Probabilistic Spatial Meteorological Estimates for Alaska and the Yukon

Andrew J Newman[1], Martyn P. Clark[2], Andrew W Wood[1], and Jeffrey Richard Arnold[3]

[1]National Center for Atmospheric Research (UCAR)
[2]University of Saskatchewan at Canmore
[3]US Army Corps of Engineers

November 22, 2022

Abstract

Alaska and the Yukon are a challenging area to develop observationally based spatial estimates of meteorology. Complex topography, frozen precipitation undercatch, and extremely sparse observations all limit our capability to accurately estimate historical conditions. In this environment it is useful to develop probabilistic estimates of precipitation and temperature that explicitly incorporate spatiotemporally varying uncertainty and bias corrections. In this paper we exploit recently-developed ensemble Climatologically Aided Interpolation (eCAI) systems to produce daily historical observations of precipitation and temperature across Alaska and the Yukon territory at a 2 km grid spacing for the time period 1980-2013. We extend the previous eCAI method to include an ensemble correction methodology to address precipitation gauge undercatch and wetting loss, which is of high importance for this region. Leave-one-out cross-validation shows our ensemble has little bias in daily precipitation and mean temperature at the station locations, with an overestimate in the daily standard deviation of precipitation. The ensemble has skillful reliability compared to climatology and significant discrimination of events across different precipitation thresholds. Comparing the ensemble mean climatology of precipitation and temperature to PRISM and Daymet v3 show large inter-product differences, particularly in precipitation across the complex terrain of SE and northern Alaska. Finally, long-term mean loss adjusted precipitation is up to 36% greater than the unadjusted estimate in windy areas that receive a large fraction of frozen precipitation.

Posted on 22 Nov 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.1002/essoar.10502438.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Probabilistic Spatial Meteorological Estimates for Alaska and the Yukon

Andrew J. Newman[1], Martyn P. Clark[1*], Andrew Wood[1], and Jeffrey R. Arnold[2]

[1]National Center for Atmospheric Research, Boulder, CO, USA.

[2]U. S. Army Corps of Engineers, Climate Preparedness and Resilience Program, Seattle, WA, USA

Corresponding author: Andrew Newman (anewman@ucar.edu)

[*]Current affiliation: University of Saskatchewan, Centre for Hydrology and Coldwater Laboratory, Canmore, AB

Key Points:

- It is difficult to estimate meterology across the Arctic due to complex topography, sparse observations, and frozen precipitation
- We developed a probabilistic approach to estimate precipitation and temperature including precipitation undercatch
- There are large differences between independent products, and correspondingly large uncertainties in our product

Abstract

Alaska and the Yukon are a challenging area to develop observationally based spatial estimates of meteorology. Complex topography, frozen precipitation undercatch, and extremely sparse *in situ* observations all limit our capability to accurately estimate historical conditions. In this environment it is useful to develop probabilistic estimates of precipitation and temperature that explicitly incorporate spatiotemporally varying uncertainty and bias corrections. In this paper we exploit recently-developed ensemble Climatologically Aided Interpolation (eCAI) systems to produce daily historical observations of precipitation and temperature across Alaska and the Yukon territory at a 2km grid spacing for the time period 1980-2013. We extend the previous eCAI method to include an ensemble correction methodology to address precipitation gauge undercatch and wetting loss, which is of high importance for this region. Leave-one-out cross-validation shows our ensemble has little bias in daily precipitation and mean temperature at the station locations, with an overestimate in the daily standard deviation of precipitation. The ensemble has skillful reliability compared to climatology and significant discrimination of events across different precipitation thresholds. Comparing the ensemble mean climatology of precipitation and temperature to PRISM and Daymet v3 show large inter-product differences, particularly in precipitation across the complex terrain of SE and northern Alaska. Finally, long-term mean loss adjusted precipitation is up to 36% greater than the unadjusted estimate in windy areas that receive a large fraction of frozen precipitation.

Plain Language Summary

Alaska and the Yukon are a challenging area to create spatial maps of precipitation and temperature. Very rugged terrain and extreme conditions, particularly snow and wind limit our ability to measure historical conditions. Because of this, it is critical to understand how uncertain our products are. Here we develop a new estimate of uncertainty for historical meteorology and include corrections to precipitation measurements for errors due to snowfall and wind. We show that our uncertainty estimates are reliable as compared to our observations, and there are large differences between several independent mapping efforts, and our precipitation corrections increase precipitation by more than 30% in some regions.

1 Introduction

Complex topography, extremely sparse *in situ* observations, and a high percentage of frozen precipitation and resultant precipitation undercatch issues limit our ability to estimate historical conditions across the Arctic (Serreze et al., 2003). The sparse observation networks across the Arctic may not be adequate, resulting in interpolation uncertainties and biases. Additional systematic biases such as wind undercatch or wetting loss can be corrected using existing empirical functions, but these correction functions are uncertain and often require additional observations not directly available (e.g., wind speed at station locations). As a consequence, historical estimates of precipitation and temperature based only on station observations in the Arctic can have significant biases and are intrinsically uncertain.

It is important to explicitly estimate the uncertainty in precipitation estimates in the Arctic. Traditional deterministic products may use complex interpolation routines to account for the impact of topographic gradients on spatial meteorological estimates (e.g. Daly et al. 1994, 2009) or implement under-catch corrections to reduce systematic precipitation biases (Adam and Lettenmaier 2003). Yet these deterministic products do not generally account for uncertainty or create ensemble estimates for the end-user community. Ad-hoc “ensembles of opportunity” from deterministic products may be created if at least a few products are available for the domain of interest (e.g. Henn et al. 2018). However, this method is of limited value if there are very few products available, particularly if the products use similar methods (e.g. the same topographic correction method). For the Alaska and Yukon region, two daily gridded *in situ* observation-based products are available as of May 2019: Daymet version 3 (https://daymet.ornl.gov/, Thornton et al. 2018), and a new product from the Pacific Climate Impacts Consortium product (Werner et al. 2019). Other remotely sensed (e.g. IMERG, Huffman et al. 2018) and model-based products available (e.g. ERA5, Hershbach and Dee 2016) that could be included in an ad-hoc or weighted ensemble (e.g. Beck et al. 2019), depending on user preferences and application needs.

In this study we modify the ensemble CAI (eCAI) methodology developed in Newman et al. (2019a) in two ways. First, we incorporate probabilistic estimates of the climatological precipitation and temperatures (Newman and Clark 2019). This method follows the general concepts of Daly et al. (1994, 2000, 2002, 2007, 2008), where a DEM provides additional spatial information by including known physical relationships between topography and meteorological variables in the statistical model. The ensemble of daily precipitation and temperature is conditioned on the probabilistic climatological estimates. Second, we include an ensemble correction methodology to address precipitation gauge undercatch and wetting loss, which is of high importance for this region. The final dataset provides daily probabilistic estimates of precipitation and temperature at 2 km spatial resolution over the time period 1980-2013; high spatial resolution data is needed in areas with large climate gradients for many applications such as streamflow forecasting.

The remainder of this paper is organized as follows. The input datasets are described in section 2, the underlying methods are presented in section 3. Section 4 contains detailed ensemble validation and comparisons to other datasets, and section 5 has description of the ensemble after loss corrections are applied. Finally, summary and data availability are found in sections 6 and 7, respectively.

2 Datasets

A variety of input data sources are used including a digital elevation model (DEM), point observations from various sources, and the gridded North American Regional Reanalysis (NARR, Mesigner et al. 2006) product. The NARR is used in estimating the default temperature lapse rates in the climatologically aided interpolation (Section 3a) and the monthly climatological wind speed for gauge undercatch estimates (Section 3b) instead of other products (e.g. the high resolution WRF simulations in Monaghan et al. 2018) because it provides data for the same temporal period as our ensemble product.

2.1 Domain

The domain of the ensemble product (Figure 1) covers nearly the entire U.S. state of Alaska and the Canadian Yukon Territory, and NW portions of British Columbia within the Yukon River watershed. This domain balances coverage over areas with observations, population, and key watersheds that flow through Alaska. The DEM is at 2 km grid spacing and is based on the Scenarios Network for Alaska and Arctic Planning (SNAP) DEM, which is a

'Barnes filtered GTOPO30 DEM obtained from the PRISM climate group' (SNAP 2019). The use of this DEM provides the opportunity to compare against the PRISM products.

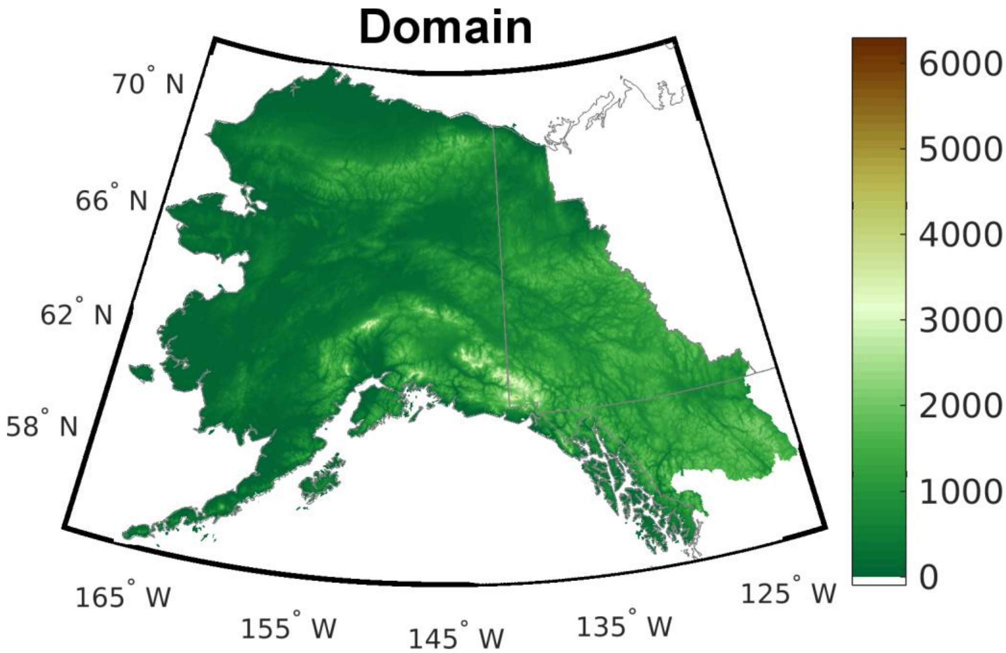

Figure 1. Spatial extent of the domain (colored pixels) with geopolitical boundaries in gray.

2.2 Input Station Data

Station observations were obtained from two sources, the Imiq Hydroclimate Database and Data Portal through the Arctic Landscape Conservation Cooperative (Cherry et al. 2016), and the adjusted daily precipitation dataset for Canada (Wang et al. 2017). The Imiq dataset contains many local network stations across Alaska not included in the GHCN-D as well as the available Alaskan stations in the GHCN-D. The adjusted daily precipitation dataset for Canada contains observations from Environment and Climate Change Canada, with extensive quality control and undercatch corrections (e.g. wind undercatch and wetting losses) applied to the data (Wang et al. 2017). Therefore, we did not use the Global Historical Climatology Network – Daily (GHCN-D) from the National Centers for Environmental Information (Menne et al. 2012a,b). Figure 2 illustrates the input precipitation and temperature station networks.

Differences in gauge type and measurement practices between the U.S. and Canada, particularly for snowfall, result in discontinuities in both amount and occurrence for precipitation estimates across the US-Canadian border irrespective of the native input station data from Canada (e.g. Adam and Lettenmaier 2003, hereafter AL03; Scaff et al. 2015). Here we attempt to minimize the differences between the two datasets through additional quality control of the Imiq data as well as producing ensemble estimates of gauge losses due to wind undercatch and wetting losses (section 3b). Gauge wind undercatch can be extreme for frozen precipitation depending on the gauge type (Goodison et al. 1998; Kochendorfer et al. 2018); thus an initial quality screening step was performed during the colder half of the year (November-April) for stations in the Imiq database with significantly less precipitation occurrence than a reference climatology. Canadian stations were not screened because Wang et al. (2017) perform extensive quality control.

The reference climatology is a high-resolution, 4 km grid spacing, WRF regional climate model (RCM) simulation shown to have good representation of the Alaskan climate (Monaghan et al. 2018). The long-term November-April probability of precipitation (PoP) from the WRF simulation (2002-2016) using a precipitation threshold of 0.5 mm was compared to the long-term PoP from all Imiq stations. Stations that had significantly smaller PoP, more than 0.15 less than the WRF estimated PoP, were removed for this half of the year (Fig. 2a). Often these stations reported little-to-no precipitation for the entire period while the WRF simulation or other higher quality stations within the same general area had PoP values ~0.2-0.3.

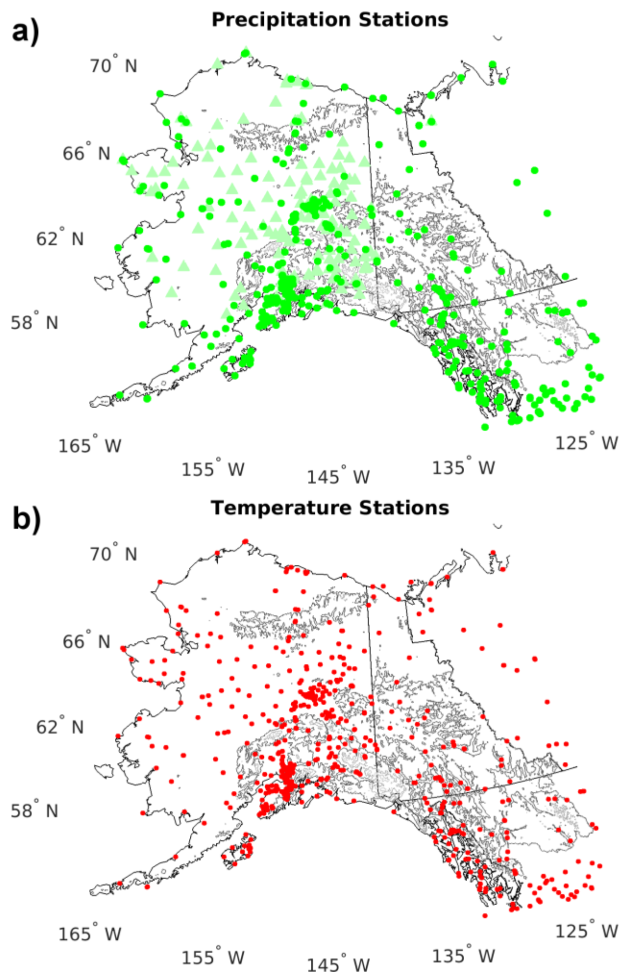

Figure 2. a) Precipitation stations and b) temperature stations included in the daily ensemble product. Precipitation stations denoted with light shaded triangles are included only for the warm season (May-October), while dark circles in both a) and b) are used over all days.

3 Ensemble Methodology

The ensemble generation methodology, the Gridded Meteorological Ensemble Tool (GMET), has been applied to a diverse range of climate conditions across the Contiguous United States (CONUS) and southern Canada (Clark and Slater 2006, Newman et al. 2015), Hawaii (Newman et al. 2019a), and now Alaska (this study). The CONUS application is a straightforward application of probabilistic interpolation methods on daily time scales (Newman et al. 2015), while the Hawaii application developed the ensemble Climatologically Aided

Interpolation (eCAI) method. In Hawaii, the climatological station network was sufficiently dense to use the same underlying methodology for both the climatology and daily steps.

In Alaska, our climatological station network is also the same as the daily network, but is of insufficient density to use the ensemble methodology for both the climatology and the daily time steps. This was determined during initial testing wherein non-physical lapse rates where estimated for many grid points. Hence we implemented knowledge-based climatology interpolation for climatological estimation within the general eCAI method. Further, because gauge undercatch from wind and wetting loss can be significant at high latitudes (e.g. AL03) an ensemble approach to estimating wind undercatch and wetting loss has been developed and is described in Section 3b. Figure 3 provides a flow chart of GMET development from the initial concept (Clark and Slater 2006) through the current ensemble gauge undercatch methodological addition.

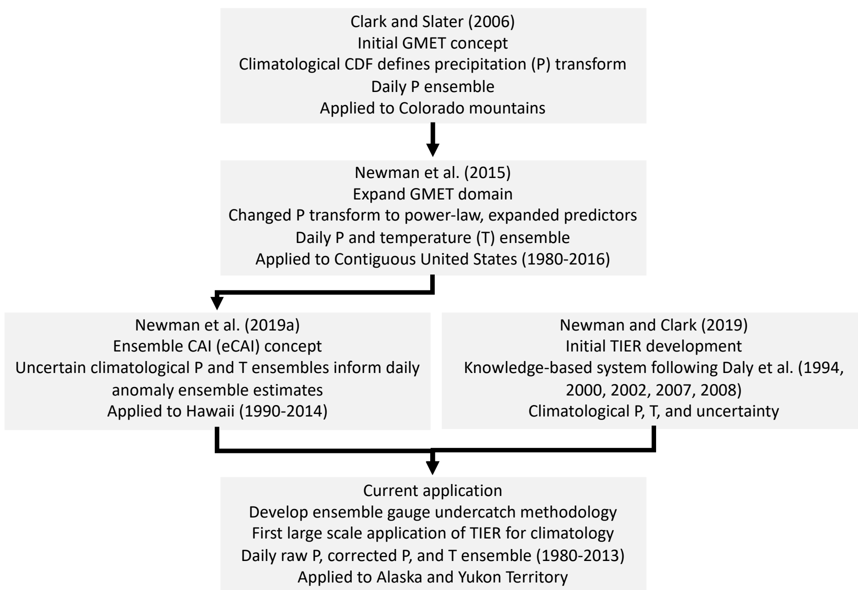

Figure 3. Development flow chart for the Gridded Meteorological Ensemble Tool (GMET) and other directly related development from the initial GMET concept paper, Clark and Slater (2006), through the current work.

3.1 Ensemble Climatologically Aided Interpolation

The ensemble generation methodology uses locally weighted multiple logistic and linear regression to estimate grid point distributions of precipitation (Clark and Slater 2006, Newman et al. 2015, 2019a). Locally weighted multiple linear regression is used to estimate mean daily temperature and the diurnal temperature range (DTR) distributions (Newman et al. 2019a). After grid point distributions are estimated, ensemble members are generated by sampling from those distributions using spatiotemporally correlated random fields. These fields include spatial and temporal correlation length scales for each variable estimated from the input data (Newman et al. 2019a). In Clark et al. (2006) and Newman et al. (2015), these probability distributions for precipitation and temperature are estimated for each grid point and day using the daily data.

In CAI, climatological grids are used to inform the spatial interpolation of climate variables (e.g. Willmott and Robeson 1995). In CAI daily ratios are interpolated for precipitation (the fraction of climatological precipitation on a given day), and anomalies are interpolated for mean temperature and the DTR (the difference between temperature and the climatological mean temperature). The interpolated daily ratios (or daily anomalies) are then multiplied (added) to the climatological values. In the eCAI system, an ensemble of climatological values is first estimated, and the ensemble mean values are used to compute the station daily ratios for precipitation or anomalies for temperature. Distributions of daily ratio or anomalies are then estimated using multiple logistic/linear regression techniques and sampled using spatiotemporally correlated random fields. Finally, the ensemble of climatological values is used in the back transformation of the estimated daily ratios or anomalies. This accounts for uncertainty in both the climatological and daily time scales (Newman et al. 2019a). Here we summarize the GMET theory and eCAI methodology following Newman et al. (2019a), then briefly review the TIER methodology (Newman and Clark 2019).

3.1.1 Theory

Precipitation is an intermittent process at many timescales, thus the probability distribution function (PDF) of precipitation usually contains a concentration at zero. Precipitation can therefore be modeled as $P(X_P = 0) = p_o$, where p_o is the probability of zero precipitation, with a CDF for the rest of the values $X_P > 0$. From Papalexiou (2018) and Newman et al. (2019a), the CDF for precipitation, X_P, can be written as

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

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

$$F_{X_T}(x_T) \tag{1b}$$

$$F_{X_D}(x_D) \tag{1c}$$

Transformation functions are used to map precipitation to a normal distribution. While Clark and Slater (2006) use the empirical CDF of precipitation derived from the historical observations, Newman et al. (2015) use a parametric power-law transformation to remove the requirement of computing and storing the empirical CDFs for each station:

$$X_P = v(Y_P) = {Y_P}^{\alpha} \tag{2a}$$

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

where α is the transform exponent and the transformation is performed on only non-zero precipitation values. Y_P is assumed to be normally distributed:

$$Y_P \sim \mathcal{N}\left(\mu_{Y_P}, \sigma^2_{Y_P}\right), \qquad \text{for } x_P > 0 \tag{3a}$$

with mean μ_{Y_P} and variance $\sigma^2_{Y_P}$. Temperature and the DTR are assumed to be Gaussian without having to perform any transformation:

$$X_T \sim \mathcal{N}(\mu_T, \sigma_T) \tag{3b}$$

$$X_D \sim \mathcal{N}(\mu_D, \sigma_D) \tag{3c}$$

Ensemble realizations of precipitation and temperature are generated by sampling from the estimated distributions of $\mathbf{Y}_{\mathrm{P}}$, $\mathbf{X}_{\mathrm{T}}$, and $\mathbf{X}_{\mathrm{D}}$ at all valid grid points and time steps through the use of spatiotemporally correlated random fields. At the first timestep, we draw from:

$$\mathbf{Z}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_Z) \tag{4a}$$

where $\mathbf{\Sigma}_Z$ is the covariance matrix:

$$\mathbf{\Sigma}_Z = [\psi_{gg'}] = \psi(|s_g - s_{g'}|) \tag{4b}$$

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. See Newman et al. (2019a) for specific details regarding the spatial correlation function.

The spatial fields at subsequent times, $\mathbf{Z}_{\mathrm{t}} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_Z)$ are defined to represent the autocorrelation for mean temperature and the DTR, as well as the instantaneous cross-correlation between the DTR and precipitation (see Newman et al. 2019a). Once $\mathbf{Z}_{(\cdot),\boldsymbol{t}}$ is computed, physical values are generated through:

$$\hat{x}_{u,g} = v\left(F_Z^{-1}(u_g)\right) \tag{5}$$

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

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

$$\hat{x}_{u,P} = \begin{cases} 0 & ,0 \le u_P \le \hat{p}_o \\ v\left(F_P^{-1}\left(\frac{u_P - \hat{p}_o}{1-\hat{p}_o}\right)\right), & \hat{p}_o < u_P \le 1 \end{cases} \tag{6a}$$

$$\hat{x}_{u,T} = F_T^{-1}(u_T) \tag{6b}$$

$$\hat{x}_{u,D} = F_D^{-1}(u_D) \tag{6c}$$

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

To implement CAI into the ensemble framework, monthly climatological values and an ensemble of monthly climatological values are computed first following Eq. 6:

$$\hat{x}_{u_PC} = \begin{cases} 0 & ,0 \le u_{PC} \le \hat{p}_{oC} \\ v\left(F_{PC}^{-1}\left(\frac{u_{PC} - \hat{p}_{oC}}{1-\hat{p}_{oC}}\right)\right), & \hat{p}_{oC} < u_{PC} \le 1 \end{cases} \tag{7a}$$

$$\hat{x}_{u_TC} = F_{TC}^{-1}(u_{TC}) \tag{7b}$$

$$\hat{x}_{u_DC} = F_{DC}^{-1}(u_{DC}) \tag{7c}$$

The daily ratio and daily anomalies for precipitation, temperature, and DTR are then computed using the estimated climatological ensemble means for the closest grid point to each station. Ensemble realizations of daily precipitation and temperature at each time step and grid point follow Equations (4-5), with the inclusion of transforming the anomalies back to physical space. Precipitation is computed as:

$$\hat{x}_{u_{PA}} = \begin{cases} 0 & , \ 0 \leq u_{PA} \leq \hat{p}_{oA} \\ v\left(F_{Y_{PA}}^{-1}\left(\frac{u_{PA}-\hat{p}_{oA}}{1-\hat{p}_{oA}}\right)\right), & \hat{p}_{oA} < u_{PA} \leq 1 \end{cases} \tag{8}$$

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

$$\hat{x}_{u_P} = \hat{x}_{u_{PC}}^{(i)} \hat{x}_{u_{PA}}^{(i)} \tag{9}$$

where *(i)* is the *i*th ranked value of $\hat{x}_{u_{PA}}$. Mean daily temperature and DTR are simple because they are continuous, assumed Gaussian, and the anomalies have physical units. Thus, the simulated mean temperature and DTR are simply:

$$\hat{x}_{u_T} = F_T^{-1}(\hat{u}_T) \tag{10a}$$

$$\hat{x}_{u_D} = F_D^{-1}(\hat{u}_D) \tag{10b}$$

3.1.2 Topographically Informed Regression

One of the underlying assumptions of the ensemble system is that the observation station density is sufficient to resolve gradients given the chosen predictor set. GMET uses the simple least squares estimate and no bounds are placed on the regression coefficients to enforce physical realism in the final regression equation. In Hawaii, the station density assumption was not met for the daily station network, necessitating development of eCAI. In Alaska, this assumption is not met for either the climatology or daily time steps. We therefore applied the Topographically InformEd Regression (TIER) model to generate the climatological precipitation, mean temperature, and DTR, and their corresponding uncertainty estimates for input into the climatological ensemble generation step.

It is well known that surface characteristics influence the climatological distribution of precipitation and temperature (e.g. Alter 1919; Spreen 1947; Chua and Bras 1982; Phillips et al. 1992; Daly et al. 1994; Clark and Slater 2006). The Topographically InformEd Regression (TIER) model is a knowledge-based, meteorological variable-elevation locally weighted linear regression model. It incorporates our knowledge of atmospheric physics through modification of the weights assigned to each observation for each grid point in the linear regression model following Daly et al. (1994, 2000, 2002, 2007, 2008) (Newman and Clark 2019). Specifically, Daly et al. (1994) developed a knowledge based regression system focused on topographic facets. A facet is defined as a continuous area with similar aspects, or slope orientation, using a smoothed DEM (Daly et al. 1994). Additional studies added complexity to this base model through further additions to account for physics related to new variables (e.g. Temperature) and other processes responsible for spatial gradients such as distance from the coast or valley inversions (e.g. Daly et al. 2008).

Before spatial estimates of precipitation and temperature are generated, several preprocessing steps are performed with the DEM to derive the necessary topographical attributes. First, the DEM is smoothed to remove high frequency, microscale topographic features. Once the DEM is smoothed, topographic facets are created. Currently, TIER has five (5) facets: 1) North (aspect > 315°, aspect ≤ 45°); 2) East (45° < aspect ≤ 135°); 3) South (135° < aspect ≤ 225°); 4) West (225° < aspect ≤ 315°); and 5) Flat. Flat aspects are areas with terrain gradients (slopes) less than a user-specified gradient value. Next, the topographic position of each grid cell and the corresponding placement within an idealized two layer atmosphere are computed. These two attributes are focused on temperature as they identify valleys and areas commonly within inversion layers in an idealized two-layer atmosphere (Daly et al. 2002). Finally, the distance to the coast is determined for all valid grid points in the domain.

The core of the TIER model is the locally varying station weight vector defined at each grid point (Newman and Clark 2019):

$$\mathbf{W} = \mathbf{W_d W_f W_l W_t W_p} \tag{11}$$

Where $\mathbf{W}$ is the final weight vector, $\mathbf{W_d}$ are the distance dependent weights, $\mathbf{W_f}$ are the facet weights, $\mathbf{W_l}$ are the atmospheric layer weights, $\mathbf{W_t}$ are the topographic position weights, and $\mathbf{W_p}$ are the coastal proximity weights. For precipitation, only $\mathbf{W_d}$, $\mathbf{W_f}$, and $\mathbf{W_p}$ are used in the final $\mathbf{W}$, while all five component weights are used for calculating $\mathbf{W}$ for temperature. Once the station weight vector is defined, a base grid point estimate is developed as the weighted average of nearby stations up to *n* stations:

$$\hat{\mu}_{(\cdot)_o} = \sum_{i=1}^{n_s} y_i * W_i \tag{12}$$

where $\hat{\mu}_{(\cdot)_o}$ is the base grid point estimate of precipitation ($\hat{\mu}_{P_o}$), mean temperature ($\hat{\mu}_{T_o}$), or DTR ($\hat{\mu}_{D_o}$), and y_i and W_i are the observed station value and the station weight for station *i*, respectively. Next, the meteorological field-elevation lapse rate is estimated, and then the grid point value is estimated as:

$$\hat{\mu}_{(\cdot)} = \hat{\mu}_{(\cdot)_o} + \hat{\beta}_{1,(\cdot)}\Delta E, \qquad \beta_{1M,(\cdot)} \leq \hat{\beta}_{1,(\cdot)} \leq \beta_{1X,(\cdot)} \tag{13}$$

where $\hat{\beta}_{1,(\cdot)}$ is the regression estimated lapse rate for any variable, ΔE is the difference between the smoothed DEM elevation and the $\mathbf{W}$ weighted station elevation using the smoothed DEM station elevations, and $\beta_{1M,(\cdot)}$ and $\beta_{1X,(\cdot)}$ are the user defined, variable specific minimum and maximum valid regression lapse rates.

In TIER, the uncertainty of $\hat{\mu}_{(\cdot)_o}$ is estimated as the standard deviation of the leave-one-out estimates, which is all possible combinations of n_r-1 stations, $\binom{n_r}{n_r-1}$:

$$\hat{\sigma}_{(\cdot)_o} = \sqrt{\frac{\sum_{i=1}^{n_r}\left(\hat{\mu}_{(\cdot)_o,-1}\right)^2}{n_r-1}} \tag{14}$$

where n_r is the subset of stations that are both within a user defined distance threshold and on the same facet as the current grid cell, $\hat{\sigma}_{(\cdot)_o}$ is the estimated standard deviation of $\hat{\mu}_{(\cdot)_o}$, and $\hat{\mu}_{(\cdot),-1}$ is the estimated value when the *i*-th station is withheld. The standard deviation of all valid lapse rate estimates from the leave-one-out estimates, $\binom{n_r}{n_r-1}$, is used as the uncertainty estimate of $\hat{\beta}_{1,(\cdot)}$:

$$\hat{\sigma}_{\beta_{1,(\cdot)}} = \sqrt{\frac{\sum_{i=1}^{n_r}\left(\hat{\beta}_{1,(\cdot),-1}\right)^2}{n_r-1}} \tag{15}$$

where $\hat{\sigma}_{\beta_{1,(\cdot)}}$ is the estimated standard deviation of $\hat{\beta}_{1,(\cdot)}$ and $\hat{\beta}_{1,(\cdot),-1}$ is the estimated value when the *i*-th station is withheld. For precipitation, the lapse rates are normalized after the regression estimates are made; this reduces the large spatial variability due to inherently large differences in climatological precipitation amounts, which allows for lapse rate bounds to be applied domain-wide (Daly et al. 1994). Then, several post-processing steps are undertaken including filtering and feathering on the full precipitation field to reduce any remaining non-physical gradients in the precipitation field (e.g. Daly et al. 1994). For temperature, a spatial filter is used to smooth the lapse rate before the final temperature field is estimated (Newman and Clark 2019).

Finally, uncertainty estimates, $\hat{\sigma}_P$, $\hat{\sigma}_T$, and $\hat{\sigma}_D$ are computed as the combined standard deviation of the two-component uncertainty estimates in Eq. (14-15):

$$\hat{\sigma}_{(\cdot)} = \hat{\sigma}_{(\cdot)_o} + \hat{\sigma}_{\beta 1,(\cdot)} + 2\sqrt{\mathrm{cov}\left(\hat{\sigma}_{(\cdot)_o}, \hat{\sigma}_{\beta,(\cdot)}\right)} \tag{16}$$

Equation (16) accounts for any covariance between the two component uncertainties. The covariance is computed locally with a user-defined 2-D window of points around the current grid point. The results of Eq. (13) and (16) are used as inputs into the monthly climatological ensemble generation step (Eq. 6).

To summarize, the TIER approach differs from the climatological interpolation step in Newman et al. (2019a) in two important ways. First, TIER uses knowledge of the physical system to enforce spatial consistency in the interpolated fields (see the discussion at the beginning of this section). Second, TIER estimates the uncertainty in the interpolation using an estimate of the variance in the slope and intercept of the regression coefficients (Eqs. 14-16), while Newman et al. (2019a) uses the cross-validation error of the regression equation. Lastly, TIER uses geophysical attributes to estimate the spatial patterns of precipitation and temperature, which is similar to Newman et al. (2019a), except that TIER uses a simple linear regression formulation while Newman et al. (2019a) uses multiple linear regression. Exploration of how these differences could manifest in a final precipitation or temperature estimate is the subject of future work.

3.2 Ensemble Gauge Loss Corrections

Since gauge losses can be substantial in the snow-dominated environments in Alaska and the Yukon, we developed a gauge loss correction methodology in order to explicitly represent uncertainty in precipitation undercatch. Following AL03, we focus only on wetting loss and wind-induced undercatch. To account for uncertainty in the gauge undercatch terms, we use separate Gaussian distributions for wetting loss and wind undercatch:

$$X_a \sim \mathcal{N}(\bar{a}, \sigma_a) \tag{17a}$$

$$X_R \sim \mathcal{N}(C_R, \sigma_R) \tag{17b}$$

Where $\bar{a}$, and C_R are the grid point mean wetting loss and wind undercatch estimates with corresponding standard deviations σ_a and σ_R. We sample from Eq. (17) to generate an ensemble of monthly climatological correction factors. Note that this approach represents spatial variability in undercatch and assumes that temporal variability in undercatch is constant for each month.

Wetting loss, where precipitation is underestimated due to moisture wetting the gauge surfaces, is taken as the average wetting loss per event. We base our wetting loss estimates, $\bar{a}$, on the study of Sevruk and Hammon (1984), who calculated wetting losses for multiple gauge types in locations around the world. Sevruk and Hammon (1984) found negligible uncertainty in the estimate of $\bar{a}$ for a particular gauge, but larger differences between gauges, possibly as high as 0.2-0.3 mm per event between gauges. While Sevruk and Hammon (1984) define $\bar{a}$ for the US standard 8" non-recording gauge, this gauge type is not used at every site in Alaska. Because the ensemble system uses many gauges in each grid point estimate, a mix of gauge types is likely for each grid point. Assuming a blend of half U.S. standard gauges and half tipping bucket gauges, an ad-hoc estimate of wetting loss can be given for the Imiq gauges as:

$$\bar{a}_I = 0.13 \pm 0.05 \text{ mm} \tag{18}$$

This estimate is based on $\bar{a} = 0.2$ mm for the US standard gauges (Sevruk and Hammon 1984) and $\bar{a} = 0.06 \pm 0.02$ mm from qualitative inspection of the results in Niemczynowicz (1986) and Fankhauser (1998). The subscript I denotes the Imiq network stations. Eq. (18) accounts for the small uncertainty in an individual gauge estimate combined with the uncertainty of gauge type for an individual grid point.

Wind undercatch results from the gauge orifice disturbing the airflow, and creating updrafts that are able to divert some hydrometeors from entering the gauge (Groisman and Legates 1994; Nespor and Sevruk 1999). Gauge wind undercatch is particularly acute for low mass hydrometeors that respond quickly to the flow, such as snowfall, for which undercatch can be 20-50% or more depending on gauge configuration (Goodison et al. 1998; Kochendorfer et al. 2018). Typically, a power law relationship using wind speed as the explanatory variable is used for a specific gauge configuration to estimate the gauge catch efficiency; in the case of the U.S. standard 8" gauge with no shield for frozen precipitation, this is (Goodison et al. 1998):

$$C = \exp(a - bm^c) \tag{19}$$

where C is the catch efficiency in percent (0-100+), m is the wind speed at gauge height, and the coefficients are found through fitting experimental data. Goodison et al. (1988) used the coefficients a=4.61, b=0.16 and c=1.28. For the ensemble in this study we convert C into a catch ratio $C_{R,I} = \max\left(\frac{100}{C}, 1\right)$, the multiplicative factor used to correct gauge undercatch. Examination of Goodison et al. (1998) reveals significant scatter around the best estimate of C for a given wind speed for all gauge types.

Final grid point $\bar{a}$ and $C_R$values in Eq. (17) are the weighted average of the estimates for Imiq and Canadian stations as:

$$\bar{a} = w_I \bar{a}_I + (1 - w_I)\bar{a}_C \tag{20a}$$

$$C_R = w_I C_{R,I} + (1 - w_I) C_{R,C} \tag{20b}$$

where w is the fractional contribution of stations from a given network to the total for a grid point, and the subscripts I and C denote the Imiq and Canadian networks, respectively. $\bar{a}_I$ is determined from Eq. (18), $C_{R,I}$ from Eq. (19), and $\bar{a}_C = 0$, and $C_{R,C} = 1$ because the Canadian stations are already deterministically adjusted.

Lastly, σ_a and σ_R in Eq. (17) are determined in the same fashion using Eq. (20):

$$\sigma_a = w_I \sigma_{a,I} + (1 - w_I)\sigma_{a,C} \tag{21a}$$

$$\sigma_R = w_I \sigma_{R,I} + (1 - w_I)\sigma_{R,C} \tag{21b}$$

where $\sigma_{a,I}$ is taken from Eq. (18), $\sigma_{a,C}$ is set to 0.05 mm to account for unknown gauge distributions and correction methodologies in Canada. $\sigma_{R,C}$ and $\sigma_{a,I}$ are set to 0.05 from a qualitative examination of Goodison et al. (1998).

For this domain, there are very few wind measurements with significant record length, and fewer spanning a majority of the ensemble generation period. Therefore, we use the 1980-2013 monthly NARR 10-m climatological wind speed reduced to gauge height (1.1 m, Sevruk and Klemm (1989)) using a logarithmic wind profile with roughness lengths of 0.01 and 0.03 m for the cold and warm seasons respectively (Golubev et al. 1992), and a threshold wind speed of 6.5 m s^{-1} (Yang et al. 1998) as input to Eq. (19). Figure 4 illustrates the NARR climatological wind field for January and the corresponding climatological wind undercatch values using Eq. (19).

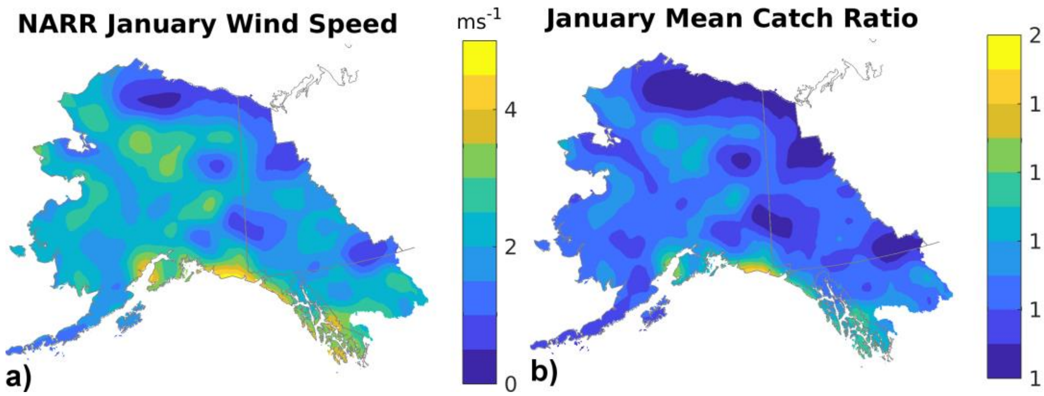

Figure 4. a) NARR 1980-2013 climatological January wind speed (m s^{-1}), and b) the corresponding January catch ratio (C_R) using only Eq. (19).

The grid point estimate for one day is a blended estimate using the estimated fractional liquid/frozen precipitation type as (Legates and Willmott 1990; AL03):

$$\hat{x}_{u,P'} = (1 - S)\kappa_r\left(\hat{x}_{u,P} + X_a\right) + SX_R\left(\hat{x}_{u,P} + X_a\right) \tag{22}$$

where $\hat{x}_{u,P'}$ is the final adjusted precipitation, and $\hat{x}_{u,P}$ is the raw current ensemble member precipitation from Eq. (6a), S is the fraction of precipitation falling as snow, and κ_r is the wind undercatch correction factor for rainfall and is set to 1 here. *S* is estimated at each grid point and day following Froidurot et al. (2014) using only daily mean air temperature:

$$S = 1 - \frac{1}{1+e^{a_0+a_1\hat{x}_{u,T}}} \tag{23}$$

where $\hat{x}_{u,T}$ is the current ensemble member mean daily 2 m air temperature from Eq. (6b), and a_0 and a_1 are fitted coefficients set here as a_0 = 2.2347 and a_1 = -1.7108.

4 Comparisons and Validation

In this section, comparisons between PRISM, Daymet, the WRF RCM simulation, and the ensemble product are presented along with in-depth deterministic and probabilistic leave-one-out validation statistics. Note that all precipitation comparisons and verification discussion in this section is in reference to the unadjusted precipitation.

4.1. Climatological Comparisons and Validation

The ensemble (hereafter Ensemble) mean 1980-2013 precipitation (mm yr^{-1}) and mean temperature (°C) are compared to PRISM, Daymet v3 (Thornton et al. 2018), and the WRF RCM. PRISM is a 1971-2000 climatology, while the Ensemble and Daymet are the 1980-2013 mean values from the daily fields, and the WRF RCM simulation spans 2002-2016. There is a general maximum in precipitation across Alaska along coastal zones, and the SE portion of the state in particular. Areas of complex terrain are also locations of relative maxima in precipitation (Fig. 5a-d). Of note are the large inter-product differences seen in precipitation, particularly across the complex terrain of SE and northern Alaska. Mean temperatures generally decrease from South to North across the domain, with the warmest areas near the southern coastline and the coldest areas along the northern edge of the state (Fig. 5e-h). All products have local minima in temperature at higher elevations, and PRISM, Daymet, and WRF are colder than the Ensemble mean (Fig. 5e-h). This suggests the temperature-elevation lapse rate in the Ensemble may need further investigation.

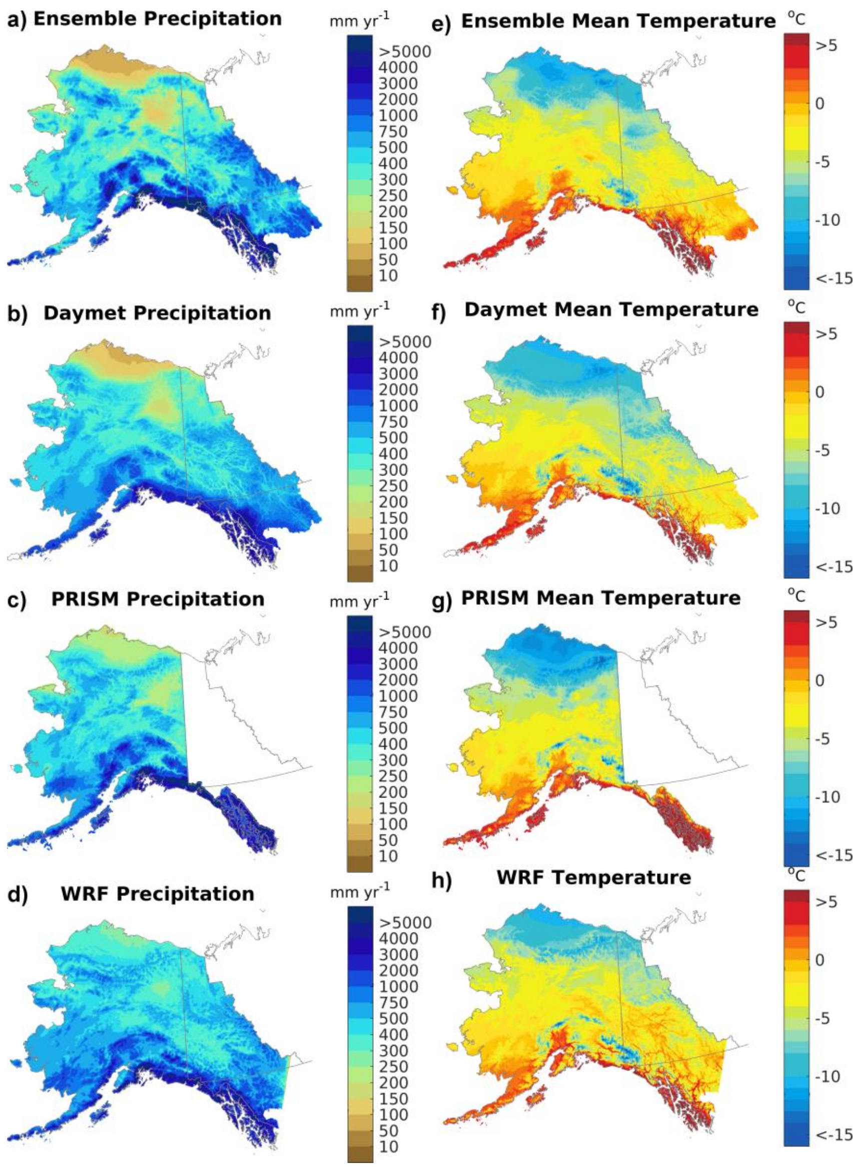

Figure 5. Long-term mean daily precipitation (mm yr^{-1}) along the left column for a) the Ensemble mean (1980-2013), b) Daymet (1980-2013), c) PRISM (1971-2000), and d) WRF (2002-2016). Long term mean temperature (°C) along the right column for a) the Ensemble mean (1980-2013), b) Daymet (1980-2013), c) PRISM (1971-2000), and d) WRF (2002-2016).

Figure 6 highlights differences in precipitation across Alaska between PRISM, Daymet, WRF RCM, and the Ensemble products. Differences in precipitation range from $\pm$ 50 % up to 120% in a few isolated areas across the domain. Overall, the Ensemble is the wettest observation based product using domain average precipitation with 2.73, 2.53, 2.76 mm day^{-1} or 997, 925, 1008 mm yr^{-1} for PRISM, Daymet, and the Ensemble mean respectively, while WRF has a domain average of 2.72 mm day^{-1} (994 mm yr^{-1}). However, PRISM is the wettest using grid cell comparisons with 51%, 31%, and 18% of the common domain having PRISM, Daymet, and the Ensemble mean as the maximum precipitation product. Finally WRF is wetter than all three observation based products in northern portions of the domain where gauge undercatch can be more severe (section 5). Correspondingly, the Ensemble estimated climatological relative uncertainty using the Ensemble standard deviation is greater than 20% for 65% of the domain, which empirically supports large differences between distinct products (not shown).

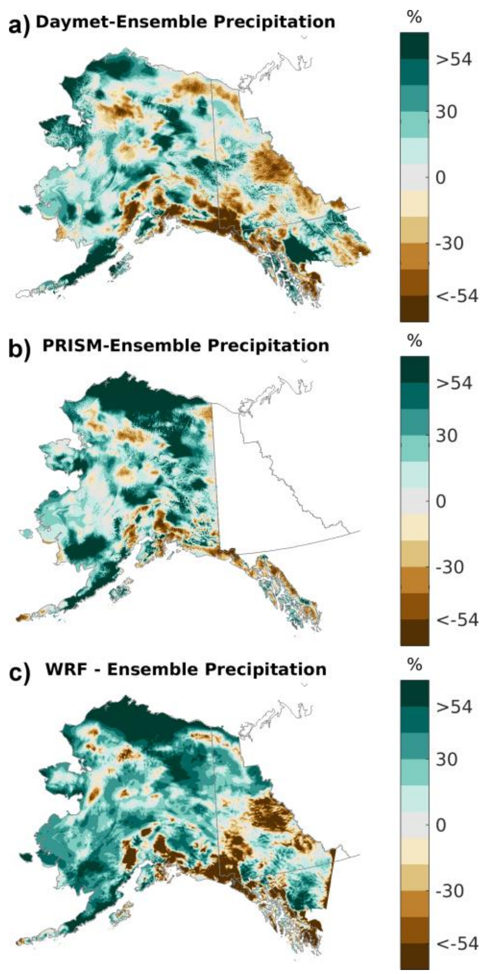

Figure 6. Relative (%) precipitation differences between a) PRISM – Ensemble mean, b) Daymet – Ensemble mean, and c) WRF – Ensemble mean.

Finally, Daymet and Ensemble climatological PoP values are compared in Figure 7. The Ensemble has higher PoP across most of the domain than Daymet with the largest differences occurring over the Canadian portion of the domain. Across the state of Alaska, the PoP differences are smaller in an absolute sense, but in most of interior Alaska, climatological PoP is ~0.2 or less. Specifically for the Ensemble product, leave-one-out cross validation (LOOCV) statistics for PoP show the Ensemble is nearly unbiased in Figure 8 when compared to the observations included in the Ensemble product, while Daymet PoP is underestimated for nearly all of the same observation sites, particularly for stations with less precipitation (Fig. 8b).

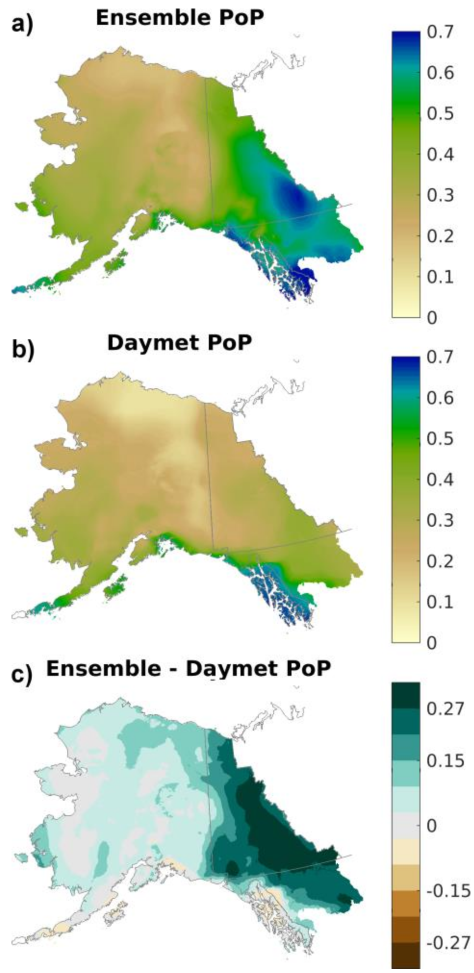

Figure 7. Climatological probability of precipitation (PoP) for a) Daymet, b) Ensemble, and c) Ensemble – Daymet difference field.

Because the PoP differences between Daymet and the Ensemble arise from either methodological or input data, or a combination, we can attempt to separate the influence of the two to identify the primary contributor. For this comparison, the inputs are significantly different. Daymet uses only GHCN-D data, which has not been quality checked for severe wind undercatch in winter, while the Ensemble discards these stations across Alaska and uses the adjusted daily precipitation dataset for Canada (Wang et al. 2017), which are the stations used for comparisons here, implying that the Daymet comparison may have many out-of-sample stations, while the Ensemble comparison uses LOOCV. Additionally, there is a conditional bias in the Daymet data with precipitation amount, which indicates a constant threshold for occurrence is problematic (e.g. Newman et al. 2019b). Thus, both input station differences and methodological decisions are playing a role in the PoP differences shown here.

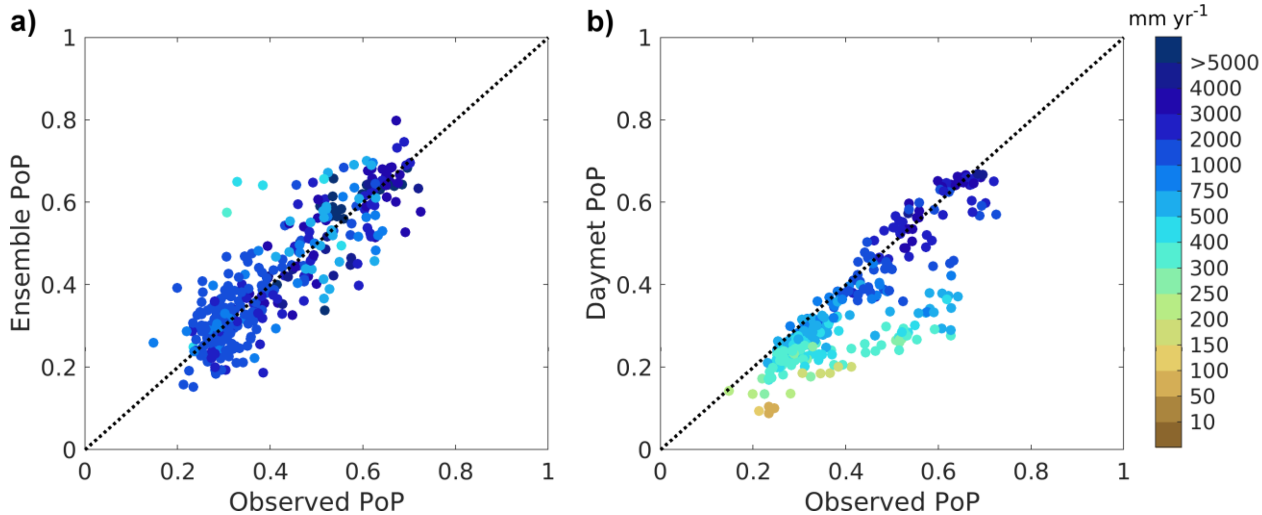

Figure 8. a) Leave-one-out cross-validation probability of precipitation (PoP) bias as compared to observations for the Ensemble, and b) PoP bias for Daymet using the same stations.

Overall, differences in precipitation and temperature may arise from many factors. First, PRISM is generated using a different climatological period. Next, each product is generated using different methodologies and input station networks. For example, the three products estimate the elevation lapse rate of precipitation and temperature in different ways. PRISM and the Ensemble use similar methods, but have many subtle differences in model parameters and methodological decisions. Furthermore, PRISM includes additional manual adjustments to more closely represent observed conditions in specific locations through a process of local expert review (Daly et al. 2009). Unfortunately, it is difficult to disentangle the precise choice(s) underlying these differences, even for PoP in this case, and is an area of active research (e.g. Newman et al. 2019b, Newman and Clark 2019).

4.2 Deterministic Daily Validation

Deterministic cross-validation comparisons to observation locations for precipitation amount and variability (represented as the standard deviation of the daily precipitation time series) shows that the Ensemble is nearly unbiased (Table 1) overall. The Ensemble has a small conditional bias for precipitation as seen in Figure 9a with a fitted slope of 1.09, indicating underestimation at low rainfall rates and overestimation at high rainfall rates. The conditional bias is worse for variability (slope of 1.46) with overestimation of daily variability for the stations with daily variability >5 mm day (Fig. 9b). The Ensemble mean normalized MAE decreases with increasing rainfall rate and has increasing Spearman rank correlation with increasing rainfall rate (Fig. 9c-d). There is a tendency for higher MAE and lower correlation at higher elevation stations, primarily because those stations tend to have lower rainfall rates (Fig. 9c-d).

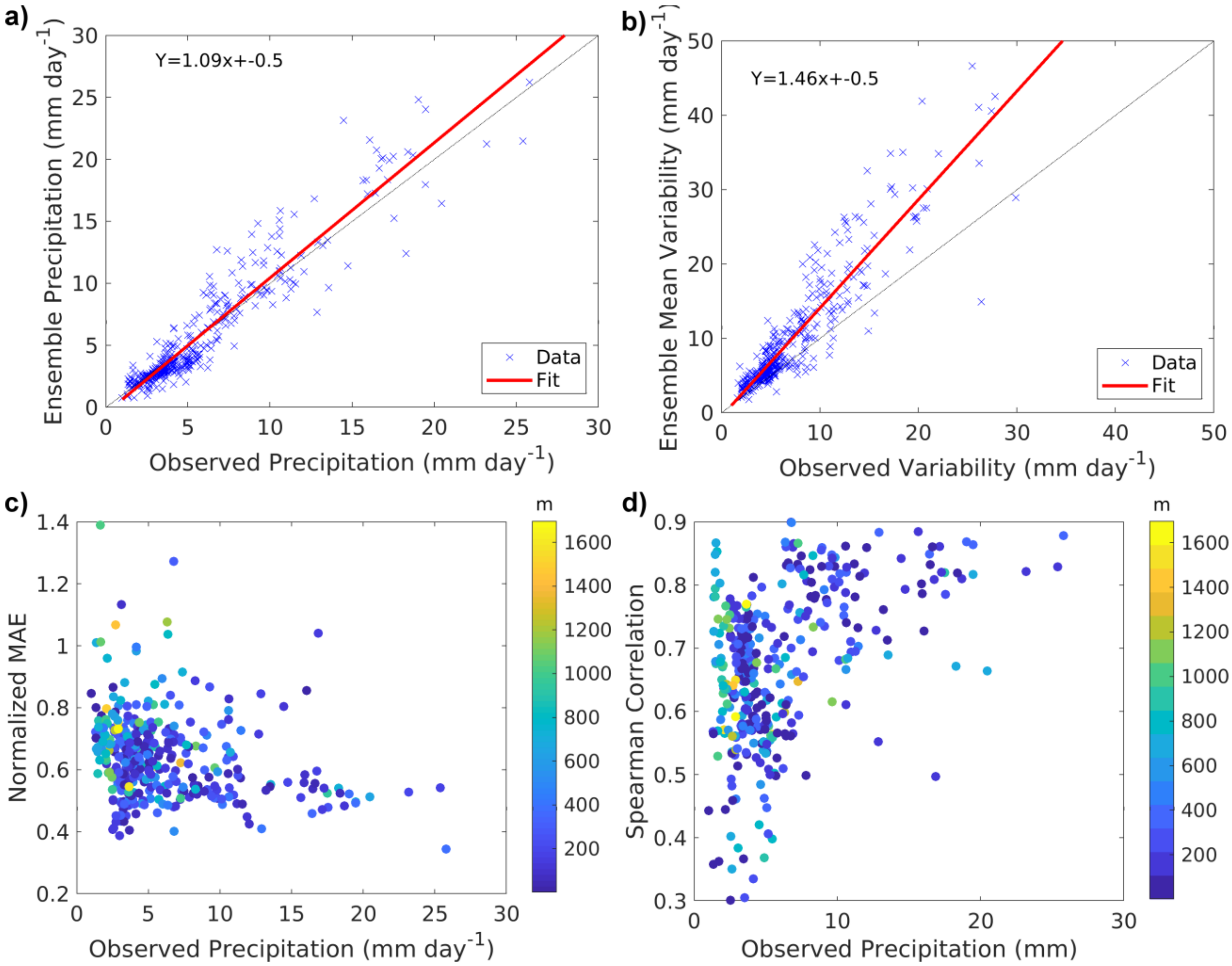

Figure 9. Ensemble deterministic leave-one-out cross-validation for a) mean daily precipitation (mm day^{-1}), b) precipitation variability (mm day^{-1}) as the standard deviation of the daily precipitation time series, c) normalized mean absolute error (MAE), and d) spearman rank correlation of Ensemble and observed precipitation.

For temperature, the Ensemble mean is essentially unbiased for mean daily temperature and DTR (Table 1). There is little spatial pattern to any non-zero bias, but more isolated stations tend to have higher MAE values because less information is available nearby for interpolation (not shown). (not shown). The Ensemble has little conditional bias (Figure 10a) for mean temperature with a fitted slope of 0.96 as compared to the observations. A larger conditional bias is present for DTR (Fig. 10b) with a slope of 0.9, signifying overestimation of DTR for small observed DTR and underestimation of DTR for large observed DTR. Mean temperature MAE values are about two-thirds of DTR (Table 1, Fig. 10a-b), which is expected considering DTR implicitly models maximum and minimum temperatures. Somewhat unexpectedly, the Pearson correlation of DTR is significantly lower than mean temperature with a drop in the median correlation of around 0.2 (Fig. 10c).

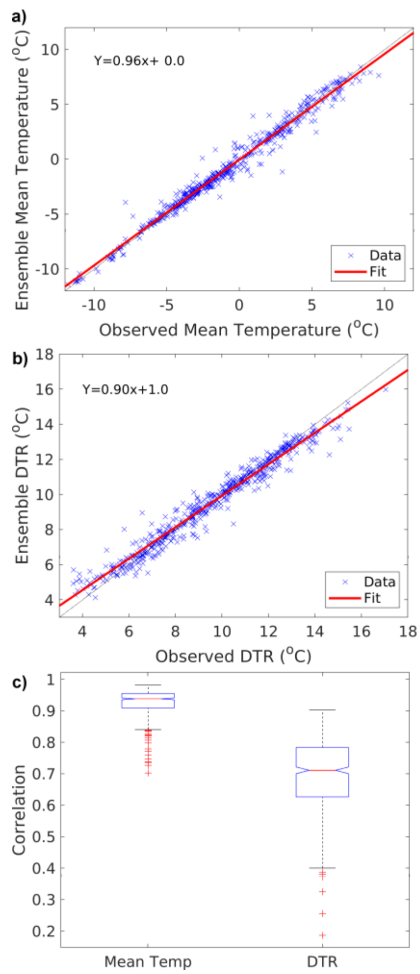

Figure 10. Ensemble deterministic leave-one-out cross-validation for a) Ensemble mean daily temperature (°C), b) Ensemble mean daily diurnal temperature range (DTR) (°C), and c) Pearson correlation of daily Ensemble mean temperature and DTR.

Table 1. Summary deterministic cross-validation statistics for Ensemble daily values of precipitation, precipitation variability (standard deviation), mean temperature, and diurnal temperature range. 90% bootstrapped (1000 iterations) confidence intervals are in parentheses.

	Precipitation	Precipitation Variability	Mean Temperature	Diurnal Range
Mean Bias	0.1 (-0.1 – 0.24) mm day^{-1}	2.9 (2.6 – 3.3) mm day^{-1}	0.0 (-0.1 – 0.1) K	0 (-0.1 – 0.1) K
Median Bias	0.5 (0.4 – 0.7) mm day^{-1}	1.6 (1.5 – 1.8) mm day^{-1}	0.0 (-0.1 – 0.1) K	-0.1 (-0.1 – 0.1) K

Mean Absolute Error	3.7 (3.4 – 3.9) mm day^{-1}	3.3 (3.0 – 3.6) mm day^{-1}	1.6 (1.5 – 1.6) K	2.3 (2.3 – 2.4) K

4.3 Probabilistic Daily Validation

The Ensemble product probabilistic daily leave-one-out cross validation (Figure 11 and Table 2) indicates a reliable product compared to climatology and the ability to discriminate between events and non-events . For all non-zero precipitation days the Ensemble has an underestimation at low estimated probabilities and an overestimation at high estimated probabilities, or overconfidence (lack of resolution) in the predicted probabilities. For other event thresholds, the Ensemble overestimates event probabilities, particularly at low observed probabilities (Fig. 11a-d). For all non-zero precipitation days, the Ensemble has significant discrimination as noted by distributions that have minimal overlap and likelihood distribution mean values for events and non-events of 0.70 and 0.22, respectively. Ensemble event discrimination decreases with increasing event threshold, but maintains significant distributional separation for all thresholds (Wilks 2006).

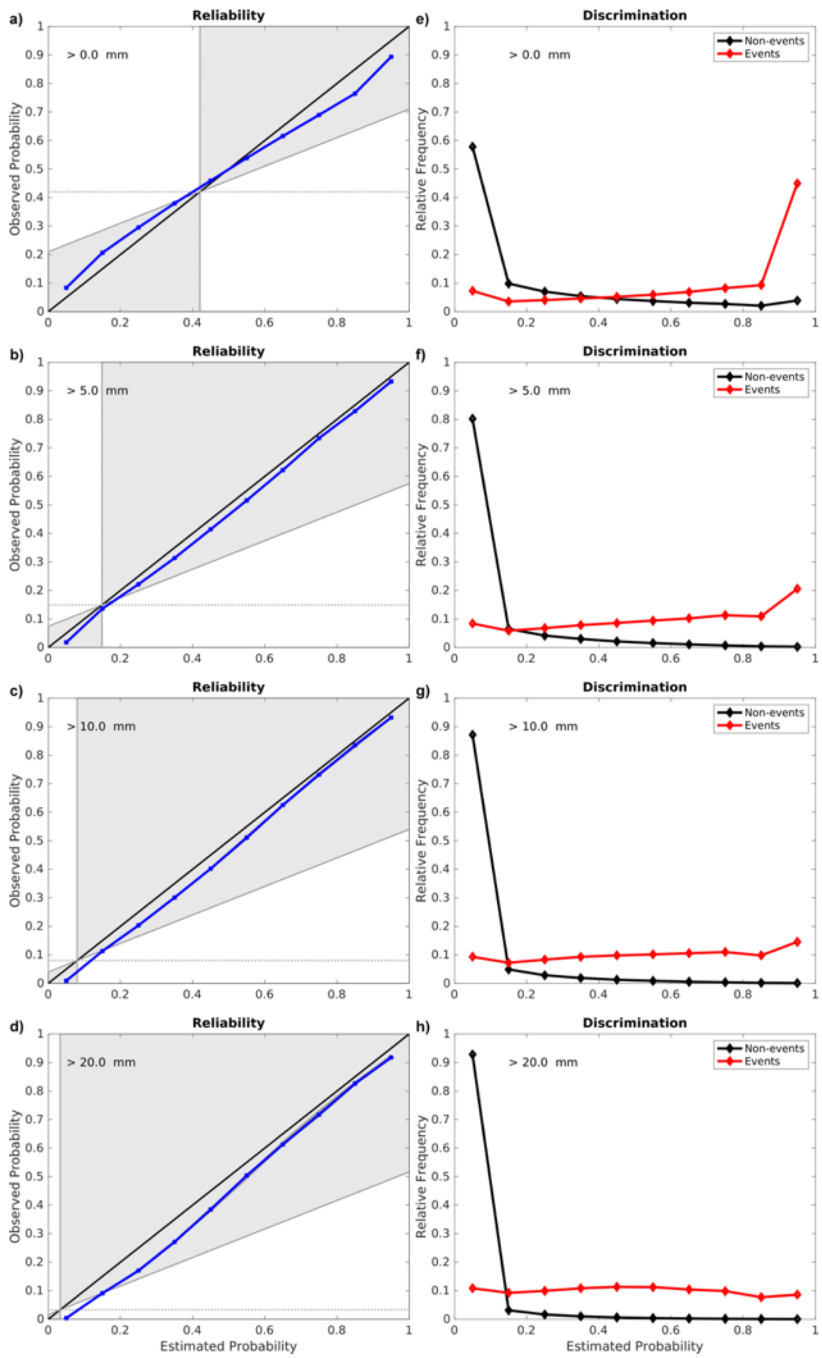

Figure 11. a-d) Reliability diagrams for 0, 5, 10, and 20 mm event thresholds, and e-h) discrimination plots for the same thresholds. The light gray shaded areas in a-d) indicate skillful reliability above climatology and the black and red lines in e-h) denote non-events and event probability distributions respectively. Areas with dark gray shading in any panel indicate uncertainty bounds using bootstrapping (1000 samples); where no dark gray shading is present, the sampling uncertainty is less than the plotted line thickness.

Table 2. Ensemble cross-validation non-event and event distribution mean probabilities across precipitation thresholds.

Event threshold	Non-event Probability	Event Probability
0 mm	0.22	0.70
5 mm	0.11	0.59
10 mm	0.08	0.54
20 mm	0.07	0.48

5 Loss Correction

Systematic loss adjustments from wind induced undercatch of frozen precipitation and gauge wetting are applied to each day and ensemble member uniquely using the specific ensemble member daily precipitation, temperature, and monthly climatological wind speed. The 1980-2013 Ensemble mean precipitation relative adjustment is shown in Figure 12. The Ensemble mean adjustment is zero for grid points with a Canadian station weight of 1 because these stations have already been deterministically adjusted. However, there is spread across the ensemble members for any given day due to the estimated uncertainty in Eq. (17) as shown in an example distribution of undercatch correction ratios for one grid point for one day (Fig. 12b). Elsewhere, relative adjustments of several percent are common in climatologically less windy locations, while areas such as the northern coastal regions of Alaska and the high, windy, coastal glacial areas of southern Alaska with high fractions of frozen precipitation and higher wind speeds can have precipitation adjustments of greater than 30%.

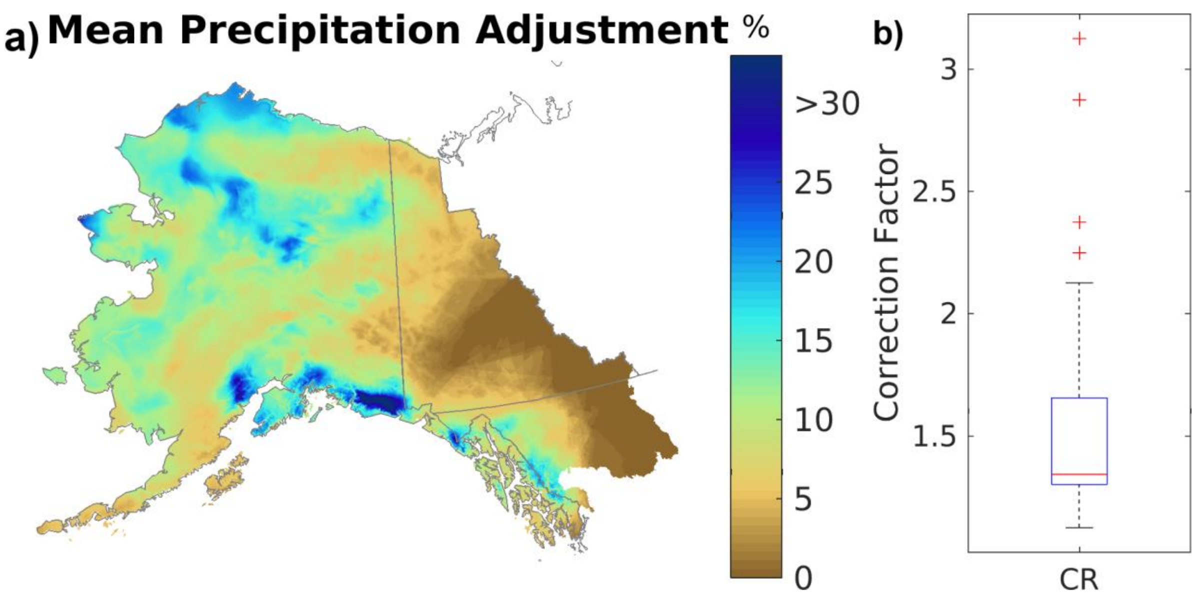

Figure 12. a) Mean relative precipitation adjustment (%), and b) example distribution undercatch correction ratios (CRs) for one grid point for one day.

6 Conclusions

Here we present probabilistic precipitation and temperature estimates for Alaska and the Canadian portion of the Yukon River watershed using an ensemble Climatologically Aided Interpolation (eCAI) system (Newman et al. 2019a, c). This region is a challenging environment to develop gridded observation estimates given the complex topography, sparse observations, and large fraction of frozen precipitation and associated measurement errors. The ensemble estimates developed here includes spatiotemporally varying uncertainty loss correction estimates.

Overall, the Ensemble represents the spatial complexity of temperature and precipitation as compared to PRISM, Daymet, and a WRF RCM simulation (Fig. 5). Qualitatively there are sometimes significant differences between the products, particularly in precipitation, with the climatological component of the eCAI system estimating relative uncertainty greater than 20% over a majority of the domain, and at high elevations in southern Alaska. The Ensemble probability of precipitation (PoP) is generally higher than Daymet and much higher across Canada. This is most likely due to the different input data sources across the two products. Daymet uses only global historical climatology network daily (GHCN-D) observations while the Ensemble uses Canadian data with more extensive quality control, additional observations across Alaska not available in GHCN-D, and data with additional quality control of Alaskan observations (section 2b). However, differences in climatological precipitation and temperature are harder to diagnose without more systematic comparisons (Newman et al. 2019b).

Leave-one-out deterministic cross-validation shows the Ensemble has little bias in daily precipitation and mean temperature but overestimates the daily standard deviation of precipitation given these specific input observations used (Fig. 9 Table 1). The MAE of diurnal temperature range (DTR) is 50% larger than that of mean temperature because DTR is more difficult to estimate as it includes both maximum and minimum temperature (Fig. 10, Table 1). The Ensemble has skillful reliability compared to climatology, and significant discrimination of events across many precipitation thresholds (Fig. 11). At higher event thresholds the Ensemble creates a wet bias, indicating an overprediction of higher threshold events and supporting the positive bias of daily precipitation standard deviation.

The major advance for this Alaska-Yukon effort is the simple ensemble loss correction methodology. Currently we consider only wetting loss and gauge wind undercatch for frozen precipitation. A distance-weighted blend of the deterministically adjusted Canadian stations with unadjusted US stations was created to account for spatially varying station contributions across the grid. Then monthly ensemble loss multipliers are developed to account for seasonality in wind speed and uncertainty in wind speed, gauge type, precipitation fraction, and undercatch estimates. The adjusted precipitation is up to 36% greater than the unadjusted estimate in windy areas also having a frozen precipitation majority, primarily along the southern and northern Alaskan coastlines (Fig. 12).

This product was not developed with trend analysis in mind, and so should not be used for trend analysis for several reasons. First, is length of the time series is limited to 34 years. Second, the input station data underwent various levels of QA/QC, but not all station data are homogenized and tested for measurement discontinuities. Finally, missing data were filled using quantile mapping, which relies on the assumption of stationary distributions in time. Nevertheless, we hope that this product will be useful for the community for many other impact and data-based studies.

7 Data Availability

This dataset is freely available at https://doi.org/10.5065/hsbv-b152 and was generated using National Center for Atmospheric Research (NCAR) high-performance computing resources (CISL, 2017). The ensemble files include daily unadjusted and adjusted precipitation (mm day^{-1}), daily mean temperature (°C), diurnal temperature range (°C), fraction of precipitation falling as snow, and the wetting loss and wind undercatch correction factors. GMET is available at https://github.com/NCAR/GMET and TIER is available at https://doi.org/10.5281/zenodo.3234938. The input station data used to generate the gridded ensemble is available Wang et al. (2017) for Canada and the Imiq database (http://imiq-map.gina.alaska.edu), and also from the corresponding author.

Acknowledgments

The US Army Corps of Engineers (USACE) Climate Preparedness and Resilience program funded this work. 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/). We gratefully acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

References

Adam, J. C., and Kettenmaier, D. P., 2003: Adjustment of global gridded precipitation for systematic bias. *J. Geophys. Res.* **108**(D9), 4257, doi:10.1029/2002JD002499.

Alter, J. C.: Normal precipitation in Utah. *Mon. Wea. Rev.*, **47**, 633-636, 1919.

Beck, H.E., E.F. Wood, M. Pan, C.K. Fisher, D.G. Miralles, A.I. van Dijk, T.R. McVicar, and R.F. Adler, 2019: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bull. Amer. Meteor. Soc.,* **100**, 473–500, https://doi.org/10.1175/BAMS-D-17-0138.1

Chua, S. C. and Bras, R. L.: Optimal estimator of mean areal precipitation in regions of orographic influence. *J. of Hydrology*, **57**, 23-48, 1982.

Cherry, J., Jacobs, A., Heinrichs, T., Fisher, W., Delamere, J., Haase, C., Martin, P., Bradley, J., Jenkins, J., Grunblatt, J., and Balogh, G.: Imiq hydroclimate database at the University of Alaska Fairbanks. http://arcticlcc.org/projects/imiq/, accessed on 7 October 2016.

Clark, M. P. and Slater, A. G.: Probabilistic quantitative precipitation estimation in complex terrain. *J. Hydrometeorl.*, **7**(1), 3-22, 2006.

Daly, C., Neilson, R. P., and Phillips, D. L.: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain, *J. Appl. Meteor.*, **33,** 140-158, 1994.

Daly, C., Taylor, G., Gibson, W., Parzybok, T., Johnson, G., & Pasteris, P: High-quality spatial climate data sets for the United States and beyond. *Transactions of the ASAE*, **43**, 1957.3, 2000.

Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P.: A knowledge-based approach to the statistical mapping of climate. *Clim. Res.* **22**: 99–113, doi: 10.3354/cr022099, 2002.

Daly, C., Smith, J. W., Smith, J. I., and McKane, R. B: High-resolution spatial modeling of daily weather elements for a catchment in the Oregon Cascade Mountains, United States. *J. Appl. Meteorol. Climatol.*, **46**, 1565-1586, 2007.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. A.: Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *Int. J. Climatol.* **28**: 2031–2064, doi: 10.1002/joc.1688, 2008.

Daly, C., Smith. J., and Halbleib, M.: 1971-2000 High-resolution temperature and precipitation maps for Alaska final report. PRISM Climate Group, Oregon State University, available upon request. 2009

Fankhauser, R., 1998: Influence of systematic errors from tipping bucket rain gauges on recorded rainfall data. Water Science and Technology, 37(11), pp.121-129.

Froidurot, S., Zin, I., Hingray, B., and Gautheron, A., 2014: Sensitivity of precipitation phase over the Swiss Alps to different meteorological variables. *J. Hydrometeorol.*, **15**, 685-696.

Golubev, V. S., P. Y. Groisman, and R. G. Quayle, An evaluation of the United States standard 8-in. nonrecording raingauge at the Valdai polygon, Russia, J. *Atmos. Oceanic Technol.*, **9**, 624–629, 1992.

Goodison, B. E., P. Y. T. Louie, and D. Yang, WMO solid precipitation intercomparison, *Final Rep., WMO/TD-872*, 212 pp., World Meteorol. Organ., Geneva, 1998.

Groisman, P. Y., and D. R. Legates, The accuracy of United States precipitation data, *Bull. Am. Meteorol. Soc.*, **75(**2), 215–227, 1994.

Henn, B., Clark, M. P., Kavetski, D., Newman, A. J., Hughes, M., McGurk, B., and Lundquist, J. D.: Spatiotemporal patterns of precipitation inferred from streamflow observations across the Sierra Nevada mountain range. *J. Hydrology*, **556**, 993-1012, doi:10.1016/j.jhydrol.2016.08.009, 2018.

Hershbach, H., and D. Dee, 2016: ERA5 reanalysis in production. ECMWF Newsletter No. 147, p. 7.

Huffman, G. J., D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E. J. Nelkin, S. Sorooshian, J. Tan, and P. Xie, 2018: Algorithm Theoretical Basis Document (ATBD) Version 5.2, NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). *NASA/GSFC Code 612*, Greenbelt MD 20771, https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2.pdf

Kochendorfer, J., Nitu, R., Wolff, M., Mekis, E., Rasmussen, R., Baker, B., Earle, M. E., Reverdin, A., Wong, K., Smith, C. D., Yang, D., Roulet, Y.-A., Meyers, T., Buisan, S., Isaksen, K., Brækkan, R., Landolt, S., and Jachcik, A., 2018: Testing and development of transfer functions for weighing precipitation gauges in WMO-SPICE. *Hydrol. Earth Syst. Sci.*, **22**, 1437-1452, doi:10.5194/hess-22-1437-2018.

Legates, D. R., and C.J. Willmott, 1990a: Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.*, **41**, 11-21.

Legates, D.R., and C.J. Willmott, 1990b: Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.*, **10**, 111-127.

Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and T.G. Houston, 2012a: 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.

Menne, M.J., I. Durre, B. Korzeniewski, S. McNeal, K. Thomas, X. Yin, S. Anthony, R. Ray, R.S. Vose, B.E.Gleason, and T.G. Houston, 2012b: Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.22 NOAA National Climatic Data Center. http://doi.org/10.7289/V5D21VHZ, accessed 10 July 2016.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E.H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343

Monaghan, A.J., M.P. Clark, M.P. Barlage, A.J. Newman, L. Xue, J.R. Arnold, and R.M. Rasmussen, 2018: High-Resolution Historical Climate Simulations over Alaska. *J. Appl. Meteor. Climatol.*, **57**, 709–731, https://doi.org/10.1175/JAMC-D-17-0161.1

Nespor, V., and B. Sevruk, Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation, *J. Atmos. Oceanic Technol.*, **16**, 450–464, 1999.

Newman, A.J., Clark, M. P., Craig, J., Nijssen, B., Wood, A., Gutmann, E., Mizukami, N., Brekke, L., and Arnold, J. R.: Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States. *J. Hydrometeor.*, **16**, 2481–2500, doi: 10.1175/JHM-D-15-0026.1, 2015.

Newman, A. J., M. P. Clark, R. J. Longman, E. Gilleland, T. W. Giambelluca, and J. R. Arnold, 2019a: Use of daily station observations to produce high-resolution gridded probabilistic precipitation and temperature time series for the Hawaiian Islands. *J. Hydrometeorology*, **20**, 509-529.

Newman, A. J., Clark, M. P., Longman, R. L., and Giambelluca, T. W.: Methodological Inter-Comparison of Gridded Precipitation and Temperature Products. J. Hydrometeorol., **20,** 531-547, 2019b.

Newman, A. J., and Clark, M. P., 2020: TIER Version 1.0: An open-source Topographically InformEd Regression (TIER) model to estimate spatial meteorological fields. *Geosci. Model Dev. Discuss*., https://doi.org/10.5194/gmd-2019-162, accepted, in copyediting.

Niemczynowicz, J., 1986: The Dynamic calibration of tipping-bucket rain gauges. Nordic Hydrology 17, 203-214.

Papalexiou, S. M., 2018: Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. *Adv. Water Resources*, **115**, 234-252, doi:10.1016/j.advwatres.2018.02.013.

Phillips, D. L., Dolph, J., and Marks, D.: A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agric. For. Meteor.*, **58**, 119-141, 1992.

Scaff, L., Yang, D., Li, Y., and Mekis, E.: Inconsistency in precipitation measurements across the Alaska-Yukon border. *The Cryosphere*, **9**, 2417-2428, doi:10.5194/tc-9-2417-2015, 2015.

Scenarios Network for Alaska + Arctic Planning (SNAP),: SNAP Data – Elevation, http://ckan.snap.uaf.edu/dataset/elevation, accessed on 4 January 2017.

Serreze, M.C., Clark, M.P. and Bromwich, D.H., 2003. Monitoring precipitation over the Arctic terrestrial drainage system: Data requirements, shortcomings, and applications of atmospheric reanalysis. *Journal of Hydrometeorology*, *4*(2), pp.387-407.

Sevruk, B., and W. R. Hamon, 1984: International comparison of national precipitation gauges with a reference pit gauge. *World Meteorological Organization*, Instrument and observing methods report No. 17. WMO/TD-No. 38.

Sevruk, B., and S. Klemm 1989: Types of standard precipitation gauges. *World Meteorological Organization*, Instrument and observing methods report No. 48., WMO/TD-No.328, edited by B. Sevruk, pp. 227–232.

Spreen, W. C: A determination on the effect of topography upon precipitation. *Trans. Am. Geophys. Union.*, **28**, 285-290, 1947.

Thornton, P.E., M.M. Thornton, B.W. Mayer, Y. Wei, R. Devarakonda, R.S. Vose, and R.B. Cook. 2018. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1328

Wang, X. L., Hong, X., Qian, B., Feng, Y., and Mekis, E.: Adjusted daily rainfall and snowfall data for Canada. *Atmosphere-Ocean*, **55**:3, 155-168, doi:10.1080/07055900.2017.1342163, 2017.

Werner, A. T. et al. A long-term, temporally consistent, gridded daily meteorological dataset for northwest North America. Sci. Data. 6:180299 doi: 10.1038/sdata.2018.299 (2019).

Willmott, C. J., and Robeson, S. M.: Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatol.*, **15**, 221–229, 1995.

Yang, D., B. E. Goodison, C. B. Benson, and S. Ishida, 1998: Adjustment of daily precipitation at 10 climate stations in Alaska: Application of WMO intercomparison results. *Water Resour. Res.*, **34**(2), 241–256.