ERDC/CHL TR-20-10

Coastal and Hydraulics Laboratory

US Army Corps
of Engineers®
Engineer Research and
Development Center

Spatial Analysis of Precipitation and Snow Water Equivalent Extremes for the Columbia River Basin

Brian E. Skahill, Angela M. Duren, Luciana Cunha, and Chris Bahner June 2020

The US Army Engineer Research and Development Center (ERDC) solves the nation's toughest engineering and environmental challenges. ERDC develops innovative solutions in civil and military engineering, geospatial sciences, water resources, and environmental sciences for the Army, the Department of Defense, civilian agencies, and our nation's public good. Find out more at www.erdc.usace.army.mil.

To search for other technical reports published by ERDC, visit the ERDC online library at https://erdclibrary.on.worldcat.org/discovery.

ERDC/CHL TR-20-10
June 2020

Spatial Analysis of Precipitation and Snow Water Equivalent Extremes for the Columbia River Basin

Brian E. Skahill

Coastal and Hydraulics Laboratory
US Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Angela M. Duren

US Army Corps of Engineers, Northwestern Division
PO Box 2870
Portland, Oregon 97208-2870

Luciana Cunha

WEST Consultants
101 Parkshore Drive
Folsom, California 95630

Chris Bahner

WEST Consultants
2601 25th St. SE
Suite 450
Salem, Oregon 97302

Final report

Prepared for US Army Corps of Engineers, Northwestern Division
Portland, OR 97208-2870

Under Funding Acct Code U4363312; AMSCO Code 031398

Abstract

Recent advances in the spatial statistics of extremes and model calibration were applied to develop and deliver areal-exceedance estimates for precipitation (PREC), by season and duration, and snow water equivalent (SWE), by cool season month and for the water year, for 758 delineated sub-basins of the Columbia River Basin (CRB), which correspond to a new CRB hydrology model watershed delineation. Understanding that future US Army Corps of Engineers, Northwestern Division, mission requirements may change, project execution also included the development and delivery of an application guidance document to credibly compute areal-exceedance estimates, including uncertainty, for PREC or SWE for any arbitrary area within the CRB. R, a free software environment for statistical computing and graphics (https://www.r-project.org/), and QGIS, a free and open source geographic information system (https://qgis.org/en/site/index.html), were the primary tools used for product development and delivery. The following R software packages were primarily used during project execution: evd, Glmnet, maps, raster, rgdal, SDMTools, sp, and SpatialExtremes.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for the US Army Corps of Engineers, Northwestern Division, under Funding Account Code U4363312, AMSCO Code 031398.

The work was performed by the Hydrologic Systems Branch of the Flood and Storm Protection Division, US Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory (ERDC-CHL). At the time of publication of this report, Dr. Hwai-Ping Cheng was Branch Chief; Dr. Cary A. Talbot was Division Chief; and Dr. Julie Rosati was the Technical Director for Flood Risk Management. The Deputy Director of ERDC-CHL was Mr. Jeffrey R. Eckstein, and the Director was Dr. Ty V. Wamsley.

The Commander of ERDC was COL Teresa A. Schlosser, and the Director was Dr. David W. Pittman.

1 Introduction

1.1 Background

The Northwestern Division (NWD) of the US Army Corps of Engineers (USACE) is developing a new hydrologic modeling system for the Columbia River Basin (CRB). The new CRB hydrologic modeling system consists of integrated models, studies, and infrastructure to support the water management, dam and levee safety, and planning missions of the USACE. The objective of the CRB hydrology studies[1] is to provide a cohesive hydrologic dataset and system that support all three programs (water management, dam and levee safety, and planning) and, to the extent practicable, minimizes duplicative and/or diverging methodologies, models, datasets, and results. The technical approach of the CRB hydrology studies includes seven key tasks, and one involves the development of regional frequency analysis (RFA) estimates for precipitation (PREC) and snow water equivalent (SWE) data for the CRB.

The National Weather Service provides point precipitation frequency estimates for the United States; however, current updated estimates are unavailable for the states of Washington, Oregon, Idaho, Montana, and Wyoming (https://hdsc.nws.noaa.gov/hdsc/pfds/index.html). In addition, approximately 39,700 mi^2 of the CRB 258,000 mi^2 total drainage area also intersects the Canadian province of British Columbia.[2,3] Moreover, the deployment of an L-moments-based RFA requires the study area to be decomposed into homogeneous sub-regions, and this is problematical for the CRB hydrology studies due to the CRB basin size and the physiographic and climatological variability that exists within the CRB.

---

[1] USACE (US Army Corps of Engineers). Unpublished draft. *Columbia River Basin Hydrology Studies*. Angela Duren, P.E., P.H., USACE-NWP-EC-H-HY, Technical Lead, and Pete Dickerson, P.E., P.H., D.WRE, USACE-NWD, Senior Advisor. Prepared for US Army Corps of Engineers, Northwest Division, July.

[2] For a full list of the spelled-out forms of the units of measure used in this document, please refer to *US Government Publishing Office Style Manual*, 31st ed. (Washington, DC: US Government Publishing Office, 2016), 248-52, https://www.govinfo.gov/content/pkg/GPO-STYLEMANUAL-2016/pdf/GPO-STYLEMANUAL-2016.pdf.

[3] For a full list of the unit conversions used in this document, please refer to *US Government Publishing Office Style Manual*, 31st ed. (Washington, DC: US Government Publishing Office, 2016), 345-7, https://www.govinfo.gov/content/pkg/GPO-STYLEMANUAL-2016/pdf/GPO-STYLEMANUAL-2016.pdf.

Furthermore, the CRB hydrology studies also require the credible computation of complex areal-based assessments of risk such as

$$\Pr\left\{\int_{\mathcal{B}} \Upsilon(x)dx > z_{crit}\right\} \quad (1)$$

where $\Upsilon(x)$, $\mathcal{B}$, and z_{crit} denote the joint distribution for an arbitrary hydrometeorological parameter (e.g., PREC and SWE), any arbitrary area within the analysis domain (e.g., a sub-basin), and a critical quantity greater than zero, respectively (see Ribatet et al. [2015] and references cited within). An L-moments-based RFA would not satisfy this requirement.

Recent advances have demonstrated the capacity to efficiently, flexibly, and credibly model spatial extremes of pointwise maxima (e.g., Davison and Gholamrezaee 2011; Reich and Shaby 2012; Olinda et al. 2014; Stephenson et al. 2016; Nicolet et al. 2015; Blanchet and Davison 2011). Each of these studies involved application of max-stable processes, the stochastic process extension of extreme value theory (EVT). With their application, one can not only compute pointwise return level maps, but also by modeling the joint distribution, more complex areal-based assessments of risk such as Equation 1, all the while conforming with EVT whereupon there is an assurance of a mathematically credible extrapolation. The estimation of integrals such as Equation 1, for any hydrometeorological variable and for any $\mathcal{B}$, is relatively straightforward to estimate via simulation from max-stable spatial models. The procedure avoids the need to develop and apply empirical depth-area reduction factors to convert point to areal estimates.

1.2 Objective

Pointwise and areal-based exceedance probabilities of observed PREC, by season and duration, and SWE, by cool season month and for the water year (WY), are calculated for the CRB using, for each specific application, a spatial max-stable process model and observed pointwise maxima data. Each max-stable modeling analysis leverages EVT, at-site estimates of extreme PREC/SWE, physiographic and climatological covariate data, and recent advances in model calibration (Friedman et al. 2010; Tibshirani et al. 2010; Simon et al. 2011; Tibshirani et al. 2012).

1.3 Approach

Each max-stable process modeling analysis involved the estimation of areal-based exceedance probabilities for 758 delineated sub-basins of the CRB, which corresponds to the new CRB hydrology model watershed delineation. Future USACE-NWD mission may require the evaluation of Equation 1 for sub-areas different from the noted CRB hydrology model watershed delineation. An application guidance document was developed to support the computation of Equation 1 for any arbitrary area within the CRB (see Appendix A: Instructions on How to Obtain Exceedance Probability Estimates for a Specific Area [e.g., sub-watershed]).

The remainder of this report summarizes the data used, the methodology employed for the spatial analysis of extreme PREC and SWE for the CRB, details of the modeling analyses for PREC and SWE, and the specific products developed and delivered to the USACE-NWD to successfully satisfy the RFA task requirements of the CRB hydrology studies.

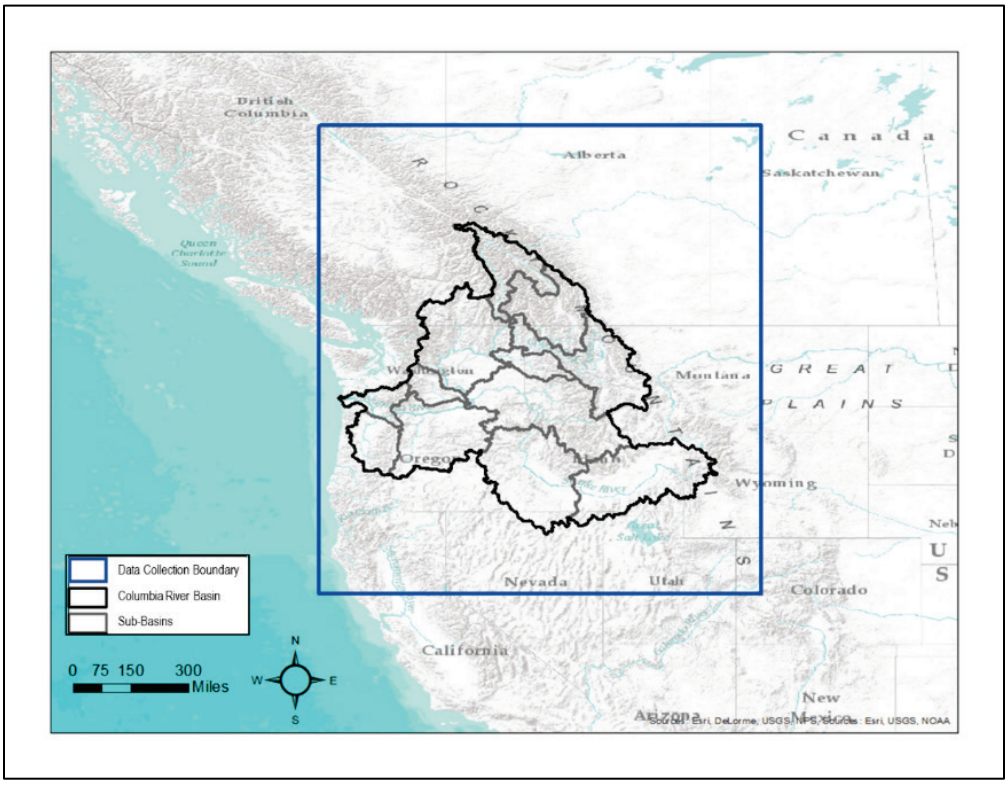
2 Data

Study data requirements included observed PREC and SWE time-series data associated with gage sites located within and/or surrounding the CRB and gridded covariates that characterize the climatology and physiography of the CRB. The SWE modeling analyses also considered temporal covariates — in particular, climate indices that have previously been shown to influence temperature and precipitation across the CRB (see Beckers et al. [2016]; Pytlak et al. [2016]; and references cited within).

2.1 Precipitation (PREC) data

PREC time-series data were collected for gage stations located within the CRB and also for a box region surrounding the CRB as is shown in Figure 1. Time series of annual maximum precipitation (AMP) were processed for defined "warm" (May to September) and "cool" (October to April) seasons for the following durations (1-, 2-, 3-, 5-, 10-, 30-, 60-, 90- and 120-day). A detailed description of the precipitation data collection and processing procedure is provided in Appendix B: Precipitation Data Collection, Quality Control and Processing.

Figure 1. CRB PREC data collection boundaries.



2.2 Snow water equivalent (SWE) data

SWE time-series data were collected from the Snow Telemetry (SNOTEL) network maintained by the US Department of Agriculture Natural Resources Conservation Service, National Water and Climate Center (https://www.wcc.nrcs.usda.gov/snow/) and automated snow weather stations' data from Canada (https://www2.gov.bc.ca/gov/content/environment/air-land-water/water/water-science-data/water-data-tools/snow-survey-data/automated-snow-weather-station-data). While other sources of SWE data do exist in the United States and Canada (e.g., snow course sites), their observation reporting frequency was identified to be inadequate to support extremal analyses.

2.3 Spatial covariate data

Gridded datasets of physiographic and climatological covariates were collected and processed to support trend surface modeling of extreme PREC, by duration and season, and SWE, by cool season month or for the WY. Table 1 is a summary listing of the collected and processed gridded covariate datasets and their sources. Monthly, annual, and seasonal products were collected and processed for the climatology datasets listed in Table 1 (PRECIP, TMEAN, TMAX, TMIN, DLWRF, DSWRF, EWIND, NWIND). Each covariate dataset was aggregated or resampled to a 30 arc second grid cell resolution to support the spatial modeling analysis. Each spatial analysis for PREC, by season and duration, and SWE, by cool season month and for the WY, used normalized spatial covariates.

Table 1. Summary listing of the collected and processed covariate datasets.

Covariates	Description	Resolution
Source: Climate data for Western North America (https://sites.ualberta.ca/~ahamann/data.html)		
PRECIP	Total precipitation	1 km(Lambert Conformal Conic projection)
TMAX	Maximum temperature	
TMIN	Minimum temperature	
TMEAN	0.5*(TMAX+TMIN)	
PAS	Precipitation as snow	
Source: National elevation dataset (Mosaic of 1 x 1 degree tiles) (US Geological Survey)		
X, Y, Z (from digital elevation model [DEM])	Coordinates and elevation for each raster grid cell	1 arc second

Covariates	Description	Resolution
Source: Use of the Geographic Resources Analysis Support System, Geographic Information System, function r.slope.aspect. (https://grass.osgeo.org/grass70/manuals/r.slope.aspect.html)		
SLOPE	Slope	30 arc seconds
ASPECT	Aspect	
PROFILE CURVATURE	Profile curvature	
TANGENTIAL CURVATURE	Tangential curvature	
Source: Tree canopy cover circa 2010 (https://landcover.usgs.gov/glc/TreeCoverDescriptionAndDownloads.php)		
TREE COVER	Percent maximum (peak of growing season) tree canopy cover	1 arc second
Source: North American Land Data Assimilation System (https://hydro1.gesdisc.eosdis.nasa.gov/data/NLDAS/README.NLDAS2.pdf)		
DLWRF	Downward longwave radiation flux	.125 degree
DSWRF	Downward shortwave radiation flux	
Source: National Centers for Environmental Prediction Climate Forecast System Reanalysis dataset (https://app.climateengine.org/)		
EWIND	Eastward wind component	.2 degree
NWIND	Northward wind component	

2.4 Temporal covariate data

Temporal covariate data were collected and processed to support trend surface modeling of extreme SWE by month and for the WY. Eight climate indices were considered, viz., the Arctic Oscillation (https://www.ncdc.noaa.gov/teleconnections/ao/), Oceanic Nino Index (http://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php), Multivariate El Nino Southern Oscillation index (https://www.esrl.noaa.gov/psd/enso/mei/), Pacific-North America pattern (https://www.ncdc.noaa.gov/teleconnections/pna/), Pacific Decadal Oscillation (https://www.ncdc.noaa.gov/teleconnections/pdo/), Southern Oscillation Index (https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/), Quasi Biennial Oscillation (https://www.esrl.noaa.gov/psd/data/correlation/qbo.data), and the East-North Pacific Index (ftp://ftp.cpc.ncep.noaa.gov/wd52dg/data/indices/epnp_index.tim) (see Beckers et al. [2016]; Pytlak et al. [2016]; and references cited within).

3 Modeling Analysis Methodology

This section is a succinct practice-oriented summary of max-stable processes for the calculation of spatial extremes. Ribatet (2017, 2013), Ribatet et al. (2015), Davison et al. (2012), and Cooley et al. (2012) provide comprehensive technical descriptions of max-stable processes. From univariate extreme value theory, it is known that if normalizing sequences exist for a sequence of independent and identically distributed random variables such that the rescaled variables have a nondegenerate limiting distribution, then the distribution is max-stable and, moreover, that a distribution is max-stable if and only if it is the generalized extreme value (GEV) distribution (Coles 2001). Mathematical nondegenerate limit law expressions of max-stability, comparable in form to the univariate case, that involve the existence of normalizing sequences also exist in the multivariate and stochastic process settings (Ribatet 2017). Univariate EVT results guarantee that the marginal distribution of max-stable processes are GEV distributed with possibly different parameters by location (Ribatet et al. 2015; Ribatet 2017).

As previously mentioned, this study applies a max-stable spatial process wherein just as with univariate EVT this limiting process is used to model the partial maxima. It is convenient and without any loss of generality to define a simple max-stable process, denoted by $\{Z(x): x \in \mathcal{X} \subset \mathbb{R}^d, d \geq 1\}$, to be with fixed unit Fréchet, rather than spatially variable GEV, margins. In this case, $\Pr\{Z(x) \leq z\} = \exp(-1/z)\ \forall x \in \mathcal{X}$ and $z > 0$. de Haan (1984) introduced the spectral representation of a simple max-stable process. It states that $\{Z(x): x \in \mathcal{X}\} \equiv \left\{\max_{i \geq 1} \zeta_i Y_i(x): x \in \mathcal{X}\right\}$, where $\{\zeta_i: i \geq 1\}$ is a Poisson point process on $(0, \infty)$ with intensity measure $\zeta^{-2} d\zeta$ and $\{Y_i(x): x \in \mathcal{X}, i \geq 1\}$ a sequence of independent copies of a non-negative stochastic process $\{Y(x): x \in \mathcal{X}\} \ni \mathbb{E}\{Y(x)\} = 1\ \forall x \in \mathcal{X}$ ($\equiv$ indicates equality in distribution). The spectral representation of a max-stable process introduced by de Haan (1984) has resulted in the subsequent development of usable parametric models for spatial extremes, with different distributional assumptions for $\{Y(x): x \in \mathcal{X}\}$ resulting in different max-stable models.

The famous Smith[1] process, also known as the Gaussian extreme value process, is given by $\{Z(x): x \in \mathcal{X}\} \equiv \left\{\bigvee_{i\geq 1}^{max} \zeta_i \varphi(x - U_i; 0, \Sigma): x \in \mathcal{X}\right\}$, where $\varphi(\cdot; 0, \Sigma)$ is the d-variable Gaussian density with zero mean and covariance Σ and $\{(\zeta_i, U_i): i \geq 1\}$ are the points of a Poisson process on $(0, \infty) \times \mathbb{R}^d$ with intensity measure $\zeta^{-2} d\zeta du$ (Ribatet et al. 2015; Ribatet 2017). The Schlather process (Schlather 2002), also known as the extremal Gaussian process, is given by $\{Z(x): x \in \mathcal{X}\} \equiv \left\{\sqrt{2\pi} \bigvee_{i\geq 1}^{max} \zeta_i max\big(0, \varepsilon_i(x)\big): x \in \mathcal{X}\right\}$, where $\{\varepsilon_i(x): x \in \mathcal{X}\}$ are independent copies of a stationary Gaussian process with correlation function ρ (Ribatet et al. 2015; Ribatet 2017). The R software package SpatialExtremes implements the Bessel, Cauchy, generalized Cauchy, powered exponential, and Whittle-Matern correlation functions (Ribatet 2018). The Brown-Resnick process (Kabluchko et al. 2009; Brown and Resnick 1977) is given by $\{Z(x): x \in \mathcal{X}\} \equiv \left\{\bigvee_{i\geq 1}^{max} \zeta_i exp\big(\varepsilon_i(x) - \gamma(x)\big): x \in \mathcal{X}\right\}$, where $\{\varepsilon_i(x): x \in \mathcal{X}\}$ are independent copies of a zero mean Gaussian process with stationary increments and semi-variogram $\gamma(x) = Var\{\varepsilon(x + h) - \varepsilon(x)\}/2$ (Ribatet et al. 2015; Ribatet 2017). The Smith model is in fact a special case of the Brown-Resnick process. The extremal-t process (Opitz 2013) is given by $\{Z(x): x \in \mathcal{X}\} \equiv \left\{c_\nu \bigvee_{i\geq 1}^{max} \zeta_i max\big(0, \varepsilon_i(x)\big)^\nu: x \in \mathcal{X}\right\}$, where $\{\varepsilon_i(x): x \in \mathcal{X}\}$ are independent copies of a stationary Gaussian process with correlation function ρ and $c_\nu = \sqrt{\pi} 2^{-(\nu-2)/2} \Gamma\left(\frac{\nu+1}{2}\right)^{-1}$ where Γ is the gamma function (Ribatet et al. 2015; Ribatet 2017). The Schlather process is a special case of the extremal-t process with $\nu = 1$ (Ribatet et al. 2015; Ribatet 2017).

Figure 2 includes separate plots of the Smith process with variance equal to 2, Schlather process with $\rho(h) = exp\{-(h/3)^1\}$, Brown-Resnick process with $\gamma(x) = (h/3)^1$, and extremal-t process with $\nu = 4$ and $\rho(h) = exp\{-(h/3)^1\}$, with each plot depicting five independent realizations of each respective process on $\mathcal{X} = [0,10]$ (Ribatet 2015). Figure 3 depicts one realization of the Smith process with Σ given by the identity matrix, Schlather process with $\rho(h) = exp\{-(h/2)^{1.5}\}$, and the extremal-t process, with $\nu = 5$ and $\rho(h) = exp\{-(h/2)^{1.5}\}$, for a 250 x 250 grid of $\mathcal{X} = [0,10] \times [0,10]$ (Ribatet et al. 2015). The Smith process realizations depicted in Figure 2 and Figure 3 underscore the reports that it produces artificial surfaces that are too smooth for many practical applications (Ribatet et al. 2015; Ribatet 2017). The Schlather process realizations

[1] Smith, R. L. Unpublished manuscript. *Max-Stable Processes and Spatial Extremes*. http://www.stat.unc.edu/faculty/rs/papers/RLS_Papers.html

depicted in Figure 2 and Figure 3 underscore summaries emphasizing that it tends to produce, unrealistically, due to a limitation, larger areas with the largest values in comparison with the extremal-t and Brown-Resnick processes (Ribatet et al. 2015). The Brown-Resnick process is known to be difficult to work with (Ribatet et al. 2015). Nicolet et al. (2015) reported better performance for the extremal-t process relative to other available max-stable processes in their study of the dependence structure of extreme snowfall in the French Alps.

Figure 2. Five independent realizations of the Smith, Schlather, Brown-Resnick, and Extremal-t processes on $\mathcal{X} = [\mathbf{0}, \mathbf{10}]$ (Ribatet 2015).

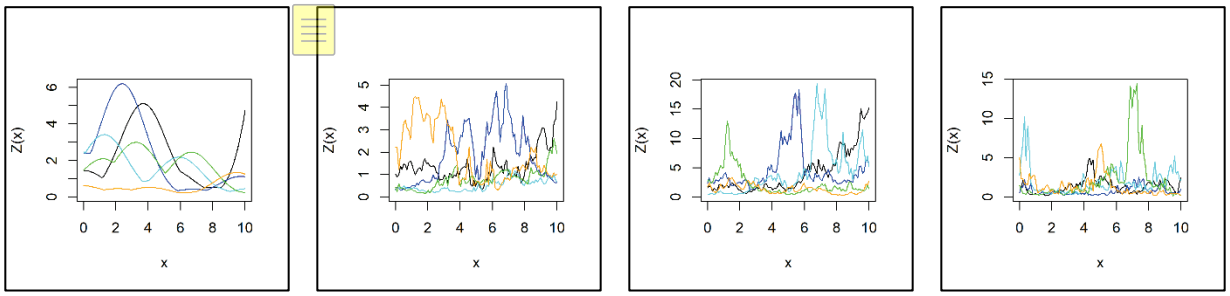
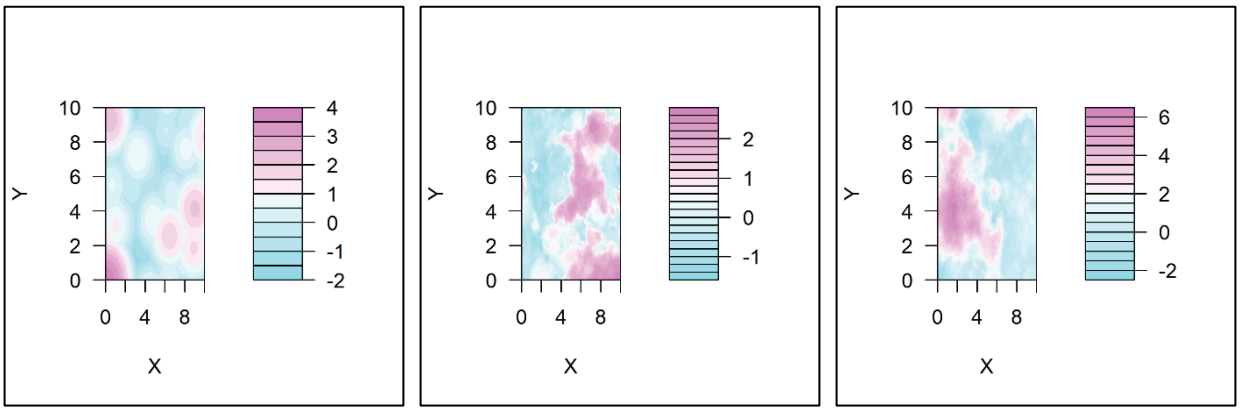


Figure 3. One realization of the Smith process with Σ given by the identity matrix, Schlather process with $\rho(\boldsymbol{h}) = \boldsymbol{exp}\{-(\boldsymbol{h}/\mathbf{2})^{\mathbf{1.5}}\}$, and the extremal-t process, with $\boldsymbol{\nu} = 5$ and $\rho(\boldsymbol{h}) = \boldsymbol{exp}\{-(\boldsymbol{h}/\mathbf{2})^{\mathbf{1.5}}\}$, on a 250 x 250 grid (Ribatet et al. 2015).



Max-stable processes are fit using their bivariate distributions and pairwise likelihood estimation for reasons of computational practicality (Davison and Gholamrezaee 2011; Ribatet 2013; Ribatet et al. 2015; Ribatet 2017). Model selection among competing max-stable processes can be based on the composite likelihood information criterion, an adaption of the Takeuchi Information Criterion (Takeuchi 1976), due to application of the composite pairwise likelihood formulation (Ribatet et al. 2015; Ribatet

2017). However, the use of qualitative evaluations of model fit are also recommended (Ribatet 2017; Ribatet 2015).

The extremal coefficient is a useful measure for summarizing dependence among extreme data (Davison and Gholamrezaee 2011; Ribatet et al. 2015; Ribatet 2017). From multivariate extremes, for a max-stable random vector with fixed unit Fréchet margins $Z = (Z_1, \dots, Z_k), k \geq 2$, $Pr(Z \leq z) = exp\big(-V(z_1, \dots, z_k)\big)$ for all $z = (z_1, \dots, z_k) \in (0, \infty)^k$ where V, the exponent function, satisfies $V(\lambda z_1, \dots, \lambda z_k) = \lambda^{-1} V(z_1, \dots, z_k)\ \forall\ \lambda > 0, V(\infty, \dots, \infty, z_d, \infty, \dots, \infty) = \frac{1}{z_d}$ for each $d \in (1, k)$, with two bounding cases, viz., at independence, $V(z_1, \dots, z_k) = \frac{1}{z_1} + \cdots + \frac{1}{z_k}$, and at dependence, $V(z_1, \dots, z_k) = min\left(\frac{1}{z_1}, \dots, \frac{1}{z_k}\right)$ (Ribatet 2017; Davison et al. 2012; Davison and Gholamrezaee 2011). Hence, $\forall\ z > 0$, the $(Z_1 \leq z, \dots, Z_k \leq z) = exp\big(-V(z, \dots, z)\big) = exp\left(\frac{-V(1,\dots,1)}{z}\right) = exp\left(\frac{-\theta}{z}\right)$, where θ, the extremal coefficient, is a quantity independent of the level of z and a summary measure of the degree of dependence among the extreme data. The extremal coefficient varies between one, when the observations are fully dependent, and k, when they are independent (Ribatet 2017; Davison et al. 2012; Davison and Gholamrezaee 2011). Results from the spectral representation of a max-stable process combined with the understanding that observations of $\{Z(x): x \in \mathcal{X} \subset \mathbb{R}^d, d \geq 1\}$ are at discrete locations in $\mathcal{X}$ yield $\theta = \mathbb{E}\left\{\underset{j=1,\dots,k}{max}\, Y\big(x_j\big)\right\}$ (Ribatet et al. 2015; Ribatet 2017; Davison and Gholamrezaee 2011). Hence, the extremal coefficient can be approximated from simulation (Davison and Gholamrezaee 2011). However, for spatial extremes, the extremal coefficient function is a convenient summary measure of dependence among the extreme data (Schlather and Tawn 2003). In the bivariate case, assuming isotropy, the extremal coefficient is a function of the Euclidean distance, h, between any two observation sites and the $Pr(Z(x) \leq z, Z(x+h) \leq z) = exp\big(-V(z,z)\big) = exp\left(\frac{-V(1,1)}{z}\right) = exp\left(\frac{-\theta(h)}{z}\right) = \mathbb{E}\big[\max\big(Y(x), Y(x+h)\big)\big]$ with $\theta(h) \in [1,2]$. There exists a unique straightforward mapping between the extremal coefficient function, $\theta(h)$, and the F-madogram (Cooley et al. 2006), a well-defined measure of dependence among extreme data. Moreover, there exist empirical estimators for the F-madogram (Ribatet et al. 2015; Cooley et al. 2012). In summary, inspection of plots of the extremal coefficient function for a model and its data is a recommended qualitative evaluation of a fitted simple max-stable model. Using its bivariate distribution and properties of correlation functions for isotropic fields, it can be shown that the Schlather

process is unable to model spatial independence (Davison and Gholamrezaee 2011). This is the previously mentioned limitation of the Schlather process and the reason it produced larger areas with the largest values in Figure 2 and Figure 3 when compared with the more flexible extremal-t and Brown-Resnick models.

The discussion so far has focused on the description, fitting, selection, and qualitative evaluation of max-stable processes with fixed unit Fréchet margins. Assuming fixed margins is not realistic for most practical applications. As previously mentioned, the marginal distributions of a max-stable process are GEV distributed, possibly varying by location. The pairwise likelihood used to fit a simple max-stable process can be readily adapted to accommodate the simultaneous estimation of trend surface and dependence parameters (Ribatet et al. 2015; Ribatet 2017; Ribatet 2015). Trend surfaces are functions defined on $\mathcal{X}$ which use covariates to model the spatial variation of the location, $\mu(x)$, scale, $\sigma(x)$, and shape, $\xi(x)$, parameters of the known GEV marginal distributions. For example, linear trend surfaces are of the form $\mu(x) = \eta_{\mu,0} + \eta_{\mu,1}cov_{\mu,1} + \cdots + \eta_{\mu,n_\mu}cov_{\mu,n_\mu}$, $\sigma(x) = \eta_{\sigma,0} + \eta_{\sigma,1}cov_{\sigma,1} + \cdots + \eta_{\sigma,n_{\mu\sigma}}cov_{\sigma,n_\sigma}$, $\xi(x) = \eta_{\xi,0} + \eta_{\xi,1}cov_{\xi,1} + \cdots + \eta_{\xi,n_\xi}cov_{\xi,n_\xi}$, where $\eta_{.,i}$ and $cov_{.,i}$ are the parameters and covariates of the linear trend surface for $\mu(x)$, $\sigma(x)$, and $\xi(x)$, respectively. Potential covariates include, for example, gridded physiographic (e.g., such as x-location, y-location, elevation, slope, aspect, curvature) and climatological (e.g., such as mean annual/monthly temperature, precipitation, wind, solar radiation) data. Ribatet (2017) underscored the importance of correctly modeling the spatial variation of the marginal parameters by carefully "building relevant trend surfaces including any relevant covariable." Poor characterizations for $\mu(x)$, $\sigma(x)$, and $\xi(x)$ complicates the dependence parameterization (Blanchet 2009; Ribatet 2017). In this study, linear trend surfaces for the marginal parameters were effectively and efficiently developed by applying previously mentioned theory from spatial extremes and recent advances for regularizing general linear models (Friedman et al. 2010; Tibshirani et al. 2010; Simon et al. 2011; Tibshirani et al. 2012).

Regularization of the linear trend surface for $\mu(x)$, $\sigma(x)$, and $\xi(x)$ is needed given the number of available and presumed relevant gridded covariate datasets and acknowledgment that there are $2^{n.}$ possible models that involve subsets of $n.$ predictors where $n. = n_\mu, n_\sigma,$ or n_ξ (Gareth et al. 2013). If $n. = 10/100$, then there are 1,024/1.267651e+30 possible models

to be considered for a given marginal parameter trend surface. Zou and Hastie (2005) introduced the elastic-net penalty as a compromise between ridge (Hoerl and Kennard 1970; Tikhonov 1943) and lasso (Tibshirani 1996) regression. Given observations $y_i, i = 1, \dots, n$, an $n \times m$ matrix of normalized covariates X, and an assumed linear model $y_i = \eta_0 + \eta_1 x_{i,1} + \cdots + \eta_m x_{i,m}$, the elastic net minimizes $\frac{1}{2n}\sum_{i=1}^{n}(y_i - \eta_0 - \eta x_i^T)^2 + \lambda \sum_{j=1}^{m}\left[\frac{1}{2}(1-\alpha)\eta_j^2 + \alpha\left|\eta_j\right|\right]$, where λ is a non-negative regularization parameter that is tuned to weight the overall strength of the penalty and $\alpha \in [0,1]$ is specified to control the penalty term to vary from ridge regression at $\alpha = 0$ to lasso regression at $\alpha = 1$ (Friedman et al. 2010). Ridge regression yields smooth solutions that include all the predictor, whereas application of lasso regression results in automatic variable selection (i.e., sparse, much more easily interpretable solutions [Gareth et al. 2013]). The elastic net mixes the two methods. As α increases from 0 to 1 for a fixed λ, the number of zero-valued η_j increases from 0 to the sparsity of the lasso (Friedman et al. 2010). In this study, variable selection was preferred, and α was specified close to 1 for numerical stability (Friedman et al., 2010). Cross validation was applied to ensure that the minimizing value for λ was properly located for each elastic net application.

Independently derived elastic net application results for $\mu(x)$, $\sigma(x)$, and $\xi(x)$ guided subsequent spatial GEV model formulation and evaluation. The log-likelihood of the spatial GEV model, which assumes independence among the sample observation sites, is given by $l(\eta_\mu, \eta_\sigma, \eta_\xi) = \sum_{i=1}^{nsite}\sum_{j=1}^{nobs}\left\{-log\sigma_i - \left(1 + \xi_i \frac{y_{i,j}-\mu_i}{\sigma_i}\right)^{-1/\xi_i} - \left(1 + \frac{1}{\xi_i}\right) log\left(1 + \xi_i \frac{y_{i,j}-\mu_i}{\sigma_i}\right)\right\}$, where μ_i, σ_i, and ξ_i are the GEV parameters for the i-th site and $y_{i,j}$ is the j-th observation for the i-th site (Ribatet 2009). For $\mu(x)$, $\sigma(x)$, and $\xi(x)$, predictors associated with a set of elastic-net derived linear trend surface models, ranging in dimension from 1 to the dimension of the trend surface at the minimizing λ, were used to define and evaluate a series of spatial GEV models. For example, if 10, 5, and 3 elastic-net directed trend surface models were extracted from applications of elastic net for $\mu(x)$, $\sigma(x)$, and $\xi(x)$, respectively, then 150 unique spatial GEV models were defined, fitted, and evaluated. Spatial GEV model selection and evaluation was based on information criterion scores and comparisons of the spatial GEV model parameter estimates with their at-site counterparts (Blanchet 2009; Ribatet 2009).

The max-stable process deployment herein was a two-step procedure involving trend surface and simple max-stable process model selection, with each step assuming independence among the extremes and fixed margins, respectively. The results from these two separate steps were subsequently combined to build a general max-stable process wherein the trend surface and dependence parameters were simultaneously fitted. Thereafter, pointwise and areal-based return levels can be computed for any arbitrary area within the model domain. Areal-based exceedances are obtained by simulating multiple independent copies of the fitted max-stable process (Ribatet 2009, 2013, 2015, 2017). Schlather (2002) introduced the basis for simulating an independent realization of a simple max-stable process with on only a finite number of replicates.

4 Model Development

The development of a max-stable process model for PREC or SWE involves two key steps:

1. accounting for the inter-site dependence among the extreme data (i.e., simple max-stable model selection)
2. characterizing the spatial/spatiotemporal variation of the extreme data (i.e., trend surface model selection).

This section highlights salient features associated with these two steps necessary for the deployment of each max-stable process model for PREC by season and duration and for SWE by cool season month or for the WY.

4.1 PREC

Table 2 lists the number of observation sites that were used for the calibration and validation of each PREC model by season and duration. Figure 17 depicts the locations of the calibration and validation sites that were used for each PREC model by season and duration (see Appendix C: Site Locations for Calibration and Validation of PREC Models by Season and Duration).

Location (X and Y), XY, elevation, monthly, annual, and seasonal summaries of the eight gridded climatological datasets summarized in Table 1 (excluding precipitation as snow), and all of their squares, constitute the entire set of 232 covariables considered during the trend surface modeling analysis for each PREC model deployment.

Figure 4 depicts plots of the extremal coefficient function (Schlather and Tawn 2003), a succinct measure of spatial dependence as a function of distance, for each of the fitted simple max-stable models for PREC by duration and season.

Table 3 and Table 4 list by duration computed correlations for each of the fitted PREC trend surface models with their at-site estimates for cool and warm season, respectively. For perspective, Figure 5 includes scatter plots, only for the calibration sites, which compare two of the fitted trend surface models with their at-site estimates.

Table 2. Number of sites used for calibration and validation of PREC models.

Cool Season PREC			Warm Season PREC		
	Calibration	Validation		Calibration	Validation
	Count			Count	
1-day	1139	126	1-day	1159	128
2-day	1131	125	2-day	1162	129
3-day	1134	126	3-day	1163	129
5-day	1126	125	5-day	1157	128
10-day	1128	125	10-day	1156	128
30-day	1123	124	30-day	1121	124
60-day	1106	122	60-day	831	92
90-day	1072	119	90-day	328	36
120-day	966	107	120-day	NA	NA

Figure 4. Fitted simple max-stables for PREC by season and duration (h is normalized distance).

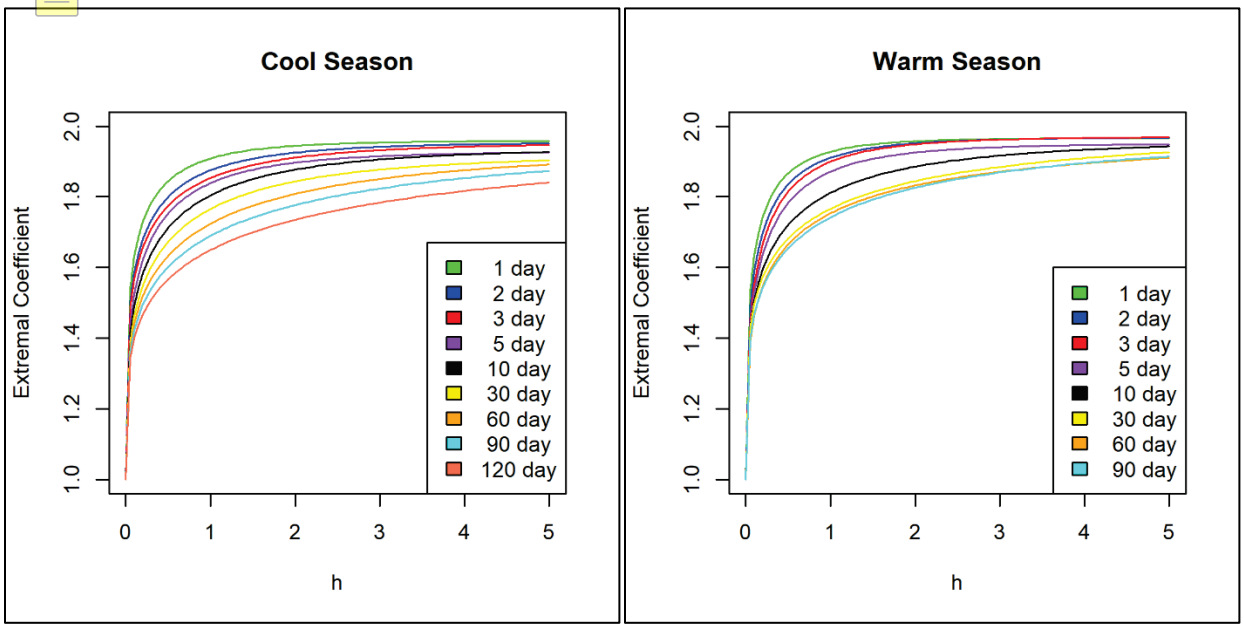


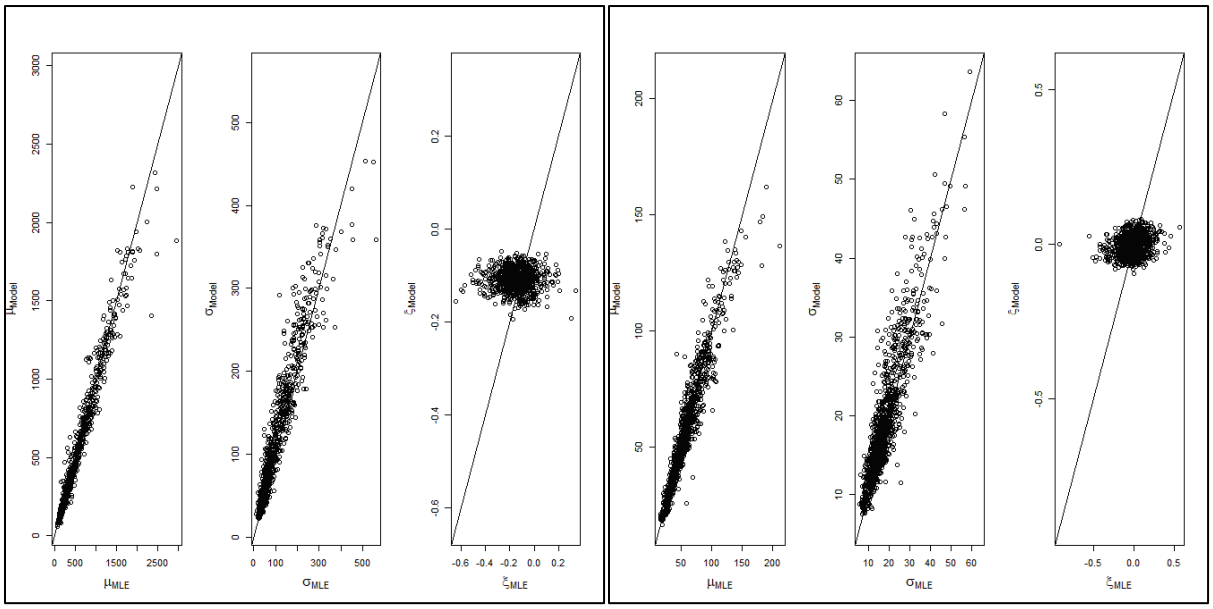
Table 3. Correlation of the fitted cool season trend surface models, by duration, with their at-site estimates for the calibration and validation sites.

Cool Season PREC						
	Calibration			Validation		
	Location	Scale	Shape	Location	Scale	Shape
1-day	0.972	0.931	0.231	0.957	0.851	0.261
2-day	0.975	0.930	0.281	0.979	0.945	0.116
3-day	0.976	0.933	0.196	0.975	0.940	0.108
5-day	0.978	0.931	0.206	0.983	0.952	0.051
10-day	0.981	0.940	0.267	0.986	0.937	0.403
30-day	0.981	0.954	0.281	0.982	0.940	0.356
60-day	0.981	0.955	0.275	0.985	0.961	0.426
90-day	0.981	0.953	0.197	0.986	0.950	0.150
120-day	0.981	0.951	0.008	0.978	0.961	-0.056

Table 4. Correlation of the fitted warm season trend surface models, by duration, with their at-site estimates for the calibration and validation sites.

Warm Season PREC						
	Calibration			Validation		
	Location	Scale	Shape	Location	Scale	Shape
1-day	0.938	0.836	0.227	0.938	0.841	0.212
2-day	0.943	0.858	0.228	0.933	0.854	0.190
3-day	0.953	0.889	0.281	0.952	0.854	0.409
5-day	0.950	0.897	0.257	0.937	0.879	0.124
10-day	0.955	0.904	0.255	0.955	0.908	0.171
30-day	0.964	0.922	0.284	0.961	0.907	0.410
60-day	0.966	0.892	0.145	0.944	0.860	0.243
90-day	0.969	0.876	0.262	0.970	0.832	0.116
120-day	NA	NA	NA	NA	NA	NA

Figure 5. Comparisons of the fitted trend surface model with corresponding at-site estimates for (a) 120-day duration cool season and (b) 10-day duration warm season (μ= GEV location; σ = GEV scale; ξ = GEV shape; Model = spatial model; MLE = at-site maximum likelihood estimate).



(a) (b)

4.2 SWE

Table 5 lists the number of observation sites that were used for the calibration and validation of each SWE model by month and for the WY. Figure 18 depicts the locations of the calibration and validation sites that were used for each SWE model (see Appendix D: Site Locations for Calibration and Validation of SWE Models by Month and for the WY).

The eight physiographic covariates, precipitation as snow, and monthly, annual, and seasonal summaries of the eight gridded climatological datasets summarized in Table 1, and all of their squares, constitute the entire set of 242 covariables considered during the trend surface modeling analysis for each SWE model deployment.

Figure 6 depicts plots of the extremal coefficient function (Schlather and Tawn 2003), a succinct measure of spatial dependence as a function of normalized distance, for each of the fitted simple max-stable models for SWE.

Table 6 lists computed correlations for each of the fitted SWE trend surface models with their at-site estimates. For perspective, Figure 7

includes scatter plots, only for the calibration sites, that compare the fitted trend surface model for March SWE with their at-site estimates.

Table 5. Number of sites used for calibration and validation of SWE models.

	SWE	
	Calibration	Validation
	Count	
Nov	290	32
Dec	316	35
Jan	319	35
Feb	315	35
Mar	311	34
Apr	280	31
May	162	17
Jun	61	6
WY	366	40

Figure 6. Fitted simple max-stables for SWE (h is normalized distance; May and WY plots were identical).

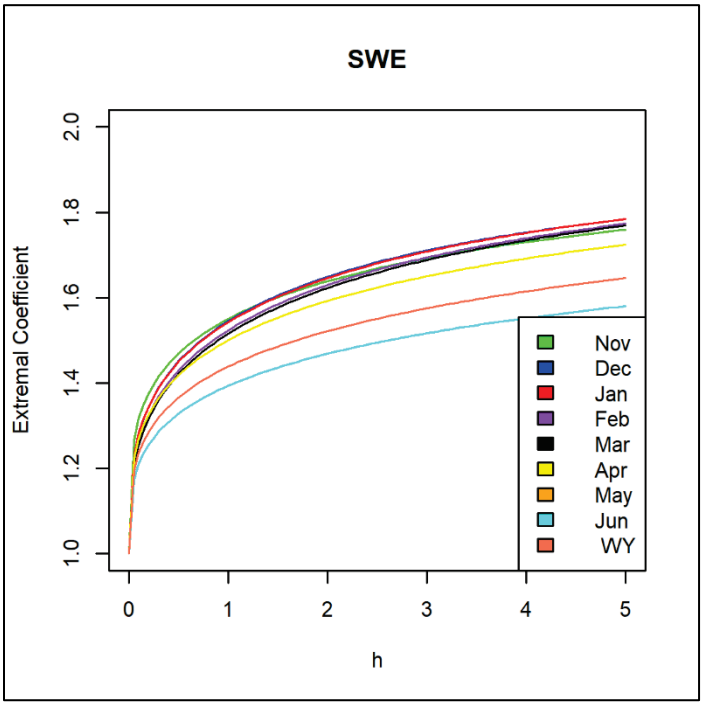
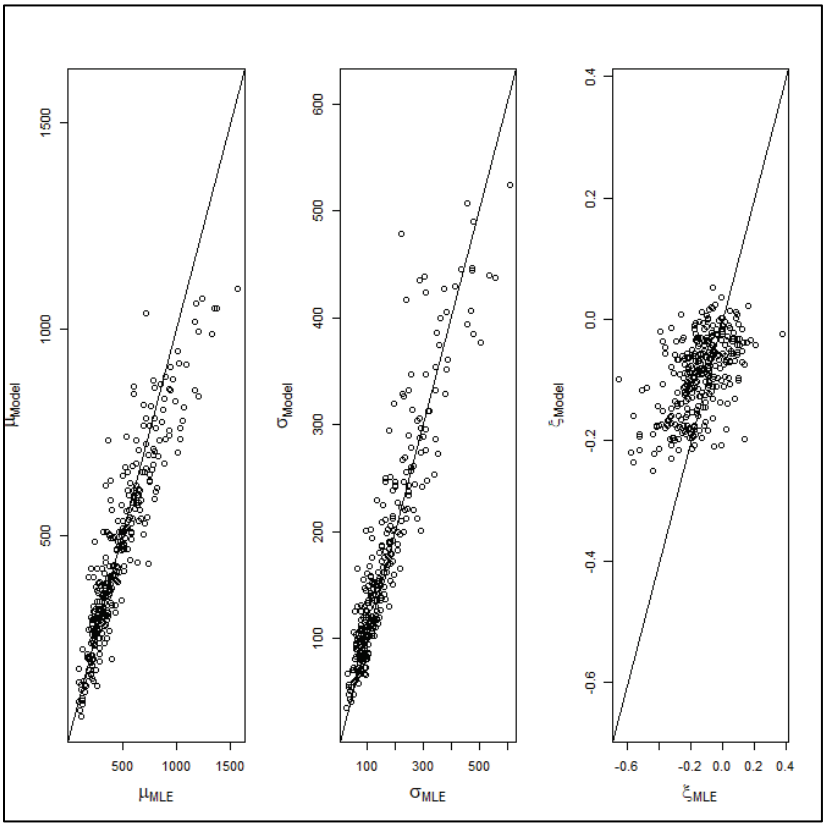


Table 6. Correlation of the fitted SWE trend surface models with their at-site estimates for the calibration and validation sites.

SWE						
	Calibration			Validation		
	Location	Scale	Shape	Location	Scale	Shape
Nov	0.899	0.879	0.571	0.871	0.853	0.643
Dec	0.927	0.918	0.349	0.860	0.894	0.362
Jan	0.925	0.912	0.472	0.878	0.925	0.374
Feb	0.930	0.932	0.572	0.927	0.951	0.553
Mar	0.921	0.923	0.545	0.942	0.932	0.515
Apr	0.933	0.925	0.435	0.811	0.911	0.329
May	0.936	0.882	NA	0.894	0.786	NA
Jun	0.913	0.933	0.476	0.533	0.706	0.314
WY	0.922	0.861	0.385	NA	NA	NA

Figure 7. Comparisons of the fitted trend surface model with corresponding at-site estimates for March SWE (μ= GEV location; σ = GEV scale; ξ = GEV shape; Model = spatial model; MLE = at-site maximum likelihood estimate).



5 Results

The models developed were applied to estimate areal-exceedances for PREC and SWE for the 758 delineated sub-basins of the CRB that correspond to the new CRB hydrology model watershed delineation. These estimates were delivered to the USACE-NWD in individual Environmental Systems Research Institute (ESRI) shapefiles, for PREC by season and duration and for SWE by cool season month and for the WY. Confidence is assigned to the areal exceedance estimates in each of the individual ESRI shapefile products only for associated return periods less than or equal to 200 years. However, the application guidance document that was mentioned in the Introduction, to support the computation of Equation 1 for any arbitrary area within the CRB, can also be used to refine, as needed, the areal exceedance estimates for any of the 758 delineated sub-basins of the CRB that correspond to the new CRB hydrology model watershed delineation. The guidance document can be further used to quantify the uncertainty of estimates for any arbitrary area within the CRB (see Appendix A: Instructions on How to Obtain Exceedance Probability Estimates for a specific area [e.g., sub-watershed]). Of course, with any of the developed models, traditional pointwise return level maps can also be produced. Representative work products are shown in the following appendices:

1. 100-year pointwise return level maps for cool season PREC by duration (See Appendix E).
2. 100-year pointwise return level maps for warm season PREC by duration (See Appendix F).
3. 100-year pointwise return level maps for SWE by month and for the WY (See Appendix G).
4. Mean areal exceedances for cool season PREC for 100-year return period for 758 delineated sub-basins of the CRB (See Appendix H).
5. Mean areal exceedances for warm season PREC for 100-year return period for 758 delineated sub-basins of the CRB (See Appendix I).
6. Mean areal exceedances for SWE for 100-year return period for 758 delineated sub-basins of the CRB (See Appendix J).

References

Beckers, J. V. L., A. H. Weerts, E. Tijdeman, and E. Wells. 2016. "ENSO-Conditioned Weather Resampling Method for Seasonal Ensemble Streamflow Prediction." *Hydrol. Earth Syst. Sci.* 20: 1–11. https://doi.org/10.5194/hess-20-3277-2016

Blanchet, J. 2009. "Max-Stable Processes and Annual Maximum Snow Depth." *6th International Conference on Extreme Value Analysis*, June 23–26, 2009, Fort Collins, CO, United States.

Blanchet, Juliette, and Anthony C. Davison. 2011. "Spatial Modeling of Extreme Snow Depth." *Ann. Appl. Stat.* 5(3): 1699–1725. doi:10.1214/11-AOAS464. https://projecteuclid.org/euclid.aoas/1318514282

Brown, B. M., and S. I. Resnick. 1977. "Extreme Values of Independent Stochastic Processes." *Journal of Applied Probability* 14: 732–739.

Coles, Stuart. 2001. *An Introduction to Statistical Modeling of Extreme Values*. London, UK: Springer-Verlag.

Cooley D., J. Cisewski, R. Erhardt, S. Jeon, E. Mannshardt, B. Omolo, and Y. Sun. 2012. "A Survey of Spatial Extremes: Measuring Spatial Dependence and Modeling Spatial Effects." *RevStat* 10: 135–165.

Cooley, D., P. Naveau, and P. Poncet. 2006. "Variograms for Spatial Max-Stable Random Fields. *Lect. Notes Statist.* 187 373–390. New York: Springer.

Davison, A. C., and M. M. Gholamrezaee. 2011. *Proc. R. Soc. A 2011*. doi: 10.1098/rspa.2011.0412

Davison, A., S. Padoan, and M. Ribatet. 2012. "Statistical Modelling of Spatial Extremes." *Statistical Science* 27(2): 161–186.

de Haan, L. 1984. "A Spectral Representation for Max-Stable Processes." *The Annals of Probability* 12(4): 1194–1204.

Friedman, Jerome, Trevor Hastie, and Rob Tibshirani. 2010. "Regularization Paths for Generalized Linear Models via Coordinate Descent." *Journal of Statistical Software* 33(1): 1–22.

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning: with Applications in R*. New York: Springer.

Hoerl, A. E., and R. W. Kennard. 1970. "Ridge Regression: Biased Estimation for Nonorthogonal Problems." *Technometrics* 12(1): 55–67.

Kabluchko, Z., M. Schlather, and L. deHaan. 2009. "Stationary Max-Stable Fields Associated to Negative Definite Functions." *Annals of Probability* 37(5): 2042–2065.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston. 2012. "An Overview of the Global Historical Climatology Network-Daily Database." *Journal of Atmospheric and Oceanic Technology* 29: 897–910. doi:10.1175/JTECH-D-11-00103.1

NCDC (National Climatic Data Center). 2003. Data documentation for data set 3240 (DSI-3240). National Climatic Data Center, 151 Patton Ave, Asheville, NC 28801-5001, USA. ftp://ftp.ncdc.noaa.gov/pub/data/hourly_precip-3240/dsi3240.pdf

Nicolet, Gilles, Nicolas Eckert, Samuel Morin, and Juliette Blanchet. 2015. "Inferring Spatio-Temporal Patterns in Extreme Snowfall in the French Alps Using Max-stable Processes." *Procedia Environmental Sciences* 26. https://doi.org/10.1016/j.proenv.2015.05.018.

Olinda, R. A., J. Blanchet, C. A. C. dos Santos, V. A. Ozaki, and P. J. Ribeiro Jr. "Spatial Extremes Modeling Applied to Extreme Precipitation Data in the State of Paraná." *Hydrol. Earth Syst. Sci. Discuss*. 11: 12731–12764. https://doi.org/10.5194/hessd-11-12731-2014

Opitz, T. 2013. "Extremal-T Process: Elliptical Domain of Attraction and a Spectral Representation." *Journal of Multivariate Analysis* 122: 409–413.

Pytlak, E., A. McManamon, S. P. Hughes, R. A. Van Der Zweep, P. Butcher, C. Karafotias. J. Beckers, and E. Welles. 2016. "Objective Use of Climate Indices to Inform Ensemble Streamflow Forecasts in the Columbia River Basin - An Initial Review." *AGU Fall Meeting Abstracts*. Session H43C-1430. http://adsabs.harvard.edu/abs/2016AGUFM.H43C1430P

Reich, Brian J., and Benjamin A. Shaby. 2012. "A Hierarchical Max-Stable Spatial Model for Extreme Precipitation." *Ann. Appl. Stat*. 6(4): 1430–1451. doi:10.1214/12-AOAS591. https://projecteuclid.org/euclid.aoas/1356629046

Ribatet, M. 2013. "Spatial Extremes: Max-Stable Processes at Work." *Journal of Société Française de Statistique* (special edition on extreme value theory) 154(2): 156–177.

Ribatet, M. 2015. "Modelling Spatial Extremes with the Spatial Extremes Package." *9th International Conference on Extreme Value Analysis: EVA 2015*. 15 June - 19 June, Ann Arbor, Michigan.

Ribatet, M. 2017. "Modelling Spatial Extremes Using Max-Stable Processes." *Nonlinear and Stochastic Climate Dynamics*. Edited by C. Franzke and T. O'Kane. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316339251

Ribatet, M., C. Dombry, and M. Oesting. 2015. "Spatial Extremes and Max-Stable Processes." *Extreme Value Modeling and Risk Analysis: Methods and Applications*. Edited by D. K. Dey and J. Yan. Boca Raton: Taylor & Francis.

Schlather, M. 2002. "Models for Stationary Max-Stable Random Fields." *Extremes*, 5(1): 33–44.

Schlather, M., and J. Tawn. 2003. "A Dependence Measure for Multivariate and Spatial Extremes: Properties and Inference." *Biometrika* 90(1):139–156. https://www.jstor.org/stable/30042025

Simon, Noah, Jerome Friedman, Trevor Hastie, and Rob Tibshirani. 2011. "Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent." *Journal of Statistical Software* 39(5): 1–13. https://www.jstatsoft.org/article/view/v039i05

Stephenson, A. G., E. A. Lehmann, and A. Phatak. 2016. "A Max-Stable Process Model for Rainfall Extremes at Different Accumulation Durations." *Weather and Climate Extremes* 13: 44–53. https://doi.org/10.1016/j.wace.2016.07.002

Takeuchi, K. 1976. "Distribution of Informational Statistics and a Criterion of Fitting" [in Japanese]. *Suri-Kagaku* 153: 12–18.

Tibshirani, R. 1996. "Regression Shrinkage and Selection via the Lasso." *J. Royal. Statist. Soc B*. 58(1): 267–288.

Tibshirani, Robert, Jacob Bien, Jerome Friedman, Trevor Hastie, Noah Simon, Jonathan Taylor, and Ryan J. Tibshirani. 2010. "Strong Rules for Discarding Predictors in Lasso-type Problems." *Journal of the Royal Statistical Society: Series B* (Statistical Methodology) 74(2): 245–266. https://arxiv.org/abs/1011.2234

Tikhonov, A. N. 1943. "On the Stability of Inverse Problems. *Doklady Akademii Nauk SSSR* 39(5): 195–198.

Zou, H., and T. Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society B* 67(2): 301–320.

Appendix A: Instructions on How to Obtain Exceedance Probability Estimates for a Specific Area (e.g., sub-watershed)

Purpose

Provide instructions on how to generate a frequency curve for PREC (for a specific season and duration) or SWE (for a month (November – June) or the WY) for an arbitrary subarea within the CRB.

Software requirements

R-Studio (download at https://www.rstudio.com/products/rstudio/download/) or R (https://cran.r-project.org/)

R-Studio libraries requirements

The required libraries will be installed the first time the procedure is executed in R. The following libraries are required: (1) SpatialExtremes, (2) maps, (3) raster, (4) SDMTools, (5) rgdal, (6) sp, (7) fitdistrplus, (8) VGAM, (9) lmomco, (10) actuar, (11) FAdist, (12) extremeStat, (13) gamlss, and (14) plotrix..

Input requirements

A polygon ESRI shapefile, in geographic coordinates (Geographic Coordinate System GCS_North_American_1983, and datum D_North_American_1983), for a single area of interest in the CRB (example provided: "UpperTouchetRiver.shp"). The shapefile attribute table must contain one column with value equal to 1 for the area of interest. An example is shown in Figure 8. Figure 8 shows the attribute table as seen in ArcGIS. In that case, column "OBJECTID" indicates a value equal to 1 for the polygon of interest.

Figure 8. Shapefile attribute table for example file.

UpperTouchetRiver

	FID	Shape	OBJECTID	TNMID
▸	0	Polygon	1	{5E68D593-E491-4737-BAFC-8EDDA2AD69CC}

Directory/File structure

The following file structure is provided in the "HowToFinal.zip" file:

HowToFinal/

/ Instructions.docx: Contains instructions on how to generate a frequency curve for PREC or SWE for a specific sub-area specified by the user.

/Code/ CRB_QuantilesUnce.r.r: Code to be modified by the user.

/Example/: An example of run for the user defined sub-watershed, "UpperTouchetRiver".

/Data/: Directory containing the datasets required for the simulations for each season and duration. Do not modify these files.

Run the code to generate sub-basin quantiles

1. Open RStudio or R.
2. Open R code at /Code/CRB_QuantilesUnce.r.
3. Save code with a different name.
4. Modify lines 35 to 46 in the code.
   a. Line 35 – DirSubbasin: Directory containing shapefile for input
   b. Line 36 – shpname: Shapefile name. Example: "UpperTouchetRiver.shp "
   c. Line 37 – Column: Column in shapefile attribute table that contains value equal to 1 for the polygon of interest. Example: "OBJECTID"
   d. Line 38 – PREC equal to one if running precipitation analysis, zero otherwise
   e. Line 39 – Season: if running precipitation, define the season for the analysis (cool or warm)
   f. Line 42 – Duration: 1,2,3,5,10,30,60,90,120-day for cool and 1,2,3,5,10,30,60,90-day for warm season
   g. Line 40 – SWE equal to one if running snow analysis, zero otherwise
   h. Line 41 – Month: if running snow, define the month for the analysis or winter for all winter months. (NOV,DEC,JAN,FEB,MAR,APR,MAY,JUN,WY)

i. Line 43 – NSIMS: Number of simulations (realizations from a max-stable random field). As a rule of thumb, NSIMS= RP (years)/2 where RP is the maximum return period for which high accuracy is required.
j. Line 44 – Prob: Probabilities for the quantiles calculation. Note the following:

$$RP = \frac{1}{1 - Prob}$$

where RP is the return period in years and Prob is the probability for the calculation of the return period. Standard values are 0.9999, 0.999, 0.998, 0.995, 0.99, 0.98, 0.95, 0.9, 0.8, 0.5, 0.2, 0.1, which correspond to return periods of 10000-, 1000-, 500-, 200-, 100-, 50-, 20-, 10-, 5-, 2-, 1.25-, 1-year.

5. Verify line 46: Verify if the path to the /Data/ directory in your computer is correctly specified. Use double slash (\\) to specify path to directories.
6. Select all the lines in the code.
7. Run the script. The run is over when the symbol ">" is shown in the R or RStudio Console.
8. Verify if the output was generated. See the list of output in the "Output" session.

Output

Multiple distributions are fit to the simulation results. The best fit distribution is then selected and used to estimate quantiles. The outputs of the simulations and quantiles estimation are saved at the DirSubbasin directory. Precipitation and SWE quantiles are in millimeters. The following files are generated as the result of the simulation:

1. Quantile_CI.txt: This file contains the following:
   a. The distribution that best fits the data. The best fitting distribution is selected among the following distributions:
      (1) lnorm
      (2) pois
      (3) exp
      (4) gamma
      (5) binom

(6) nbinom
(7) beta
(8) cauchy
(9) unif
(10) logis
(11) weibull
(12) gev
(13) burr
(14) pareto
(15) kappa4

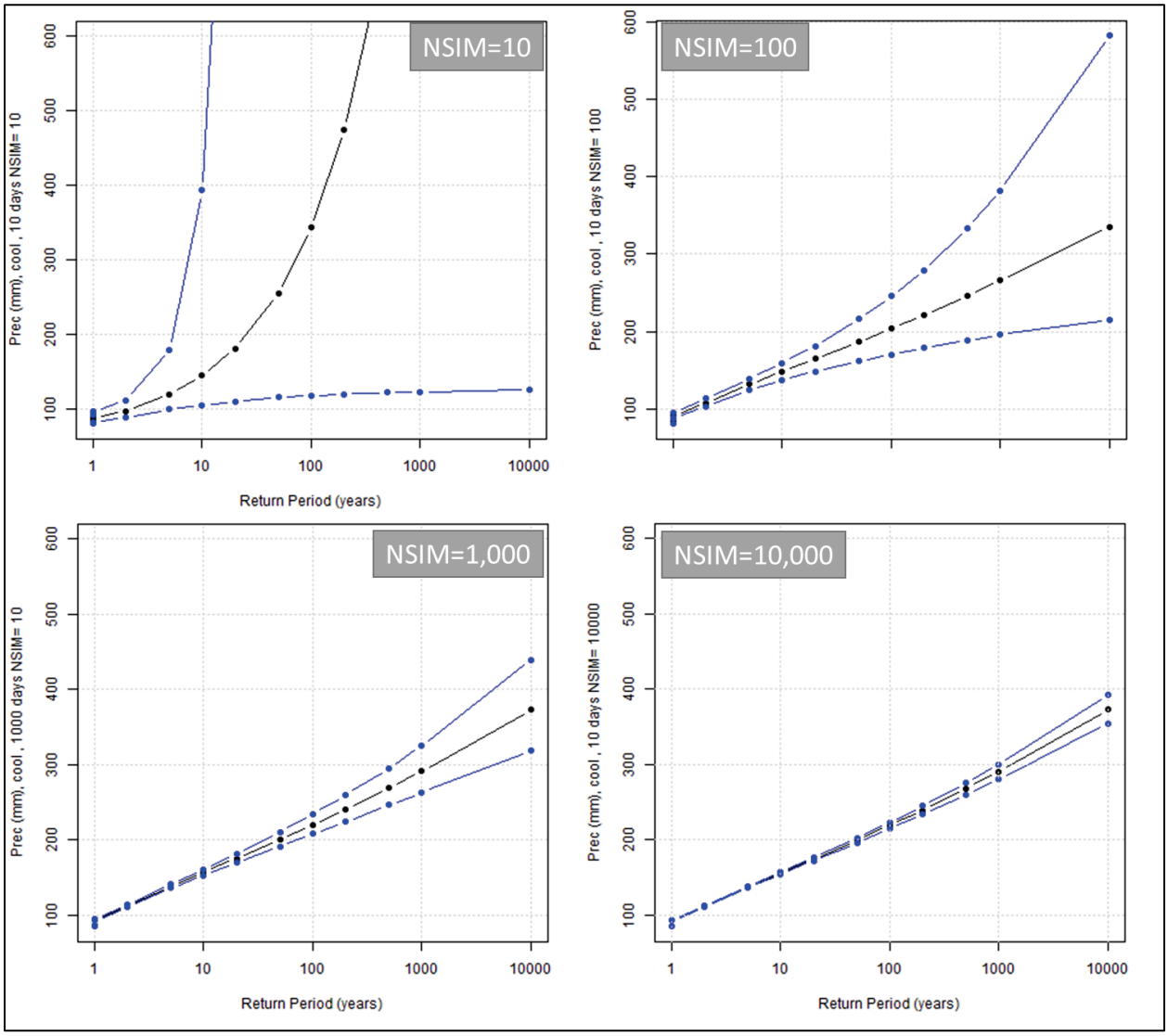
b. The parameters of the selected distribution and the 95% percentile confidence interval for the parameters.
c. Quantiles
(1) Calculate based on the original distribution parameters.
(2) Calculated based on the median bootstrap estimates.
(3) Calculated based on the 95% percentile confidence intervals (2.5% and 97.5%).

2. CDF_ALL.png: a plot containing the empirical and theoretical cumulative distribution function (CDF) for all distributions tested;
3. Result_CI.png: a plot containing precipitation estimates for multiple return periods with confidence interval.
4. R data with the parameters for the calculation and final results (e.g., if shpname <- "UpperTouchetRiver.shp"; column <- "OBJECTID"; PREC<-1; Season<-"cool"; Duration<-"10", then the file name would be UpperTouchetRiver_cool_10.RData).

Number of simulations and uncertainty

Generally, NSIMS = RP (years)/2 where RP is the maximum return period for which high accuracy is required. To guide the user on the application of the results of this analysis, a bootstrap method is applied to generate confidence intervals for simulation. Figure 9 shows how the precipitation estimates and confidence intervals change with changes in the NSIMS.

Figure 9. Uncertainty as a function of the NSIMS.



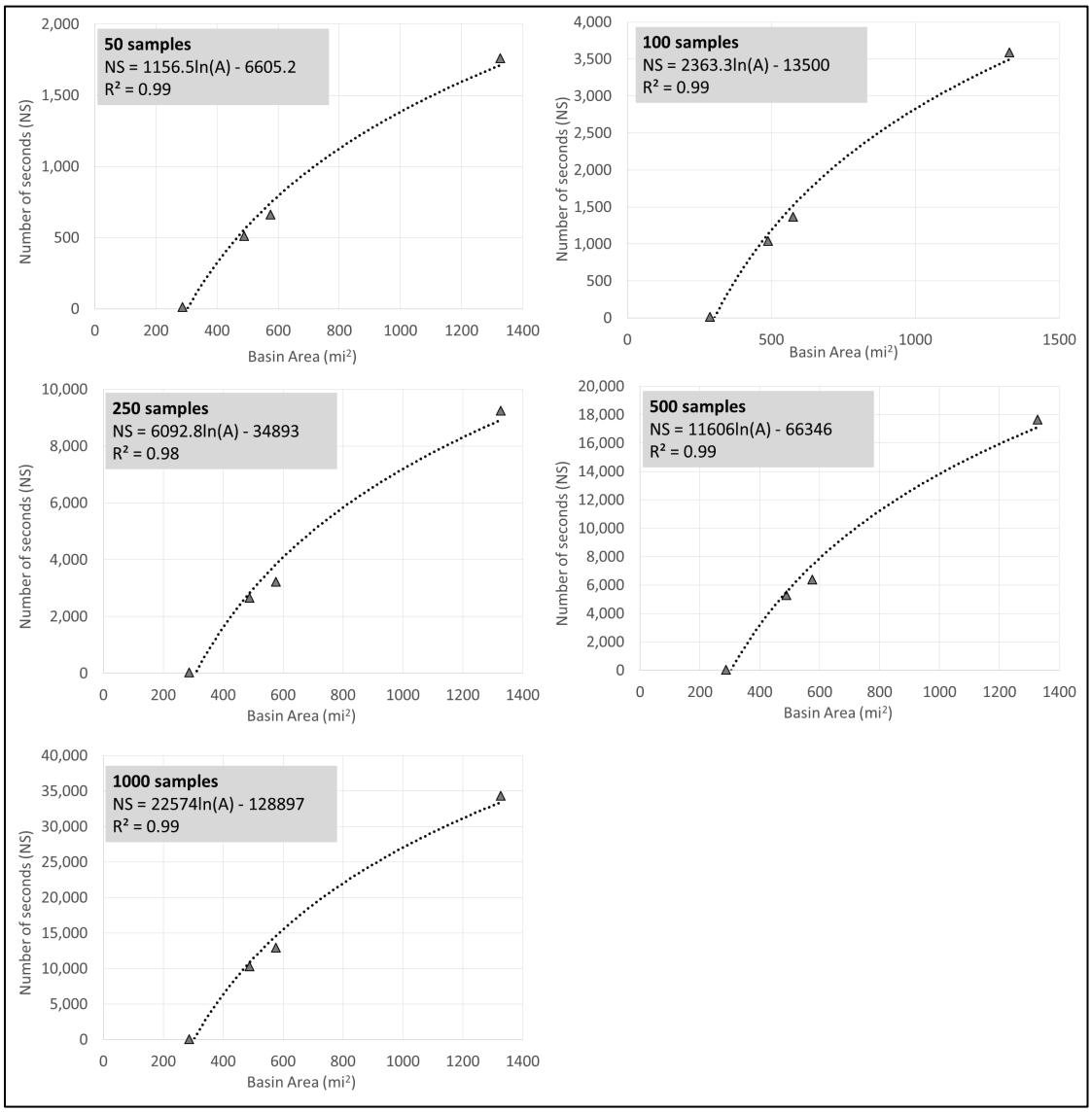
Computation time for simulations

The computation time required to generate multiple independent samples from any one of the max-stable models developed for PREC and SWE depends on the size of the basin. Figure 10 shows how computational time increases as the basin size increases for the generation of 50, 100, 250, 500, and 1,000 samples. Logarithmic curves are fitted to the data, and equations to estimate the run time (number of seconds) as a function of the basin size in square miles are provided. These plots provide an approximate estimate for computation time. The exact time to run any given simulation may vary depending on the computer configuration. The

information presented in Figure 10 was generated using a Windows computer with the following specifications:

- Processor: Intel(R) Xeon(R) W-2155 CPU @ 3.30GHz, 3312 Mhz, 10 Core(s), 20 Logical Processor(s)
- Installed Physical Memory (RAM): 32.0 GB
- Total Physical Memory: 31.7 GB
- Available Physical Memory: 22.5 GB
- Total Virtual Memory 36.4 GB
- Available Virtual Memory: 24.6 GB.

Figure 10. Computational time (seconds) as a function of the basin area (square mile).



Appendix B: Precipitation Data Collection, Quality Control, and Processing

This appendix describes the procedures in the development of the precipitation data set for the CRB Precipitation Time Series Task Order. This task order consists of obtaining and processing existing precipitation time series data collected at precipitation gaging stations located within the entire CRB and a buffer region surrounding the CRB.

Time series of AMP was generated for the warm (May to September) and cool (October to April) seasons for various durations (1-, 3-, 6-, 12-, 24-, 48-, 72-hour, 5-, 10-, 30-, 60-, 90- and 120-day).

Throughout this document, *station* refers to a unique sensor that belongs to a specific network and is identified by a unique station identification (ID). Over time, weather stations may be installed, removed or moved, or the IDs used to describe the station might change. Data from the same weather stations might also be reported by multiple networks. As a consequence, the data measured at a specific location might be recorded by gauges identified by different station names and IDs. To overcome this issue, co-located stations were combined in a cluster of stations and their records were combined into a single time series. This procedure avoids redundancy and allows the creation of longer records that correspond to gauges for a specified location. In this Appendix, *site* refers to a single location which time series might be made of one or more co-located stations. When a cluster of co-located stations is reduced to a single representative point, it becomes a site.

The following deliverables are included as part of this appendix:

- Inventory of stations
- List of duplicate stations
- List of co-located stations
- Inventory of sites, which correspond to the list of stations without duplicated stations and including only one of the co-located stations
- List of closest 40 sites to each site, which will be used to check outliers
- Time series of *raw* precipitation data at the highest temporal resolution available (TXT and DSS format)

- Time series of AMP for the warm and cool season for all station (durations ranging from 1 hour to 120 days)
- Quality control time series of AMP for the warm and cool season for all sites and durations
- Constrained observations correction factors
- AMP removed for each station (due to large percentage of missing data or due to outlier detection)
- Stationarity evaluations.

Project and data collection area

Data were collected for an area larger than the CRB boundaries to guarantee that stations with similar hydro-meteorological patterns to the stations in the CRB boundary were included. The area for which data were collected is shown in Figure 11.

Datasets

Time series of precipitation were obtained from the following sources.

01 – Global Historical Climatology Network (GHCN) Daily Data (National Climatic Data Center Daily [NCDCDaily])

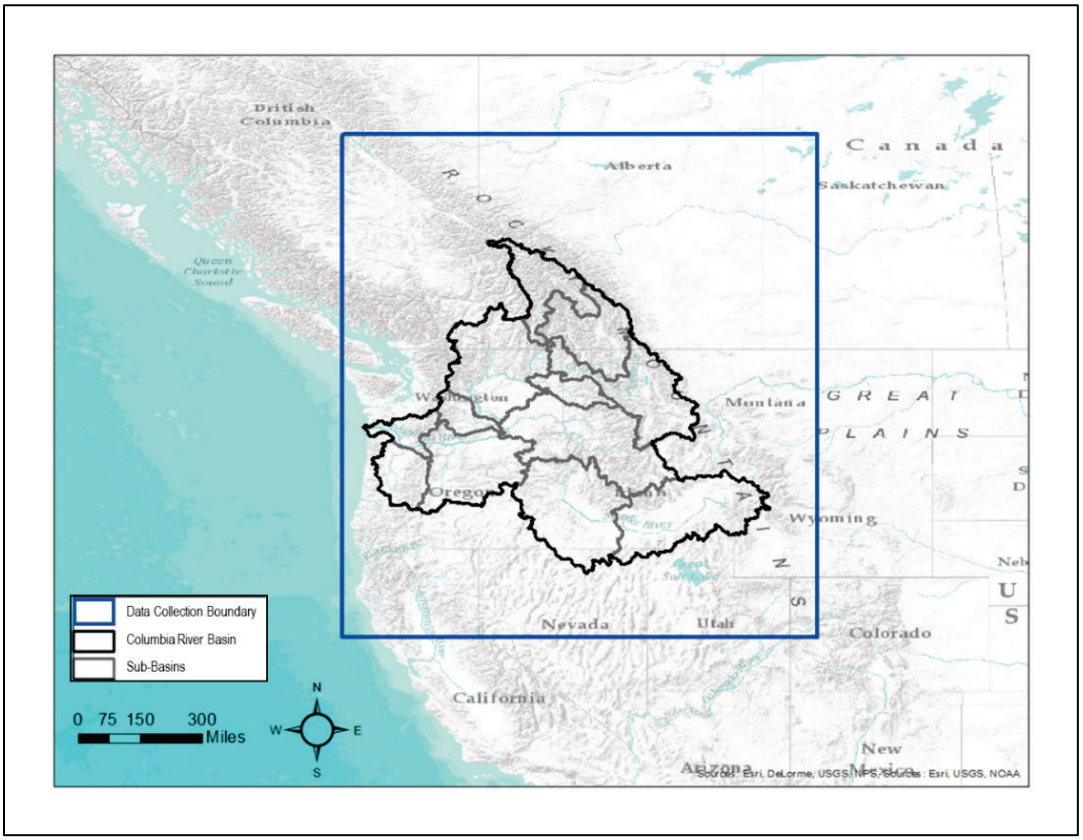
GHCN is a comprehensive database of global climatological daily data which contains records from over 80,000 stations in 180 countries and territories. A description of the database is provided by Menne et al. (2012). One of the main advantages of using this dataset is that daily data has been previously quality controlled based on a standard procedure. This dataset compiles data from multiple sources, including the following:

- United States Historical Climatology Network
- US Cooperative Summary of the Day (NCDC DSI-3200)
- CDMP Cooperative Summary of the Day (NCDC DSI-3206)
- US Cooperative Summary of the Day – Transmitted via WxCoder3 (NCDC DSI-3207)
- US Automated Surface Observing System (ASOS) real-time data (since January 1, 2006)
- US ASOS data for October 2000-December 2005 (NCDC DSI-3211)

02 –NCDC Hourly Data (NCDCHourly)

Hourly precipitation data from the DSI-3240 dataset (hourly precipitation for primarily US locations) was also included in this study. The hourly data are available at ftp://ftp.ncdc.noaa.gov/pub/data/hpd/auto/v2/beta/.The precipitation values provided in this dataset are checked and edited by automated and manual quality control procedures. More information about this dataset is provided by NCDC (2003).

Figure 11. CRB data collection boundaries.



03 – NCDC 15–Minutes Data (NCDC15min):

The 15-minute observations that are used to compute hourly precipitation totals in the National Centers for Environmental Information Cooperative Observers Program (COOP) Hourly Precipitation Dataset. These data are provided for the years 2014 through the present. Therefore, these data were only used to calculate the correction for constrained observations, such as daily observations.

04 – SNOTEL

The SNOTEL daily precipitation data are available at http://www.wcc.nrcs.usda.gov/snow/. Precipitation values in this dataset are reported as accumulations in inches for the climatological year starting October 1. The SNOTEL network is installed, operated, and maintained by the Natural Resources Conservation Service. More information about SNOTEL can be found at http://www.wcc.nrcs.usda.gov/factpub/sntlfct1.html and NCDC (2002).

05 – USACE

These data are stored in the Northwestern Division Corps Water Management System database at http://www.nwd-wc.usace.army.mil/dd/common/dataquery/www/. For this network, data are provided in temporal resolution of 15-minute, 1-hour, 6-hour, and 1-day. The ID for the gauges and location sometimes are the same, but the period of record varies. For example, data for the ARK station were provided at the 1-hour resolution for the period of Aug 1992 to Dec 2017; at 6-hour from August 1992 to August 2018; and at 1-day from November 1974 to December 2017. To avoid losing any data, a network ID number was attributed to each temporal resolution: (1) 51 was assigned for 15-minute; (2) 52 was assigned for 1-hour; (3) 53 was assigned for 6-hour; and (4) 54 was assigned for 1-day. Initially, each time series will be processed independently. However, sites that have the same site ID and location were classified as co-located during the screening process, and the data from multiple temporal resolutions were combined in a unique time series for the gauge site.

06 – Southwest Climate and Environmental Information Collaborative (SCENIC)

Data collected through the SCENIC database. This database includes data from the following networks:

- Eather-Bureau-Army-Navy
- COOP
- Federal Aviation Administration
- World Meteorological Organization
- International Civil Aviation Organization
- GHCN
- National Weather Service Location Identifier
- ThreadEx.

07 – Environmental Canada Daily Data (CADaily)

Daily data provided by the Environment and Climate Change Canada (ECCC), Government of Canada.

08 – Environmental Canada Hourly Data (CAHourly)

Hourly data provided by Meteorological Services of Canada ECCC, Government of Canada.

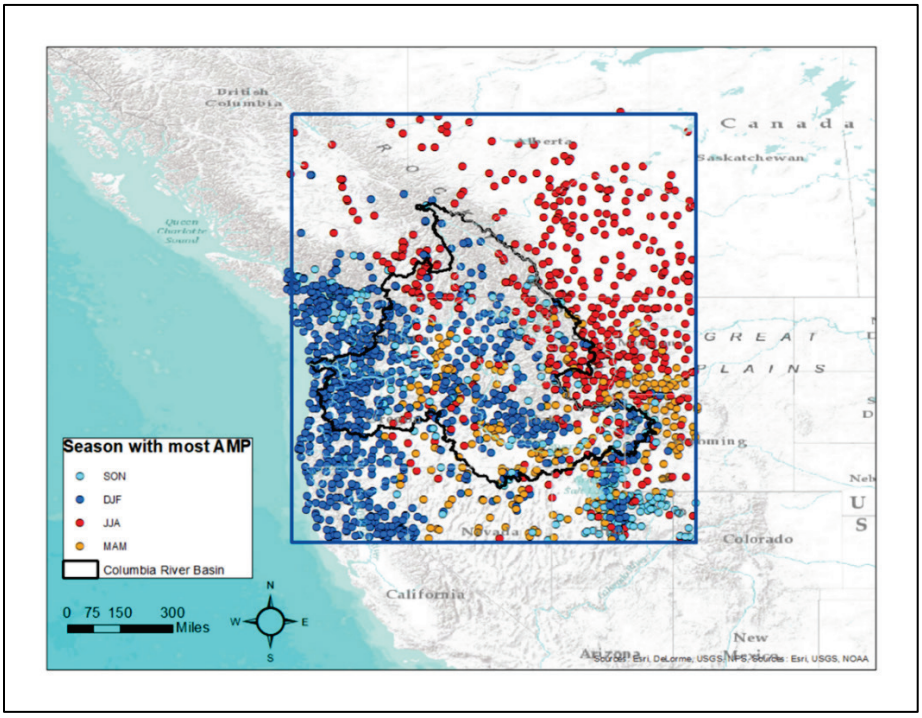
Each station is identified by the number of the network and the station ID. For example: 01-CA1BC000011 refers to a gauge from the NCDCDaily network, with Station ID CA1BC000011.

Seasonality

The following seasons were defined based on the analysis of monthly normal precipitation, seasonality of annual maximum precipitation (see Figure 12), and the seasonality of temperature:

1. Warm Season (May to September). Annual maximum events during these months occur mainly on the eastern part of the basin.
2. Cool Season (October to April). Annual maximum events during these months are mainly associated with atmospheric rivers and are most common in the western part of the basin.

Figure 12. CRB data collection boundaries and seasonality of AMP.



Data collection and processing

The following tasks were performed to collect and process the AMP data.

Data download and formatting

For each network, a Python code was developed to process the original data and to save the data into a standard TXT format with standard units (millimeters) and the inclusion of data flags. The PREC time series were also saved in DSS format. However, note that DSS does allow the inclusion of data flags. The code also generates an inventory of stations for each network (see flowchart in Figure 13).

1. **Station and Site Inventory:** All stations (all networks) were evaluated to identify duplicated and co-located stations. Duplicates occur when two stations are located at the same geospatial location (latitude and longitude), have the same period of record, and at least 80% of the AMP data from one station are the same as the data from the other station (see flowchart in Figure 14). Since different networks store latitude and longitude with different precision, sites located close to each other were also checked to guarantee they do not have the same data and period of record. If the distance between the stations is less than 1.5 km apart, then the station is considered co-located if both gauges present different data and is considered duplicated if the period of record is the same, and at least 80% of the AMP data from one station is the same as the data from the other station data. Duplicated and co-located stations were represented by one station in the final inventory of sites. Files were created containing the station ID of duplicated and co-located sites. If the original network does not contain elevation data for the station, elevation data were calculated based on the HydroSHED (15 sec) dataset available at https://www.hydrosheds.org/. This dataset was selected since it covers Canada and the United States, and the data for both countries were produced in a consistent way. HydroSHEDS is based on high-resolution elevation data obtained during a Space Shuttle flight for the NASA Shuttle Radar Topography Mission. A file containing the station ID for the closest 40th stations to each station was also generated. This file was used to evaluate outliers in AMP time series (see flowchart in Figure 14).
2. **Calculate Constrained Observations Correction Factors:** 15- and 60-minutes gauges were used to calculate the constrained

observation correction factors. The need for constrained observation correction arises from using rainfall observations that are usually based on a fixed accumulation time period (e.g., 00:00 to 23:59) to calculate AMP. For these time series, it is possible that a larger value of AMP could have been observed if a fixed beginning and ending period were not considered. To account for this likely underestimation of AMP, a correction factor is applied for AMP values obtained from data with temporal resolution larger than 15 minutes. To calculate the correction factors, the 15-minutes and 60-minutes resolution data were used to calculate AMP based on two methods: (1) Unconstrained, which is based on the moving average sums and guarantees that the highest value is selected, and (2) Constrained, which is based on accumulations based on a pre-defined beginning and end time that underestimates the AMP. The ratio between unconstrained and constrained AMP were calculated, and the mean, median, 75th and 90th percentile of the ratio were recorded for gauges that have at least 5 years of data. Different correction factors were calculated for each season.

3. **Data Processing 1:** In this task, AMP time series were generated for all stations. Auxiliary data that supports the quality assurance/quality control was also computed, including (1) percent of missing data in season (MDS), (2) number of days with rain > 0 (NDR), (3) total seasonal rainfall (TSR), (4) percent of missing values for the time of occurrence of the AMP Missing Data Time (MDT), (5) flags, and (6) percent of values flagged for the time of occurrence of the AMP (see flowchart in Figure 15).

Figure 13. PREC time-series download and re-formatting flowchart.

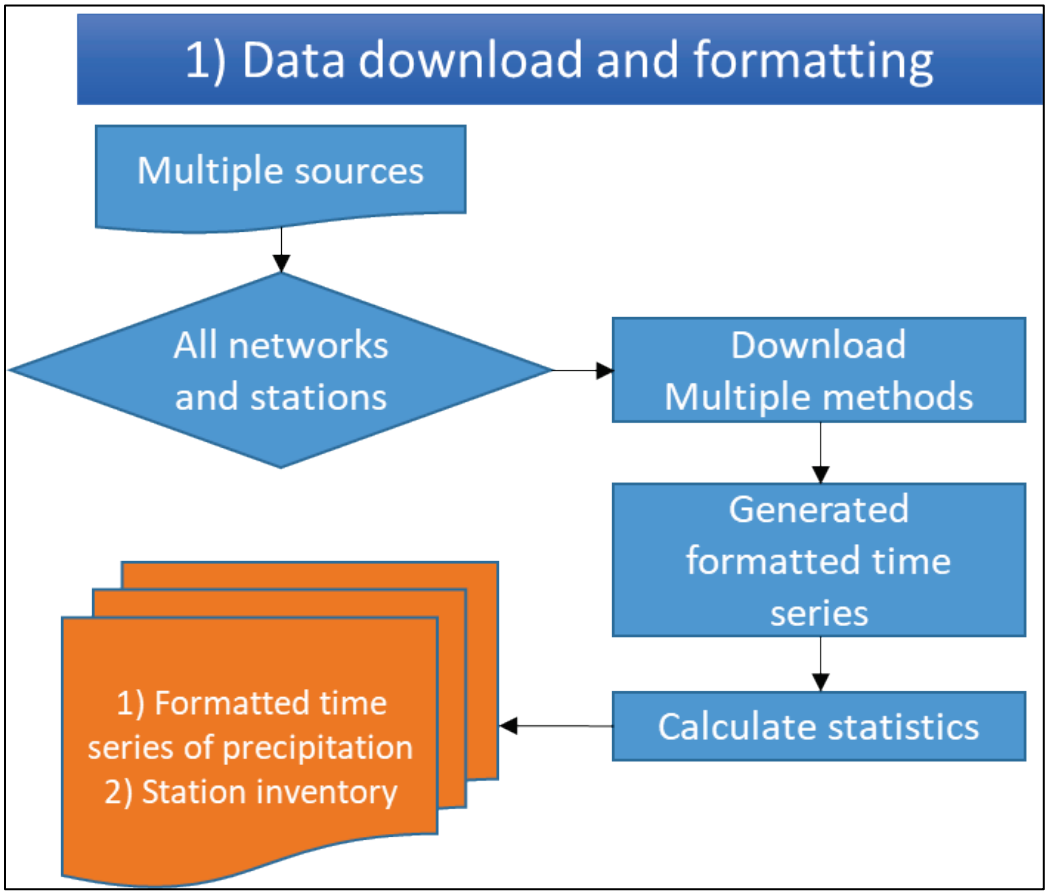


Figure 14. Station screening flowchart.

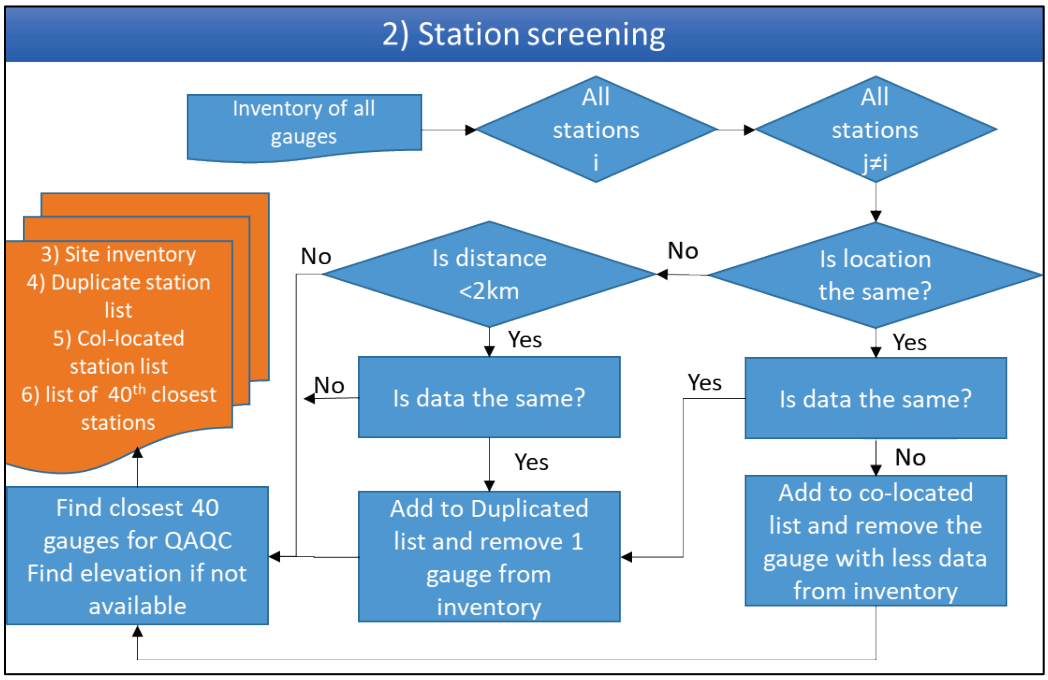
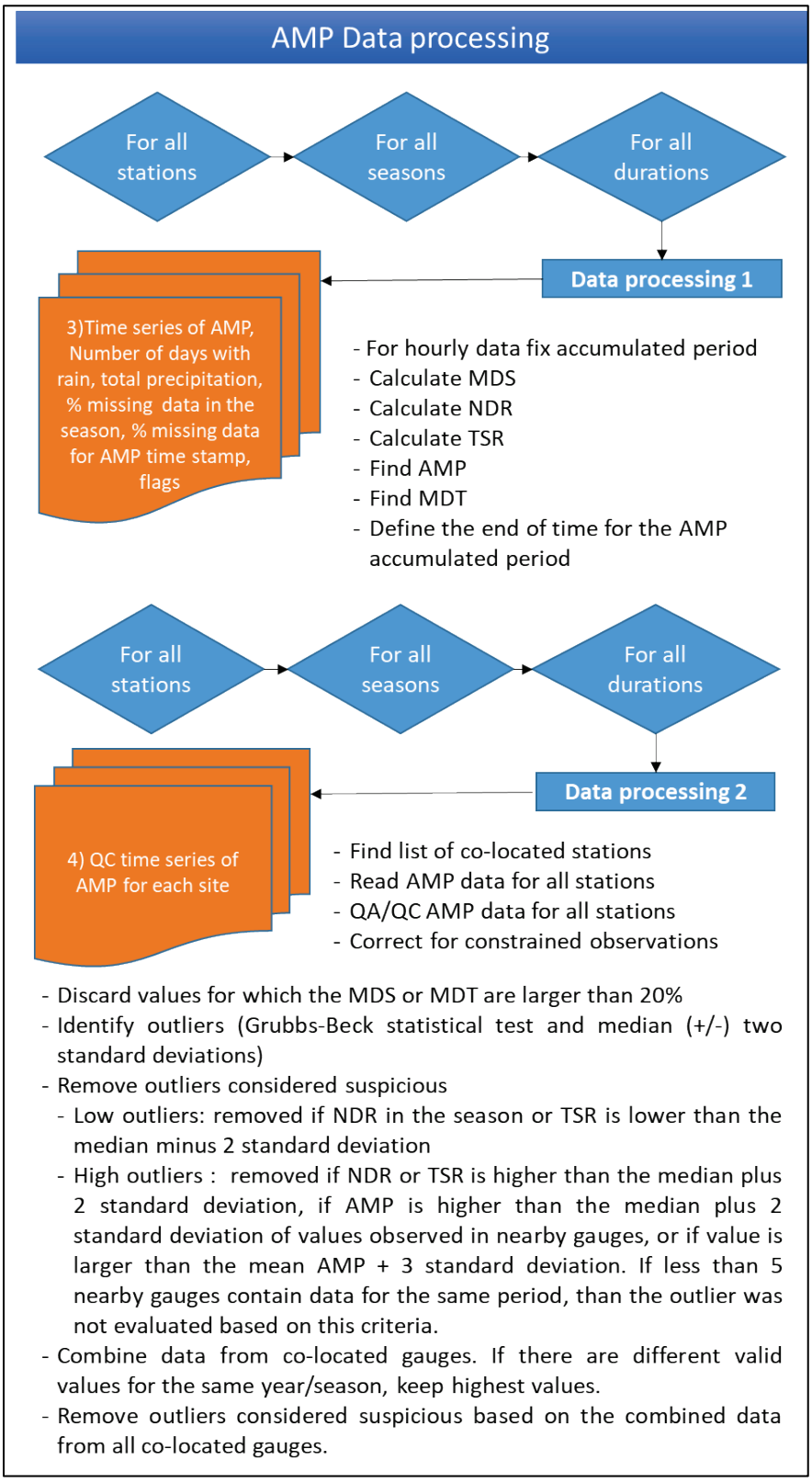


Figure 15. AMP processing and quality control flowchart.



4. **Data Processing 2:** In this task, quality-controlled AMP time series were generated for all sites (see flowchart in Figure 15). Co-located stations were first identified. The AMP for each station was corrected for constrained observations and quality control to identified and remove outliers. The correction for constrained observation applied the median correction factor for the closest station for the specific season. Only one factor is applied for hourly data that correspond to the correction from 15 minutes to 60 minutes. Two factors were applied for data resolution greater than 60 minutes: (1) Factor 1: corrects from 15 minutes to 60 minutes; and (2) Factor 2: corrects from 60 minutes to the resolution of the data. The quality control starts with removing all values for which MDS or MDT are more than 20%. Within the remaining values, outliers are detected based on the Grubbs-Beck statistical test and the median (+/-) two standard deviations thresholds. The identified outliers were then checked by consistency with surrounding gauges and with the seasonal information (number of rainy days and season total rainfall). Outliers were removed in case the value also failed either the Lower and High Outlier Checks:

    a. Low Outliers
        (1) NDR for the season that the outlier occurs is less than the median NDR for the gauge minus two standard deviations.
        (2) TSR for the season that the outlier occurs is less than the median TSR for the gauge minus two standard deviations.
    b. High Outliers
        (1) NDR for the season that the outlier occurs is greater than the median NDR for the gauge plus two standard deviations.
        (2) TSR for the season that the outlier occurs is greater than the median TSR for the gauge plus two standard deviations.
        (3) Collected AMP for the 40 closest stations for the same year (when available). The point is removed if the AMP for the station is larger than the median AMP plus two standard deviations for all 40 stations. If fewer than five nearby gauges are available, the outlier is not removed based on this criteria.
        (4) AMP is larger than the median AMP plus three standard deviations.

Files organization

The following folder and files are included as deliverables for this submittal:

Appendix: Contains a digital copy (Word format) of this appendix.

2-PrecData

Contains the collected time series of precipitation. Each network provides data in different formats. The precipitation data obtained from each network were converted into a unique format for all networks. The data were saved in TXT, which also contains flags, and DSS file format, which only contains precipitation values. All precipitation data are in millimeters.

3-AMP

Contains the output of the "Data Processing 1" routine, which corresponds to the time series of AMP for all sites and networks for the cool (directory CoolS) and warm (directory WarmS) seasons. No quality control of the data was performed in this step. The file name corresponds to the duration (e.g., 1 hour, 1 day) and the station ID (e.g., 01-CA1AB000001, 06-US1ORLA0065). The files contain the following information:

- Year: Year of the AMP.
- AMP(millimeters): AMP in millimeters
- EndDate: This corresponds to the end date of the accumulation period for the AMP. This is especially important for larger accumulation (e.g., 5 days or more).
- PorMissingSeason: Percentage of missing values in the season.
- PorMissingAccum: Percentage of missing values in the accumulation period. This value is more relevant for larger accumulation periods (e.g., 5 days or more).
- RainyDays: Number of rainy days in the season.
- TotalRain: Total accumulated rain in the season.
- Flags: For the original duration, the file shows the original flag. For other accumulations, the percentage of the values that are flagged in the accumulation period is provided.

Since these files contain more information than only the time series of AMP, these files were saved in comma separated values (CSV) format (and not in DSS format). The file name structure is the following:

[Season: CoolS or WarmS] +"/" + [duration]+"_"+[Station ID]+".csv"

This folder also contains the ConstrainsCorrection.csv file, which lists the constrained observation correction factor for all 15-minutes and 60-minutes gauges for multiple duration and for the cool and warm season.

5-StationScreening

Contains the results of the station screening processes. The following files are included:

- "Colocated.csv": List of collocated gauges.
- "Duplicated.csv": List of duplicated gauges.
- "Eliminated.csv": List of eliminated gauges because of lack of valid data (number of AMP < 4).
- "InvNoDuplicates.csv": Inventory of sites (Stations that are not duplicated and only one station representing cluster of co-located stations).
- "NearbyGauge.csv": List of 40th closest stations to each site.
- "Summary.csv": Summary of the screening process. A total of 10,954 stations are available in the area. Seventy-eight stations were duplicated, 2,799 stations were co-located. A total of 6,543 sites were identified.

5-QCAMP

Contains the quality-controlled AMP for all sites and seasons. The following files are included:

- *[Season: CoolS or WarmS]* + [duration] + "AMPMatrix.csv": all valid AMP data stored in a matrix of site versus year of data. The file includes information for the site: Station ID, Lat, Long, Elevation and number of valid AMP (NValYears).
- *[Season: CoolS or WarmS]* + [duration] + "AllAMP.csv": all valid AMP data stored in sequence containing all the auxiliary information listed in 3-AMP: Year, AMP(millimeter), EndDate, PorMissingSeason, PorMissingAccum, RainyDays,TotalRain, and Flags. This file can be

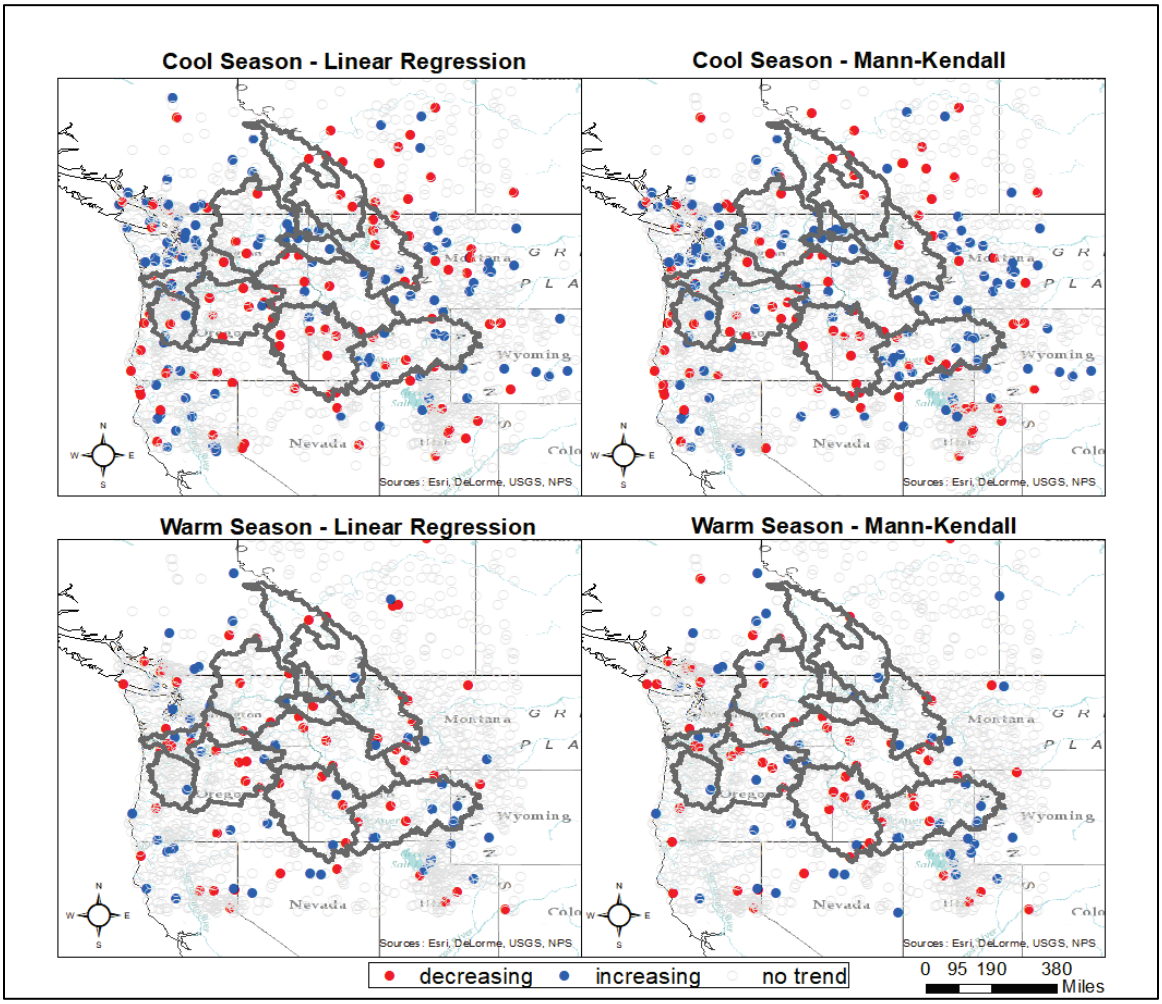
used to evaluate the data; for example: evaluate specific events (sort data by date) or verify high and low values of AMP (sort by AMP values).

- *[Season: CoolS or WarmS] + [duration] + "AMPConstObsCorr.csv":* contain the correction factor for constrained observations. Initially, two constrained correction factors were calculated: (1) Coef15: Correction from 15 minutes to 1 hour, calculated based on 15-minutes precipitation data. Only 11 stations had enough data to calculate this coefficient; (2) Coef: Correction from 1 hour to the duration being evaluated calculated based on hourly data. For each station, the median value of the ratio between non-constrained and constrained AMP was calculated and used in the analysis. The coefficient for the closest station to the site being evaluated was used. The Coef value was used to correct daily AMP. The maximum correction factor applied to the data is 1.22 for the Cool season and 1.25 for the Warm season.
- *[Season: CoolS or WarmS]* + [duration] + "Stationarity.csv": stationarity tests for the mean were performed for sites with more than 30 years of valid data. Stationarity test for the variance was performed for sites with more than 60 years of data. Results include the following:
  - LRSlope: Slope of the linear regression model.
  - LRrvalue: Intercept of the linear regression model.
  - LRpvalue: p-value of the linear regression model. Trend is observed if p-value<0.05.
  - LRtrend: no trend, increasing or decreasing trend.
  - MKtrend: Mann-Kendall test result. False indicates no Trend.
  - MKh: no trend, increasing or decreasing trend for Mann-Kendall.
  - MKpmk: p value of the significance test for Mann-Kendall.
  - MKz: normalized test statistics.
  - LeWv: results of the Levene test. The Levene test tests the null. hypothesis that all input samples are from populations with equal variances.
  - Lepv: p-value of the significance test for the Levene test. P-value lower than 5% indicates trend in the variance.

Figure 16 shows the results of the stationarity test. The large majority of the sites do not present any trend. No clear spatial pattern is observed for the sites that trends were identified. Trends were not removed for the calculation of AMP.

- *[Season: CoolS or WarmS]* + [duration] + "ValidInv.csv": Inventory of sites with valid data.
- *[Season: CoolS or WarmS]* + [duration] + "EntireInv.csv": complete inventory of stations, including duplicated and co-located stations.
- *[Season: CoolS or WarmS]* + [duration] + "RemovedAMP.csv": AMP values that were excluded of the time series due to any of the quality control procedures, including exceeding 20% missing data in the season or 20% missing data in the accumulation period, or values that were identified as outliers and failed the validation test.

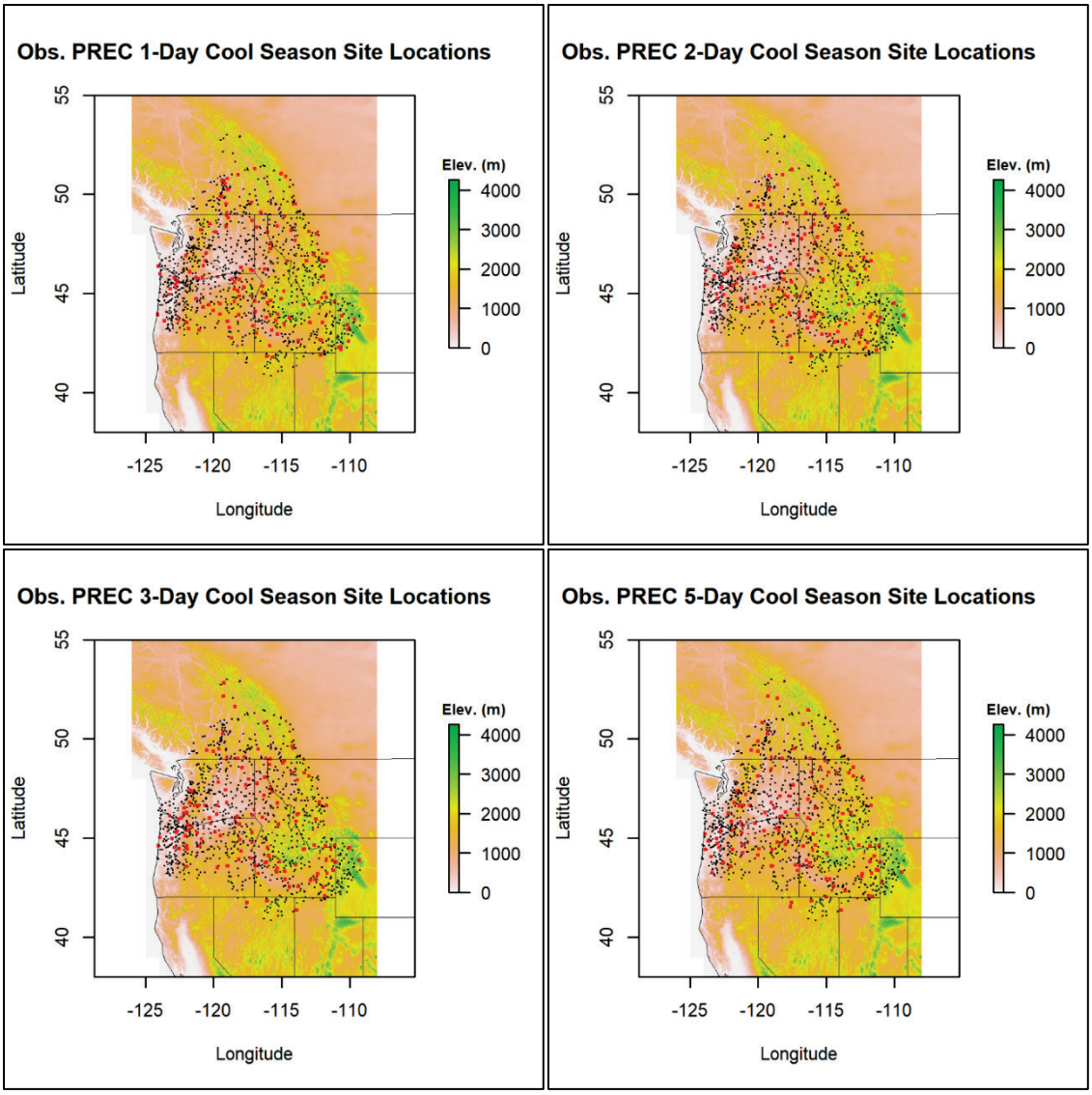
Figure 16. Trends in daily AMP for the cool and warm season based on linear regression and Mann-Kendall.



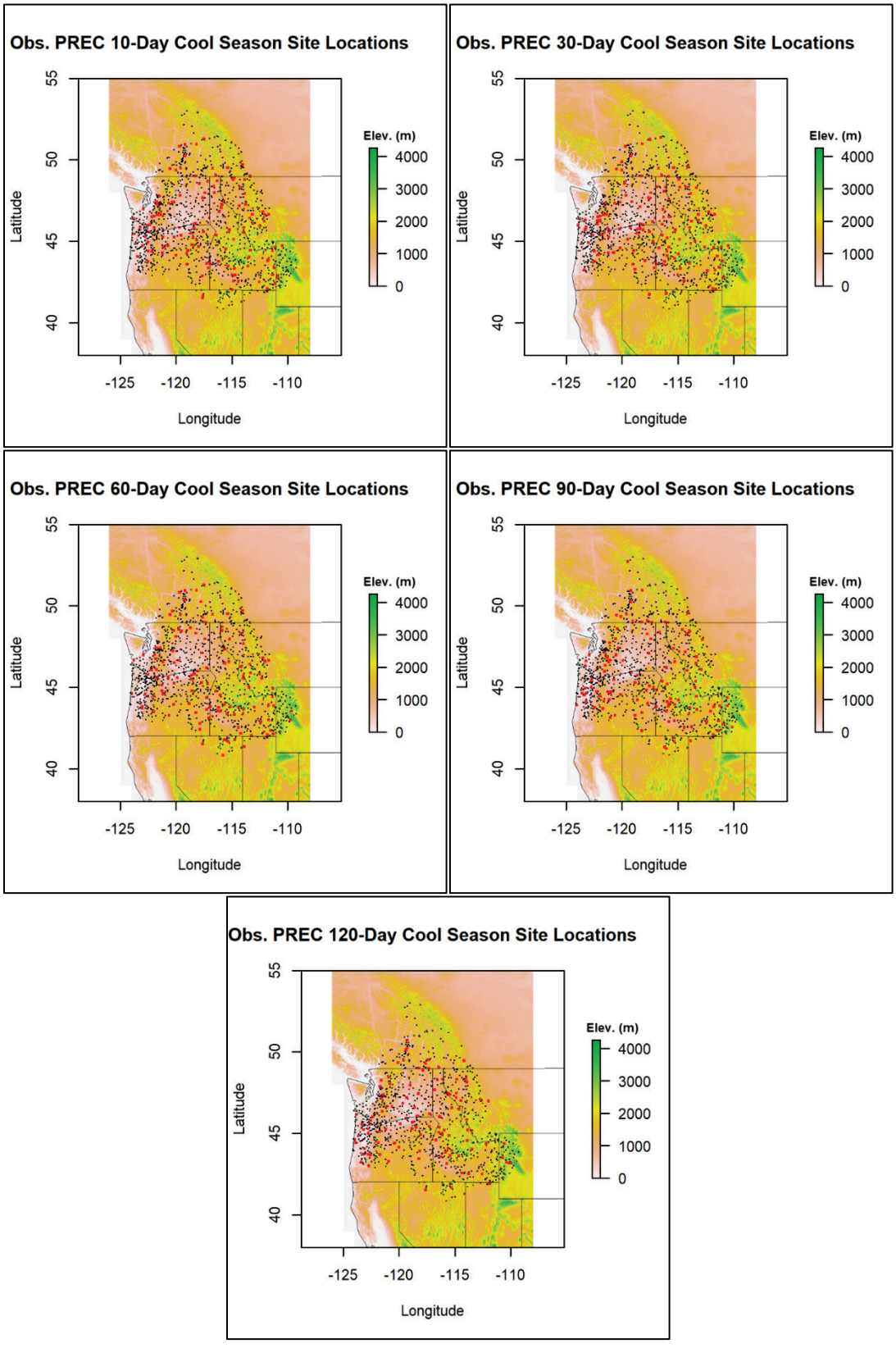
Appendix C: Site Locations for Calibration and Validation of PREC Models by Season and Duration

Figure 17. DEM, state line, and PREC observation site location data by season and duration (▪=Calibration, ▪=Validation).

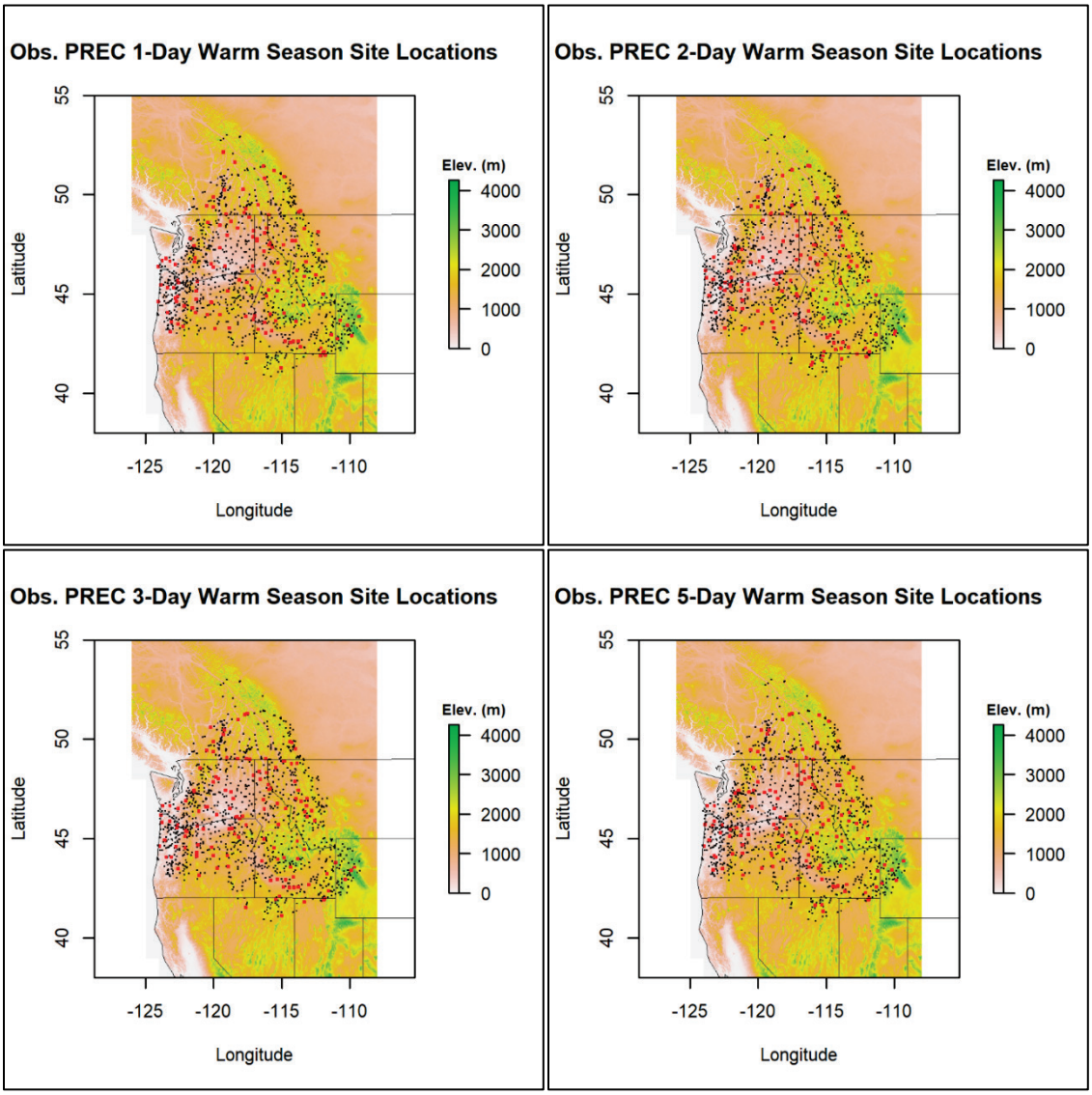
▪=Calibration, ▪=Validation



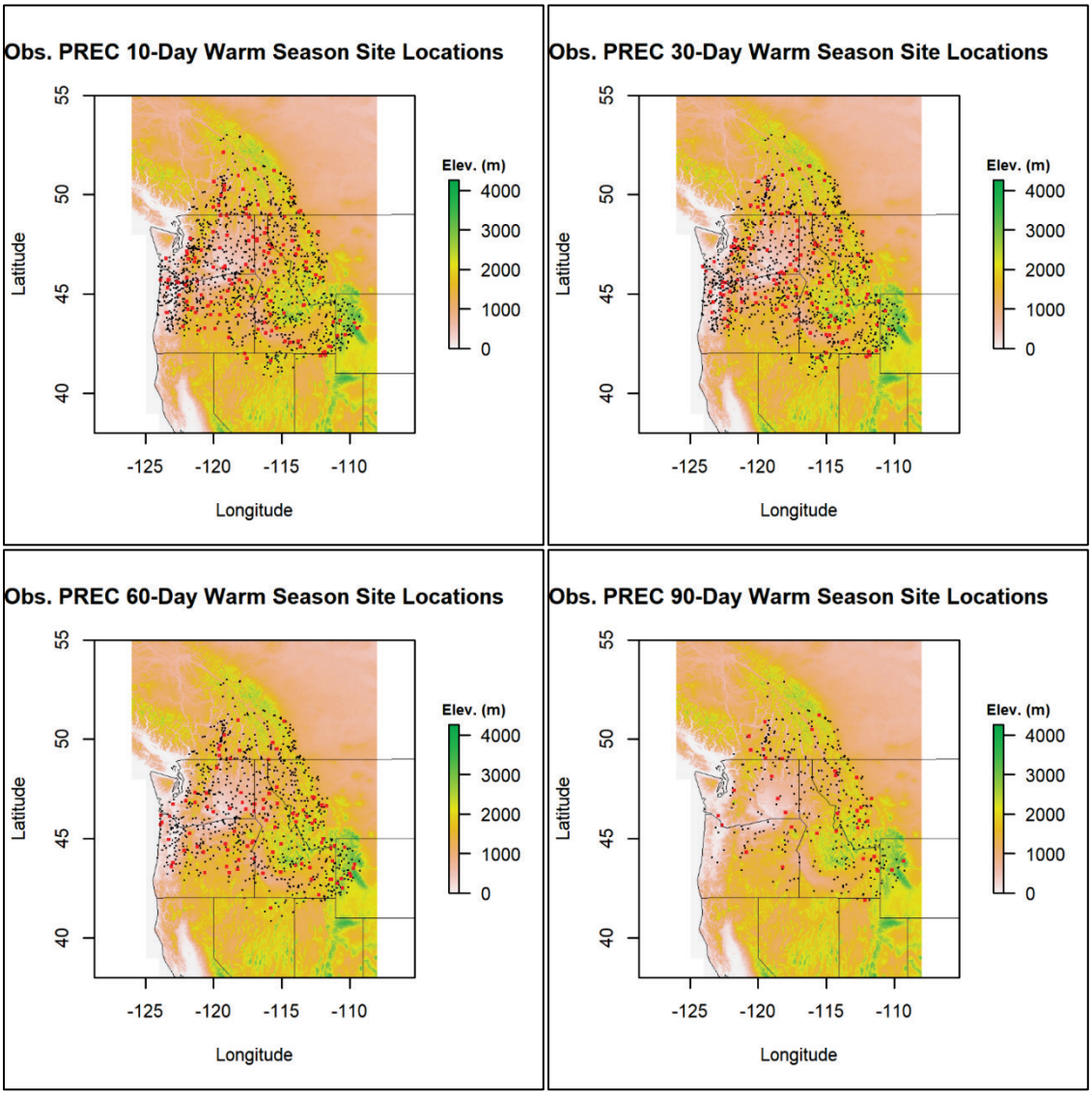
▪=Calibration, ▪=Validation



▪=Calibration, ▪=Validation



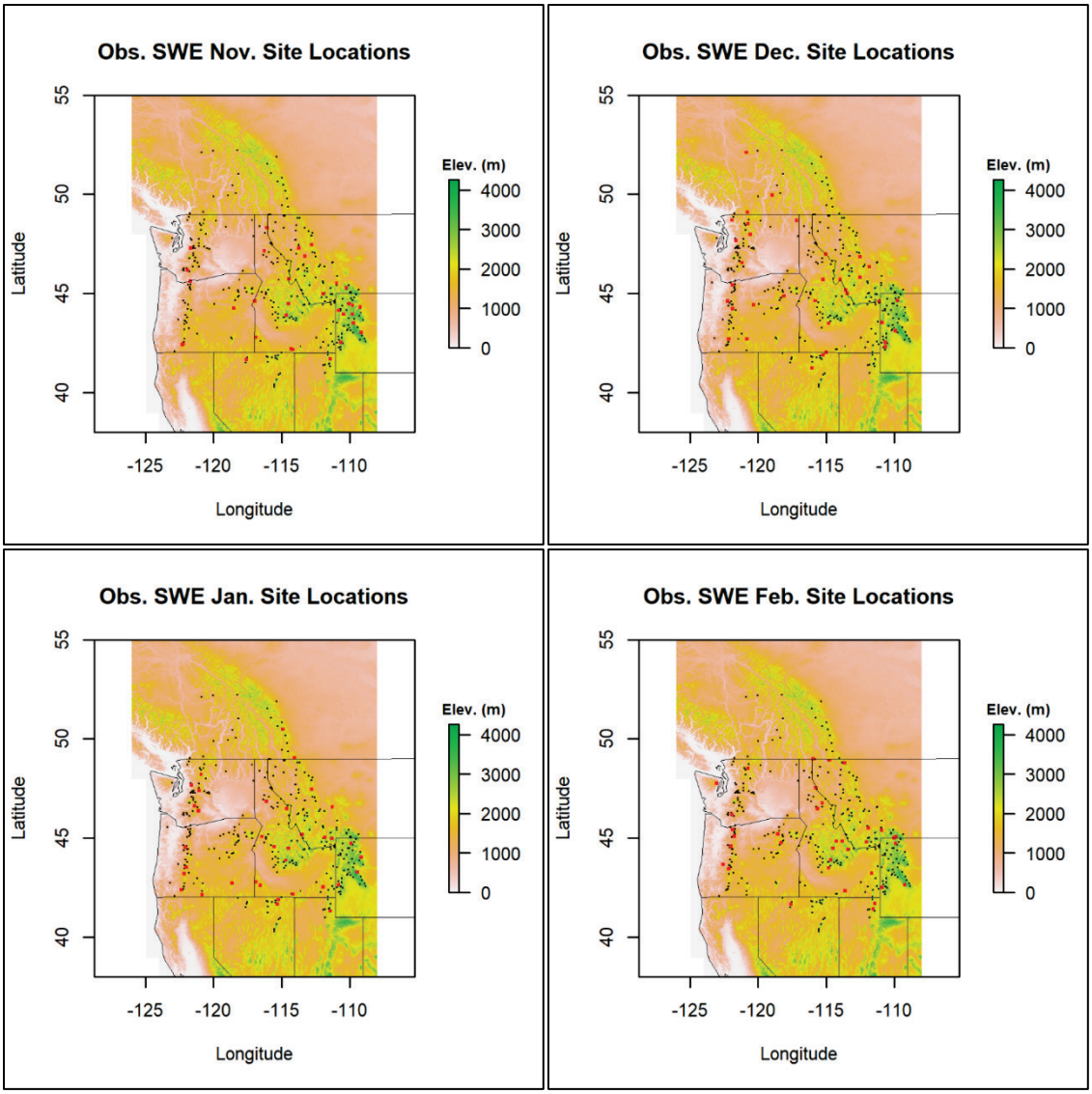
▪=Calibration, ▪=Validation



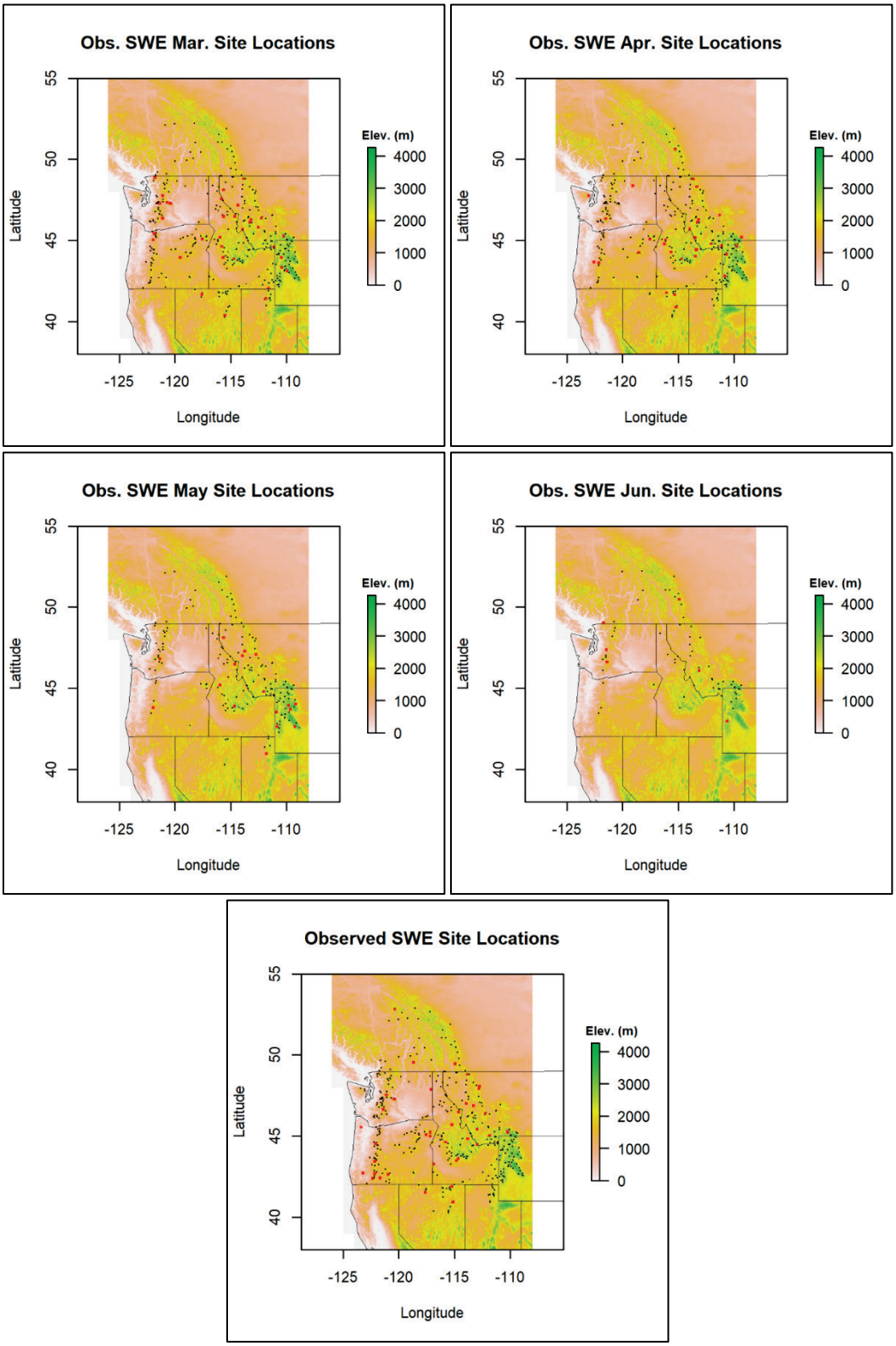
Appendix D: Site Locations for Calibration and Validation of SWE Models by Month and for the WY

Figure 18. DEM, state line, and SWE observation site location data by month and for the WY (▪=Calibration, ▪=Validation).

▪=Calibration, ▪=Validation

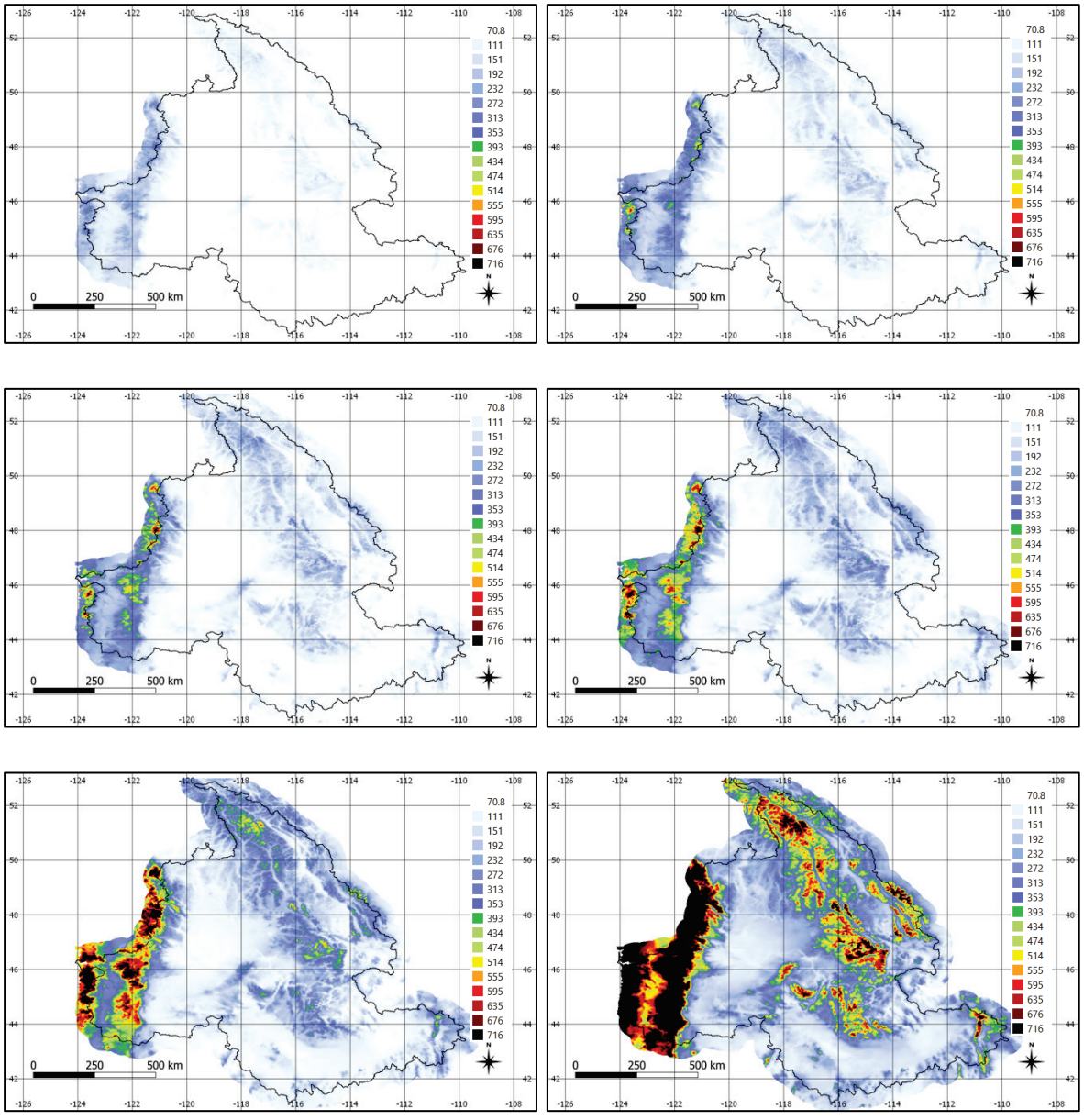


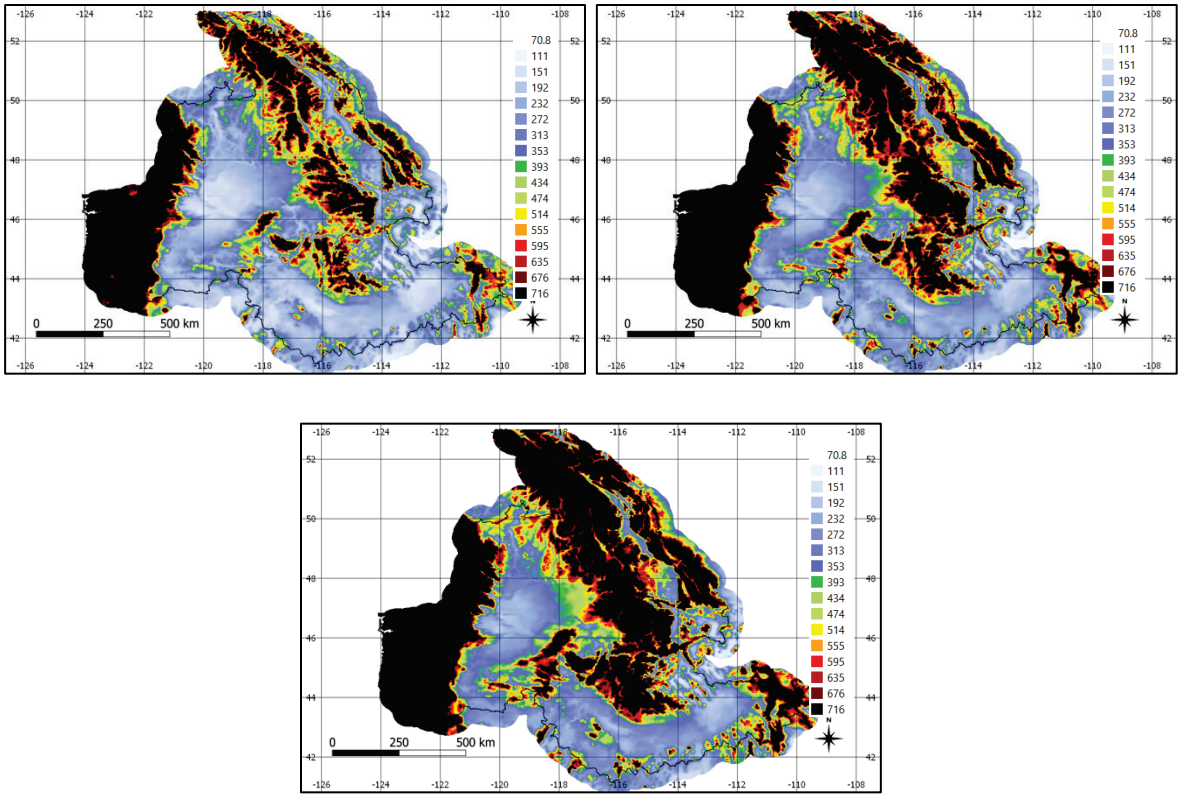
▪=Calibration, ▪=Validation



Appendix E: Cool Season PREC Pointwise 100-Year Return Levels for the CRB

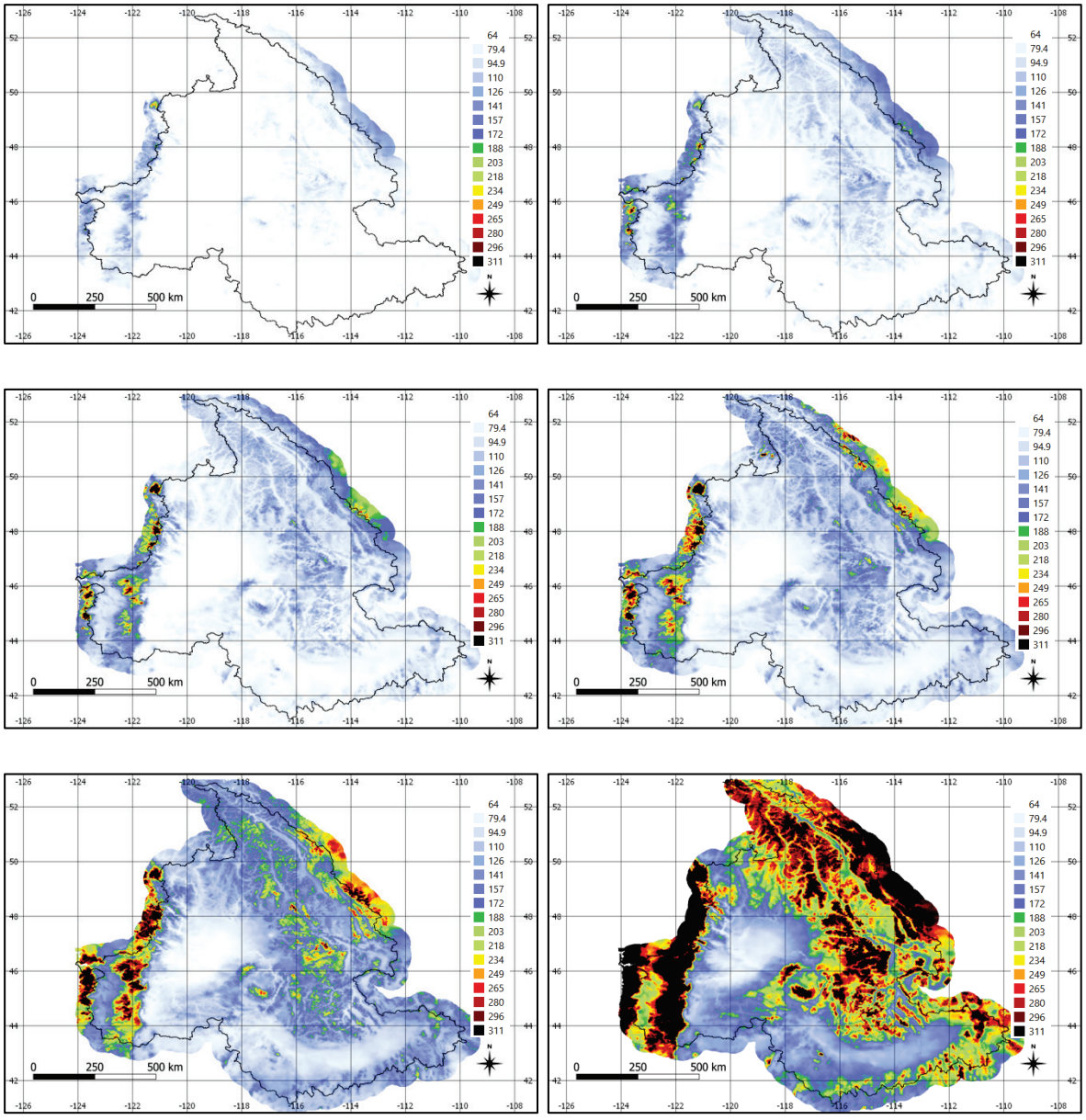
Figure 19. Cool season PREC pointwise 100-year return levels for the CRB (for 1-, 2-, 3-, 5-, 10-, 30-, 60-, 90-, and 120-day durations) (values in millimeters).

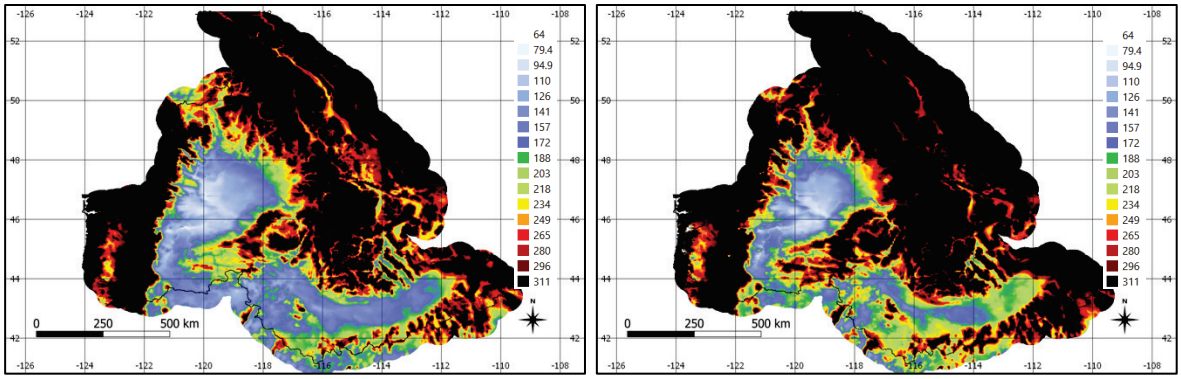



70.8
111
151
192
232
272
313
353
393
434
474
514
555
595
635
676
716
0
250
500 km
N

Appendix F: Warm Season PREC Pointwise 100-Year Return Levels for the CRB

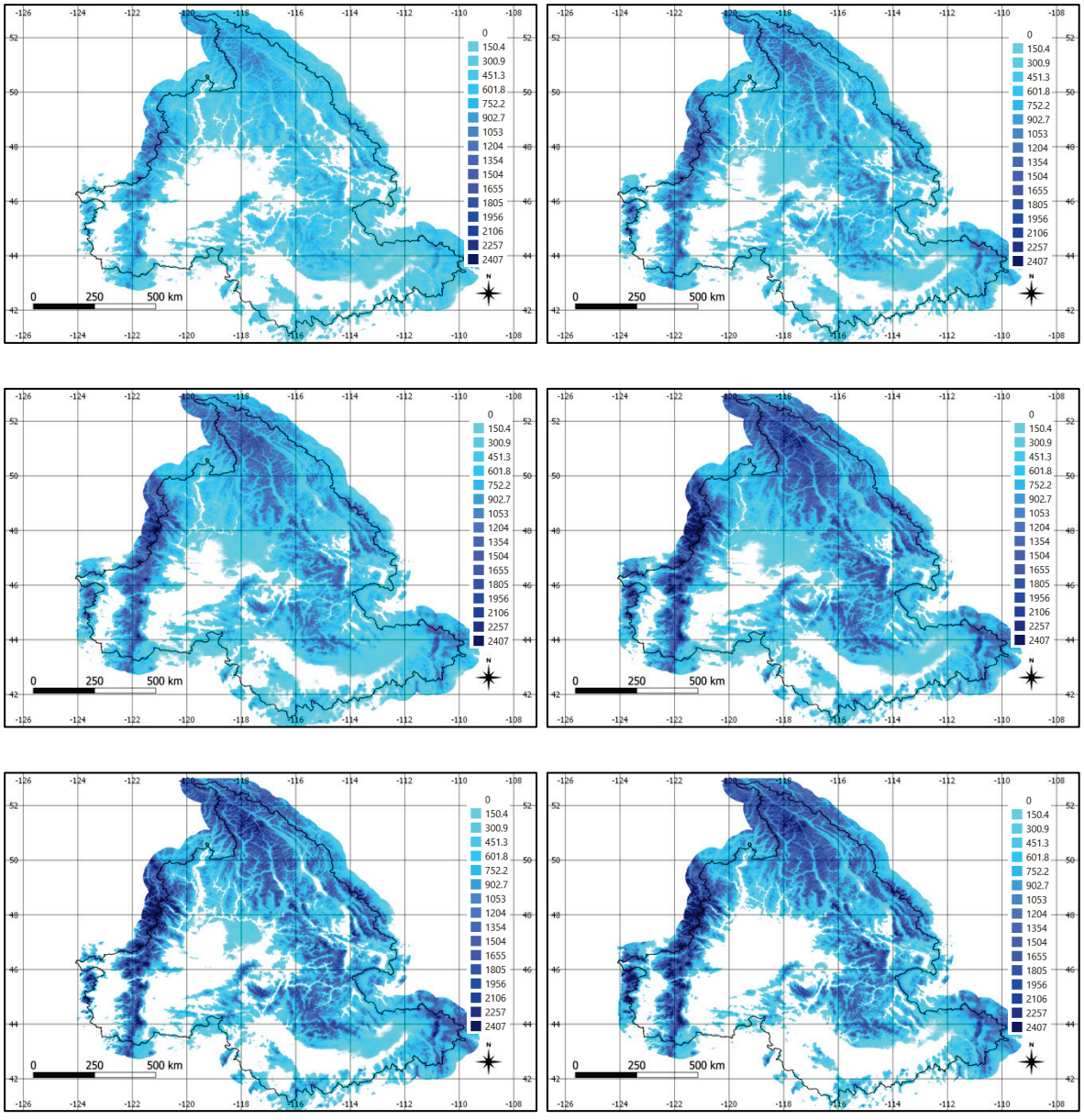
Figure 20. Warm season PREC pointwise 100-year return levels for the CRB (for 1-, 2-, 3-, 5-, 10-, 30-, 60-, and 90-day durations) (values in millimeters).

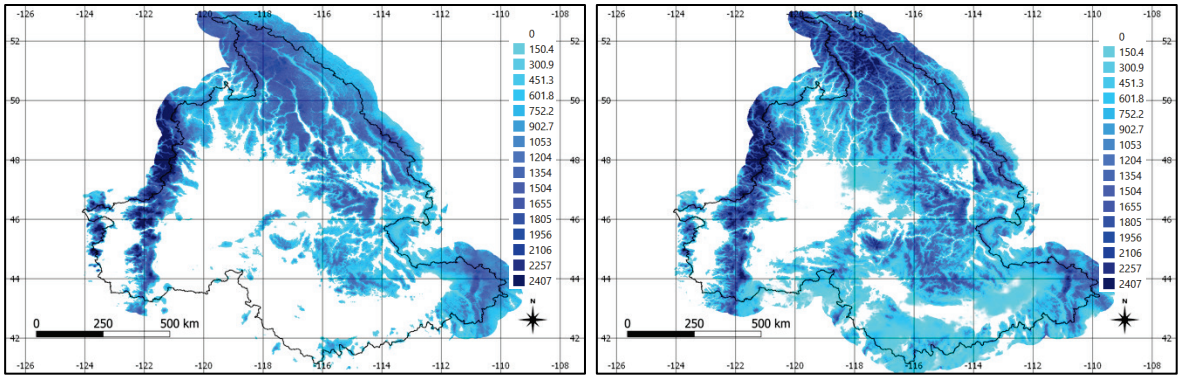



-126 -124 -122 -120 -118 -116 -114 -112 -110 -108
52 50 48 46 44 42
64
79.4
94.9
110
126
141
157
172
188
203
218
234
249
265
280
296
311
0 250 500 km
N

Appendix G: Monthly (Decemeber–June) and WY SWE Pointwise 100-Year Return Levels for the CRB

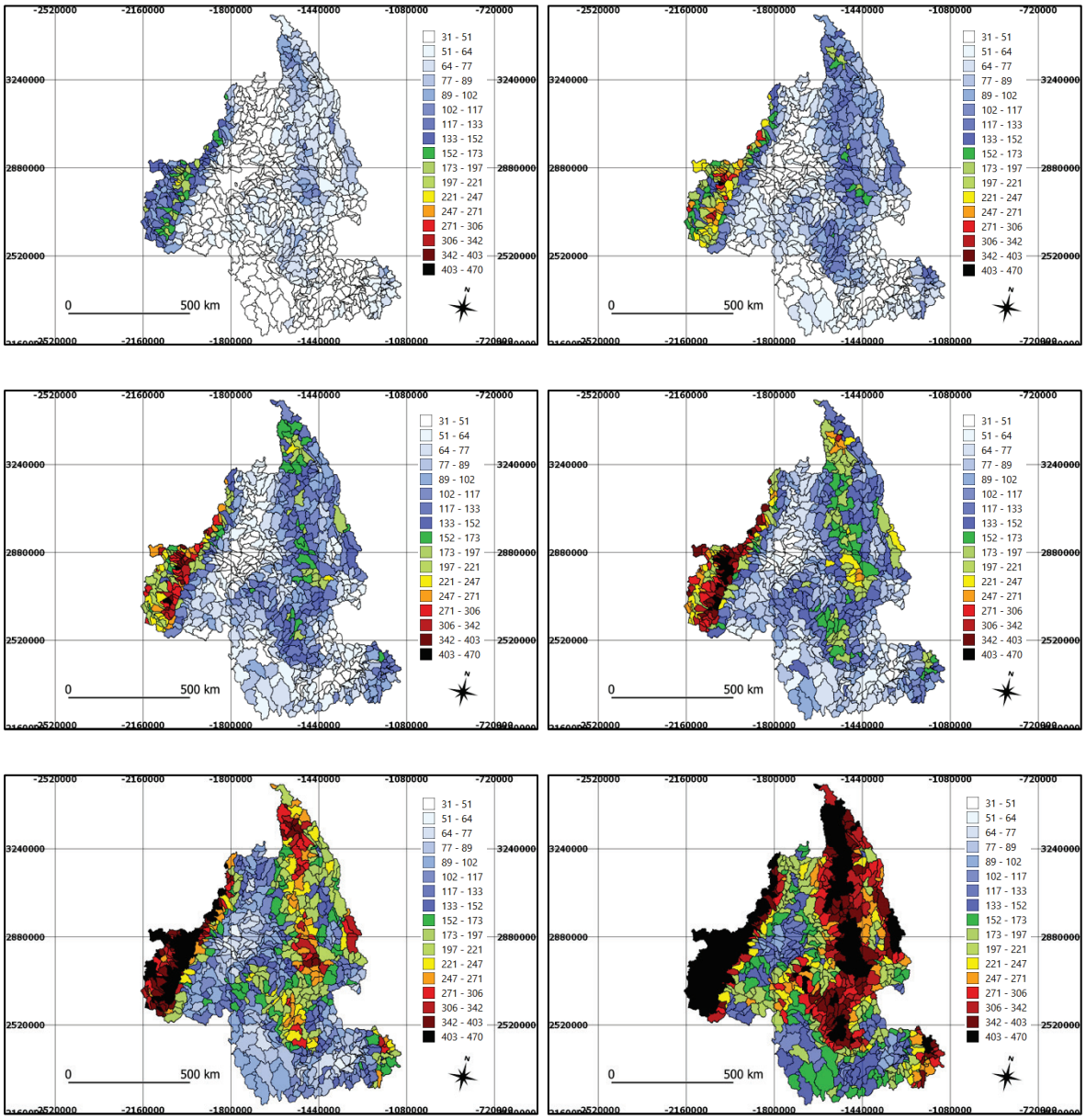
Figure 21. Monthly (December – June) and WY SWE pointwise 100-year return levels for the CRB (values in millimeters).

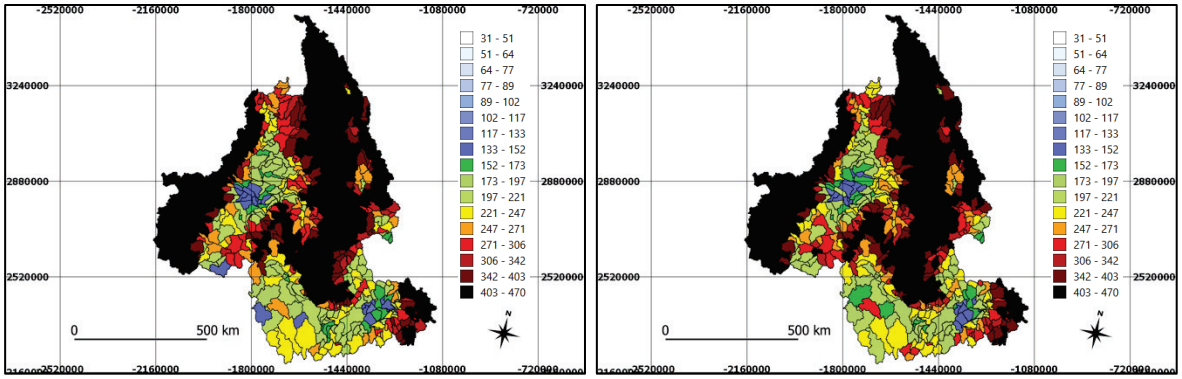


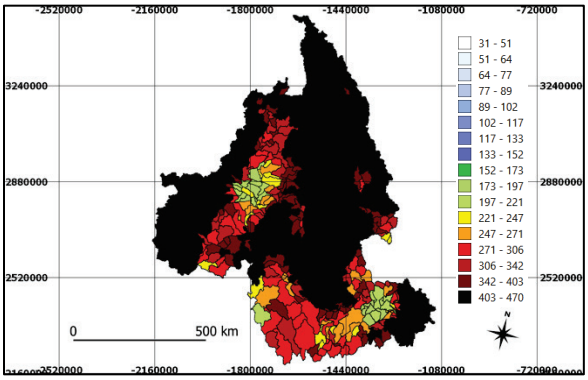

-126 -124 -122 -120 -118 -116 -114 -112 -110 -108
52 50 48 46 44 42
0
150.4
300.9
451.3
601.8
752.2
902.7
1053
1204
1354
1504
1655
1805
1956
2106
2257
2407
0 250 500 km
N

Appendix H: Cool Season PREC Mean Areal Exceedance Probabilities for 100-Year Return Period for 758 Delineated Sub-Basins of the CRB

Figure 22. Cool season PREC mean areal exceedance probabilities for 100-year return period for 758 delineated sub-basins of the CRB (for 1-, 2-, 3-, 5-, 10-, 30-, 60-, 90-, and 120-day durations) (values in millimeters).

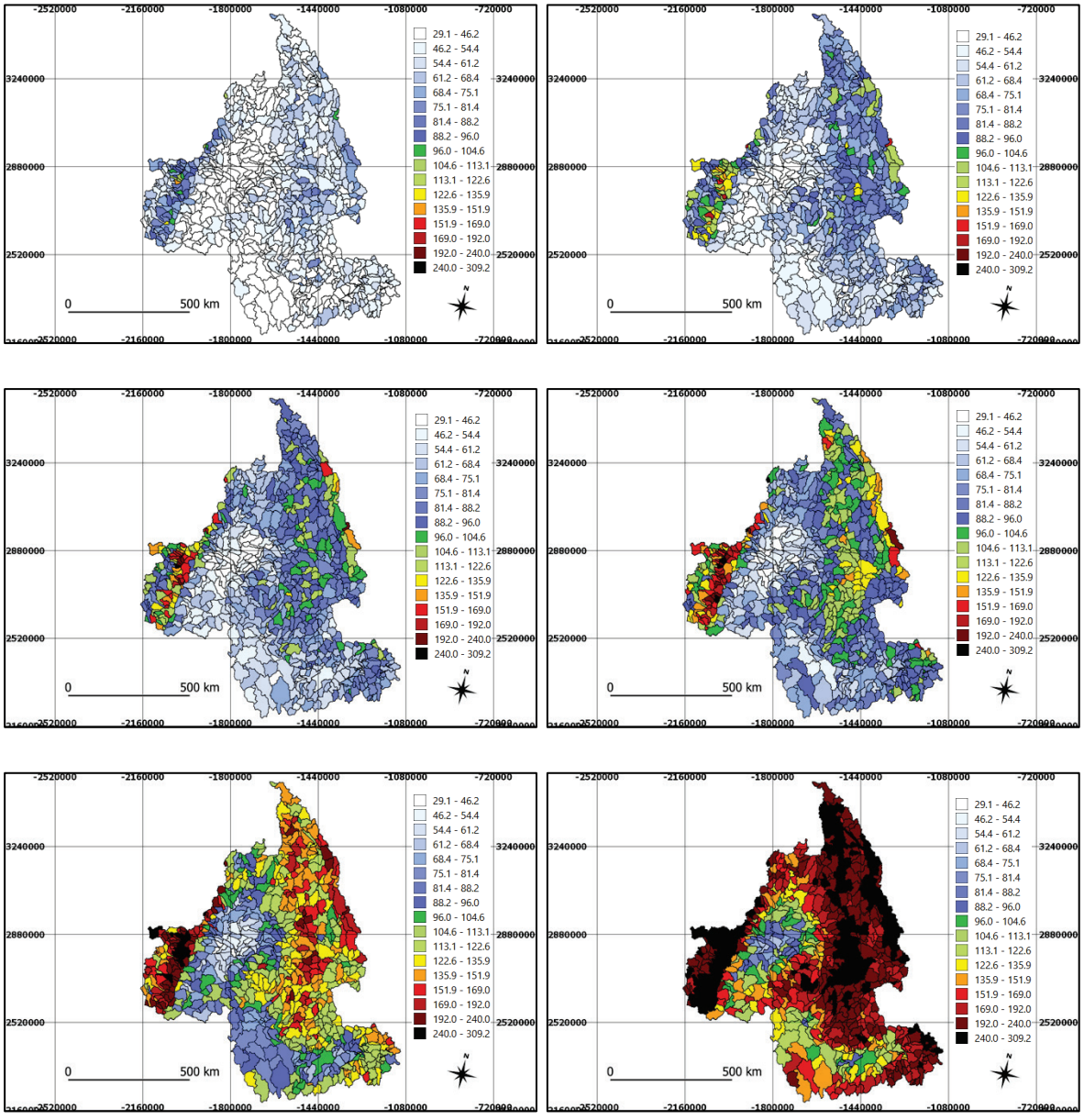


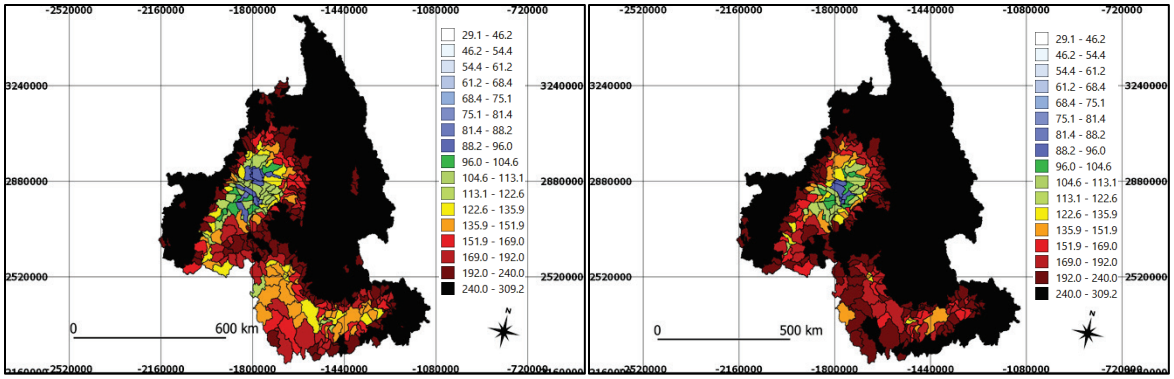

-2520000
-2160000
-1800000
-1440000
-1080000
-720000
3240000
2880000
2520000
31 - 51
51 - 64
64 - 77
77 - 89
89 - 102
102 - 117
117 - 133
133 - 152
152 - 173
173 - 197
197 - 221
221 - 247
247 - 271
271 - 306
306 - 342
342 - 403
403 - 470
0
500 km
N


-2520000
-2160000
-1800000
-1440000
-1080000
-720000
3240000
2880000
2520000
31 - 51
51 - 64
64 - 77
77 - 89
89 - 102
102 - 117
117 - 133
133 - 152
152 - 173
173 - 197
197 - 221
221 - 247
247 - 271
271 - 306
306 - 342
342 - 403
403 - 470
0
500 km
N

Appendix I: Warm Season PREC Mean Areal Exceedance Probabilities for 100-Year Return Period for 758 Delineated Sub-Basins of the CRB

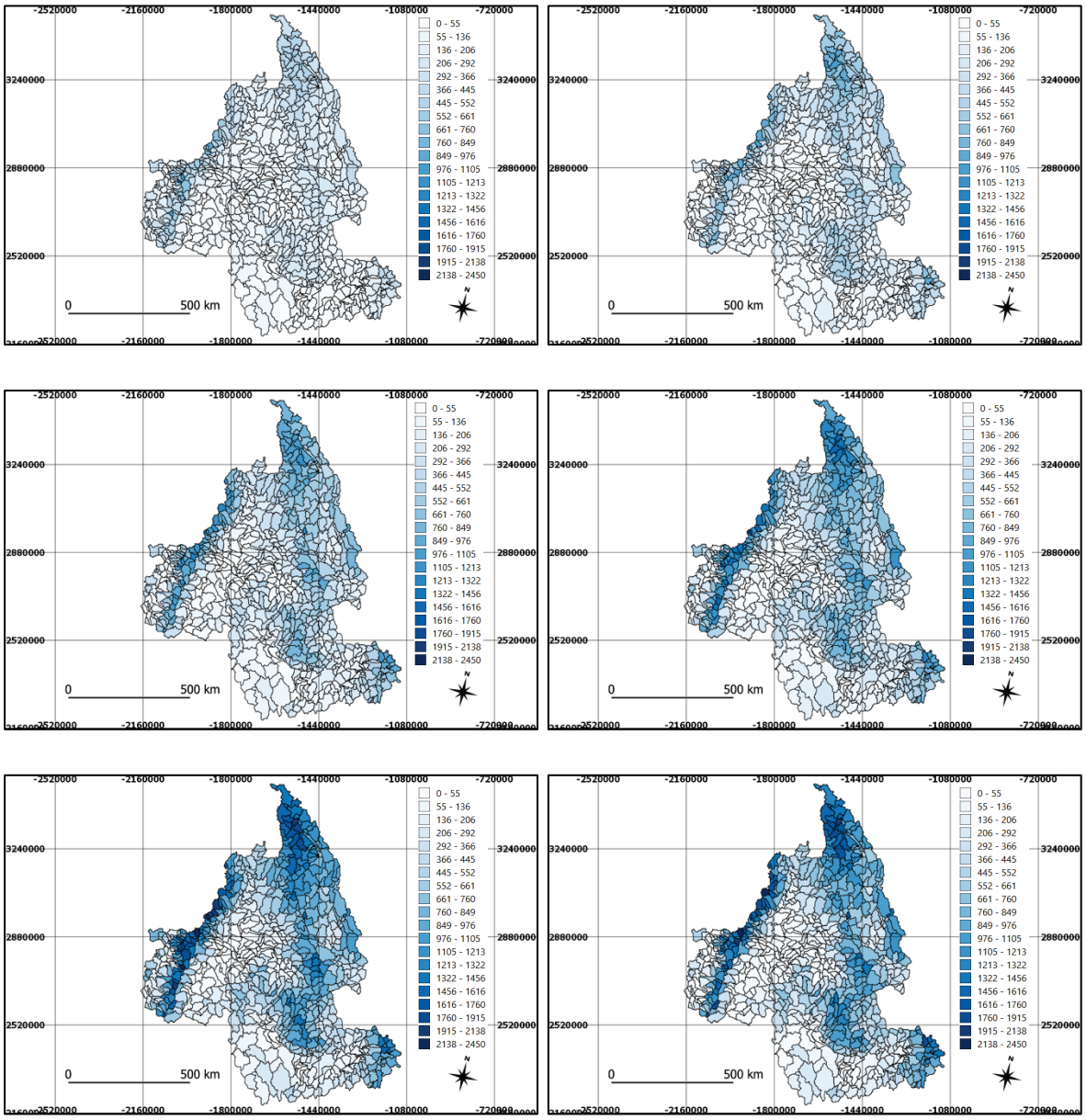
Figure 23. Warm season PREC mean areal exceedance probabilities for 100-year return period for 758 delineated sub-basins of the CRB (for 1-, 2-, 3-, 5-, 10-, 30-, 60-, and 90-day durations) (values in millimeters).

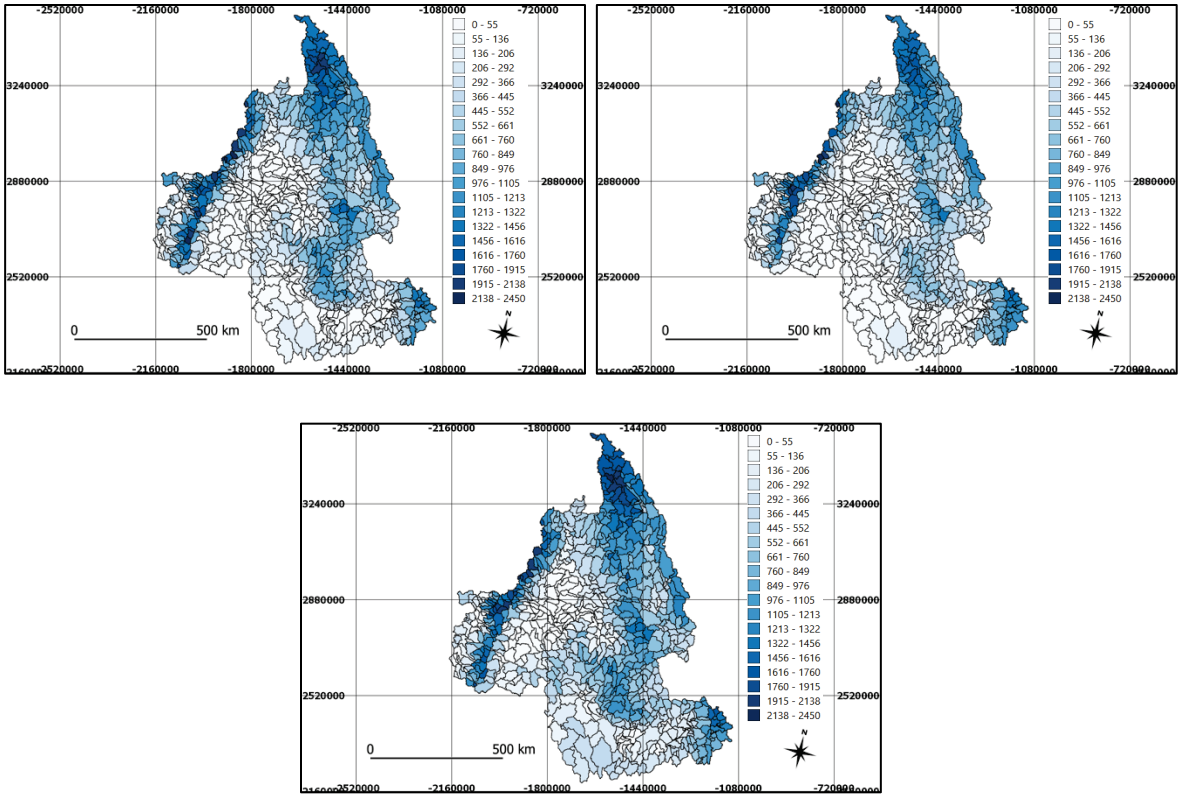



-2520000
-2160000
-1800000
-1440000
-1080000
-720000
3240000
2880000
2520000
29.1 - 46.2
46.2 - 54.4
54.4 - 61.2
61.2 - 68.4
68.4 - 75.1
75.1 - 81.4
81.4 - 88.2
88.2 - 96.0
96.0 - 104.6
104.6 - 113.1
113.1 - 122.6
122.6 - 135.9
135.9 - 151.9
151.9 - 169.0
169.0 - 192.0
192.0 - 240.0
240.0 - 309.2
0
600 km
0
500 km

Appendix J: Monthly (November–June) and WY SWE Mean Areal Exceedance Probabilities for 100-Year Return Period for 758 Delineated Sub-Basins of the CRB

Figure 24. Monthly (November – June) and WY SWE mean areal exceedance probabilities for 100-year return period for 758 delineated sub-basins of the CRB (values in millimeters).




-2520000
-2160000
-1800000
-1440000
-1080000
-720000
3240000
2880000
2520000
0 - 55
55 - 136
136 - 206
206 - 292
292 - 366
366 - 445
445 - 552
552 - 661
661 - 760
760 - 849
849 - 976
976 - 1105
1105 - 1213
1213 - 1322
1322 - 1456
1456 - 1616
1616 - 1760
1760 - 1915
1915 - 2138
2138 - 2450
0
500 km
N

Acronyms and Abbreviations

AMP	annual maximum precipitation
ASOS	Automated Surface Observing System
COOP	Cooperative Observers Program
CRB	Columbia River Basin
CSV	comma separated values
DEM	digital elevation model
ECCC	Environment and Climate Change Canada
ESRI	Environmental Systems Research Institute
EVT	extreme value theory
GEV	generalized extreme value
GHCN	Global Historical Climatology Network
ID	identification
MDS	missing data in season
MDT	Missing Data Time
MLE	maximum likelihood estimate
NDR	number of days with rain
NLDAS	North American Land Data Assimilation System
NSIMS	number of simulations
NWD	Northwestern Division
PREC	precipitation
Prob	probablilites
RFA	regional frequency analysis
RP	return period
SCENIC	Southwest Climate and Environmental Information Collaborative
SNOTEL	Snow Telemetry
SWE	snow water equivalent
TSR	total seasonal rainfall
USACE	US Army Corps of Engineers
WY	water year

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED *(From - To)*
June 2020	Final Report	

4. TITLE AND SUBTITLE
Spatial Analysis of Precipitation and Snow Water Equivalent Extremes for the Columbia River Basin

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Brian E. Skahill, Angela M. Duren, Luciana Cunha, and Chris Bahner

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) (see reverse)

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC/CHL TR-20-10

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
US Army Corps of Engineers, Northwestern Division
Portland, OR 97208-2870

10. SPONSOR/MONITOR'S ACRONYM(S)
USACE NWD

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Funding Acct Code U4363312; AMSCO Code 031398

14. ABSTRACT
Recent advances in the spatial statistics of extremes and model calibration were applied to develop and deliver areal-exceedance estimates for precipitation (PREC), by season and duration, and snow water equivalent (SWE), by cool season month and for the water year, for 758 delineated sub-basins of the Columbia River Basin (CRB), which correspond to a new CRB hydrology model watershed delineation. Understanding that future US Army Corps of Engineers, Northwestern Division, mission requirements may change, project execution also included the development and delivery of an application guidance document to credibly compute areal-exceedance estimates, including uncertainty, for PREC or SWE for any arbitrary area within the CRB. R, a free software environment for statistical computing and graphics (https://www.r-project.org/), and QGIS, a free and open source geographic information system (https://qgis.org/en/site/index.html), were the primary tools used for product development and delivery. The following R software packages were primarily used during project execution: evd, Glmnet, maps, raster, rgdal, SDMTools, sp, and SpatialExtremes.

15. SUBJECT TERMS
Columbia River Watershed, Hydrologic models, Hydrology, Precipitation (Meteorology), Runoff, Snow

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Brian E. Skahill
Unclassified	Unclassified	Unclassified	SAR	73	19b. TELEPHONE NUMBER *(Include area code)* 971-804-0373

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) (continued)

Coastal and Hydraulics Laboratory
US Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

US Army Corps of Engineers, Northwestern Division
PO Box 2870
Portland, Oregon 97208-2870

WEST Consultants
101 Parkshore Drive
Folsom, California 95630

WEST Consultants
2601 25th St. SE
Suite 450
Salem, Oregon 97302