Use of Daily Station Observations to Produce High-Resolution Gridded Probabilistic Precipitation and Temperature Time Series for the Hawaiian Islands

ANDREW J. NEWMAN AND MARTYN P. CLARK

National Center for Atmospheric Research, Boulder, Colorado

RYAN J. LONGMAN

Department of Geography and Environment, University of Hawai'i at Mānoa, Honolulu, Hawaii

ERIC GILLELAND

National Center for Atmospheric Research, Boulder, Colorado

THOMAS W. GIAMBELLUCA

Department of Geography and Environment, University of Hawai'i at Mānoa, Honolulu, Hawaii

JEFFREY R. ARNOLD

Climate Preparedness and Resilience Program, U. S. Army Corps of Engineers, Seattle, Washington

(Manuscript received 31 May 2018, in final form 25 January 2019)

ABSTRACT

It is a major challenge to develop gridded precipitation and temperature estimates that adequately resolve the extreme spatial gradients present in the Hawaiian Islands. The challenge is particularly pronounced because the available station networks are irregularly spaced and sparse, creating large uncertainties in gridded spatial meteorological estimates. Here a 100-member, daily ensemble of precipitation and temperature estimates over the Hawaiian Islands for the period 1990–2014 at 1-km grid resolution is developed. First, an intermediary ensemble estimate of the monthly climatological precipitation and temperature is created, and those climatological surfaces are used to inform daily anomaly interpolation. This climatologically aided interpolation (CAI) method extends our initial ensemble system developed for the continental United States. This study demonstrates that direct interpolation of daily precipitation values is inferior to the CAI methodology, particularly over longer time periods (from years to decades). Daily interpolation performs better for short time periods (e.g., 1 month or less) or when the precipitation distribution substantially diverges from climatology. The CAI ensemble is able to reproduce observed precipitation and temperature patterns, including precipitation occurrence. Leave-one-out cross-validation results illustrate that the ensemble has 1) minimal bias for precipitation and temperature; 2) a mean absolute error of 2.5 mm day^{-1} , 1.0 K , and 2.2 K for precipitation and mean and diurnal temperature, respectively; 3) a mean absolute error of 3.3 mm day^{-1} for the standard deviation of precipitation; and 4) nearly unbiased probability distributions across multiple thresholds of precipitation intensity. Additionally, the ensemble provides estimates of uncertainty across the distributions with increasing uncertainty for higher percentiles.

1. Introduction

Gridded datasets of precipitation and temperature are an integral component of hydrometeorological research (Legates and Willmott 1990a,b; Daly et al. 1994; Maurer et al. 2002; Abatzoglou 2013; Xia et al. 2012; Livneh et al. 2013; Newman et al. 2015, hereafter N15; Thornton et al. 2016). They are used in a wide variety of models

Denotes content that is immediately available upon publication as open access.

Corresponding author: Andrew Newman, anewman@ucar.edu

in a large array of applications such as streamflow and hydropower forecasting, species evolution, and infrastructure resilience assessments (e.g., Day 1985; Franklin 1995; USBR 2012). Important requirements for datasets in these applications are 1) the ability to reproduce the spatial heterogeneity of precipitation and temperature across complex terrain and other climatic gradients, 2) a time resolution of daily or subdaily, and 3) proper representation of precipitation occurrence (wet/dry spells). An additional feature that is highly beneficial (but not widely used in applications) is the characterization of uncertainty.

While there are numerous datasets that reasonably fulfill the requirements above, they generally exist only for highly developed regions, such as the contiguous United States or Europe. Many fewer datasets are available for areas such as Hawaii where there is a dearth of publicly available gridded observation-based meteorological products. Moreover, the requirement to represent spatial heterogeneity generally precludes the use of global or regional atmospheric reanalysis products (e.g., ERA-Interim; Dee et al. 2011); satellite precipitation products (e.g., Huffman et al. 2007, 2015); or coarse, gauge-based, or combined observation and reanalysis products (Sheffield et al. 2006; Harris et al. 2014; Beck et al. 2017). These datasets are either too coarse to sufficiently resolve topographic gradients, or in the case of some satellite products, have large uncertainties in complex terrain (e.g., Derin et al. 2016).

Past gridded products of sufficient resolution for impact modeling over Hawaii have primarily consisted of long-term climatological fields (Giambelluca et al. 1986, 2013; Daly et al. 2006). Only recently has a deterministic monthly precipitation dataset been developed (Frazier et al. 2016); and even more recently, deterministic daily products have been developed [Daymet version 3, released September 2016, and the University of Hawai'i daily product introduced by Longman et al. (2019)]. Thus, there is a need to continue development of hydrometeorological forcing datasets over Hawaii for the myriad application uses.

Additionally, accounting for uncertainty in these datasets is critical for properly understanding uncertainty in any application modeling chain. Uncertainty quantification is especially necessary in complex terrain or regions with large gradients (e.g., Slater and Clark 2006; Kuczera et al. 2010; Gutmann et al. 2014; Prein and Gobiet 2017). Specifically, quantification and inclusion of uncertainty in near surface meteorology is key for a fuller understanding of model parameter uncertainty during model calibration (Kuczera et al. 2010) and any subsequent climate impact modeling (Clark et al. 2016), for specifying model state uncertainty in data assimilation

(e.g., Slater and Clark 2006; Huang et al. 2017; Kumar et al. 2017), and for assessment of atmospheric model output (Prein and Gobiet 2017). Of the aforementioned studies, only Giambelluca et al. (2013) includes an estimate of the uncertainty in the gridded climatological fields over Hawaii.

Motivated by the need for rigorous uncertainty quantification for impact assessment, we develop a daily 100-member ensemble of precipitation and temperature estimates over the Hawaiian Islands. Hawaii is characterized by extreme topographic relief spanning from sea level to over 4000 m within tens of kilometers horizontally, with extreme precipitation gradients caused by the interaction of the complex terrain with persistent northeasterly trade winds (Giambelluca et al. 2013; Frazier et al. 2016). This environment makes Hawaii an ecologically diverse region that is home to many endemic and threatened species (Myers et al. 2000; Sakai et al. 2002) (Fig. 1). It is also strongly influenced by interannual and interdecadal climate oscillations (El Niño–Southern Oscillation and the Pacific decadal oscillation) and has experienced long-term drying across most of the islands (Chu and Chen 2005; Frazier and Giambelluca 2017).

Our main contribution is ensemble gridded precipitation, mean temperature, and the diurnal temperature range (DTR) estimates for the Hawaiian Islands for the period 1990–2014 at 1-km grid resolution, to enable consistent spatiotemporal uncertainty quantification and subsequent use in the aforementioned applications. We extend the methodology of Clark and Slater (2006, hereafter CS06) and N15 to include hierarchical time-scale separation—this is effectively a probabilistic application of the deterministic climatologically aided interpolation (CAI) approach of Willmott and Robeson (1995). We choose to model daily mean temperature and the DTR because successful recreation of DTR is important to biological life cycle models (e.g., Monaghan et al. 2016) and energy balance models that use empirical relationships between temperature and radiation to provide gridded radiation estimates (e.g., MT-CLIM; Bristow and Campbell 1984; Running et al. 1987). Additionally, estimating DTR avoids any nonphysical interpolated $T_{\min} > T_{\max}$ situations.

The remainder of the paper is organized as follows: we introduce the input datasets in section 2; discuss the methodology to generate the ensemble fields, compare methodological advances, and validate the results in section 3; provide a detailed leave-one-out cross validation in section 4; specifically address validation of extremes in section 5; and finally, note summary discussion in section 6.

FIG. 1. Extent of domain with elevation contours in shaded colors.

2. In situ observational datasets

a. Monthly precipitation data

Monthly rain gauge data were compiled from many sources across Hawaii including plantations, National Weather Service cooperative gauges, and mesonetworks (Longman et al. 2018). This collection resulted in 2000 stations that were then quality controlled and gap-filled with approximately 1100 unique gauges that passed

quality control procedures such as homogeneity testing and manual screening and had sufficiently filled record length following Eischeid et al. (2000) in the 1920–2012 period (Frazier et al. 2016). Here we focus on stations with valid 1978–2007, 30-yr climate normal monthly precipitation to generate monthly climatological precipitation (see section 3 for further discussion), as in Giambelluca et al. (2013). Here valid stations for a particular month are those from the Frazier et al. (2016) dataset with at least two monthly values during the 1978–2007 timeframe. This choice was made to maximize available observations. Each month has a different number of valid observations, ranging from 86.5% (950) to 90.4% (993) of the approximately 1100 possible unique gauges. The 1978–2007 July climate normal precipitation gauge network is shown in the left panel of Fig. 2. Note that these precipitation data do not include fog drip, the interception of cloud droplets by the canopy, which can be a nonnegligible portion of total surface water input (Giambelluca et al. 2013), and in some localized areas fog drip is larger than precipitation (Juvik et al. 2011). This may impact the quality of hydrologic simulations in certain areas of Hawaii that have frequent clouds intersecting the terrain as the ensemble

FIG. 2. (left) Climatological precipitation gauge network for July, (center) daily precipitation gauge network, and (right) daily temperature station network. Dark circle markers indicate stations used in the climatological or daily analyses, and light squares are unused stations resulting from incomplete data filling. Map scale bars and north orientation arrow are in the leftmost panel. All panels and subplots in this and subsequent figures have the same orientation and scale. Elevation is given as grayscale contours with sea level (thick black), and 500-m, 1500-, and 2500-m contours transitioning from thin dark gray to thin light gray.

product will significantly underestimate total surface water inputs. Finally, an important distinction in this study is that we include only actual in situ gauge observations. Giambelluca et al. (2013) also include in their climatological observation set radar point precipitation estimates and synthetic observations from limited-area model simulations and ecologic observations. See Giambelluca et al. (2011) for additional details on station data collection, quality control, and data filling.

b. Daily data

Daily precipitation and temperature records were compiled from 16 climate networks across the Hawaiian Islands over the period from 1 January 1990 through 31 December 2014. Longman et al. (2018) discusses the various networks and initial quality control applied to the data. These daily data were then filled using the cumulative density function (CDF) matching method described in N15 to create serially complete daily precipitation and temperature data. To be considered for interpolation, a station is required to have at least three years of observations in this case (as opposed to 10 in N15) to balance robust sampling statistics and retention of as many stations as possible. Then, the nearby stations are used to fill missing data using the highest available rank correlated station, where the rank correlation is determined from a minimum of three years of overlap between the target and nearby stations. If a station is unable to be made serially complete, it is not used. The choice to fill the data was made to produce a more homogeneous input station network to avoid issues in estimating long-term means in data sparse regions, interannual variability, and trends (e.g., Henn et al. 2018; Walton and Hall 2018). By filling individual stations to be serially complete, these issues are somewhat alleviated (e.g., N15; Henn et al. 2018). However, this step uses the assumption of stationarity at each individual station, thus any long-term trend analysis will be influenced by that assumption. Initially, there were 471 precipitation and 142 temperature observations. After filling, 439 (93% of original) precipitation and 117 (83% of original) temperature stations were made serially complete and included in the final product (Fig. 2, center and right panels).

3. Methods

a. Ensemble gridding methodology

Here we estimate an ensemble of daily precipitation, daily mean temperature, and the DTR. We first describe the underlying theory and express it specifically for this case, which defines the distribution functions and parameters. Estimation of the parameters is achieved using locally

weighted regression, and ensemble realizations are generated by sampling from the estimated distributions using spatiotemporally correlated random fields.

1) THEORY

Precipitation is an intermittent process at many time scales (e.g., daily), therefore the probability distribution function (PDF) of precipitation typically contains a fraction concentrated at zero. Intermittent processes such as precipitation can be modeled as $P(X_P = 0) = p_o$, where p_o is the probability of zero precipitation, and a CDF for the rest of the values $X_P > 0$. Following Papalexiou (2018), the CDF for precipitation X_P can be written as

$$F_{X_P}(x_P) = (1 - p_o)F_{X_P|X_P>0}(x_P) + p_o, \quad \text{for } x_P \geq 0, \quad (1a)$$

where $F_{X_P|X_P>0}(\cdot)$ is the CDF of precipitation given that precipitation occurs. Because temperatures X_T and X_{DTR} are not intermittent like precipitation, the CDFs for mean temperature and DTR (subscript D) can be written as

$$F_{X_T}(x_T), \quad (1b)$$

$$F_{X_D}(x_D). \quad (1c)$$

Transformation functions are used to map precipitation to a normal distribution regardless of the form of $F_{X_P|X_P>0}(\cdot)$. CS06 define the transformation using the empirical CDF of precipitation derived from the historical observations. N15 modify the precipitation transformation to a parametric power-law form to remove the requirement of computing and storing the empirical CDFs for each station:

$$X_P = v(Y_P) = Y_P^\alpha, \quad (2a)$$

$$Y_P = v^{-1}(X_P) = X_P^{1/\alpha}, \quad (2b)$$

where α is the transform exponent (set to 4 in N15) and the transformation is performed on only nonzero precipitation values. Parameter Y_P is assumed to be normally distributed:

$$Y_P \sim \mathcal{N}(\mu_{Y_P}, \sigma_{Y_P}^2), \quad \text{for } x_P > 0, \quad (3)$$

with mean μ_{Y_P} and variance $\sigma_{Y_P}^2$.

2) PARAMETER ESTIMATION

From Eqs. (1)–(3), there are three parameters that need to be estimated for precipitation: p_o , μ_{Y_P} , and σ_{Y_P} .

The parameter p_o is estimated using locally weighted logistic regression. Following CS06, local dependence is

incorporated using distance-dependent station weights computed for each grid cell:

$$w_i = \left[1 - \left(\frac{d_i}{d_{\max}} \right)^3 \right]^3, \quad (4)$$

where w_i is the weight assigned to individual stations (used to populate a diagonal weight matrix $\mathbf{W}$), d_i is the distance of the current station to the current grid point being considered, and $d_{\max}$ is the maximum distance beyond which stations receive zero weight. N15 set the number of stations considered for each grid point n_s to 30. N15 also considers two cases using an initial search radius of 100 km:

$$d_{\max} = \begin{cases} 100 \text{ km}, & \text{for } n_r \geq n_s \\ d_{i=n_s} + 1 \text{ km}, & \text{for } n_r < n_s \end{cases}, \quad (5)$$

where $d_{i=n_s}$ is the distance to the n_s th station, and n_r is the number of stations within the search radius. For the first case the n_s closest stations are used, and in the second case the search radius is expanded until $n_r = n_s$.

At each time step, the logistic regression estimate for $\hat{p}_0$ at each grid cell g is given as

$$1 - \hat{p}_{0,g} = \frac{1}{1 + \exp(-\mathbf{A}_g \hat{\boldsymbol{\beta}}_g^o)}, \quad (6a)$$

where $\mathbf{A}_g$ is the row vector of $n_a + 1$ spatial attributes for the current grid cell (n_a is the number of spatial attributes, and the extra element accounts for the regression intercept), and $\hat{\boldsymbol{\beta}}_g^o$ is the column vector of regression coefficients, estimated iteratively as

$$\hat{\boldsymbol{\beta}}_{g,\text{new}}^o = \hat{\boldsymbol{\beta}}_{g,\text{old}}^o + (\mathbf{A}_s^T \mathbf{W} \mathbf{V} \mathbf{A}_s)^{-1} \mathbf{A}_s^T \mathbf{W} (\mathbf{P}_0 - \hat{\boldsymbol{\pi}}), \quad (6b)$$

where $\mathbf{A}_s$ is the $n_s \times n_a + 1$ design matrix of spatial attributes for the n_s nearest stations, $\mathbf{P}_0$ is the vector of binary precipitation occurrence values, $\hat{\mathbf{V}}$ is the variance matrix, $\hat{\boldsymbol{\pi}}$ is the vector of estimated occurrence values at each station, and $\hat{\boldsymbol{\beta}}_{g,\text{old}}^o$ is initialized as a vector of ones. The spatial attributes in N15 are latitude, longitude, elevation, east–west slope, and north–south slope.

Next, $\mu_{Y_{P,g}}$ and $\sigma_{Y_{P,g}}$ are estimated using locally weighted multiple linear regression:

$$\hat{\mu}_{Y_{P,g}} = \mathbf{A}_g \hat{\boldsymbol{\beta}}_g^a, \quad (7a)$$

where $\hat{\boldsymbol{\beta}}_g^a$ is the vector of multiple linear regression coefficients estimated as

$$\hat{\boldsymbol{\beta}}_g^a = (\mathbf{A}_s^T \mathbf{W} \mathbf{A}_s)^{-1} \mathbf{A}_s^T \mathbf{W} \mathbf{Y}_P, \quad (7b)$$

where $\mathbf{Y}_P$ is the vector of station values computed using Eq. (2b) and all other variables are as defined previously.

Finally, $\hat{\sigma}_{Y_{P,g}}^2$ is estimated using the sum of the squared leave-one-out cross-validation errors, compiled by applying Eq. (7) to individual stations:

$$\hat{\sigma}_{Y_{P,g}}^2 = \frac{\sum_{i=1}^{n_s} w_i (\hat{y}_{P,-i} - y_{P,i})^2}{\sum_{i=1}^{n_s} w_i}, \quad (8)$$

where $\hat{y}_{P,-i}$ is the estimated value when the i th station is withheld from the regression equation and $y_{P,i}$ is the observed value at the i th station transformed using Eq. (2b).

Temperature and the DTR are assumed to be Gaussian without having to perform any transformation:

$$X_T \sim \mathcal{N}(\mu_T, \sigma_T), \quad (9a)$$

$$X_D \sim \mathcal{N}(\mu_D, \sigma_D). \quad (9b)$$

To estimate μ_T , σ_T , μ_D , and σ_D at each grid point, we use Eqs. (7) and (8), substituting Y_P for X_T or X_D .

3) ENSEMBLE GENERATION

Ensemble realizations of precipitation and temperature are generated by sampling from the estimated distributions of $\mathbf{Y}_P$, $\mathbf{X}_T$, and $\mathbf{X}_D$ which are estimated at all valid grid points and time steps. The sampling is done using spatiotemporally correlated random fields. At the first time step, draw from

$$\mathbf{Z}_0 \sim \mathcal{N}(0, \boldsymbol{\Sigma}_Z), \quad (10a)$$

where $\boldsymbol{\Sigma}_Z$ is the covariance matrix,

$$\boldsymbol{\Sigma}_Z = [\psi_{gg'}] = \psi(|s_g - s_{g'}|); \quad (10b)$$

s_g and $s_{g'}$ are the spatial locations of grid points g and g' ; and $\psi(\cdot)$ is a spatial correlation function that depends only on the distance between two stations. The exponential correlation function is used here, given by

$$\psi(s)_{gg'} = c_o \exp\left(-\frac{h_{gg'}}{\eta}\right) \quad (10c)$$

with

$$h_{gg'} = s_g - s_{g'}, \quad (10d)$$

where c_o is the initial correlation value, η is the spatial correlation length, and the $s_g - s_{g'}$ norm is calculated using the great-circle distance. Currently c_o and η are

spatially constant, have seasonally varying values in N15, and vary for precipitation and temperature. The suitability of these choices is an area that will be examined in future work.

The spatial fields $\mathbf{Z}_t \sim \mathcal{N}(0, \Sigma_Z)$ are defined to represent the lag-1 correlation for mean temperature and the DTR (subscript D), as well as the instantaneous cross correlation between the DTR and precipitation:

$$\mathbf{Z}_{T,t} = \rho_T \mathbf{Z}_{T,t-1} + \sqrt{1 - \rho_T^2} \mathbf{W}_{T,t}, \quad (11a)$$

$$\mathbf{Z}_{D,t} = \rho_D \mathbf{Z}_{D,t-1} + \sqrt{1 - \rho_D^2} \mathbf{W}_{D,t}, \quad (11b)$$

$$\mathbf{Z}_{P,t} = \rho_{CX} \mathbf{Z}_{D,t-1} + \sqrt{1 - \rho_{CX}^2} \mathbf{W}_{P,t}, \quad (11c)$$

where ρ_T and ρ_D are the lag-1 correlation for temperature and the DTR, ρ_{CX} is the cross correlation between the DTR and precipitation, and $\mathbf{W}_{0,t} \sim \mathcal{N}(0, \Sigma_Z)$ is a new spatially correlated random field. For temperature, ρ_0 in Eqs. (11a) and (11b) is the usual lag-1 autocorrelation estimated by regression. We do not explicitly simulate the lag-1 correlation for precipitation or the cross correlation between the DTR and temperature.

Once $\mathbf{Z}_{0,t}$ is computed, physical values are generated through the following transformation:

$$\hat{x}_{u,g} = v[F_Z^{-1}(u_g)], \quad (12)$$

where $F_Z \sim \mathcal{N}(\mu, \sigma)$, u_g is the cumulative probability of $\mathbf{Z}_{0,t}$ at grid point g , and $\hat{x}_{u,g}$ is the estimated value at u_g .

Precipitation and temperature at one time step and grid point are (the subscripts t and g are dropped to simplify notation):

$$\hat{x}_{u,P} = \begin{cases} 0, & 0 \leq u_P \leq \hat{p}_o \\ v\left[F_P^{-1}\left(\frac{u_P - \hat{p}_o}{1 - \hat{p}_o}\right)\right], & \hat{p}_o < u_P \leq 1 \end{cases}, \quad (13a)$$

$$\hat{x}_{u,T} = F_T^{-1}(u_T), \quad (13b)$$

$$\hat{x}_{u,D} = F_D^{-1}(u_D), \quad (13c)$$

where $v(\cdot)$ is the transformation defined in Eq. (2).

b. Ensemble climatologically aided interpolation

Several modifications to this method are made because of the challenging hydroclimate in the Hawaiian Islands. The general workflow for the ensemble CAI after station and DEM preprocessing is presented in Fig. 3, while a summary of the methodological modifications described next is given in Fig. 4. The left side of the solid line describes the previous direct daily methodology

FIG. 3. Workflow diagram of the climatologically aided ensemble system moving from top to bottom.

(CS06; N15), while the right side describes the new ensemble CAI methodology. The CAI portion is split into the climatological and daily anomaly calculations.

On any given day, there are roughly 400 daily precipitation observations and 100–150 temperature observations available throughout the islands. Initial testing using the N15 configuration with daily precipitation lapse rates indicated the daily station densities were insufficient to resolve key climatological precipitation maxima around the island chain (see section 4). Therefore, a CAI methodology is employed (e.g., Dawdy and Langbein 1960; Peck and Brown 1962; Jones 1994; Daly et al. 1994; Willmott and Robeson 1995; Di Luzio et al. 2008; Frazier et al. 2016) to better capture long-term precipitation maxima. CAI allows for propagation of known information from better-resolved climatological gradients across temporal scales (in this case climatological to daily).

In CAI, daily values are generated by interpolating daily anomalies of precipitation or temperature, and then combining the daily anomalies with the climatological fields to generate the daily values. Daily ratio (or multiplicative) anomalies for each station are calculated using the climatological precipitation at the nearest grid point for each month. For example, if the July climatological precipitation is 93 mm at the nearest grid point to a station, the daily climatological precipitation is 93/31 or 3 mm day⁻¹, and anomalies for all July dates are the ratio of the station/climatological grid. Thus, if 1 July had 6 mm of precipitation the anomaly ratio is 6/3, or 2. For temperature, anomalies are calculated as departures from the climatological monthly mean values.

Methodological Comparison

FIG. 4. Methodological comparison between the direct daily value and CAI for the ensemble system. Black and green text denotes the base methodology with green text denoting steps taken for precipitation only, while red text highlights methodological changes in the CAI system.

To implement CAI into the ensemble framework, monthly climatological values are computed first. The CDFs for monthly climatological precipitation and temperature follow Eq. (1) with X_P substituted for X_{PC} and X_T for X_{TC} . For precipitation, the transformation in Eq. (3) is changed from a simple power law to a Box-Cox transform (Box and Cox 1964) because the back transformation is more stable:

$$X_{PC} = v(Y_{PC}) = (Y_{PC}\lambda + 1)^{1/\lambda}, \quad (14a)$$

$$Y_{PC} = v^{-1}(X_{PC}) = \frac{X_{PC}^\lambda - 1}{\lambda}, \quad (14b)$$

where λ is set to 1/3 after sensitivity testing (not shown) and Y_{PC} follows Eq. (3). Mean daily temperature and the DTR again follow Eq. (9). Then p_{oC} , $\mu_{Y_{PC}}$, $\sigma_{Y_{PC}}^2$, μ_{TC} , σ_{TC}^2 , μ_{DC} , and σ_{DC}^2 are estimated using Eqs. (4)–(8). Here we set $n_s = 25$, $d_{\max} = 10$ km, and the initial search

radius to 10 km in Eq. (5). These changes increase the high frequency responsiveness of the logistic and ordinary regressions [Eqs. (6) and (7)]. Also, the spatial attributes included in $\mathbf{A}_g$ and $\mathbf{A}_s$ are changed to latitude, longitude, elevation, and distance to the coast. This predictor set was found to represent climatological precipitation patterns on Hawaii better than the attributes used in N15 (not shown). An ensemble of monthly climatological precipitation, mean temperature, and DTR is generated using Eqs. (10)–(12) substituting $\hat{x}_{uP}$, $\hat{x}_{uT}$, and $\hat{x}_{uD}$ with $\hat{x}_{uPC}$, $\hat{x}_{uTC}$, and $\hat{x}_{uDC}$, respectively. The result [following Eq. (13)] is

$$\hat{x}_{uPC} = \begin{cases} 0, & 0 \leq u_{PC} \leq \hat{p}_{oC} \\ v \left[F_{PC}^{-1} \left(\frac{u_{PC} - \hat{p}_{oC}}{1 - \hat{p}_{oC}} \right) \right], & \hat{p}_{oC} < u_{PC} \leq 1 \end{cases}, \quad (15a)$$

$$\hat{x}_{uTC} = F_{TC}^{-1}(u_{TC}), \quad (15b)$$

$$\hat{x}_{u_{DC}} = F_{DC}^{-1}(u_{DC}). \quad (15c)$$

The daily ratio and daily anomalies for precipitation, temperature, and DTR are computed using the estimated climatological ensemble means for the closest grid point to each station, to force consistency between the daily anomalies and the estimated climatological ensemble. Daily anomalies follow Eq. (1) and their transforms follow Eq. (13) substituting X_{PC} with X_{PA} , X_{TC} with X_{TA} , and X_{DC} with X_{DA} ; and p_{oA} , $\mu_{Y_{PA}}$, $\sigma_{Y_{PA}}^2$, μ_{TA} , σ_{TA}^2 , μ_{DA} , and σ_{DA}^2 are estimated using Eqs. (4)–(8). For the daily anomaly regressions, the spatial attributes included in $\mathbf{A}_g$ and $\mathbf{A}_s$ are changed to latitude and longitude, because the climatological fields are assumed to contain the elevation and coastal gradient information.

Ensemble realizations of daily precipitation and temperature at each time step and grid point follow Eqs. (10)–(12), with the additional step of transforming anomalies back to physical space. For precipitation,

$$\hat{x}_{u_{PA}} = \begin{cases} 0, & 0 \leq u_{PA} \leq \hat{p}_{oA} \\ v \left[F_{Y_{PA}}^{-1} \left(\frac{u_{PA} - \hat{p}_{oA}}{1 - \hat{p}_{oA}} \right) \right], & \hat{p}_{oA} < u_{PA} \leq 1 \end{cases}. \quad (16)$$

Then the rank of $\hat{x}_{u_{PA}}$ is used to determine which climatological precipitation ensemble member is used to create the daily precipitation:

$$\hat{x}_{u_p} = \hat{x}_{u_{PC}}^{(i)} \hat{x}_{u_{PA}}^{(i)}, \quad (17)$$

where (i) is the i th ranked value of $\hat{x}_{u_{PA}}$. Note that we censor the distribution in Eq. (16) when $Y_{PA}\lambda$ is less than -1 to prevent negative precipitation values and set the corresponding final $\hat{x}_{u_p}$ to 0.1 mm. Mean temperature and DTR are less complicated because they are continuous, assumed Gaussian, and the anomalies have physical units. The climatological and daily anomaly means and variances are combined as

$$\hat{\mu}_T = \hat{\mu}_{TC} + \hat{\mu}_{TA}, \quad (18a)$$

$$\hat{\mu}_D = \hat{\mu}_{DC} + \hat{\mu}_{DA}, \quad (18b)$$

$$\hat{\sigma}_T^2 = \hat{\sigma}_{TC}^2 + \hat{\sigma}_{TA}^2, \quad (18c)$$

$$\hat{\sigma}_D^2 = \hat{\sigma}_{DC}^2 + \hat{\sigma}_{DA}^2, \quad (18d)$$

where $\hat{\mu}_T$ and $\hat{\mu}_D$ are the regression estimated mean temperature and DTR, and $\hat{\sigma}_T^2$ and $\hat{\sigma}_D^2$ are the estimated total variance of temperature and DTR. Note that this assumes independence between climatological and daily

estimates. The simulated mean temperature and DTR are simply [following Eq. (13)]:

$$\hat{x}_{u_T} = F_T^{-1}(\hat{u}_T), \quad (19a)$$

$$\hat{x}_{u_D} = F_D^{-1}(\hat{u}_D). \quad (19b)$$

Following Giambelluca et al. (2013) and Frazier et al. (2016), we focus on seven of the eight major Hawaiian islands: Kauai, Oahu, Molokai, Lanai, Maui, Kahoolawe, and Hawaii Island (Fig. 1). The same initial 250-m DEM is used, coarsened to 1 km such that 16 of the Giambelluca et al. (2013) grid cells fit exactly into one 1-km grid cell, thus comparisons are as direct as possible. On the 1-km grid, there are 18 009 valid grid cells. A 100-member ensemble of daily precipitation, mean temperature, and DTR for 1 January 1990 through 31 December 2014 is generated. An example application of the ensemble CAI method for one day of precipitation is given in Fig. 5.

c. Methodological comparison

To quantify the changes with use of CAI in the ensemble system we provide a brief methodological comparison of results using the previous method and the new ensemble CAI presented in section 3a. As discussed in section 2, the primary justification for moving to a CAI methodology is that there are more than 2 times as many long-term monthly precipitation observations as daily precipitation observations. The climatological station network increases the number of observations across Hawaii (Figs. 2a,b) nearly everywhere, which improves the sampling of climatological precipitation features. We examine the long-term daily mean precipitation climatology from the two methods, and two example monthly accumulation periods, January and July 2005, where the January precipitation patterns were less similar to the climatological spatial pattern than those in July. Finally, for temperature, the climatological and daily station networks are the same, so the switch to the ensemble CAI methodology is motivated primarily by the argument of consistency between the precipitation and temperature generation methodologies. Therefore, no methodological comparison was performed for temperature variables.

d. Validation methodology

Daily validation is performed using leave-one-out cross validation for ensemble predictions at the station locations. For a station location, the ensemble estimated precipitation and temperature using the next 25 nearest stations is computed for a day. This was done across all

FIG. 5. Example creation of one daily precipitation ensemble member. The (a) daily ratio anomaly (unitless) is combined with a (b) spatially correlated random field (unitless) and the (c) daily ratio uncertainty (unitless) to create a daily ratio anomaly for one ensemble member (not shown). This result is combined with the (d) ordered ensemble climatological precipitation at each grid point (mm) to create one realization of (e) daily precipitation (mm). Map orientation, scale, and elevation contours are the same as Fig. 2.

station locations for the entire time period in the subsequent analysis. We perform both deterministic validation using the ensemble mean and probabilistic validation using the full ensemble. Validation of the probability of precipitation (PoP) is important because spatial and temporal representation of occurrence effects subsequent results throughout the impact modeling chain (e.g., Mizukami et al. 2016). Note that we provide validation of PoP for the full product (essentially calibration mode, includes all stations and data).

Specifically for the deterministic precipitation validation we use all days in the observed record because the ensemble is a probabilistic product; some of the time it will inherently assign nonzero precipitation probabilities to grid cells nearest observations even when the nearest observation has zero precipitation, or assign zero (or small) precipitation probabilities when the nearest observation is reporting precipitation. This behavior leads to an underestimation of precipitation when verifying using only on days with observed nonzero precipitation in the ensemble (not shown). Finally, we define precipitation variability simply as the mean of the standard deviation of the individual ensemble member precipitation time series.

Ensemble estimates of precipitation allow for probabilistic verification; here we use reliability and discrimination diagrams. Briefly, reliability diagrams indicate the observed occurrence of an event (e.g., precipitation > 1 mm) versus the predicted occurrence of the event. Perfect predicted reliability would fall along the 1–1 line, the prediction always having the same probability as observed. Predicted probabilities to the left (right) of the line imply an underprediction (overprediction). Discrimination diagrams highlight the ability of a prediction to discriminate between events and nonevents.

The probability distributions for nonevents and events should have minimal overlap and maximize the difference between the mean values of the two distributions.

Moreover, we focus a significant portion of the validation on daily extremes.¹ We examine precipitation 99.9th (~1 day in 3 years high exceedance) and temperature 99th daily percentile (~3 days per year high exceedance) and temperature 1st percentile (3 days per year low exceedance). Ensemble 99.9th and 1st percentiles are computed for each grid box for each ensemble member using all days. The ensemble mean is then the average of the 100 member estimates.

Last, we examine the fraction of precipitation coming from extreme precipitation days using the R95pTot metric from the CLIMDEX project (<https://www.climdex.org/learn/indices/>, index 25; e.g., Gampe and Ludwig 2017; Walton and Grieco 2017) in fractional space:

$$\text{R95pTot} = \frac{\sum_{d=1}^n R_d > \text{P95}}{\sum_{d=1}^n R_d}, \quad (20)$$

where n is the number of rainy days, R_d is the accumulation on the d th day, and P95 is the value of precipitation at the 95th percentile when considering only rainy days. This is the total precipitation from days greater than the 95th percentile of rainy days over the total precipitation from all rainy days, where a rainy day is defined as a day with ≥ 1 mm accumulation.

¹ Note that we use the descriptor extreme for uncommon (e.g., 1st or 99.9th percentiles) but not exceedingly rare events.

FIG. 6. (a) Hawaii CAI annual precipitation climatology (mm yr^{-1}), (b) Daily method derived annual precipitation climatology (mm yr^{-1}), (c) relative (to CAI amount) difference field between Hawaii CAI and Daily method (%). Map orientation, scale, and elevation contours are the same as Fig. 2.

This integrated metric offers insight into behavior of the ensemble method for all extreme precipitation days and thus the behavior of the tail of the precipitation distribution, rather than a snapshot metric such as the 99th percentile.

4. Validation

a. Methodological comparison

Annual ensemble mean climatological precipitation from the previous direct daily value method applied to the contiguous United States in N15 (Daily hereafter) and new CAI methodologies are shown in Figs. 6a and 6b, respectively. The Daily methodology operates directly on the daily data, and underestimates heavy precipitation zones on the windward (northeast) sides of all the islands. Additionally, there are differences in the placement of the precipitation gradients between the two methods. For example, the Daily methodology predicts greater annual precipitation at high elevations near the center of Hawaii Island than the CAI methodology (Fig. 6). However, consistent under prediction of the Daily method (and to a lesser extent the CAI methodology) is observed along windward sides of all islands when compared to observations (Fig. 7). Substantial long-term daily mean underestimation of precipitation would negatively influence all further impact modeling through an incorrect water balance derived from the climatologically biased precipitation inputs.

The CAI methodology for Hawaii improves estimation of the spatial distribution of the long-term daily mean precipitation. Very often, where the CAI methodology is wetter than the Daily method, the biases are reduced compared to the Daily method (Figs. 6c and 7; e.g., east Molokai). Overall, the archipelago-wide biases are -228 (from -183 to -275) and -304 (from -242 to -368) mm with mean absolute errors (MAEs) of 472 (428–521) and 601 (543–682) mm for the Hawaii CAI and Daily methodologies, respectively. The values in parentheses are the 90% bootstrapped confidence intervals (bootstrapped on the resulting statistics) using 1000 iterations here and in the rest of the study. The MAE differences are statistically significant but the bias differences are not.

Using CAI and reducing the number of predictors limits the flexibility of the regression on shorter (daily to monthly) time scales, particularly in situations where the spatial distribution of precipitation deviates from the climatological pattern, for example, more precipitation in climatologically dry than wet regions. Here, one may expect some degradation in a CAI-based product (e.g., Lundquist et al. 2015). The performance between the two methodologies is similar (Table 1) for the January and July 2005 example monthly accumulations. In this case, the Daily methodology performs statistically the same as the CAI methodology for mean and median biases yet outperforms the CAI method for MAE in January of 2005, when the observed precipitation patterns deviated from climatology. In July of 2005 the

FIG. 7. (a) Hawaii CAI to observation relative difference in annual precipitation (%) and (b) Daily method to observation relative (to observed amount) difference in annual precipitation (%). Map orientation, scale, and elevation contours are the same as Fig. 2.

methods are statistically the same for the mean bias and MAE, while the CAI method is statistically better for median bias.

This result suggests that at monthly time scales, particularly when the precipitation pattern deviates from climatological patterns, CAI methodologies may be less skillful than daily interpolation approaches. For this case, the improvements in long-term daily precipitation bias and MAE at station locations, the general improvement in the climatological pattern of precipitation in unobserved locations (e.g., windward areas are wetter and leeward areas are drier in the CAI fields), and the essentially equivalent performance for daily to monthly time scales supports usage of the CAI methodology for Hawaii. More broadly, daily interpolation and CAI are end points of the interpolation continuum. Future investigation will examine the middle ground, which could

include shorter climatological integration periods (e.g., weekly, monthly), weather-type CAI, or analog CAI.

b. Deterministic precipitation

The deterministic precipitation cross validation (Fig. 8) shows that the ensemble has a bias of 0.1 mm day^{-1} across all stations (Figs. 8a,b, Table 2). Precipitation is overestimated in the windward coastlines of Hawaii, Maui, and Oahu and precipitation is underestimated in inland locations. As seen in the climatological estimates, the precipitation peaks along Kauai, Oahu, and Maui are also underestimated. The ensemble MAE is generally similar across the islands with maximum values in areas of higher precipitation; the overall MAE is 2.5 mm day^{-1} . Similar to the other results here, MAE is greatest on the eastern side of Maui, and the peaks of Kauai and Oahu (Fig. 8c).

TABLE 1. Monthly precipitation mean bias (mm), median bias (mm), and MAE (mm) for January and July 2005. The 90% bootstrapped (1000 iterations) confidence intervals are in parentheses. Bold font indicates statistically significant improvement vs the other method at the 90% level.

	January 2005		July 2005	
	CAI	Daily	CAI	Daily
Mean bias	10.7 (5.8–15.0)	5.7 (1.6–9.3)	6.9 (0.3–12.3)	11.7 (6.6–15.4)
Median bias	4.7 (1.9–9.3)	9.2 (5.6–12.2)	4.1 (3.0–5.5)	11.0 (9.3–13.6)
MAE	45.4 (41.8–48.9)	36.2 (33.9–38.9)	36.8 (32.5–42.6)	33.3 (30.3–37.1)

FIG. 8. (a) Ensemble mean daily cross-validation bias using the mean difference (mm), (b) ensemble mean daily cross-validation bias using the median difference (mm), (c) ensemble mean MAE, and (d) ensemble mean variability (standard deviation) bias (mm). The ensemble mean variability is the mean of the individual ensemble member standard deviations rather than the standard deviation of the ensemble mean precipitation. Map orientation, scale, and elevation contours are the same as Fig. 2.

The ensemble daily precipitation variability has a bias of 2.8 mm day^{-1} and an MAE of 3.3 mm day^{-1} (Fig. 8d, Table 2). Spatial error patterns are very similar to the other statistics, with a general overestimation of variability on the windward coasts and some underestimation inland. Variability of precipitation is another key statistic that should be well represented

in successful gridded precipitation products used for hydrologic impact modeling as precipitation intensity/variability affects modeled infiltration. Variability can be poorly represented in gridded products, and should be used with caution when evaluating changes in variability for climate assessments (e.g., Begueria et al. 2016).

TABLE 2. Summary deterministic cross-validation statistics for ensemble daily values of precipitation, precipitation variability (standard deviation), mean temperature, and diurnal temperature range. The 90% bootstrapped (1000 iterations) confidence intervals are in parentheses.

	Precipitation (mm day^{-1})	Precipitation variability (mm day^{-1})	Mean temperature (K)	Diurnal range (K)
Mean bias	0.1 (0–0.3)	2.8 (2.5–3.2)	0.1 (0–0.2)	0 (–0.3 to 0.3)
Median bias	0.3 (0.2–0.4)	1.6 (1.3–1.9)	0.1 (–0.1 to 0.2)	0.1 (–0.4 to 0.3)
MAE	2.5 (2.3–2.7)	3.3 (3.0–3.7)	1.0 (0.9–1.1)	2.2 (2.0–2.3)

Finally, the ensemble mean PoP is validated using the entire time series of the daily station data and ensemble members in Fig. 9. Long-term PoP for the observations and an ensemble member is simply the number of rainy days divided by the total number of days. The ensemble mean PoP is then the mean of the individual member long-term PoP values. PoP is highest in the areas of highest precipitation, along the windward slopes of the islands below the trade wind inversion (Fig. 9a), with secondary occurrence maxima along the southern and western coasts of Hawaii Island. In these regions, precipitation occurs over 80% of the time. Conversely, on the leeward coasts of Maui, Hawaii, and Kahoolawe, precipitation occurs as little as 10% of the time. The ensemble is able to reproduce the observed PoP patterns well (Figs. 9b,c), is unbiased across all the stations, and has a MAE of 0.07 (7%). This result implies that the ensemble product will be useful for energy balance modeling applications for Hawaii where the prediction of occurrence is important for surface energy balance calculations (e.g., Mizukami et al. 2016).

c. Probabilistic precipitation

The probabilistic reliability (e.g., Wilks 2006) is shown in Figs. 10a, 10c, and 10e for three different precipitation thresholds. For all events, there is an under prediction at low observed probabilities and an overprediction at high probabilities. As the threshold of event occurrence is increased, the reliability deteriorates for the highest observed probabilities of occurrence (Figs. 10c,e). The discrimination of event occurrence (Figs. 10b,d,f) shows that the ensemble can discriminate precipitation events from nonevents, with a difference between the event and nonevent distribution mean probabilities of 0.51, 0.41, and 0.36 for $>0\text{ mm}$, $>25\text{ mm}$, and $>50\text{ mm}$, respectively.

As the threshold increases, discrimination ability decreases, but results for Hawaii are better than the initial continental United States (CONUS) product of N15. This may be due to the different regions of application or due to a methodological difference. Higher threshold precipitation event frequency closely follows the mean

FIG. 9. (a) Ensemble long-term PoP, (b) observed long-term PoP, and (c) ensemble PoP bias. Map orientation, scale, and elevation contours are the same as Fig. 2.

FIG. 10. (a),(c),(e) Reliability diagrams from the daily cross-validation data for three precipitation thresholds: 0, 25, and 50 mm, respectively. The black line is the perfect 1–1 line; the blue line is the ensemble performance. (b),(d),(f) Discrimination diagrams for the daily cross-validation of the ensemble for the same three thresholds: 0, 25, and 50 mm. Black (red) lines denote the nonevent (event) probability distributions. Gray shading in the figure panels represents the 95% confidence interval due to sampling uncertainty determined via bootstrapping; no gray shading indicates the plotted colored line width is larger than the sampling uncertainty.

precipitation distribution in Hawaii, which may not happen as frequently over the CONUS. This implies that relationships between stations may be stronger in Hawaii. The addition of CAI also adds spatial predictability from

the mean into the daily fields. This combination of region and methodological difference is likely improving the probabilistic verification for Hawaii, but future work will investigate further.

FIG. 11. (a) Ensemble mean temperature cross-validation bias (K), (b) ensemble mean diurnal temperature cross-validation bias (K), (c) ensemble mean temperature MAE (K), and (d) the ensemble mean diurnal temperature MAE (K). Map orientation, scale, and elevation contours are the same as Fig. 2.

d. Deterministic temperature

The mean temperature and DTR (Fig. 11) are both statistically unbiased overall, with biases of 0.1 K for the mean T and 0 K for DTR (Table 2, Figs. 11a,b) and 90% confidence intervals encompassing zero. Mean temperature is simulated to within a MAE of 2 K at 114 of 117 stations (97%), and the DTR is simulated to within a MAE of 4 K at 109 of 117 stations (93%). The spatial pattern of bias and MAE appears random, but they do have statistically significant (90% level) increases with elevation for mean temperature bias and MAE and DTR bias and MAE of 0.12, 0.13, 0.38, and 0.35 K km⁻¹, respectively. This trend may be influenced by the relative lack of stations at high elevations.

Giambelluca et al. (2014) found that DTR is impacted by the trade wind inversion. The trade wind inversion (TWI) is a product of the global general circulation

FIG. 12. Long-term mean observed (red dots) and ensemble mean cross-validation estimated values (blue dots) diurnal temperature range (K) at the observation locations.

pattern; there is large-scale subsidence from the Hadley cell over Hawaii (e.g., Peixoto and Oort 1992). In Hawaii the TWI generally demarcates the boundary of the planetary boundary layer and the free atmosphere (Giambelluca et al. 2014). Giambelluca et al. (2014) also found that DTR increases with elevation (neglecting the influence of precipitation) to a maximum around the trade wind inversion and then decreases again above the inversion. This behavior prompted them to treat temperature using a piecewise regression framework, one equation for elevations below the inversion, and one above. We adopted a more general treatment in the ensemble system, and we are able to reproduce this general feature of DTR over Hawaii.

Climatological DTR for every observation location for the observation and ensemble mean cross-validation-derived values are given in Fig. 12. Longman et al. (2015) found that the TWI is most frequently between 2000 and 2300 m across the island chain throughout the year. There is an increase in the observed DTR from sea level to just below those elevations, with a decrease at higher elevations above the trade wind inversion. The ensemble is able to recreate the increase in DTR but shows less decrease than the observed DTR at higher elevations above the inversion.

5. Daily extremes

Daily extremes are higher-order statistics that can be difficult to capture in gridded interpolations (e.g., Gervais et al. 2014). The locally varying nature of the ensemble regression and inclusion of uncertainty for each daily realization as here may better represent

FIG. 13. (a) Mean of the 99.9th daily precipitation percentile (mm) across the individual ensemble members, (b) observed 99.9th percentile (mm), (c) ensemble - observed 99.9th percentile difference (%), and (d) the standard deviation of the 99.9th percentile value (mm) across the ensemble members. Map orientation, scale, and elevation contours are the same as Fig. 2.

extreme values than more traditional deterministic interpolation methods. A more detailed examination of this issue can be found in Newman et al. (2019). Results for the 99.9th (1 day in 3 years) precipitation percentiles across ensembles are shown in Fig. 13. The ensemble highlights areas of heavy precipitation (Fig. 13a), with areas on Hawaii, Maui, Oahu, and

Kauai receiving $>300 \text{ mm day}^{-1}$ in 3 years and confirmed by the observations (Fig. 13b). Additionally, these areas receive $>100 \text{ mm}$ for ~ 18 days a year (not shown, 95th percentile).

Generally, the ensemble overestimates the 99.9th percentile across the state of Hawaii (Fig. 13c) with a bias of 50 mm and an MAE of 53 mm, with the largest

FIG. 14. The fractional contribution of precipitation > 95 th percentile for rainy days, R95pTot. (a) Mean ensemble R95pTot, (b) observed R95pTot, and (c) the mean ensemble – observation fractional difference. Map orientation, scale, and elevation contours are the same as Fig. 2.

absolute uncertainties along the rainy windward sides of the islands of Hawaii, Maui, and Kauai. This stems from the combined impacts of the addition of noise in the ensemble sampling and the precipitation transformation. Transformations like the Box–Cox are sensitive to perturbations to large transformed values—the addition of positive perturbations and then applying the transformation widens the tails of the ensemble distribution. Improving the representation of the full precipitation distribution in the ensemble system is a topic of ongoing research.

As an example, the estimated maximum 99.9th percentile in the ensemble is 592 mm on the island of Kauai, while the maximum observed is 349 mm on Kauai, about 3.5 km southeast of the ensemble maximum. Correspondingly, absolute uncertainty in the percentile value increases as the percentile value increases with the islands of Hawaii and Kauai having the largest absolute uncertainty (Fig. 13b). As seen in Figs. 13a–c, the ensemble overestimates 99.9th percentile precipitation values at the observation locations in the heaviest precipitation areas, but in the case of the 99.9th percentile, the ensemble standard deviation (Fig. 13d) is often larger than 30–50 mm in those regions.

Additionally, the R95pTot statistic (section 3c) illustrates that the ensemble overestimates the contributions from extremes to the total precipitation across the state of Hawaii (Fig. 14). The ensemble overestimates R95pTot with a mean bias of 0.033 (0.030–0.038) and a MAE of 0.04 (0.037–0.043). Fractional overestimation is

larger in the very dry leeward portions of the islands, while on the wetter windward sides the ensemble is generally very close to observed values. The ensemble mean here is the mean of the R95pTot values across each individual ensemble member, not the R95pTot metric for the ensemble mean precipitation.

Finally, the 1st and 99th percentiles for daily mean temperature across the ensemble members and the ensemble – observation differences are highlighted in Fig. 15. Again, the strong dependency of temperature on elevation is evident, with the higher elevations on Hawaii Island having daily mean temperatures below freezing more than 3 days a year (Fig. 15a), and daily mean temperatures greater than 10° – 15° C less than 4 days a year (Fig. 15c). At low elevations near the coasts, daily mean temperatures are rarely below 10° – 12° C (in time or space) and in some areas can reach near 30° C several times a year (Fig. 15a). Comparisons to observations (Fig. 15b) show that the ensemble produces an unbiased estimate of the 1st percentile on average, -0.1 K (from -0.2 to 0.1 K) with an MAE of 0.8 K (0.7 – 0.9 K). For the 99th percentile (Fig. 15d), the ensemble performs slightly worse with a mean bias of 1.0 K (0.8 – 1.1 K) and an MAE of 1.2 K (1.1 – 1.4 K).

6. Summary discussion

In this study we developed a high-resolution (1-km grid) daily precipitation and temperature ensemble that can be used for hydrometeorological, ecological, and

FIG. 15. (a) Mean of the 1st daily mean temperature percentile ($^{\circ}\text{C}$) across the individual ensemble members, (b) ensemble – observation 1st percentile differences ($^{\circ}\text{C}$), (c) mean of the 99th daily mean temperature percentile ($^{\circ}\text{C}$) across the individual ensemble members, and (d) ensemble – observation 99th percentile differences ($^{\circ}\text{C}$). Note the different temperature ranges from (a) to (c). Map orientation, scale, and elevation contours are the same as Fig. 2.

climate downscaling/impact assessment. We extend the methodology of CS06 and N15 to include hierarchical time-scale separation—a probabilistic application of the deterministic climatologically aided interpolation (CAI) approach of Willmott and Robeson (1995) to replace the original daily orographic lapse rate determination used in N15. The CAI approach first develops long-term monthly climatologies from the climatological network that better resolve the large-scale climatic gradients with less noise than the daily data. The long-term monthly climatology is created using the ensemble framework, thus our ensemble estimates also include uncertainty in our climatological estimates. After producing the climatological ensemble, daily anomalies are generated

with uncertainty, which are combined with the climatological ensemble to produce a 100-member ensemble of daily gridded fields of precipitation and temperature at 1-km resolution.

While daily lapse rate determination allows for flexibility in the orographic lapse rates to change with different meteorological conditions (e.g., Daily method), it can introduce noise from insufficient sampling and/or measurement error. Climatological lapse rates are less noisy, but do not vary with meteorological conditions and may not represent individual precipitation event distributions properly (e.g., Lundquist et al. 2015). CAI produces a more unbiased product over the long term with smaller mean absolute error (MAE) (Figs. 6, 7, Table 1), which is critical for long-term impact modeling (e.g., climate change impact assessment). However, for shorter time scales (e.g., 1 month) the CAI methodology has essentially equivalent (or slightly lower) skill than a daily lapse rate method over Hawaii, particularly when the precipitation accumulation patterns diverge from climatology. The users of these types of datasets should consider methodological choices such as this one, as they affect which products are most suitable for a specific application. More broadly, these two methodological options for lapse rate determination—purely daily or long-term climatology—are end points along a continuum of lapse rate definition. Future work should be directed at examining the tradeoffs of these endpoints and developing and understanding techniques that lie in between (e.g., aggregation along contiguous temporal blocks, temporal aggregation via weather typing).

Cross-validation examination here highlights that long-term precipitation amounts are represented with little bias (Table 2, Fig. 8). The ensemble also captures the spatial variability of precipitation occurrence (Fig. 9). The probability of precipitation (PoP) is a higher-order precipitation statistic, yet it is critical for energy balance modeling, as incorrect occurrence of precipitation will influence empirical estimates of radiation and the subsequent model water budget partitioning (Mizukami et al. 2016). Another higher-order statistic important for hydrologic modeling is precipitation intensity or variability, simply represented here as the standard deviation of daily precipitation. We show that the ensemble is able to reproduce high daily precipitation variability and slightly overestimates the day-to-day variability across the domain (Fig. 8d). Rainfall intensity controls infiltration, particularly in more physically based models that have infiltration excess runoff generation mechanisms [e.g., the Variable Infiltration Capacity (VIC) model]. Traditional interpolation techniques often struggle to retain high intensity events (e.g., Gervais et al. 2014).

Future work to further improve the representation of extreme events will examine approaches such as the integrated weighted quantile square error loss proposed by Craigmire et al. (2005) or a simulation approach using Monte Carlo sampling from the ensemble daily time series (e.g., Gilleland and Nychka 2005). For the truly rare events, a simple vector generalized additive model tied to an extreme value distribution could be a straightforward way of obtaining added value (e.g., Gilleland et al. 2006; Yee and Stephenson 2007; Cooley and Sain 2010; Heaton et al. 2011) or other spatial extreme value methods, although this research area is still very active (e.g., Turkman et al. 2010; Ribatet 2013; Huser et al. 2017; de Fondeville and Davidson 2018; Huser and Wadsworth 2019).

Daily mean temperature and DTR are statistically unbiased in cross validation across the islands as a whole (Table 2, Figs. 11, 12). DTR varies appropriately with elevation and precipitation across Hawaii, with a maximum in DTR near the trade wind inversion in drier areas (Giambelluca et al. 2013). The ensemble generally reproduces these relationships (Fig. 12). DTR is also important for energy balance modeling as it too is included in downwelling radiation estimates as a second-order effect. Low temperature extremes are statistically unbiased in the ensemble while high temperature extremes have a slight positive bias (1 K) as compared to observations (section 5, Fig. 15). The behavior of the ensemble to overestimate high temperature extremes will need further study.

There are now several daily precipitation and temperature datasets (and four climatological datasets) available for Hawaii, and our companion paper examines the differences across the datasets (Newman et al. 2019). The key aspects of gridded products discussed here—PoP, extreme events and temporal variability, and DTR—are highlighted along with the primary methodological choices that lead to interproduct differences. These types of analyses are useful to understand where products differ for application purposes (e.g., Henn et al. 2018) and more broadly for methodological improvements in gridded field generation.

Acknowledgments. This dataset is freely available at <https://doi.org/10.5065/D6SB44JV> and was generated using National Center for Atmospheric Research (NCAR) high-performance computing resources (CISL 2017). The U.S. Army Corps of Engineers (USACE) Climate Preparedness and Resilience program funded this work. We thank two anonymous reviewers whose comments greatly improved the manuscript. Color tables used here are provided by Wikipedia (precipitation), ESRI (PoP), GMT (difference plots), and the GRID-Arendal project (<http://www.grida.no/>) (temperature) via the NCAR NCL

(https://www.ncl.ucar.edu/Document/Graphics/color_table_gallery.shtml) and cpt-city color table archive (<http://soliton.vm.bytemark.co.uk/pub/cpt-city/>).

REFERENCES

- Abatzoglou, J. T., 2013: Development of gridded surface meteorological data for ecological applications and modeling. *Int. J. Climatol.*, **33**, 121–131, <https://doi.org/10.1002/joc.3413>.
- Beck, H. E., A. I. J. M. van Dijk, V. Levizzani, J. Schellekens, D. G. Miralles, B. Martens, and A. de Roo, 2017: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.*, **21**, 589–615, <https://doi.org/10.5194/hess-21-589-2017>.
- Beguieria, S., S. M. Vicente-Serrano, M. Tomas-Burguera, and M. Maneta, 2016: Bias in the variance of gridded data sets leads to misleading conclusions about changes in climatic variability. *Int. J. Climatol.*, **36**, 3413–3422, <https://doi.org/10.1002/joc.4561>.
- Box, G. E., and D. R. Cox, 1964: An analysis of transformations. *J. Roy. Stat. Soc.*, **26B**, 211–252.
- Bristow, K. L., and G. S. Campbell, 1984: On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric. For. Meteorol.*, **31**, 159–166, [https://doi.org/10.1016/0168-1923\(84\)90017-0](https://doi.org/10.1016/0168-1923(84)90017-0).
- Chu, P.-S., and H. Chen, 2005: Interannual and interdecadal rainfall variations in the Hawaiian Islands. *J. Climate*, **18**, 4796–4813, <https://doi.org/10.1175/JCLI3578.1>.
- Clark, M. P., and A. G. Slater, 2006: Probabilistic quantitative precipitation estimation in complex terrain. *J. Hydrometeorol.*, **7**, 3–22, <https://doi.org/10.1175/JHM474.1>.
- , and Coauthors, 2016: Characterizing uncertainty of the hydrologic impacts of climate change. *Curr. Climate Change Rep.*, **2**, 55–64, <https://doi.org/10.1007/s40641-016-0034-x>.
- Computational and Information Systems Laboratory, 2017: Cheyenne: SGI ICE XA System (NCAR Community Computing). National Center for Atmospheric Research, <https://doi.org/10.5065/D6RX99HX>.
- Cooley, D., and S. R. Sain, 2010: Spatial hierarchical modeling of precipitation extremes from a regional climate model. *J. Agric. Biol. Environ. Stat.*, **15**, 381–402, <https://doi.org/10.1007/s13253-010-0023-9>.
- Craigmire, P. F., N. Cressie, T. J. Santner, and Y. Rao, 2005: A loss function approach to identifying environmental exceedances. *Extremes*, **8**, 143–159, <https://doi.org/10.1007/s10687-006-7964-y>.
- Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, **33**, 140–158, [https://doi.org/10.1175/1520-0450\(1994\)033<0140:ASTMFM>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2).
- , J. Smith, M. Doggett, M. Halbleib, and W. Gibson, 2006: High-resolution climate maps for the Pacific Basin Islands, 1971–2000. Rep. submitted to National Park Service Pacific West Regional Office, 59 pp., <http://www.botany.hawaii.edu/basch/uhnpscesu/pdfs/sam/Daly2006AS.pdf>.
- Dawdy, D. R., and W. B. Langbein, 1960: Mapping mean areal precipitation. *Hydrol. Sci. J.*, **5**, 16–23, <https://doi.org/10.1080/0262666009493176>.
- Day, G. N., 1985: Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.*, **111**, 157–170, [https://doi.org/10.1061/\(ASCE\)0733-9496\(1985\)111:2\(157\)](https://doi.org/10.1061/(ASCE)0733-9496(1985)111:2(157)).

- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- de Fondeville, R., and A. C. Davidson, 2018: High-dimensional peaks over threshold inference. *Biometrika*, **105**, 575–592, <https://doi.org/10.1093/biomet/asy026>.
- Derin, Y., and Coauthors, 2016: Multiregional satellite precipitation products evaluation over complex terrain. *J. Hydrometeor.*, **17**, 1817–1836, <https://doi.org/10.1175/JHM-D-15-0197.1>.
- Di Luzio, M., G. L. Johnson, C. Daly, J. K. Eischeid, and J. G. Arnold, 2008: Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. *J. Appl. Meteor. Climatol.*, **47**, 475–497, <https://doi.org/10.1175/2007JAMC1356.1>.
- Eischeid, J. K., P. A. Pasteris, H. F. Diaz, M. S. Plantico, and N. J. Lott, 2000: Creating a serially complete, national daily time series of temperature and precipitation for the western United States. *J. Appl. Meteor.*, **39**, 1580–1591, [https://doi.org/10.1175/1520-0450\(2000\)039<1580:CASCND>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<1580:CASCND>2.0.CO;2).
- Franklin, J., 1995: Predictive vegetation mapping: geographical modelling of biospatial patterns in relation to environmental gradients. *Prog. Phys. Geogr.*, **19**, 474–499, <https://doi.org/10.1177/030913339501900403>.
- Frazier, A. G., and T. W. Giambelluca, 2017: Spatial trend analysis of Hawaiian precipitation from 1920 to 2012. *Int. J. Climatol.*, **37**, 2522–2531, <https://doi.org/10.1002/joc.4862>.
- , —, H. F. Diaz, and H. L. Needham, 2016: Comparison of geostatistical approaches to spatially interpolate month-year precipitation for the Hawaiian Islands. *Int. J. Climatol.*, **36**, 1459–1470, <https://doi.org/10.1002/joc.4437>.
- Gampe, D., and R. Ludwig, 2017: Evaluation of gridded precipitation data products for hydrological applications in complex topography. *Hydrology*, **4**, 53, <https://doi.org/10.3390/hydrology4040053>.
- Gervais, M., L. B. Tremblay, J. R. Gyakum, and E. Atallah, 2014: Representing extremes in a daily gridded precipitation analysis over the United States: Impacts of station density, resolution, and gridding methods. *J. Climate*, **27**, 5201–5218, <https://doi.org/10.1175/JCLI-D-13-00319.1>.
- Giambelluca, T. W., M. A. Nullet, and T. A. Schroeder, 1986: Rainfall atlas of Hawai'i. Rep. R76, Hawaii Division of Water and Land Development, Department of Land and Natural Resources, 267 pp.
- , Q. Chen, A. F. Frazier, J. P. Price, Y.-L. Chen, P.-S. Chu, and J. K. Eischeid, 2011: The rainfall atlas of Hawai'i. Final Rep., Department of Geography, University of Hawai'i at Mānoa, 72 pp., <http://rainfall.geography.hawaii.edu/assets/files/PDF/RF%20Atlas%20Final%20Report.pdf>.
- , —, A. G. Frazier, J. P. Price, Y.-L. Chen, P.-S. Chu, J. K. Eischeid, and D. M. Delparte, 2013: Online Rainfall Atlas of Hawai'i. *Bull. Amer. Meteor. Soc.*, **94**, 313–316, <https://doi.org/10.1175/BAMS-D-11-00228.1>.
- , and Coauthors, 2014: Evapotranspiration of Hawai'i. Final Rep. submitted to the U.S. Army Corps of Engineers–Honolulu District and the Commission on Water Resource Management, 178 pp., <http://evapotranspiration.geography.hawaii.edu/assets/files/PDF/ET%20Project%20Final%20Report.pdf>.
- Gilleland, E., and D. W. Nychka, 2005: Statistical models for monitoring and regulating ground-level ozone. *Environmetrics*, **16**, 535–546, <https://doi.org/10.1002/env.720>.
- , —, and U. Schneider, 2006: Spatial models for the distribution of extremes. *Hierarchical Modelling for the Environmental Sciences: Statistical Methods and Applications*, J. S. Clark and A. Gelfand, Eds., Oxford University Press, 170–183.
- Gutmann, E., T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen, 2014: An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.*, **50**, 7167–7186, <https://doi.org/10.1002/2014WR015559>.
- Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2014: Updated high-resolution grids of monthly climatic observations – The CRU TS3.10 dataset. *Int. J. Climatol.*, **34**, 623–642, <https://doi.org/10.1002/joc.3711>.
- Heaton, M. J., M. Katzfuss, S. Ramachandar, K. Pedings, E. Gilleland, E. Mannshardt-Shamseldin, and R. L. Smith, 2011: Spatio-temporal models for large-scale indicators of extreme weather. *Environmetrics*, **22**, 294–303, <https://doi.org/10.1002/env.1050>.
- Henn, B., A. J. Newman, B. Livneh, C. Daly, and J. D. Lundquist, 2018: An assessment of differences in gridded precipitation datasets in complex terrain. *J. Hydrol.*, **556**, 1205–1219, <https://doi.org/10.1016/j.jhydrol.2017.03.008>.
- Huang, C., A. J. Newman, M. P. Clark, A. W. Wood, and X. Zheng, 2017: Evaluation of snow data assimilation using the ensemble Kalman filter for seasonal streamflow prediction in the western United States. *Hydrol. Earth Syst. Sci.*, **21**, 635–650, <https://doi.org/10.5194/hess-21-635-2017>.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multi-satellite Precipitation Analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. *J. Hydrometeor.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E. J. Nelkin, and P. Xie, 2015: NASA Global Precipitation Measurement Integrated Multi-satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 4.5, 30 pp., http://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.5.pdf.
- Huser, R., and J. L. Wadsworth, 2019: Modeling spatial processes with unknown extremal dependence class. *J. Amer. Stat. Assoc.*, <https://doi.org/10.1080/01621459.2017.1411813>, in press.
- , T. Opitz, and E. Thibaud, 2017: Bridging asymptotic independence and dependence in spatial extremes using Gaussian scale mixtures. *Spat. Stat.*, **21**, 166–186, <https://doi.org/10.1016/j.spa.2017.06.004>.
- Jones, P. D., 1994: Hemispheric surface air temperature variations: A reanalysis and an update to 1993. *J. Climate*, **7**, 1794–1802, [https://doi.org/10.1175/1520-0442\(1994\)007<1794:HSATVA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1794:HSATVA>2.0.CO;2).
- Juvik, J. O., J. K. DeLay, K. M. Kinney, and E. W. Hansen, 2011: A 50th anniversary reassessment of the seminal 'Lanai fog drip study' in Hawai'i. *Hydrol. Processes*, **25**, 402–410, <https://doi.org/10.1002/hyp.7803>.
- Kuczera, G., B. Renard, M. Thyer, and D. Kavetski, 2010: There are no hydrological monsters, just models and observations with large uncertainties! *Hydrol. Sci. J.*, **55**, 980–991, <https://doi.org/10.1080/02626667.2010.504677>.
- Kumar, S. V., J. Dong, C. D. Peters-Lidard, D. Mocko, and B. Gomez, 2017: Role of forcing uncertainty and background model error characterization in snow data assimilation. *Hydrol. Earth Syst. Sci.*, **21**, 2637–2647, <https://doi.org/10.5194/hess-21-2637-2017>.
- Legates, D. R., and C. J. Willmott, 1990a: Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.*, **41**, 11–21, <https://doi.org/10.1007/BF00866198>.
- , and —, 1990b: Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.*, **10**, 111–127, <https://doi.org/10.1002/joc.3370100202>.

- Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Climate*, **26**, 9384–9392, <https://doi.org/10.1175/JCLI-D-12-00508.1>.
- Longman, R. J., H. F. Diaz, and T. W. Giambelluca, 2015: Sustained increases in lower tropospheric subsidence over the central tropical North Pacific drives a decline in high elevation precipitation in Hawaii. *J. Climate*, **28**, 8743–8759, <https://doi.org/10.1175/JCLI-D-15-0006.1>.
- , and Coauthors, 2018: Compilation of climate data from heterogeneous networks across the Hawaiian Islands. *Sci. Data*, **5**, 180012, <https://doi.org/10.1038/sdata.2018.12>.
- , and Coauthors, 2019: High-resolution gridded daily precipitation and temperature for the Hawaiian Islands (1990–2014). *J. Hydrometeorol.*, **20**, 489–508, <https://doi.org/10.1175/JHM-D-18-0112.1>.
- Lundquist, J. D., M. Hughes, B. Henn, E. D. Gutmann, B. Livneh, J. Dozier, and P. Neiman, 2015: High-elevation precipitation patterns: Using snow measurements to assess daily gridded datasets across the Sierra Nevada, California. *J. Hydrometeorol.*, **16**, 1773–1792, <https://doi.org/10.1175/JHM-D-15-0019.1>.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *J. Climate*, **15**, 3237–3251, [https://doi.org/10.1175/1520-0442\(2002\)015<3237:ALTHBD>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2).
- Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. Da Fonseca, and J. Kent, 2000: Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858, <https://doi.org/10.1038/35002501>.
- Mizukami, N., and Coauthors, 2016: Implications of the methodological choices for hydrologic portrayals of climate change over the contiguous United States: Statistically downscaled forcing data and hydrologic models. *J. Hydrometeorol.*, **17**, 73–98, <https://doi.org/10.1175/JHM-D-14-0187.1>.
- Monaghan, A. J., and Coauthors, 2016: On the seasonal occurrence and abundance of the Zika virus vector mosquito *Aedes Aegypti* in the contiguous United States. *PLOS Currents Outbreaks*, <https://doi.org/10.1371/currents.outbreaks.50dfc7f46798675fc63e7d7da563da76>.
- Newman, A. J., M. P. Clark, J. Craig, B. Nijssen, A. Wood, E. Gutmann, N. Mizukami, L. Brekke, and J. R. Arnold, 2015: Gridded ensemble precipitation and temperature estimates for the contiguous United States. *J. Hydrometeorol.*, **16**, 2481–2500, <https://doi.org/10.1175/JHM-D-15-0026.1>.
- , —, R. L. Longman, and T. W. Giambelluca, 2019: Methodological intercomparisons of station-based gridded meteorological products: Utility, limitations, and paths forward. *J. Hydrometeorol.*, **20**, 531–547, <https://doi.org/10.1175/JHM-D-18-0114.1>.
- Papalexiou, S. M., 2018: Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. *Adv. Water Resour.*, **115**, 234–252, <https://doi.org/10.1016/j.advwatres.2018.02.013>.
- Peck, E. L., and M. J. Brown, 1962: An approach to the development of isohyetal maps for mountainous areas. *J. Geophys. Res.*, **67**, 681–694, <https://doi.org/10.1029/JZ067i002p00681>.
- Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate*. 1st ed. American Institute of Physics, 520 pp.
- Prein, A., and A. Gobiet, 2017: Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. *Int. J. Climatol.*, **37**, 305–327, <https://doi.org/10.1002/joc.4706>.
- Ribatet, M., 2013: Spatial extremes: Max-stable processes at work. *J. Soc. Fr. Stat.*, **154** (2), 156–177.
- Running, S. W., R. R. Nemani, and R. D. Hungerford, 1987: Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evaporation and photosynthesis. *Can. J. For. Res.*, **17**, 472–483, <https://doi.org/10.1139/x87-081>.
- Sakai, A. K., W. L. Wagner, and L. A. Mehrhoff, 2002: Patterns of endangerment in the Hawaiian flora. *Syst. Biol.*, **51**, 276–302, <https://doi.org/10.1080/10635150252899770>.
- Sheffield, J., G. Goteti, and E. F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate*, **19**, 3088–3111, <https://doi.org/10.1175/JCLI3790.1>.
- Slater, A. G., and M. P. Clark, 2006: Snow data assimilation via an ensemble Kalman filter. *J. Hydrometeorol.*, **7**, 478–493, <https://doi.org/10.1175/JHM505.1>.
- Thornton, P. E., M. M. Thornton, B. W. Mayer, Y. Wei, R. Devarakonda, R. S. Vose, and R. B. Cook, 2016: Daymet: Daily Surface Weather Data on a 1-km Grid for North America, version 3. Time period: 1980-01-01 to 2015-12-31, Spatial Range: Full North America and Hawaii domains, ORNL DAAC, accessed 29 September 2016, <https://doi.org/10.3334/ORNLDAAAC/1328>.
- Turkman, K. F., M. A. A. Turkman, and J. M. Pereira, 2010: Asymptotic models and inferences for extremes of spatio-temporal data. *Extremes*, **13**, 375–397, <https://doi.org/10.1007/s10687-009-0092-8>.
- USBR, 2012: Colorado River basin water supply and demand study. U.S. Bureau of Reclamation, 1697 pp., <https://www.usbr.gov/lc/region/programs/crbstudy.html>.
- Walton, D., and M. Grieco, 2017: Evaluating extreme precipitation in gridded datasets: A case study over California. *2017 Fall Meeting*, New Orleans, LA, Amer. Geophys. Union, Abstract A41C-2295.
- , and A. Hall, 2018: An assessment of high-resolution gridded temperature datasets over California. *J. Climate*, **31**, 3789–3810, <https://doi.org/10.1175/JCLI-D-17-0410.1>.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Vol. 100, Academic Press, 627 pp.
- Willmott, C. J., and S. M. Robeson, 1995: Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatol.*, **15**, 221–229, <https://doi.org/10.1002/joc.3370150207>.
- Xia, Y., and Coauthors, 2012: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**, D03109, <https://doi.org/10.1029/2011JD016048>.
- Yee, T., and A. Stephenson, 2007: Vector generalized linear and additive extreme value models. *Extremes*, **10**, 1–19, <https://doi.org/10.1007/s10687-007-0032-4>.