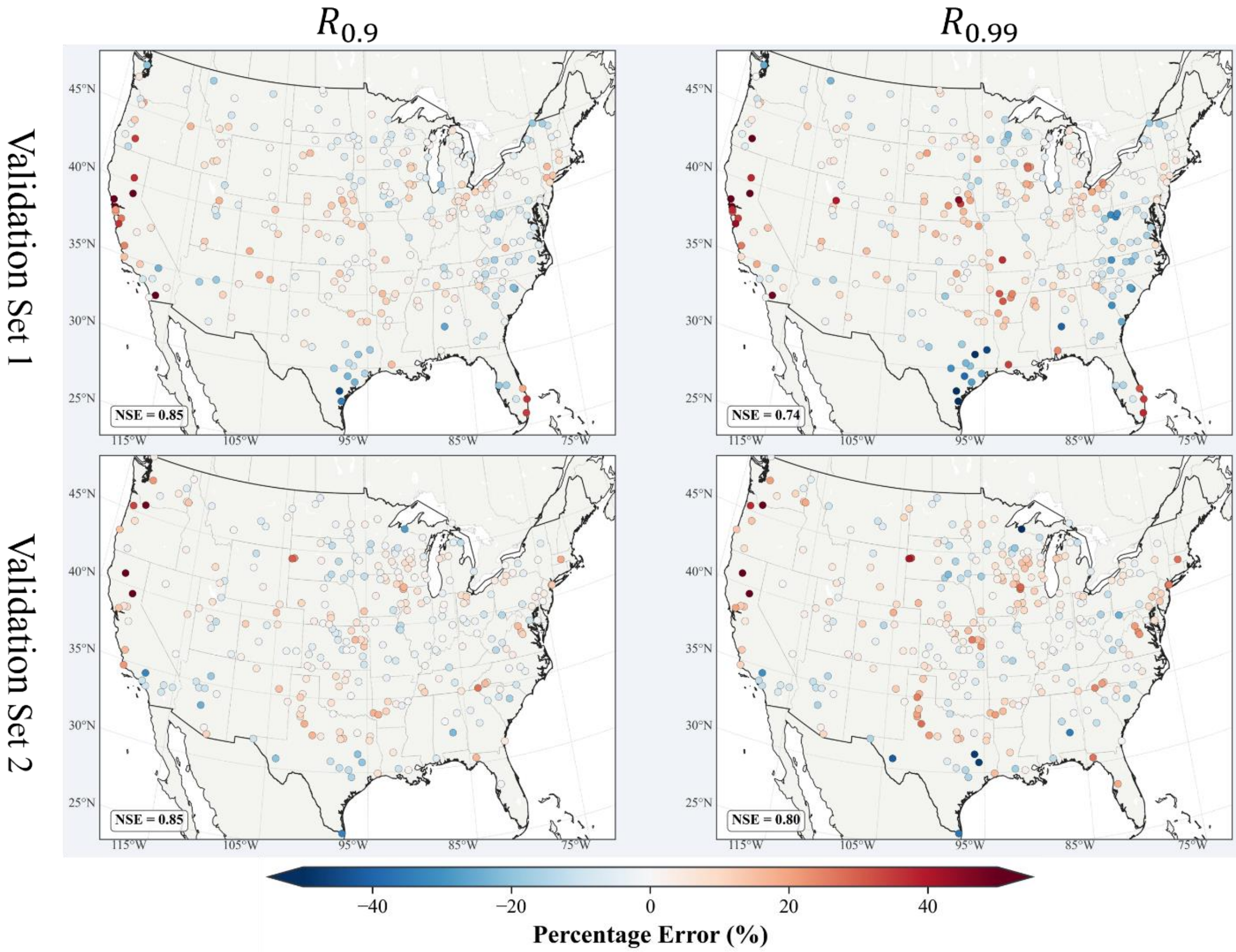


Figure 4. Maps show percentage error between return-level estimates obtained from conditional spatial simulation and reference estimates obtained by fitting M4 directly to the annual maxima at held-out validation subsets. Results are shown for the mean 10-year and 100-year return levels, $\mathbf{R_{0.90}}$ and $\boldsymbol{R}_{\mathbf{0.99}}$, across two validation subsets.

The percentage errors were also generally larger for $R_{0.99}$ than for $R_{0.90}$, indicating that predictive accuracy decreases for less frequent events. This is expected because high-return-period estimates are more sensitive to uncertainty in the GEV scale and shape parameters, whereas lower-return-period estimates are more strongly controlled by the location parameter. Therefore, small differences in estimated scale and shape can lead to larger discrepancies in $R_{0.99}$. Figure 4 shows that errors for the 100-year return level were larger and more spatially variable than those for the 10-year return level, indicating greater sensitivity of long-return-period estimates to uncertainty in the GEV scale and shape parameters. The maps show larger positive errors concentrated along parts of the

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Pacific Coast. In the northern Pacific Coast region, comparatively sparse calibration-gauge coverage and strong orographic gradients in extreme precipitation limit the local information available for spatial prediction, causing the conditionally simulated fields to be smoothed toward broader regional behaviour and resulting in overestimation at some sites. Overestimation at northern California sites and underestimation in southern California and the southern United States, particularly Texas at longer return periods, highlight regional heterogeneity in performance under a single continental-scale spatial model.

To further evaluate the performance of M4 in estimating time-invariant return levels relative to an existing precipitation-frequency dataset, posterior-mean return levels for the 2-, 5-, and 10-year return periods were compared with NOAA Atlas 14 estimates using empirical quantiles of the observed annual maxima as a reference. Figure 5 compares M4 and NOAA Atlas 14 return-level estimates against the empirical quantiles at both calibration and held-out sites, and their agreement was evaluated using MAE.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

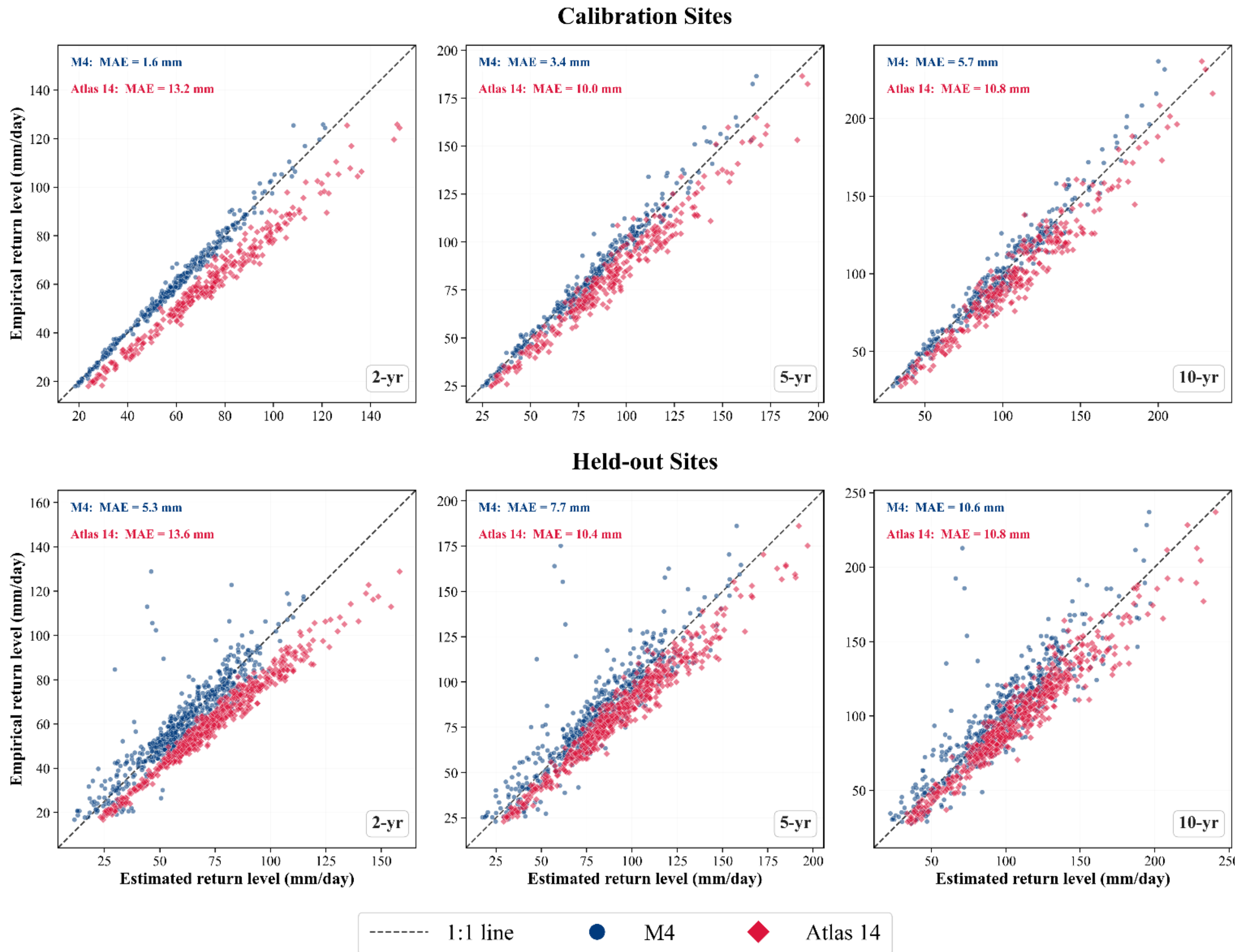


Figure 5. Comparison of the 2-, 5-, and 10-year return-level estimates from M4 and NOAA Atlas 14 against empirical quantiles of the observed annual maxima at calibration sites (top row) and held-out sites (bottom row). M4 values represent posterior-mean return-level estimates; at held-out sites, these were obtained through conditional simulation of the spatial parameter fields. NOAA Atlas 14 estimates, derived using regional frequency analysis based on L-moments, are included as an external benchmark. The dashed line indicates perfect agreement with the empirical quantiles, and the mean absolute error (MAE) for M4 and Atlas 14 is reported in each panel.

Both NOAA Atlas 14 and M4 showed generally good agreement with the empirical quantiles, while M4 provided closer agreement across the evaluated return periods. At the calibration sites, the MAE of M4 was reduced by 87.8%, 66.0%, and 47.1% relative to NOAA Atlas 14 for the 2-, 5-, and 10-year return

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

periods, respectively. At the held-out sites, the corresponding reductions were approximately 61%, 26%, and 2%. The performance at the held-out sites further demonstrates the ability of the spatially pooled M4 formulation to estimate return levels at unobserved locations, as these estimates were obtained through conditional simulation without using observations from the held-out sites for model fitting. The MAE of M4 increased with return period at both calibration and held-out sites, consistent with the increasing influence of sampling variability on upper-tail estimation from finite observational records. In addition to posterior-mean return levels, M4 provides posterior distributions that characterize uncertainty in the estimated precipitation frequencies. Overall, the comparison indicates that M4 provides more accurate time-invariant precipitation-frequency estimates than NOAA Atlas 14 for most of the evaluated cases, while maintaining the ability to generate probabilistic estimates at locations without observations. While these estimates characterize the time-invariant precipitation-frequency distribution, the occurrence probability of extreme events can vary from year to year in response to large-scale climate variability. The following section, therefore, evaluates the climate-conditioned occurrence model used to represent this temporal variability.

3.3 Climate-conditioned occurrence model evaluation

Bernoulli occurrence models were constructed using return-level thresholds from M4, the best-performing annual-maximum formulation. The SCI combination was then determined using the stepwise DIC-based selection procedure described in Sect. 2.4. Table 2 summarizes the DIC values for the candidate occurrence-model structures.

Table 2. DIC values for candidate Bernoulli occurrence models evaluated during the stepwise climate-index selection procedure.

AMO	NAO	MEI	PDO	DIC	Number of SCIs
✗	✗	✗	✗	**14394.03**	0
✓	✗	✗	✗	14363.73	
✗	✓	✗	✗	14359.00	

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

✗	✗	✓	✗	**14355.99**	1
✗	✗	✗	✓	14371.25	
✓	✗	✓	✗	14328.44	
✗	✓	✓	✗	**14316.99**	2
✗	✗	✓	✓	14355.30	
✓	✓	✓	✗	**14301.80**	
✗	✓	✓	✓	14319.27	3
✓	✓	✓	✓	14302.77	4

The DIC-based selection was conducted using exceedances of the site-specific posterior-mean 10-year return-level threshold, $R_{0.90}$, estimated from M4. As shown in Table 2, including MEI, NAO, and AMO in N2 produced the largest reduction in DIC relative to the baseline model N1 among the candidate index combinations. Including MEI in N1 resulted in the greatest improvement in model fit among the single-index models, while adding NAO to the MEI-based model provided a further improvement and outperformed the other two-index specifications. Adding AMO further improved model fit, whereas adding PDO provided no additional improvement. The selected N2 structure was retained for the remaining thresholds to maintain a consistent model specification across the released products. To further evaluate whether this improvement translated into better probabilistic classification, N1 and N2 were compared using AUC and Brier score across multiple exceedance thresholds at calibration and held-out sites.

Table 3. Comparison of the baseline occurrence model N1 and the selected climate-informed occurrence model N2 using AUC and Brier score across return-level exceedance thresholds at the calibration and held-out sites.

Sites	**AUC**		**Brier Score**		**Threshold**
	N1	N2	N1	N2	

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Calibration	0.530	0.617	0.0861	0.0849	
Held out	0.542	0.550	0.1034	0.1030	$R_{0.9}$
Calibration	0.545	0.618	0.0338	0.0336	
Held out	0.557	0.572	0.0490	0.0489	$R_{0.96}$
Calibration	0.560	0.630	0.0164	0.0164	
Held out	0.617	0.618	0.0283	0.0282	$R_{0.98}$

Table 3 shows that N2 consistently improved AUC relative to N1 across all evaluated thresholds, with gains ranging from 12% to 16% at the calibration sites. In contrast, the Brier score improvements were relatively small. Improvements were also observed at the held-out sites, although the magnitude of improvement was smaller than at the calibration sites. Improvements in AUC and Brier scores indicate that including SCIs enhanced the model's ability to distinguish between exceedance and non-exceedance years. The moderate AUC values likely reflect the low frequency of return-level exceedances and the limited number of positive events available in the 70-year observational record, particularly at higher thresholds. This data limitation increases uncertainty in estimating climate–exceedance relationships. Nevertheless, the consistently higher AUC values for N2 indicate that incorporating climate indices improved the ranking of exceedance years relative to non-exceedance years, although the improvement remained modest.

The occurrence models used posterior mean M4 return levels as fixed thresholds. Hence, uncertainty in the GEV parameters and M4 return-level estimates was not propagated into the estimated climate-conditioned exceedance probabilities, effective return periods, or climate-conditioned return levels. The uncertainty associated with these products is therefore conditional on the selected M4 thresholds. A joint hierarchical formulation could propagate uncertainty across both modeling stages, although it would require substantially greater computational resources.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

To assess the spatial influence of each climate index on extreme-precipitation occurrence, conditional spatial simulation was used to estimate exceedance probabilities at held-out sites. For each return-level threshold, baseline probabilities were computed with all standardized SCIs set to zero. Each SCI was then increased individually to +1, representing a one-standard-deviation increase, while the others remained at zero. Figure 6 shows the resulting percentage change in exceedance probability relative to the baseline.

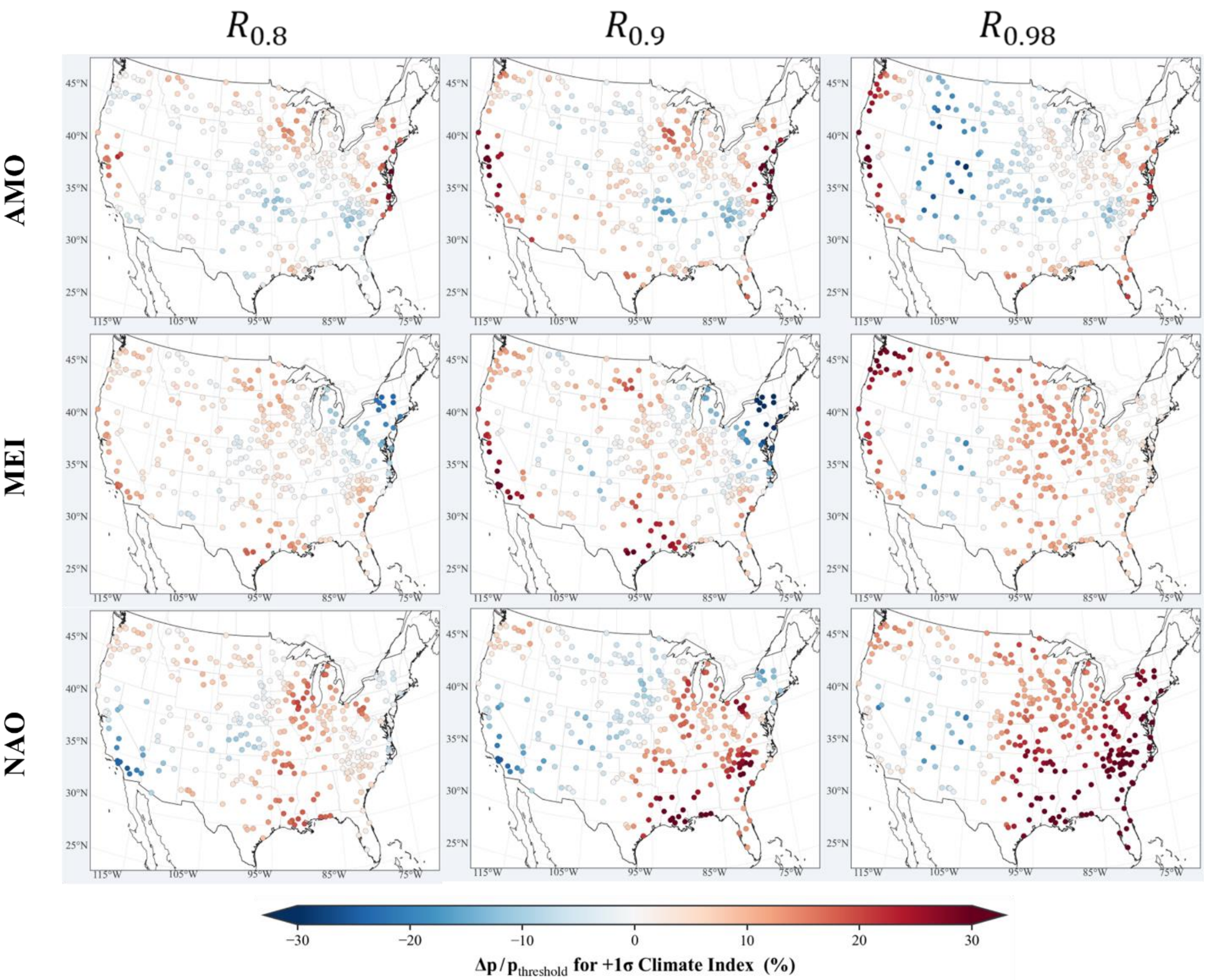


Figure 6. Relative change in exceedance probability associated with one-standard-deviation increase in an individual climate index. Rows correspond to AMO, MEI, and NAO, while columns correspond to the 5-, 10-, and 50-year reference thresholds. At each held-out gauge, the indicated climate index was increased from its mean to +1 standard deviation, while all other standardized indices were held at zero. The resulting posterior mean exceedance probability was compared with the baseline probability obtained when all standardized climate indices were set to zero. Positive (red) and negative (blue) values indicate higher and lower posterior mean exceedance probabilities, respectively, relative to the baseline.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Figure 6 shows that the relative changes are generally larger for less frequent events, such as $R_{0.98}$, indicating that climate variability has a stronger effect on high-threshold exceedance probabilities than on more frequent events. Positive values indicate that the reference precipitation threshold is more likely to be exceeded, whereas negative values indicate a lower exceedance probability. Modes of climate variability altered the probability of extreme-precipitation occurrence by more than 40% in some regions, with NAO exerting its strongest influence over the eastern U.S., MEI primarily affecting the Pacific Coast region, and AMO showing comparatively weaker but more geographically widespread effects. An additional benefit of integrating N2 with M4, as described in Sect. 2.5, is that the climate-conditioned exceedance probabilities provide regional adjustments to the M4 return-level estimates. In the Pacific Southwest, exceedance probabilities above the nominal values implied by M4 indicate that the reference thresholds occur more frequently, reducing their effective return periods and increasing the corresponding return-level estimates. This adjustment may reduce the underestimation observed in the region. In contrast, lower exceedance probabilities across the Great Basin and Rocky Mountains increase the effective return periods and reduce the corresponding return-level estimates, potentially correcting regional overestimation.

The spatial patterns estimated by N2 were broadly consistent with previous studies, which have shown that El Niño conditions are associated with more frequent extreme precipitation across parts of California, the Southwest, and the southern tier, although the response remains regionally heterogeneous (Bracken et al., 2015; Cayan et al., 1999; Shang et al., 2011). N2 also indicated a pronounced NAO influence across the eastern U.S., consistent with previous evidence that NAO effects on extreme winter precipitation are strongest over eastern North America (Whan and Zwiers, 2017). However, the results presented here are based on statistical inference from the occurrence model. A detailed physical attribution of the regional responses to each climate index is beyond the scope of this study.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

3.4 Evaluation of climate-conditioned precipitation frequency products

Figure 7 presents maps of the effective return periods associated with the 5-, 10-, and 25-year thresholds for four representative years: 1971, which had the most negative MEI value; 1987, which had the most positive MEI value; 2010, characterized by strongly negative MEI and NAO conditions; and 2015, characterized by simultaneously positive MEI and NAO conditions. The N2 occurrence model was used to estimate exceedance probabilities and derive the corresponding effective return periods, with the posterior mean return levels from M4 treated as fixed thresholds. N2 therefore quantifies how the probability of exceeding each M4 threshold varies with climate conditions, whereas the nominal exceedance probability under the time-invariant M4 formulation is approximately $1/r$.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

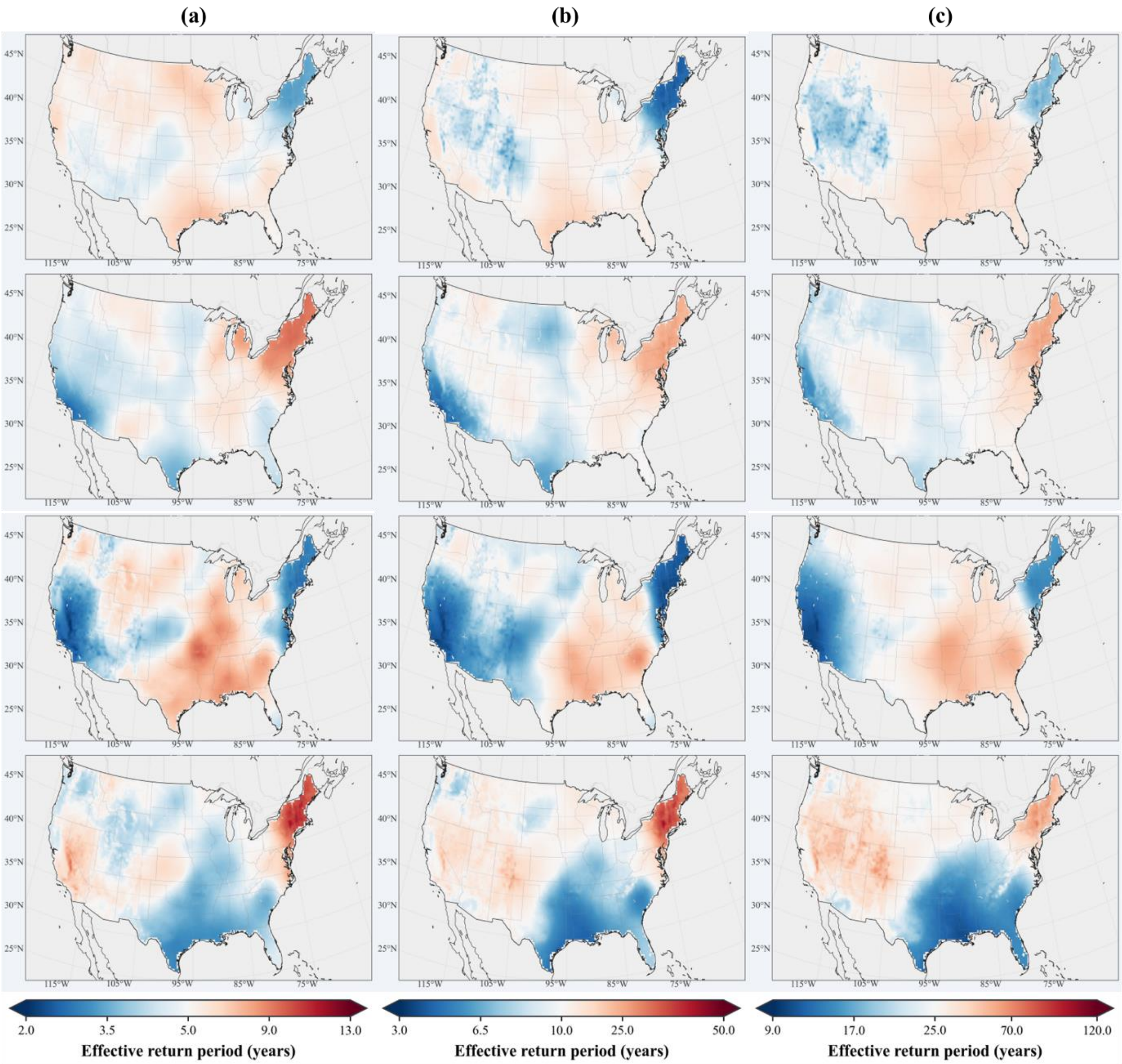


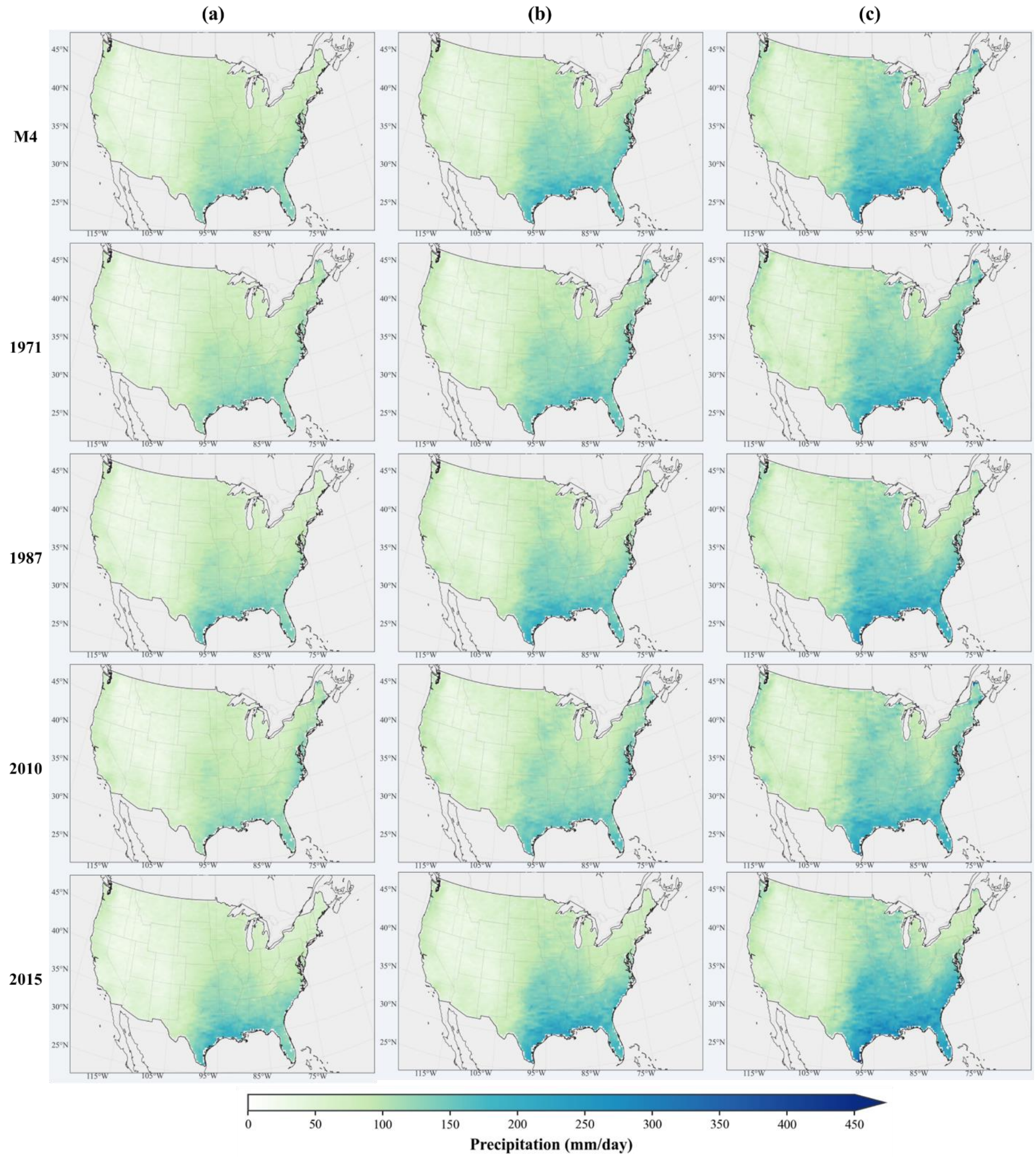
Figure 7. Spatial maps of the effective return periods derived from N2-estimated exceedance probabilities for thresholds corresponding to the (a) 5-year, (b) 10-year, and (c) 25-year return levels from M4. From top to bottom, the rows represent 1971, 1987, 2010, and 2015.

Figure 7 illustrates how climate-conditioned effective return periods vary by year, region, and threshold. An effective return period shorter than the baseline indicates more frequent exceedance of the corresponding precipitation threshold, whereas a longer effective return period indicates less frequent exceedance. For example, a precipitation magnitude corresponding to the baseline 5-year return level had an effective return period of approximately 2 years in the Southwest during 2010, compared with

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

about 13 years in the Northeast during 2015. Threshold dependence was also evident in 1987, when reduced effective return periods extended from the Southwest into the central United States at the 5-year threshold but became more concentrated in the Southwest at the 10- and 25-year thresholds. The positive MEI and NAO phases in 2015 were associated with pronounced changes across the southeastern CONUS. In 2010, despite the negative MEI phase, the occurrence model estimated pronounced changes in the Pacific Southwest, primarily associated with the strongly negative NAO phase. These spatial contrasts and the comparatively weak response in the upper Midwest are consistent with documented regional differences in the influence of climate variability on precipitation extremes (Glantz, 2001; Nizamani et al., 2025; Yu et al., 2017).

The estimated exceedance probabilities for each grid cell, year, and threshold were used to derive climate-conditioned return-level estimates. Figure 8 compares the M4 reference return levels with the corresponding climate-conditioned estimates for 1971, 1987, 2010, and 2015 at the 5-, 10-, and 25-year return periods.



590 **Figure 8.** Spatial maps of precipitation return levels for the 5-, 10-, and 25-year return periods, shown in columns (a)–(c), respectively. The top row presents the reference return-level estimates from M4, whereas the subsequent rows show the corresponding estimates accounting for climate variability in 1971, 1987, 2010, and 2015.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

The climate-conditioned maps show that precipitation return levels vary among years in response to different modes of climate variability. For example, the estimated return levels across the eastern and southeastern U.S. in 2015 were higher than the M4 estimates, consistent with the concurrent positive phases of NAO and MEI. In contrast, the 2010 estimates were generally lower than those from M4 under the concurrent negative phases of these two indices. Across return periods, the lowest estimates were primarily located over the interior western U.S., across the Great Basin and the inland Pacific Northwest, whereas the highest values occurred along the Gulf Coast, extending from coastal Texas and southern Louisiana to coastal Alabama. These high Gulf Coast return levels are consistent with the influence of tropical cyclones and their associated intense precipitation, as exemplified by Hurricane Harvey in 2017 (van Oldenborgh et al., 2017). The location of the maximum return level also shifted with return period. For longer return periods, the maxima were concentrated along the western Gulf Coast, particularly in Texas, whereas for shorter return periods they occurred farther east, especially in Alabama. This pattern indicates that incorporating N2 increased the M4 return-level estimates in regions where M4 showed greater underestimation, particularly at longer return periods.

In contrast, the southern Rocky Mountains exhibited persistently lower annual extreme-precipitation return levels than most other regions of CONUS (Guirguis and Avissar, 2008; Rupp et al., 2022), partly because western mountain ranges, including the Sierra Nevada and Cascades, intercept Pacific moisture transport associated with atmospheric rivers and limit its penetration into the interior Southwest (Rutz et al., 2014).

As climate-conditioned return levels were obtained by interpolating exceedance curves, the estimates should be interpreted within the supplied return-period range together with the accompanying posterior uncertainty and spatial-support fields. Reuse of the dataset should generally remain within the supplied return-period range unless additional validation is performed. Future extensions could consider alternative copula families, regionally varying dependence structures, joint estimation of the annual-maximum and occurrence components, and seasonally resolved climate predictors.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

3.5 Potential applications and community use

PreXFOCUS provides 0.1° gridded reference and climate-conditioned precipitation return-level estimates across the CONUS. Precipitation-frequency information has long supported engineering and water-management applications; NOAA Atlas 14, for example, provides design precipitation estimates used for stormwater, drainage, culvert, transportation, and related hydrologic applications (Bonnin et al., 2006; Perica et al., 2013, 2018). Such information can also support flood-control infrastructure design when coupled with hydrologic models (Schlef et al., 2023). The gridded products extend precipitation-frequency information beyond gauge locations, supporting applications in data-sparse and ungauged areas (Lu et al., 2025; Schlef et al., 2023). As extreme precipitation frequencies change, updated precipitation-frequency information is increasingly relevant for evaluating infrastructure designed using historical conditions (Martel et al., 2021; Sharma et al., 2021; Wright et al., 2019).

A distinguishing feature of PreXFOCUS is the provision of climate-conditioned return-level estimates that reflect variations in extreme-precipitation frequency associated with large-scale climate conditions. Climate modes such as ENSO have been shown to modify the frequency of precipitation extremes across the United States (Cayan et al., 1999; Huang and Stevenson, 2023). Therefore, the climate-conditioned products show how precipitation magnitudes associated with selected return periods vary across regions and years under prevailing climate conditions. Such information can complement conventional precipitation-frequency estimates in applications where temporal variations in extreme-event probability are relevant, including flood-risk assessment (François et al., 2019).

4 Usage notes and limitations

When using PreXFOCUS, users should consider the following practical recommendations and limitations. The GeoTIFF products use the WGS 1984 geographic coordinate system at a 0.1° resolution and can be accessed using raster-compatible GIS software, such as ArcGIS, or Python libraries such as Rasterio and GDAL. Users should select one of the six reference return periods (2, 5, 10, 25, 50, or 100 years) according to their application. The reference products provide one estimate per return period, whereas the climate-conditioned products provide annual estimates for 1951–2020. The products

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

include posterior mean estimates and the 5th and 95th posterior percentiles, representing the lower and upper bounds of the 90% credible interval. For engineering applications, the estimates should be considered alongside applicable design standards, regional validation results, and available site-specific information.

The return-level products were generated using a continental-scale spatial model and should be interpreted alongside regional validation results. Additional caution is recommended for California and the southern U.S., particularly Texas and longer return periods, where atmospheric rivers, ENSO variability, Gulf moisture, tropical cyclones, and mesoscale convection may require regionally adapted modeling (Gershunov et al., 2019; Huang and Stevenson, 2023; Schumacher and Johnson, 2005; Zhu and Quiring, 2013). The climate-conditioned products are based on annual mean values of the three selected climate indices: MEI, NAO, and AMO. Annual averaging may weaken seasonal climate signals, while correlations among indices may limit interpretation of their individual effects (Lee et al., 2023; Westra et al., 2015; Whan and Zwiers, 2017). Applications requiring seasonal resolution or additional climate information may therefore benefit from alternative predictors derived from broader atmospheric and oceanic fields, although such predictors may have lower physical interpretability (Renard et al., 2022; Renard and Thyer, 2019).

5 Data availability

The PreXFOCUS dataset is archived on Zenodo and is publicly available at

https://doi.org/10.5281/zenodo.21979711 (Takallou et al., 2026).

6 Conclusions

This study developed PreXFOCUS, a 0.1° gridded precipitation-frequency dataset providing reference and climate-conditioned estimates across the CONUS. The selected annual-maximum formulation, which spatially pooled all GEV parameters and represented data-level dependence using a Gaussian copula, outperformed formulations with no or partial spatial pooling at the latent-process level and those assuming conditional independence or independence at the data level, based on goodness-of-fit and

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

probabilistic-skill metrics, including the CRPSS. It maintained reliable posterior predictive quantiles and produced lower uncertainty than the no-pooling formulation. However, it was less sharp and more uncertain than the conditional-independence formulations, suggesting that omitting data-layer dependence may overstate predictive precision.

The spatial predictive capability of the selected model was evaluated by comparing return levels obtained through conditional spatial simulation with estimates from direct model fitting at held-out sites. The two sets of estimates showed good overall consistency, although the conditional simulations underestimated return levels along the western Gulf Coast and in the Pacific Southwest. Following validation, the selected model was used to generate return-level maps on a 0.1° grid across CONUS. The reference return levels were used to define exceedance indicators for the occurrence models. The selected climate-conditioned formulation included MEI, NAO, and AMO and produced more accurate exceedance-probability estimates than the intercept-only baseline.

The estimated climate relationships varied regionally. Positive MEI phases were associated with higher exceedance probabilities across the Pacific Southwest and southern U.S., whereas positive NAO phases were associated with increases primarily across the eastern and southern U.S. Their combined influence was strongest across the southern and southeastern regions. The occurrence model also increased return-level estimates along the Texas Gulf Coast, where M4 tended to underestimate longer-period return levels, and reduced estimates across parts of the Rocky Mountains, where overestimation was observed.

Finally, PreXFOCUS integrates the reference return-level estimates with climate-conditioned exceedance probabilities to provide publicly available climate-conditioned return-level maps. These maps showed that the locations of the largest return-level estimates varied with return period, shifting toward the western Gulf Coast at longer return periods and toward the eastern Gulf Coast at shorter return periods. Together, the reference and climate-conditioned products provide precipitation-frequency information that can support infrastructure planning, flood-risk assessment, and adaptive water-resource management. They also provide a benchmark for evaluating alternative spatial extreme-value models, identifying climate-sensitive regions, and coupling precipitation-frequency estimates with hydrologic and flood-inundation models. Future studies may extend the framework to additional climate predictors and independent time periods while more fully propagating uncertainty across the annual-maximum and occurrence-modeling stages.

Appendix A: Statistical model specifications

A1 Gaussian-process formulation

As summarized in Sect. 2.2, spatial variation in the GEV parameters and occurrence-model coefficients was represented using latent GPs. For any spatially varying quantity $\psi(s)$ modeled using a GP, the finite-dimensional representation at N locations is

$$\boldsymbol{\psi} = (\psi(s_1), \dots, \psi(s_N))^\top \sim MVN\left(\boldsymbol{m}_\psi, \boldsymbol{\Sigma}_\psi\right), \tag{A1}$$

where $\boldsymbol{m}_\psi$ is the GP mean vector and $\boldsymbol{\Sigma}_\psi$ is the spatial covariance matrix. When site-level covariates are included, the GP mean is specified as

$$\boldsymbol{m}_\psi = \boldsymbol{X}\boldsymbol{\alpha}_\psi, \tag{A2}$$

where $\mathbf{X}$ is the site-descriptor matrix consisting of an intercept and standardized longitude, latitude, and elevation, and $\boldsymbol{\alpha}_\psi$ is the corresponding regression-coefficient vector.

Spatial covariance was represented using an exponential kernel:

$$\left[\boldsymbol{\Sigma}_\psi\right]_{i,j} = \gamma_{\psi,1}^2 \exp\left(-\frac{d_{ij}}{\gamma_{\psi,2}}\right), \tag{A3}$$

where $\gamma_{\psi,1}^2$ is the sill parameter, $\gamma_{\psi,2}$ is the range parameter, and d_{ij} denotes the Euclidean distance between sites s_i and s_j, computed from UTM coordinates and scaled in units of 10 km.

A2 Annual-maximum model formulations

The four annual-maximum formulations summarized in Sect. 2.2 differ in the extent of spatial pooling among the GEV parameters and in whether dependence among annual maxima is represented at the data layer.

M1 estimates site-specific GEV location, scale, and shape parameters independently without latent spatial pooling.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

M2 introduces spatial pooling by assigning the GP structure in Eqs. (A1)–(A3) to the GEV location and log-scale parameters while retaining a spatially homogeneous shape parameter. This formulation is consistent with regional Bayesian extreme-value approaches that share information spatially across the location and scale parameters (Boumis et al., 2023; Lu et al., 2025; Najafi and Moradkhani, 2015; Najafi and Moradkhani, 2013; Yan and Moradkhani, 2015, 2016).

M3 additionally allows the GEV shape parameter to vary spatially, following hierarchical extreme-value formulations that account for regional variation in tail behavior (Cooley and Sain, 2010; Lima et al., 2016; Wu et al., 2018, 2019). At site s_i, the shape parameter is represented through a transformed latent spatial process,

,

$$\boldsymbol{\xi}(s_i) = \text{softplus}\,(\eta_\xi(s_i))\,, \tag{A4}$$

where

$$\eta_\xi(s_i) = \bar{\xi} + \boldsymbol{\varepsilon}_\xi(s_i), \tag{A5}$$

$\bar{\xi}$ denotes the domain-wide latent baseline, and $\varepsilon_\xi(s_i)$ denotes the spatial deviation at site s_i. Across all sites,

$$\boldsymbol{\varepsilon}_\xi = \left(\varepsilon_\xi(s_1), \ldots, \varepsilon_\xi(s_N)\right)^\top \sim MVN\,(0, \boldsymbol{\Sigma}_\xi), \tag{A6}$$

where $\boldsymbol{\Sigma}_\xi$ follows the exponential covariance structure in Eq. (A3). The softplus transformation was used to constrain the GEV shape parameter to positive values while allowing it to vary spatially (Koutsoyiannis, 2004; Papalexiou and Koutsoyiannis, 2013).

M2 and M3 assume that annual maxima are conditionally independent across gauges given the latent spatial processes for the GEV parameters. M4 retains the spatial GEV parameter structure of M3 and additionally represents dependence among annual maxima at the data layer through the Gaussian-copula formulation described in Appendix A3.

A3 Gaussian copula spatial dependence

For year t, let

$$\mathbf{Y}_t = \left(Y(s_1, t), \ldots, Y(s_N, t)\right)^\top, \tag{A7}$$

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

denote the vector of annual maximum precipitation across N gauges. Under M4, the joint distribution is represented as

$$\mathbf{Y}_t \sim C_N\left[\mathbf{R}_{cop}; F_1\left(.;\boldsymbol{\theta}(s_1)\right), \dots, F_N\left(.;\boldsymbol{\theta}(s_N)\right)\right], \tag{A8}$$

where C_N denotes an N-dimensional Gaussian copula, $\mathbf{R}_{cop}$ is the copula dependence matrix, and $F_i\left(.;\boldsymbol{\theta}(s_i)\right)$ is the GEV marginal distribution at gauge s_i.

The Gaussian copula is defined as

$$C_{\mathbf{R}_{cop}}(v_1, \dots, v_N) = \Phi_{\mathbf{R}_{\mathrm{cop}}}[\Phi^{-1}(v_1), \dots, \Phi^{-1}(v_N)], \tag{A9}$$

where

$$v_i = F_i\left(y_i;\boldsymbol{\theta}(s_i)\right),$$

$\Phi^{-1}(.)$ is the inverse standard normal CDF, and $\Phi_{\mathbf{R}_{\mathrm{cop}}}(.)$ is the multivariate normal CDF with dependence matrix $\mathbf{R}_{\mathrm{cop}}$.

The dependence matrix is specified as a function of inter-gauge distance using an exponential dependogram:

$$\left[\mathbf{R}_{\mathrm{cop}}\right]_{ij} = \exp\left(-\frac{d_{ij}}{\rho_{\mathrm{cop}}}\right), i \neq j, \qquad \left[\mathbf{R}_{\mathrm{cop}}\right]_{ii} = 1. \tag{A10}$$

Here, d_{ij} is the distance between gauges s_i and s_j, and ρ_{cop} is the copula range parameter controlling the rate at which dependence decays with distance. This structure allows M4 to represent dependence among annual maxima beyond that induced by the spatially varying GEV parameters (Bracken et al., 2016; Renard, 2011).

A4 Bernoulli occurrence-model formulation

As summarized in Sect. 2.2, the selected annual-maximum model was used to define site-specific thresholds corresponding to return periods $r \in \{2,5,10,25,50,100\}$. Let $z_r\left(s_i, t_j\right)$ denote the binary indicator of whether the annual maximum precipitation at location s_i and year t_j exceeds the estimated r-year return-level threshold. The indicator is defined as

$$z_r\left(s_i, t_j\right) = \begin{cases} 1, & Y\left(s_i, t_j\right) > q_r(s_i), \\ 0, & Y\left(s_i, t_j\right) \leq q_r(s_i), \end{cases} \tag{A11}$$

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

where $Y(s_i, t_j)$ is the annual maximum precipitation and $q_r(s_i)$ is the estimated r-year return level at gauge s_i. For each return-period threshold, the resulting binary responses can be represented as

$$\boldsymbol{Z}_r = [z_r(s_i, t_j)]_{i=1:N_s, j=1:N_t}, \tag{A12}$$

where N_s and N_t denote the numbers of gauges and years, respectively.

For each threshold, exceedance occurrence is modeled as

$$z_r(s_i, t_j) \sim \text{Bernoulli}\left(p_r(s_i, t_j)\right), \tag{A13}$$

where $p_r(s, t)$ denotes the probability of exceeding the r-year return level threshold at location s_i in year t_j. Climate dependence is introduced through the logit link,

$$\text{logit}\left[p_r(s_i, t_j)\right] = \lambda_{0,r}(s) + \sum_{k=1}^{K} \lambda_{k,r}(s).SCI_k(t), \tag{A14}$$

where $\lambda_{0,r}(s)$ is the spatially varying intercept, representing the baseline log-odds of exceedance for the r-year threshold, $\lambda_{k,r}(s)$ is the spatially varying coefficient associated with the kth climate index, and $\text{SCI}_k(t)$ is its annual value in year t_j. The spatially varying regression coefficients were represented using the GP formulation described in Appendix A1.

Author contributions

A.T.: Conceptualization, Methodology, Software, Formal analysis, Data curation, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing. N.S.: Validation, Supervision, Writing – review & editing. H.M.: Conceptualization, Methodology, Validation, Supervision, Project administration, Funding acquisition, Resources, Writing – review & editing.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgement

Partial Funding for this project was provided by National Science Foundation, Award #2223893.

References

Aryal, S. K., Bates, B. C., Campbell, E. P., Li, Y., Palmer, M. J., and Viney, N. R.: Characterizing and modeling temporal and spatial trends in rainfall extremes, J. Hydrometeorol., 10, 241–253, 2009.

Atyeo, J. and Walshaw, D.: A region-based hierarchical model for extreme rainfall over the UK, incorporating spatial dependence and temporal trend, Environmetrics, 23, 509–521, 2012.

Banerjee, S., Gelfand, A. E., Finley, A. O., and Sang, H.: Gaussian predictive process models for large spatial data sets, J. R. Stat. Soc. Series B Stat. Methodol., 70, 825–848, 2008.

Beck, N., Genest, C., Jalbert, J., and Mailhot, M.: Predicting extreme surges from sparse data using a copula-based hierarchical Bayesian spatial model, Environmetrics, 31, e2616, 2020.

Bonnin, G. M., Martin, D., Lin, B., Parzybok, T., Yekta, M., and Riley, D.: NOAA Atlas 14 Precipitation-Frequency Atlas of the United States Volume 2 Version 3.0: Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, New Jersey, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia, West Virginia, Silver Spring, Maryland: National Oceanic and Atmospheric Administration.[Available online: http://www. nws. noaa. gov/oh/hdsc/PF_documents/Atlas14_Volume2. pdf], 2006.

Boumis, G., Moftakhari, H. R., and Moradkhani, H.: Storm surge hazard estimation along the US Gulf Coast: A Bayesian hierarchical approach, Coastal Engineering, 185, 104371, 2023.

Bracken, C., Rajagopalan, B., Alexander, M., and Gangopadhyay, S.: Spatial variability of seasonal extreme precipitation in the western United States, Journal of Geophysical Research: Atmospheres, 120, 4522–4533, 2015.

Bracken, C., Rajagopalan, B., Cheng, L., Kleiber, W., and Gangopadhyay, S.: Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain, Water Resour. Res., 52, 6643–6655, 2016.

Bracken, C., Holman, K. D., Rajagopalan, B., and Moradkhani, H.: A Bayesian hierarchical approach to multivariate nonstationary hydrologic frequency analysis, Water Resour. Res., 54, 243–255, 2018.

Castro-Camilo, D. and Huser, R.: Local likelihood estimation of complex tail dependence structures, applied to US precipitation extremes, J. Am. Stat. Assoc., 115, 1037–1054, 2020.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Castruccio, S., Huser, R., and Genton, M. G.: High-order composite likelihood inference for max-stable distributions and processes, Journal of Computational and Graphical Statistics, 25, 1212–1229, 2016.

Cayan, D. R., Redmond, K. T., and Riddle, L. G.: ENSO and hydrologic extremes in the western United States, J. Clim., 12, 2881–2893, 1999.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P.: An introduction to statistical modeling of extreme values, Springer, 2001.

Cooley, D. and Sain, S. R.: Spatial hierarchical modeling of precipitation extremes from a regional climate model, J. Agric. Biol. Environ. Stat., 15, 381–402, 2010.

Cooley, D., Nychka, D., and Naveau, P.: Bayesian spatial modeling of extreme precipitation return levels, J. Am. Stat. Assoc., 102, 824–840, 2007.

Davison, A. C., Padoan, S. A., and Ribatet, M.: Statistical modeling of spatial extremes, 2012.

England Jr, J. F., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas Jr, W. O., Veilleux, A. G., Kiang, J. E., and Mason Jr, R. R.: Guidelines for determining flood flow frequency—Bulletin 17C, US Geological Survey, 2018.

François, B., Schlef, K. E., Wi, S., and Brown, C. M.: Design considerations for riverine floods in a changing climate–a review, J. Hydrol. (Amst)., 574, 557–573, 2019.

García, J. A., Martín, J., Naranjo, L., and Acero, F. J.: A Bayesian hierarchical spatio-temporal model for extreme rainfall in Extremadura (Spain), Hydrological Sciences Journal, 63, 878–894, 2018.

Gershunov, A., Shulgina, T., Clemesha, R. E. S., Guirguis, K., Pierce, D. W., Dettinger, M. D., Lavers, D. A., Cayan, D. R., Polade, S. D., and Kalansky, J.: Precipitation regime change in Western North America: the role of atmospheric rivers, Sci. Rep., 9, 9944, 2019.

Ghosh, S. and Mallick, B. K.: A hierarchical Bayesian spatio-temporal model for extreme precipitation events, Environmetrics, 22, 192–204, 2011.

Glantz, M. H.: Currents of change: impacts of El Niño and La Niña on climate and society, Cambridge University Press, 2001.

Glenn, W. B.: Verification of forecasts expressed in terms of probability, Mon. Weather Rev., 78, 1–3, 1950.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Guirguis, K. J. and Avissar, R.: An analysis of precipitation variability, persistence, and observational data uncertainty in the western United States, J. Hydrometeorol., 9, 843–865, 2008.

Gumbel, E. J.: The return period of flood flows, The annals of mathematical statistics, 12, 163–190, 1941.

Hoffman, M. D. and Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo., J. Mach. Learn. Res., 15, 1593–1623, 2014.

Hu, L., Nikolopoulos, E. I., Marra, F., and Anagnostou, E. N.: Sensitivity of flood frequency analysis to data record, statistical model, and parameter estimation methods: An evaluation over the contiguous United States, J. Flood Risk Manag., 13, e12580, 2020.

Huang, X. and Stevenson, S.: Contributions of climate change and ENSO variability to future precipitation extremes over California, Geophys. Res. Lett., 50, e2023GL103322, 2023.

Hurrell, J. W. and Deser, C.: North Atlantic climate variability: the role of the North Atlantic Oscillation, Journal of marine systems, 79, 231–244, 2010.

Knight, D. B. and Davis, R. E.: Contribution of tropical cyclones to extreme rainfall events in the southeastern United States, Journal of Geophysical Research: Atmospheres, 114, 2009.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations, Hydrological Sciences Journal, 49, 2004.

Kwon, H., Khalil, A. F., and Siegfried, T.: Analysis of extreme summer rainfall using climate teleconnections and typhoon characteristics in South Korea 1, JAWRA Journal of the American Water Resources Association, 44, 436–448, 2008.

Lee, J. H., Julien, P. Y., and Lee, S.: Teleconnection of ENSO extreme events and precipitation variability over the United States, J. Hydrol. (Amst)., 619, 129206, 2023.

Lima, C. H. R. and Lall, U.: Hierarchical Bayesian modeling of multisite daily rainfall occurrence: Rainy season onset, peak, and end, Water Resour. Res., 45, 2009.

Lima, C. H. R., Lall, U., Troy, T. J., and Devineni, N.: A climate informed model for nonstationary flood risk prediction: Application to Negro River at Manaus, Amazonia, J. Hydrol. (Amst)., 522, 594–602, 2015.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Lima, C. H. R., Lall, U., Troy, T., and Devineni, N.: A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates, J. Hydrol. (Amst)., 541, 816–823, 2016.

Lu, Y., Seiyon Lee, B., and Doss-Gollin, J.: Bayesian spatiotemporal nonstationary model quantifies robust increases in daily extreme rainfall across the Western Gulf Coast, Environmental Research: Climate, 4, 035016, 2025.

Martel, J.-L., Brissette, F. P., Lucas-Picher, P., Troin, M., and Arsenault, R.: Climate change and rainfall intensity–duration–frequency curves: Overview of science and guidelines for adaptation, J. Hydrol. Eng., 26, 03121001, 2021.

Miniussi, A. and Marra, F.: Estimation of extreme daily precipitation return levels at-site and in ungauged locations using the simplified MEV approach, J. Hydrol. (Amst)., 603, 126946, 2021.

Mishra, A. K. and Singh, V. P.: Changes in extreme precipitation in Texas, Journal of Geophysical Research: Atmospheres, 115, 2010.

Montanari, A. and Koutsoyiannis, D.: Modeling and mitigating natural hazards: Stationarity is immortal!, Water Resour. Res., 50, 9748–9756, 2014.

Najafi, M. R. and Moradkhani, H.: Multi-model ensemble analysis of runoff extremes for climate change impact assessments, J. Hydrol. (Amst)., 525, 352–361, 2015.

Nash, J. E. and Sutcliffe, J. V: River flow forecasting through conceptual models part I — A discussion of principles, J. Hydrol. (Amst)., 10, 282–290, https://doi.org/https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Neal, R. M.: MCMC using Hamiltonian dynamics, Handbook of markov chain monte carlo, 2, 2, 2011.

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., Mantua, N. J., Miller, A. J., Minobe, S., and Nakamura, H.: The Pacific decadal oscillation, revisited, J. Clim., 29, 4399–4427, 2016.

Nizamani, M. M., Hughes, A. C., Wang, Y., Zhang, H.-L., and Lai, Z.: Climate extremes and socioeconomic impact of El Niño and La Niña events, Environ. Dev., 56, 101276, 2025.

Ntzoufras, I.: Bayesian modeling using WinBUGS, John Wiley & Sons, 2011.

Obeysekera, J. and Salas, J. D.: Frequency of recurrent extremes under nonstationarity, J. Hydrol. Eng., 21, 04016005, 2016.

van Oldenborgh, G. J., van der Wiel, K., Sebastian, A., Singh, R., Arrighi, J., Otto, F., Haustein, K., Li, S., Vecchi, G., and Cullen, H.: Attribution of extreme rainfall from Hurricane Harvey, August 2017, Environmental Research Letters, 12, 124009, https://doi.org/10.1088/1748-9326/aa9ef2, 2017.

Papalexiou, S. M. and Koutsoyiannis, D.: Battle of extreme value distributions: A global survey on extreme daily rainfall, Water Resour. Res., 49, 187–201, 2013.

Papalexiou, S. M., AghaKouchak, A., and Foufoula-Georgiou, E.: A diagnostic framework for understanding climatology of tails of hourly precipitation extremes in the United States, Water Resour. Res., 54, 6725–6738, 2018.

Patel, R. and Schneider, T.: How short climate records affect GEV-based estimates of rare event probabilities, Environmental Research: Climate, 2026.

Perica, S., Martin, D., Pavlovic, S., Roy, I., St Laurent, M., Trypaluk, C., Unruh, D., Yekta, M., and Bonnin, G.: Precipitation-Frequency Atlas of the United States, Volume 9 Version 2.0: Southeastern States, Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, 2013.

Perica, S., Pavlovic, S., St Laurent, M., Trypaluk, C., Unruh, D., and Wilhite, O.: Precipitation-frequency atlas of the United States, volume 11 version 2.0: Texas, 2018.

Phan, D., Pradhan, N., and Jankowiak, M.: Composable effects for flexible and accelerated probabilistic programming in NumPyro, arXiv preprint arXiv:1912.11554, 2019.

Prosdocimi, I. and Kjeldsen, T.: Parametrisation of change-permitting extreme value models and its impact on the description of change, Stochastic Environmental Research and Risk Assessment, 35, 307–324, 2021.

Reich, B. J. and Shaby, B. A.: A hierarchical max-stable spatial model for extreme precipitation, Ann. Appl. Stat., 6, 1430, 2012.

Renard, B.: A Bayesian hierarchical approach to regional frequency analysis, Water Resour. Res., 47, 2011.

Renard, B. and Thyer, M.: Revealing hidden climate indices from the occurrence of hydrologic extremes, Water Resour. Res., 55, 7662–7681, 2019.

https://doi.org/10.5194/essd-2026-726
Preprint. Discussion started: 22 September 2026

Open Access Earth System Science Data Discussions

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, Water Resour. Res., 46, 2010.

Renard, B., Thyer, M., McInerney, D., Kavetski, D., Leonard, M., and Westra, S.: A hidden climate indices modeling framework for Multivariable space-time data, Water Resour. Res., 58, e2021WR030007, 2022.

Reza Najafi, M. and Moradkhani, H.: Analysis of runoff extremes using spatial hierarchical Bayesian modeling, Water Resour. Res., 49, 6656–6670, 2013.

Ribatet, M., Cooley, D., and Davison, A. C.: Bayesian inference from composite likelihoods, with an application to spatial extremes, Stat. Sin., 813–845, 2012.

Rupp, D. E., Hawkins, L. R., Li, S., Koszuta, M., and Siler, N.: Spatial patterns of extreme precipitation and their changes under~ 2° C global warming: a large-ensemble study of the western USA, Clim. Dyn., 59, 2363–2379, 2022.

Rutz, J. J., Steenburgh, W. J., and Ralph, F. M.: Climatological characteristics of atmospheric rivers and their inland penetration over the western United States, Mon. Weather Rev., 142, 905–921, 2014.

Scarrott, C. and MacDonald, A.: A review of extreme value threshold estimation and uncertainty quantification, REVSTAT-Statistical journal, 10, 33–60, 2012.

Schlef, K. E., Kunkel, K. E., Brown, C., Demissie, Y., Lettenmaier, D. P., Wagner, A., Wigmosta, M. S., Karl, T. R., Easterling, D. R., and Wang, K. J.: Incorporating non-stationarity from climate change into rainfall frequency and intensity-duration-frequency (IDF) curves, J. Hydrol. (Amst)., 616, 128757, 2023.

Schliep, E. M., Cooley, D., Sain, S. R., and Hoeting, J. A.: A comparison study of extreme precipitation from six different regional climate models via spatial hierarchical modeling, Extremes (Boston)., 13, 219–239, 2010.

Schneider, D. P., Deser, C., Fasullo, J., and Trenberth, K. E.: Climate data guide spurs discovery and understanding, Eos, Transactions American Geophysical Union, 94, 121–122, 2013.

Schumacher, R. S. and Johnson, R. H.: Organization and environmental properties of extreme-rain-producing mesoscale convective systems, Mon. Weather Rev., 133, 961–976, 2005.

Serinaldi, F. and Kilsby, C. G.: Stationarity is undead: Uncertainty dominates the distribution of extremes, Adv. Water Resour., 77, 17–36, 2015.

Shang, H., Yan, J., and Zhang, X.: El Niño–Southern Oscillation influence on winter maximum daily precipitation in California in a spatial model, Water Resour. Res., 47, 2011.

Sharkey, P. and Winter, H. C.: A Bayesian spatial hierarchical model for extreme precipitation in Great Britain, Environmetrics, 30, e2529, 2019.

Sharma, S., Lee, B. S., Nicholas, R. E., and Keller, K.: A safety factor approach to designing urban infrastructure for dynamic conditions, Earths Future, 9, e2021EF002118, 2021.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A.: Bayesian measures of model complexity and fit, J. R. Stat. Soc. Series B Stat. Methodol., 64, 583–639, 2002.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Linde, A.: The deviance information criterion: 12 years on, J. R. Stat. Soc. Series B Stat. Methodol., 76, 485–493, 2014.

Stephenson, A. G., Lehmann, E. A., and Phatak, A.: A max-stable process model for rainfall extremes at different accumulation durations, Weather Clim. Extrem., 13, 44–53, 2016.

Sun, X. and Lall, U.: Spatially coherent trends of annual maximum daily precipitation in the United States, Geophys. Res. Lett., 42, 9781–9789, 2015.

Sun, X., Renard, B., Thyer, M., Westra, S., and Lang, M.: A global analysis of the asymmetric effect of ENSO on extreme precipitation, J. Hydrol. (Amst)., 530, 51–65, 2015.

Takallou, A., Samal, N., and Moradkhani, H.: PreXFOCUS: Daily Precipitation-Frequency Dataset, https://doi.org/10.5281/zenodo.21979711, August 2026., 2026.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C.: Rank-normalization, folding, and localization: An improved Rˆ for assessing convergence of MCMC (with discussion), Bayesian Anal., 16, 667–718, 2021.

Watanabe, S. and Opper, M.: Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory., Journal of machine learning research, 11, 2010.

Westra, S., Renard, B., and Thyer, M.: The ENSO–precipitation teleconnection and its modulation by the interdecadal Pacific oscillation, J. Clim., 28, 4753–4773, 2015.

Whan, K. and Zwiers, F.: The impact of ENSO and the NAO on extreme winter precipitation in North America in observations and regional climate models, Clim. Dyn., 48, 1401–1411, https://doi.org/10.1007/s00382-016-3148-x, 2017.

Wolter, K. and Timlin, M. S.: El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI. ext), International Journal of Climatology, 31, 1074–1087, 2011.

Wright, D. B., Bosma, C. D., and Lopez-Cantu, T.: US hydrologic design standards insufficient due to large increases in frequency of rainfall extremes, Geophys. Res. Lett., 46, 8144–8153, 2019.

Wu, Y., Lall, U., Lima, C. H. R., and Zhong, P.: Local and regional flood frequency analysis based on hierarchical Bayesian model: application to annual maximum streamflow for the Huaihe River basin, Hydrology and Earth System Sciences Discussions, 2018, 1–21, 2018.

Wu, Y., Xue, L., and Liu, Y.: Local and regional flood frequency analysis based on hierarchical Bayesian model in Dongting Lake Basin, China, Water Science and Engineering, 12, 253–262, 2019.

Yan, H. and Moradkhani, H.: A regional Bayesian hierarchical model for flood frequency analysis, Stochastic Environmental Research and Risk Assessment, 29, 1019–1036, 2015.

Yan, H. and Moradkhani, H.: Toward more robust extreme flood prediction by Bayesian hierarchical and multimodeling, Natural Hazards, 81, 203–225, 2016.

Yu, L., Zhong, S., Heilman, W. E., and Bian, X.: A comparison of the effects of El Niño and El Niño Modoki on subdaily extreme precipitation occurrences across the contiguous United States, Journal of Geophysical Research: Atmospheres, 122, 7401–7415, 2017.

Zhu, L. and Quiring, S. M.: Variations in tropical cyclone precipitation in Texas (1950 to 2009), Journal of Geophysical Research: Atmospheres, 118, 3085–3096, 2013.