Gridded daily weather data for North America with comprehensive uncertainty quantification – Daymet version 4

Peter E. Thornton[1] (0000-0002-4759-5158), Rupesh Shrestha[1] (0000-0002-3140-6623), Michele Thornton[1] (0000-0002-6533-6328), Shih-Chieh Kao[1] (0000-0002-3207-5328), Yaxing Wei[1] (0000-0001-6924-0078), and Bruce E. Wilson[1] (0000-0002-1421-1728)

[1] Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

corresponding authors: Peter E. Thornton (thorntonpe@ornl.gov) and Michele Thornton (thorntonmm@ornl.gov)

Abstract

Access to daily high-resolution gridded surface weather data based on direct observations and over long time periods is essential for many studies and applications including vegetation, wildlife, soil health, hydrological modelling, and as driver data in Earth system models. We present Daymet V4, a 40-year daily meteorological dataset on a 1 km grid for North America, Hawaii, and Puerto Rico, providing temperature, precipitation, shortwave radiation, vapor pressure, snow water equivalent, and day length. The dataset includes an objective quantification of uncertainty based on strict cross-validation analysis for temperature and precipitation results. The dataset represents several improvements from a previous version, and this data descriptor provides complete documentation for updated methods. Improvements include: reductions in the timing bias of input reporting weather station measurements; improvement to the three-dimensional regression model techniques in the core algorithm; and a novel approach to handling high elevation temperature measurement biases. We show cross-validation analyses with the underlying weather station data to demonstrate the technical validity of new dataset generation methods, and to quantify improved accuracy.

Notice: This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

SUBMITTED

Background & Summary

Gridded weather products are important historical references to support ecological, agricultural, water resources management, and climate change studies, particularly in regions with sparse weather stations and/or intermittent historical meteorological observations. The gridded products provide a spatially and temporally consistent approach to assimilate available weather station data, taking into account the changes in temperature, precipitation, downwelling radiation, and humidity caused by factors such as elevation, prevailing winds, storm tracks, and proximity to large water bodies. Daymet (https://daymet.ornl.gov) is one such gridded weather product, which provides daily minimum and maximum temperature (Tmin and Tmax), precipitation (Prcp), vapor pressure (VP), shortwave radiation (Srad), snow water equivalent (SWE), and day length on a 1 km x 1 km gridded surface for North America and Hawaii from 1980–2019, and for Puerto Rico for 1950-2019. Maintained by the Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC), Daymet is presently updated annually, as the previous year's weather station data becomes available and reaches a status of archive quality.

Daymet was first developed as a research project to provide daily weather driver data for terrestrial biogeochemical modelling applications. The intermountain West in the Conterminous US (CONUS) was used as the study area to develop and test the gridded data product [1]. A CONUS data product (Daymet V1) was developed from that early model and subsequent algorithm improvements [2,3]. The multi-agency North American Carbon Program (NACP) later supported an update of Daymet V2 [4] which included more years and a larger spatial extent. Due to data and algorithm limitations at the time, Daymet V2 was only available for CONUS, Hawaii, Puerto Rico, Mexico, and southern Canada up to 52 degrees North. With the inclusion of additional weather stations and further algorithm enhancement, the spatial coverage of Daymet V3 [5] was expanded to include all of North America.

The new Daymet V4 dataset [6] presented here provides effective solutions to known issues while taking advantage of the latest station observation datasets. Biases in station observations are identified and corrected, including inconsistencies among stations in time of observation for both temperature and precipitation, and errors related to temperature sensor bias. Independent validation with radar-based precipitation estimates are used to examine the timing of Daymet V4 precipitation. Cross validation analysis is used to quantify and correct biases related to temperature sensors at high elevation stations. Algorithm improvements address issues of out-of-range regression estimates in both temperature and precipitation, and provide increased accuracy and precision in the gridded data products.

Methods

I. Overview

The workflow which results in final Daymet V4 data records consists of five main steps:

1. collection and filtering of input weather station observations and gridded terrain data,
2. generation of primary output variables,
3. generation of secondary output variables,
4. generation of cross-validation statistics, and
5. data file standardization for archiving and data service.

Each of these workflow steps is described in its own section below. While many steps and sub-steps include automated or code-based workflows, there are also significant human interventions between steps, designed for quality assurance / quality control (QA/QC).

SUBMITTED

II. Daymet input data

A. Daily weather stations – source and general methods

The primary Daymet inputs are daily observations of near-surface maximum and minimum air temperature and daily total precipitation from weather stations. Before Daymet V3, it was necessary to retrieve and combine observations from multiple primary data sources. Since Daymet V4, all weather station inputs can now be acquired from the National Centers for Environment Information Global Historical Climate Network Daily database (GHCND; [7]). This simplified workflow was possible due to the large expansion of GHCND to cover multiple networks in the US, Canada, and Mexico with consistent QA/QC across all input weather stations [8]. We used GHCND V3.26, released in April 2019 in Daymet V4.

A preliminary screening of GHCND is performed to identify all available stations in the North American domain. The screening was conducted at each station for each of the primary input variables (Tmax, Tmin, and Prcp) in each year. A station-year is removed if more than 180 days of data are missing within a calendar year for a given variable. A station-year removed for one variable might still be included for other variables if the missing days threshold is not exceeded for those variables. Data with reported GHCND quality flags were also considered as missing. For a few stations with identical locations but different station identifiers, duplicates were removed. The total number of stations remaining after this screening is higher for Prcp than for Tmax and Tmin over the period 1980–2019 (Figure 1a), with increases in the Prcp station count after year 2000 due to the growth of the Community Collaborative Rain, Hail and Snow (CoCoRaHs) network, and decreasing number of temperature stations in the recent decade due to a sharp drop in stations from networks in Mexico as well as a decline in the number of stations in the US Cooperative Observer Program (COOP) network. In terms of the frequency distribution of number of missing days per station (Figure 1b), most of the station-years have low numbers of missing days, with secondary peaks at intervals associated with a pattern of whole months being marked as missing within a station-year (Figure 1b). The fraction of station-year records with no missing data is higher for Prcp (about 40%) than for Tmax or Tmin (about 30%), while 90% of station-years have fewer than 80 or few than 62 missing days for Prcp and Tmax/Tmin, respectively (Figure 1c).

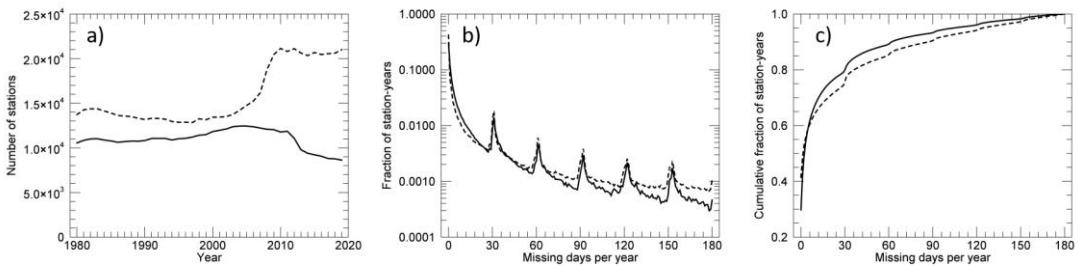

Figure 1. (a) Number of stations included in Daymet V4 by year, (b) frequency distribution of number of missing days per station-year, and (c) cumulative frequency of station-years with increasing numbers of missing days. Frequency distributions are shown as aggregated for all years. In all plots the solid line is for Tmax and the dashed line is for Prcp. The Tmax and Tmin data are nearly indistinguishable in these plots, so for clarity only Tmax is plotted.

The density of stations varies greatly over the geographic domain, and changes in station networks over time cause shifting patterns of station density (Figure 2). A primary challenge for the Daymet algorithm is to accommodate spatial and temporal shifts in station density while maintaining as much estimation accuracy as possible across the spatial and temporal domain. Given the geographic separation among North America, Hawaii and Puerto Rico, these three sub-domains were processed independently, using identical methods for each region.

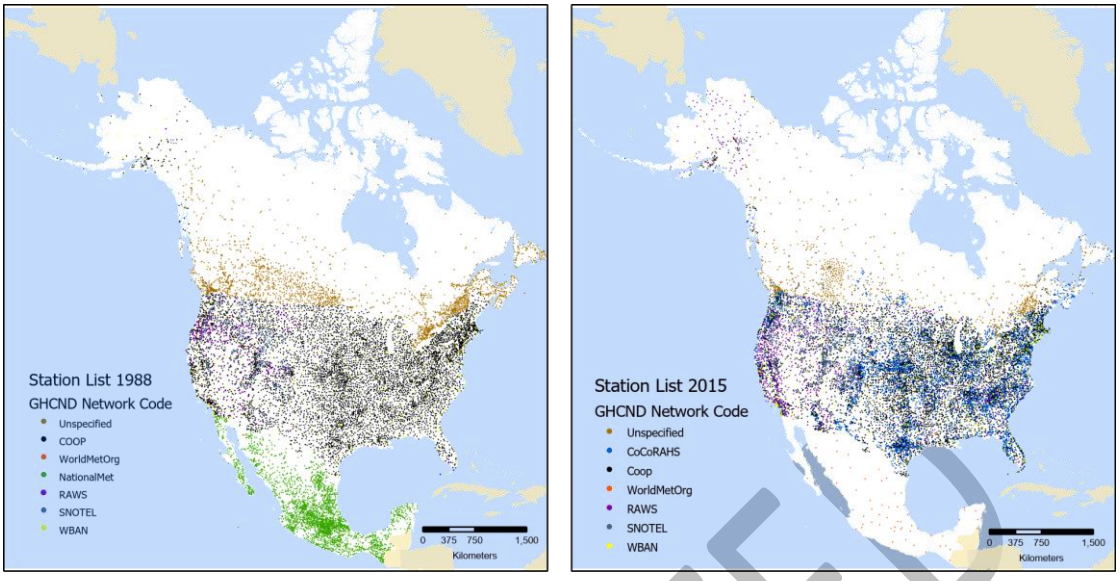

Figure 2. Spatial distribution of annual station data for years 1988 and 2015 based on GHCND defined station-level Network Code. Further source information is available within the daily data files where embedded flags provide information at the individual station level. The SFLAG1, the source flag, provides up to 30 different values. Most Canadian stations, labeled here as 'Unspecified' have an SFLAG1 identifying the source as Environment Canada.

SUBMITTED

B. Time-of-observation bias corrections

The issue of time-of-observation (TOO) arises from inconsistent report timing among stations. For some stations, automated sub-daily readings are aggregated to a midnight-to-midnight time frame. Most stations, such as those in the COOP and CoCoRaHS networks, report at some time other than midnight. Combining stations with different reporting times to make estimates at an unmeasured location can lead to estimation biases. These biases are likely to affect the estimation of Tmax since an observation time before noon will usually assign Tmax from the previous 24 hours to the current calendar day. Daily minimum temperature is less likely to be shifted by an observation time before noon since most daily minimum temperature values occur between midnight and sunrise. Daily total precipitation presents a more difficult challenge since any observation period other than midnight-to-midnight will result in misplacement of some amount of precipitation between days. While this timing bias does not affect the overall distribution and long-term climatology and was not explored in earlier versions of Daymet, it may create issues for more time-sensitive applications such as flood and heat wave analysis. Therefore, in Daymet V4 we assessed TOO for each station, as recorded in GHCND or inferred from other network metadata records. Corrections to the station-level inputs for TOO biases are performed for Tmax and Prcp

Maximum daily temperature (Tmax)

For Tmax, we posited that a bias exists in the actual day of maximum temperature for readings that have a TOO before local noon [9]. Previous studies [10,11] have demonstrated that a bias exists due to the discrepancy in the observation time and actual occurrence of the daily maximum temperature. It is conceivable that for most days, a maximum temperature recorded in the morning is likely the maximum temperature from the previous day. To correct this bias, for GHCND stations (such as most COOP) that have a TOO prior to noon local standard time (LST), we shifted the daily Tmax values to the previous day. Although TOO is not available for Canadian GHCND stations, [12] described that the majority of manual Canadian climate stations dataset (called the ordinary climate stations) typically have a morning reporting time at 0700 LST and the recorded values have been assigned to the

previous calendar day [10] before the data are incorporated in GHCND. We therefore did not apply additional TOO bias corrections to the Canadian stations. In addition, all other stations without reported TOO, including stations from Mexican networks, were not adjusted.

To determine if the time adjustment improved Tmax estimates, we compared the cross-validation estimates of the adjusted maximum temperature values to estimates generated from the same V4 algorithm, but with the input observed maximum temperature value day unaltered. We found that shifting the maximum temperature data in this way resulted in substantial reductions in cross-validation error for the daily estimates, and therefore adopted the approach for Daymet V4 processing. See the *Technical Validation* section for detailed results of this analysis.

Daily total precipitation (Prcp)

Similar to Tmax, observations of precipitation that are taken once per day but that are not reporting from midnight-to-midnight LST can result in all or part of a precipitation event being attributed to the wrong calendar day. However, unlike Tmax, there is no obvious way to correct the timing issue of Prcp. Our objective was to improve the overall timing of daily precipitation amounts without compromising the frequency of daily precipitation occurrence. We explored two alternative TOO corrections for Prcp, assessing error and bias in estimated precipitation amount and errors in the estimated frequency of wet and dry days for the two approaches. The first method was to move a fraction of the observed daily total precipitation to the previous calendar day, based on TOO at each station. The second method was to shift the entire daily precipitation amount to the previous calendar day based on TOO (same with Tmax). For the first method, the fraction shifted to the previous day was given by the fraction of the 24-hour day elapsed between TOO and the subsequent midnight. The precipitation value of the next calendar day would receive the same treatment in which any fractional amount from this calendar day would be added to the shifted value. For the second method, the total recorded precipitation amount was shifted back one day for weather stations with TOO earlier than noon LST. We obtained TOO information from the metadata of GHCND, which is available for a majority of stations in the US COOP, Weather-Bureau-Army-Navy (WBAN), Remote Automatic Weather Stations (RAWS), and Snow Telemetry (SNOTEL) networks. Although TOO is not carried into GHCND distribution for CoCoRaHS stations, the CoCoRaHS training material [13] indicates that most CoCoRaHS stations have a morning TOO, with 84% of stations recording at 0700 LST. We hence included all CoCoRaHS stations in the TOO adjustment. Given the same reason with Tmax, we did not apply correction to Canadian and Mexican stations, as well as those US stations without TOO information.

We found that the first method (splitting precipitation totals across days) gave the lowest daily errors for precipitation amount but resulted in a large increase in over-prediction of wet-day frequency. The second method (moving entire daily amounts across days) gave an improved daily error for recent years compared to the uncorrected data, and maintained the very low bias in wet-day frequency estimation characteristic of the Daymet methods applied to station data without TOO biases. We therefore selected the second method for Prcp TOO bias correction. Detailed results supporting this choice are provided in the *Technical Validation* section.

C. High-elevation temperature sensor bias correction

Since Daymet V1, gridded estimates of temperature and precipitation in mountainous regions have been aided by the inclusion of the SNOTEL network of high elevation autonomous stations [14]. A data quality issue associated with replacement of temperature sensors beginning in the mid-1990s has been reported [15-17]. The difference between old and new sensors has been shown to be a related to the observation temperature, with the new

SUBMITTED

sensors recording higher temperature than the old sensors under colder conditions, and the new sensors recording lower temperature than the old sensors under warmer conditions[16]. Oyler et al.[17] used pair-wise comparisons to perform homogenization of temperature records before and after sensor updates, based on methods by Menne and Williams[18].

We used the existing Daymet cross-validation framework (described in the *Daymet Algorithms* section, below) to assess patterns of estimation bias in the SNOTEL network before and after the installation of new sensors, using the date of sensor updates for each station from [17]. We found that the pattern of bias was in qualitative agreement with the expected bias patterns previously described. We derived a quantitative relationship between observed temperature and sensor bias using data aggregated across the entire SNOTEL network, and applied that relationship to the observations from older sensors to reduce bias and improve temporal and spatial network homogeneity. The bias correction takes the form of a piece-wise linear function of original temperature measurement ($T_{original}$), and is applied uniformly to Tmax and Tmin observations for the period prior to sensor update for each SNOTEL station, as follows:

$$T_{corrected} = \begin{cases} T_{original} + 0.4143, & T_{original} < -20°\mathrm{C} \\ 1.0245\, T_{original} + 0.9043, & -20°\mathrm{C} \leq T_{original} < -5.3°\mathrm{C} \\ 0.9523\, T_{original} + 0.5217, & -5.3°\mathrm{C} \leq T_{original} < 30°\mathrm{C} \\ T_{original} - 0.9093, & T_{original} \geq 30°\mathrm{C} \end{cases} \quad (1)$$

Detailed results leading to this correction are provided in the *Technical Validation* section.

D. Additional gridded surface data inputs and geographic projection

Additional Daymet inputs are a gridded digital elevation model (DEM) and a corresponding land/water mask defining the Daymet domain. The DEM used in Daymet is processed from the National Aeronautics and Space Administration (NASA) Shuttle Radar Topography Mission (SRTM) near-global 30 arc second DEM V2.1. The SRTM DEM was first projected from a geographic coordinate system (GCS_WGS_84) to a Lambert Conformal Conic projection [6] for Daymet projection parameters) and resampled to each 1 km x 1 km Daymet grid. The resampling method used a cubic convolution interpolation with a 1,000 m output cell size. Slope and aspect grids are derived from the DEM within the Daymet algorithm. Horizon files used for radiation estimation were generated separately with the r.horizon model using the GRASS GIS software [19].

The land/water mask was derived from the MODIS 250 m Land-Water Mask MOD44W_v2 [20]. Similar to the DEM, the 250 m MODIS land/water mask was reprojected and resampled to the 1 km x 1 km Daymet grid. Inland water bodies are considered as land in the Daymet domain, retaining only the coastline as the Daymet land/water interface.

III. Daymet algorithms

A. Spatial and temporal interpolation

The main algorithm to estimate primary Daymet variables (Tmax, Tmin, and Prcp) at each Daymet grid is based on a combination of interpolation and extrapolation, using inputs from multiple weather stations and weights that reflect the spatial and temporal relationships between a Daymet gridcell (measured from its center) and the surrounding weather stations. The approximate number of weather stations used at each Daymet gridcell is defined as a parameter for each Daymet variable. In Daymet V1 this parameter (average number of stations, or *ANS*) was used as an input to an iterative station density estimation algorithm,

SUBMITTED

which produced a search radius that was specific to each gridcell [1]. After a series of algorithm modifications intended to improve robustness in regions of very low weather station density, the Daymet V4 algorithm replaces the iterative station density calculation with a pre-calculated search radius (r_{search}) at each gridcell which is sized to capture *ANS* based on arrays of station distance. In regions with relatively high station density, such as most of the CONUS, the search radius approach is nearly identical to the iterative station density approach. In regions with very sparse observation networks, for example in the Arctic regions of Alaska and Canada, the modified approach eliminates artifacts associated with iterations in the previous density calculation, improving the stability of Daymet estimates and reducing mean cross-validation error.

At each Daymet gridcell, a truncated Gaussian convolution kernel is used to assign weights for all surrounding stations identified by r_{search}. These weights are held constant until the station list changes, which can happen at the beginning of each year due to weather station data availability. The shape of the Gaussian kernel is defined by a single parameter (the Gaussian shape parameter, or *GSP*). The normalized weight for a given station (with index *i*) in the input weather station list (w_i) is given as:

$$w_i = \frac{exp\left(-GSP\left(\frac{d_i}{r_{search}}\right)^2\right) - exp(-GSP)}{\sum_{i=1}^{ANS} w_i} \quad (2)$$

where d_i is the horizontal distance from a Daymet gridcell to the station. Normalization here ensures that the weights for all stations in the station list for a Daymet gridcell sum to 1.

Because the number and distribution of observations can differ significantly for different meteorological variables, and because the optimal interpolation parameters (ANS and GSP) can differ among variables [1], we specify a unique pair of ANS and GSP parameters for each of the primary variables (Tmax, Tmin, and Prcp). Station lists and associated weights are calculated at each Daymet gridcell in each year, and for each primary variable. These lists and weights are accessed as inputs to the subsequent workflow steps.

Because the horizontal location information recorded for some stations is not precise enough to assign unique locations within the 1 km grid, we allow for automatic adjustment of station location within +/- 1 km to minimize difference between recorded station elevation and gridded terrain data, as detailed in Thornton et al. [1].

B. Generation of primary output variables

Given the pre-processed weather station inputs and pre-calculated station lists and weights at each Daymet gridcell, two separate workflows are used to produce the primary Daymet output variables: one for Tmax and Tmin and another for Prcp.

Daily temperature estimation

Since the workflows for Tmax and Tmin are identical, here we describe them using a generic daily temperature variable T. First, the horizontal coordinates (meters from the origin) at the center of each Daymet gridcell are calculated from the Daymet projected coordinate system [6]. The vertical coordinate, or elevation, (meters above mean sea level) of a target gridcell is obtained from the pre-processed DEM. Horizontal coordinates and elevation of stations are derived from station metadata records for all selected weather stations. Based on the identified weather station list at each Daymet gridcell, the distance (km) and elevation difference (m) between Daymet gridcells and weather stations are calculated.

SUBMITTED

The estimation of temperature is based on a weighted multivariate regression model which uses the available observations to estimate spatial gradients in the observed temperature in three orthogonal spatial dimensions: two horizontal, and one vertical. This is a modification from the univariate regression adopted in Daymet V1 [1], which included estimates only for the vertical gradient in observations (in this case the temperature lapse rate). By including both horizontal and vertical gradients in a multivariate regression framework, we may obtain additional information about horizontal gradients due to short-term events such as the passage of frontal systems, and due to persistent geographic features, such as nearby water bodies, urban areas, and interaction with large-scale terrain features and prevailing wind directions. Accounting for horizontal gradients removes some aliasing of these effects onto the vertical gradient estimates and improves predictions in regions of both flat and complex terrain.

Because the realization of horizontal and vertical temperature gradients can vary over time, we estimate the gradients separately for each day. To maximize the use of information about spatial gradients provided by the station observations, we use a paired-difference approach as described by [1] to form the inputs for regressions at each Daymet gridcell. We use a weighted regression model that considers the interpolation weights for each station in the paired difference. Because the paired-difference approach is designed to have a near-zero intercept for the resulting regression equations, we do not include the intercept term in the formulation. Detailed implementation at a single Daymet gridcell is given by the following series of equations.

Paired differences in the station horizontal and vertical positions are given as:

$$\begin{cases} dx_p = s_p(x_i - x_j) \\ dy_p = s_p(y_i - y_j) \\ dz_p = s_p(z_i - z_j) \end{cases} \tag{3}$$

where *p* is an index from 1 to *np*, with *np* as the number of unique station pairings possible given the station list at each Daymet gridcell, and $np = (ANS^2 - ANS)/2$. The variables *dx, dy, and dz* are the differences in horizontal distances (*x* and *y*) and vertical distance (*z*) between the pair of stations, and the subscripts *i* and *j* denote the index values of each station in the unique pairing from the full station list at each Daymet gridcell. To avoid bias in the mean values of *dx, dy,* and *dz* which can occur if the stations in the list have non-random order with respect to their *x, y,* and *z* coordinates, the variable s_p is introduced, with a value of either 1 or -1 for a given unique pairing *p*, and that value alternating between the two values for each value of *p* from 1 to *np*. As a further guard against biased mean differences, the initial value of s_p for $n = 1$ is switched between 1 and -1. The arrays *dx, dy,* and *dz* form the matrix of independent variables in the multivariate regression model. Since the station list is fixed for all days in a given year, these independent variables are calculated once per year.

At each Daymet gridcell, the regression weight associated with each unique pairing, rw_p, is calculated as:

$$rw_p = w_i w_j \tag{4}$$

where w_i and w_j are the interpolation weights at the estimation point for the two stations making up the unique pairing, from Eq (2).

SUBMITTED

Given these regression components, which are the same for each day in a given year at a given Daymet gridcell, temperature difference is calculated for each day as:

$$dT_{p,d} = s_p\left(T_{i,d} - T_{j,d}\right) \tag{5}$$

where $T_{i,d}$ and $T_{j,d}$ are the daily temperature measurements (either Tmax or Tmin) for two stations (*i* and *j*) in a unique pairing, for a given day, *d*.

Based on the inputs defined in Eqs (2-5), the vector of least squares regression parameters is given in the standard matrix representation for multivariate regression [21]:

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{y} \tag{6}$$

where the four elements of column vector $\hat{\beta}$ represent the regression intercept (β_0) and the regression coefficients for two orthogonal horizontal temperature gradients (β_1 and β_2 for gradients in the *x* and *y* directions, respectively) and one vertical (*z* direction) temperature gradient (β_3). $\boldsymbol{X}$ in Eq (6) is the (*np* x 4) matrix of independent variables, given as:

$$\boldsymbol{X} = \begin{bmatrix} 1 & dx_1 & dy_1 & dz_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & dx_{np} & dy_{np} & dz_{np} \end{bmatrix} \tag{7}$$

$\boldsymbol{y}$ is the column vector (*np* x 1) of dependent variables for a given day *d*, given as:

$$\boldsymbol{y} = \begin{bmatrix} dT_{1,d} \\ \vdots \\ dT_{np,d} \end{bmatrix} \tag{8}$$

$\boldsymbol{W}$ is the (*np* x *np*) diagonal weighting matrix constructed from the weights in Eq (4), as:

$$\boldsymbol{W} = \begin{bmatrix} rw_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & rw_{np} \end{bmatrix} \tag{9}$$

$\boldsymbol{X}'$ is the transpose of $\boldsymbol{X}$, and $()^{-1}$ indicates the inverse matrix.

Finally, the temperature at the Daymet gridcell location and for the given day *d*, $T_{est,d}$, is given as:

$$T_{est,d} = \sum_{i=1}^{ANS} w_i \left(T_{i,d} + \beta_1(x_{est} - x_i) + \beta_2(y_{est} - y_i) + \beta_3(z_{est} - z_i)\right) \tag{10}$$

SUBMITTED

where x_{est} and y_{est} are the Daymet gridcell coordinates and z_{est} is the Daymet gridcell elevation. All geographic units are expressed in meters from map origin or meters above mean sea level, and temperatures are expressed in degrees Celsius.

Two constraints are placed on the estimations generated by Eq (10). First, β_3 (the vertical temperature gradient) is restricted to the range (-0.012 to 0.001) °C/m. This means that normal temperature lapse rates are limited to at most a 12°C decrease and 1°C increase in temperature per 1000 m elevation difference. This constraint reduces spurious estimations in regions of strong topographic relief and very sparse station networks, especially in the far northern extent of the Canadian Rocky Mountains. The second constraint caps any daily temperature estimate from Eq (10) to no more than 10°C warmer than the warmest observed temperature in the station list for the given day. This constraint prevents spurious horizontal temperature gradients from causing excessively warm temperatures in regions with very sparse and horizontally skewed station distributions, for example on the southern extremity of the Baja peninsula.

Daily precipitation estimation

For Prcp, we also started by calculating the coordinates and distance of all selected weather stations. Since the interpolation parameters ANS and GSP are different for temperature and precipitation, the list of selected stations by Eq. (2) can be different. All daily precipitation observations and gridded estimates are described here in water equivalent units of mm/day.

The precipitation estimation for a given location on a given day is performed in two steps. First, an estimation is made for daily precipitation occurrence (wet vs. dry). Next, for wet days an estimation is made for daily precipitation total. Estimated precipitation occurrence for a given day *d* ($PO_{est,d}$) is calculated as a binomial variable following Thornton et al. [1]:

$$POP_{est,d} = \sum_{i=1}^{ANS} w_i PO_{i,d} \tag{11}$$

$$PO_{i,d} = \begin{cases} 0; & P_{i,d} = 0 \\ 1; & P_{i,d} > 0 \end{cases} \tag{12}$$

$$PO_{est,d} = \begin{cases} 0; & POP_{est,d} < POP_{crit} \\ 1; & POP_{est,d} \geq POP_{crit} \end{cases} \tag{13}$$

where $POP_{est,d}$ is the estimated probability (range 0 to 1) of precipitation occurrence at a given estimation location on a given day, w_i is the station interpolation weight given by Eq (2), $PO_{i,d}$ is the observed precipitation occurrence for a given station *i* in the station list on day *d*, $P_{i,d}$ is the observed precipitation amount for that same station and day, and POP_{crit} is a threshold parameter for occurrence estimation.

Contingent on $PO_{est,d}$ = 1, the weighted multivariate regression framework of Eq (6) is used to estimate horizontal and vertical gradients in precipitation. For the purpose of estimating spatial precipitation gradients, a temporal smoothing filter is applied to the daily precipitation observations at each station. The filter has a width of 5 days and is centered on the day of estimation with relative filter weights from days *d*-2 to *d*+2 as [1, 2, 3, 2, 1]. Only wet days (and weights associated with those days) within the filter time window are included in the smoothed value at day *d*. For example, the smoothed value for a filter window with daily precipitation values given by [0, 0.1, 0.2, 2.3, 0.5] would be (2*0.1 + 3*0.2 + 2*2.3 + 1*0.5) / (2 + 3 + 2 + 1) = 0.7375. The filter window width is truncated when the day of estimation is within the first or last two days of the year. Note that the smoothed precipitation values are used to estimate gradients, but final daily precipitation predictions are based on the un-smoothed station observations.

SUBMITTED

Compared to the regression matrices for temperature gradients, there are two differences for precipitation gradients. First, the regression weight for a unique pairing of stations is set to zero if either or both of the smoothed precipitation inputs for that day are zero:

$$rw_p = \begin{cases} w_i w_j; \; PS_{i,d} \; and \; PS_{j,d} > 0 \\ 0; \; otherwise \end{cases} \tag{14}$$

where $PS_{i,d}$ and $PS_{j,d}$ are the smoothed precipitation values at stations *i* and *j* on day *d*. Second, the column vector $\boldsymbol{y}$ is made up of differences in these unique pairs of filtered precipitation observations as follows:

$$\boldsymbol{y} = \begin{bmatrix} dPS_{1,d} \\ \vdots \\ dPS_{np,d} \end{bmatrix} \tag{15}$$

$$dPS_{i,d} = \begin{cases} s_p\left(PS_{i,d} - PS_{j,d}\right); \; PS_{i,d} \; and \; PS_{j,d} > 0 \\ 0; \; otherwise \end{cases} \tag{16}$$

where s_p is the sign switching mechanism as described for Eq (3).

Still contingent on $PO_{est,d} = 1$, the daily total precipitation $P_{est,d}$ at a Daymet gridcell and day *d* is estimated using the weighted sum of wet stations as:

$$P_{est,d} = \sum_{i=1}^{ANS} \begin{cases} \frac{w_i\left(P_{i,d} + \beta_1(x_{est} - x_i) + \beta_2(y_{est} - y_i) + \beta_3(z_{est} - z_i)\right)}{w_{sum,d}}; \; P_{i,d>0} \\ 0; \; P_{i,d} = 0 \end{cases} \tag{17}$$

where $w_{sum,d}$ is the sum of all w_i for wet stations on day *d*.

The daily precipitation estimates from Eq (17) are subject to several constraints. First, the gradients estimated from weighted regression can sometimes result in $P_{est,d} < 0$, and those cases are truncated to 0. Second, if $P_{est,d}$ is more than twice the highest measured precipitation value from the station list for day *d*, those values are truncated at two times of the max measured precipitation value. Third, constraints are placed on the spatial gradients as estimated from Eq (17), with β_1 and β_2 (the horizontal components of spatial precipitation gradient) constrained to the range (-0.001, 0.001) mm/day/m and β_3 constrained to the range (0, 0.02) mm/day/m. These constraints were determined empirically by examining histograms of calculated gradients and their joint distributions with anomalous precipitation outputs from Eq (17). Fourth and finally, if precipitation occurrence (via Eq. 12) is recorded for three or fewer stations in the list for a given gridcell on a given day (suggesting a more localized storm), then all values of β_1, β_2, and β_3 are forced to zero. This final constraint prevents some spurious extrapolations near the edges of localized daily precipitation events. We also note that a constraint limiting precipitation to a maximum of 200 mm/day that was in effect for Daymet V3 has been removed for Daymet V4.

C. Generation of secondary output variables

In addition to Tmax, Tmin, and Prcp, Daymet also includes estimates of other important meteorological variables that are not routinely observed, or are only available at a small

SUBMITTED

fraction of weather stations. These secondary variables include daily total shortwave radiation (Srad), daily average water vapor pressure (VP), duration of the daylight period (daylength), and a simple estimate of accumulated snowpack, measured as snowpack water equivalent (SWE). The daylength estimate is based on geographic location and time of year. Estimates of other secondary variables (Srad, VP, and SWE) are derived from the primary variables (Tmax, Tmin, and Prcp) based on atmospheric theory and empirical relationships, as described below.

The detailed equations for joint estimation of Srad and VP based on Tmax, Tmin, and Prcp inputs are described in Thornton and Running [2,22] with modifications as given in Thornton et al. [3]. Here we provide a summary of the theory and workflow. Based on the observed positive relationship between diurnal temperature range and daily total atmospheric transmittance originally described by [23] and further developed by [24] and [22], daily Daymet Tmax and Tmin are used to estimate daily total transmittance of shortwave radiation at each Daymet gridcell and each day. The parameterization of this empirical relationship is sensitive to regional and seasonal variation in mean diurnal temperature range, to variation in atmospheric attenuation due to decreased overlying air mass with increasing terrain height, diurnal variation in optical thickness due to solar angle, reduction in transmittance with increased humidity, and observed variation in empirical relationships on days with and without precipitation.

In addition to these factors influencing atmospheric transmittance, the fraction of clear-sky transmittance realized on any day is also used to estimate the fraction of incoming radiation received as direct vs. diffuse radiation. The direct beam component is used in conjunction with terrain slope, aspect, and local horizon angles to estimate beam-slope geometry for land surface in flat or complex terrain, including the variation of this geometry over the course of each day on ten-minute time steps. Digital terrain data are also used to estimate the fraction of unobstructed sky visible at each estimation location, to estimate attenuation of incoming diffuse radiation. The influence of snow cover is also considered, as snow-covered land surface interacts with cloud cover to enhance incident radiation through multiple reflection and absorption pathways.

Estimation of water vapor pressure (VP) is based on the observed correspondence between night-time minimum temperature and dewpoint temperature for many climate regimes [24 22] with improvements to that relationship for arid and semi-arid climates as described by [25]. Because aridity correction requires an estimate of potential evapotranspiration (PET) and PET requires an estimate of incoming shortwave radiation, we use an iterative approach to jointly estimate radiation and humidity, as described in Thornton et al. [3].

The iterative approach begins by assuming the dewpoint temperature is equal to daily minimum temperature, and that the dewpoint temperature (and therefore absolute humidity) remains constant through the day. The vapor pressure implied by this assumption is used to adjust an initial estimate of daily total transmittance, providing an initial estimate of daily total shortwave radiation. PET is calculated from this initial estimate of Srad, and used to estimate aridity corrections for VP, if needed. The adjusted humidity is used to provide a final revised correction to transmittance, if needed, followed by a final revised estimate of Srad.

Estimated snow water equivalent (SWE) is based on a very simple temperature driven model of snow accumulation and snowmelt, as described in Thornton et al. [3]. The sole purpose of the SWE calculation is to provide an approximate control on Srad through the multiple reflection mechanism. We make SWE data available as part of Daymet V4 so that users can accurately diagnose the influence of this snow correction on Srad estimates. We encourage researchers who require a more accurate estimate of snowpack dynamics to use the

SUBMITTED

temperature, precipitation, and potentially radiation and humidity variables from Daymet v4 to drive a more capable and sophisticated snow process model.

D. Generation of cross-validation error estimates

Since the first public release of Daymet, results from a comprehensive cross-validation analysis have accompanied each release of the gridded daily surface weather products. The purpose of the cross-validation analysis is to provide users with details needed to evaluate the fitness of Daymet for each unique application.

The cross-validation analysis treats each variable-station-year of data from the input station lists as a unique record. This means that each primary variable (Tmax, Tmin, and Prcp) is handled separately at each station location and each year. For each such record, estimates of the primary variable are made by dropping that record from the input station list and using the exact estimation methods described above to make estimates for the primary variable on each day of the year. Since each station in the input list can have missing days where the primary variable is not recorded, only days with non-missing data for the cross-validation station record are used to calculate error statistics for that record. The number of missing days in each record is used to provide appropriate weights when reporting multi-station averages or time series summaries for the cross-validation results.

A complete record of the primary variable estimates for each day with non-missing data, for each station and each year, together with the corresponding daily observed values, are made available for user download (citation below). These cross-validation records include all of the relevant meta-data associated with each station, to allow users to assess patterns of error in relation to station location, observation type, station network, year of observation, seasonal patterns, or other analyses as deemed appropriate by the users.

Many different summary statistics can be computed from these paired daily observations and estimates, such as mean absolute error, root mean squared error, or bias. Different time periods can also be evaluated, according to the user needs. In the Technical Validation section below, we focus on mean absolute error and bias for daily and annual time periods. Further details of the Daymet V4 station-level cross-validation results is provided by Thornton, Wei, et al. [26].

Data Records

The Daymet V4 data [6] are available from the NASA-sponsored ORNL DAAC. The data are available in CF compliant netCDF file format for the time period 1980–2019 for the separate spatial extents of Continental North America and Hawaii, and from 1950–2019 for the Puerto Rico/Virgin Islands spatial extent. Data are geolocated in a projected coordinate system, the Lambert Conformal Conic projection, with a spatial resolution of 1km x 1km. Data are updated on an annual schedule.

Daymet cross-validation data are also made available through the ORNL DAAC [26]. Each data file contains the daily observations extracted from GHCND and associated Daymet model predicted primary variables (Tmax, Tmin, and Prcp), based on the cross-validation methods described above, for all input stations across the entire period. Also included are the corresponding station metadata files for each variable and year, including station name, station identifier, latitude, longitude, and elevation.

The Daymet V4 data are available by direct data download as well as specialized tools and services that focus on open, interoperability, and programable access of the Daymet data are shared from the web site https://daymet.ornl.gov/.

SUBMITTED

Technical Validation

We provide technical validation details supporting all bias corrections applied to input data from the weather station networks, followed by results from the Daymet V4 cross-validation analysis. We also provide some examples illustrating the Daymet V4 data products.

I. Time-of-observation bias corrections

A. Assessing maximum temperature time-of-observation bias correction

To evaluate the influence of a TOO bias on estimates of Tmax, we examined data records from six years: three in the earlier part of the Daymet period (1988, 1989, and 1990) and three from the later part (2015, 2016, and 2017). As described above, we focused on the US COOP network for this analysis, since we have TOO metadata records available for most stations in that network. Among all COOP stations, about 50% of Tmax records in early years and about 75% of Tmax records in later years have TOO before noon LST (Table 1). Given that a significant number of stations are exposed to this source of bias, we expected that the bias correction would have an important and positive impact on the quality of daily Tmax estimates throughout the Daymet period of record.

Table 1. Summary of TOO available from GHCND during the six years of analysis. Only US stations are shown here to illustrate percentages due to widely missing Canadian and Mexican TOO information. The percent of station-days with missing data is given for each variable and each year (% nodata).

	maximum temperature*			precipitation		
year	station days	%TOO before noon	% nodata	station days	%TOO before noon	% nodata
1988	1,729,277	49.53	2.26	2,914,822	56.70	11.90
1989	1,742,252	50.46	2.61	2,942,135	56.94	12.42
1990	1,746,314	51.24	2.91	2,937,209	57.26	12.89
2015	1,693,990	74.90	2.03	6,877,546	83.57	5.50
2016	1,676,722	74.90	2.23	6,860,087	84.68	7.87
2017	1,63,5940	74.90	2.40	6,892,927	84.38	7.40

* US COOP Network Only

We used the Daymet V4 temperature prediction and cross-validation algorithms as described above to compare cross-validation errors for the six analysis years with and without the TOO bias correction. We found that with the TOO bias correction, mean absolute error (MAE) for daily Tmax estimates dropped significantly in all six years. For the early years the average daily error was 1.81 °C without correction and 1.65 °C with correction and for the later years the same error was 1.62 °C without correction and 1.23 °C with correction, for about a 9% reduction in error for the early years and about a 24% reduction for the later years. Because this TOO bias correction moves observations in the input station dataset by at most one day, we did not expect a significant impact of this TOO bias correction on cross-validation errors based on longer averaging periods such as annual mean Tmax derived from daily estimates. This expectation was correct: MAE for annual mean Tmax was the same with and without the TOO bias correction, at 0.79 °C for the earlier years and 0.55 °C for the later years. The bias in Tmax estimates as determined by daily cross-validation analysis is very small, and is not affected by TOO correction, with Daymet underestimating Tmax on a daily basis by 0.020 °C for the earlier years, and by 0.005 °C for the later years. Cross-validation results for individual years for this evaluation are shown in Table 2.

Table 2. The percent change in weighted average daily MAE (dayMAE) and period of record MAE (porMAE) for maximum daily temperature during the selected six years evaluated with the Daymet V4 algorithm through two runs; time-adjusted vs not adjusted input values.

SUBMITTED

	Maximum daily temperature (Tmax)			
year	%change mean dayMAE	%change mean porMAE	Mean Total nstns	Avg good days/stn
1988	-8.9482	-0.5013	10,666	344.13
1989	-8.8567	-0.3006	10,644	345.03
1990	-9.4025	-0.1938	10,721	345.61
2015	-24.0978	-0.2892	9,124	350.70
2016	-23.6063	-0.2885	8,968	349.75
2017	-24.6029	0.5235	8,712	350.08

B. Assessing precipitation time-of-observation bias correction

We performed a similar cross-validation analysis to evaluate the influence of TOO bias on estimates of Prcp, examining the same six years and comparing error statistics with and without correction. Among all US GHCND stations, about 56% of Prcp records in early years and about 84% of Prcp records in later years have TOO before noon LST (Table 1). Given that a significant number of stations are exposed to this source of bias, we expected that the bias correction would have an important and positive impact on the quality of daily Prcp estimates throughout the Daymet period of record.

As described in the Methods section, we tested two correction approaches; one shifted fractions of daily precipitation totals and another moved the entire daily total to the previous day. While the fractional shift method produced the lowest daily MAE in cross-validation analysis, it also caused a significant increase in precipitation bias and a small increase in annual total precipitation MAE (Table 3). Daily MAE for the fractional shift approach averaged 1.6 mm/day for earlier years, and 1.2 mm/day for later years, compared to 1.7 mm/day and 1.5 mm/day for the uncorrected data in those two periods, respectively. This represents about an 8% and 18% reduction in daily MAE for the earlier and later years, respectively, using the fractional shift approach. Bias for the fractional shift approach averaged +0.08 mm/day and +0.10 mm/day for the earlier and later years, respectively, compared to +0.04 mm/day and +0.08 mm/day for those same periods using the uncorrected data. This represents an increase in bias of about 76% and 24% for the earlier and later periods, respectively. The whole-day shift correction method daily MAE from cross-validation was 0.18 mm/day and 0.13 mm/day for earlier years and later years, respectively. This represents an *increase* in daily MAE of about 3% for the earlier years, and a decrease of about 9% for later years, compared to the use of uncorrected data. The bias statistics for the whole-day shift approach were +0.05 mm/day and +0.09 mm/day for the earlier and later years, respectively. This represents an increase in bias of about 5% and 6% for the two periods, respectively. As expected, and similar to the results for maximum temperature, cross-validation MAE for estimates of annual total precipitation derived from daily estimates was not significantly influenced by TOO bias correction, assessed at about 0.3 mm/day for all cases (uncorrected and both bias correction methods).

Table 3. The percent change in weighted average daily MAE (dayMAE), period of record MAE (porMAE), and bias for precipitation during the selected six years evaluated with the Daymet V4 algorithm through two runs; time-adjusted vs not adjusted input values. The upper section of the table shows results for shifting entire daily precipitation totals, while the lower section shows results for shifting fractions of daily precipitation.

	Precipitation (Prcp): Total daily value shift				
year	%change mean dayMAE	%change mean porMAE	%change bias	Mean Total nstns	Avg good days/stn
1988	2.8483	0.6154	0.0000	13,235	349.22
1989	3.2680	0.9585	4.1667	13,048	351.25
1990	3.9652	0.8721	4.6512	12,918	351.77
2015	-8.3990	0.0000	4.5455	20,476	329.20
2016	-7.9365	0.3861	8.0000	20,263	330.39

SUBMITTED

2017	-9.2350	-0.3774	5.1282	20,340	329.83

	Precipitation (Prcp): Fractional daily value shift				
year	%change mean dayMAE	%change mean porMAE	%change bias	Mean Total nstns	Avg good days/stn
1988	2.8483	0.6154	0.0000	13,235	349.22
1989	3.2680	0.9585	4.1667	13,048	351.25
1990	3.9652	0.8721	4.6512	12,918	351.77
2015	-8.3990	0.0000	4.5455	20,476	329.20
2016	-7.9365	0.3861	8.0000	20,263	330.39
2017	-9.2350	-0.3774	5.1282	20,340	329.83

This analysis suggests that the two bias correction methods represent a compromise, with the fractional shift approach representing a more realistic transfer of precipitation between days, but requiring some additional information to prevent an increase in number of wet days and precipitation bias when a single wet day is split across the daily measurement boundary. By not splitting daily events, the frequency of wet and dry days is retained from the original data (results not shown) and the growth in bias is mitigated, but the improvements in daily estimates are smaller and inconsistent in time. We decided to use the whole day shifting approach to avoid shifting the precipitation frequency distributions and increasing bias.

To provide further confidence that the time shifting approach for precipitation was warranted, we performed a correlation analysis of the timing of Daymet results with and without time-shifting correction using the radar-based Stage IV Quantitative Precipitation Estimate (ST4; [27]). The ST4 data merges raw radar-based precipitation estimates with automatic hourly rainfall gauge observations and is further quality controlled by several National Oceanic and Atmospheric Administration (NOAA) River Forecasting Centers (RFCs). ST4 is available at hourly time step in 4-km horizontal resolution since 2002. An evaluation performed by [28] suggests that ST4 has the highest correlation coefficient with gauge observations among various gridded precipitation products. By aggregating 24-hour ST4 at different starting time, we try to identify when a highest correlation coefficient between Daymet and ST4 can be reached and use it to indirectly assess the actual timing of daily precipitation estimates from the Daymet algorithm. This analysis goes beyond the cross-validation approach, by using the full gridded Daymet outputs as opposed to only the Daymet predictions at individual surface observing stations. We used the Daymet V3 outputs [5] to assess timing of precipitation without TOO bias correction, and the Daymet V4 outputs (using whole-day shifting approach) to assess timing of precipitation with TOO bias correction.

At each Daymet grid point, the nearest ST4 grid point was first identified. The 24-hour ST4 was then aggregated from t = -23, -22, …, 23, in which t represents the starting hour before (positive) or after (negative) midnight LST. For instance, t = 17 indicates that the 24-hour ST4 is calculated from 7AM previous day to 7AM current day, which is the reporting time for most GHCND stations. To avoid further complicating the problem, daylight saving adjustment is not considered.

Fig. 3(a) shows the correlation coefficient between Daymet V3 and 24-hour ST4 aggregated from midnight to midnight LST. Clearly, since most of the daily stations do not report from midnight to midnight, the correlation coefficient is weak and less than 0.5 for most grid points. By testing different ST4 aggregation timing, Fig. 3(c) shows the maximum correlation coefficient between Daymet V3 and 24-hour ST4, and Fig. 3(e) shows the corresponding timing. In terms of the maximum correlation, it is greater than 0.9 in the majority of eastern US. The correlation declines in the western US where we have serious radar blockage issues. It also shows low correlation in Mexico and above the Great Lakes where we may not have

SUBMITTED

high quality gauges records for both Daymet and ST4. Fig. 3(e) suggests that the timing of Daymet V3 daily precipitation is mostly earlier than midnight and peak at t = 17, which is consistent with the timing of most GHCND stations.

By repeating the analysis using Daymet V4, the corresponding results are showed in Fig. 3(b), 3(d), and 3(f). In terms of maximum correlation, the results of Fig. 3(d) is similar to Fig. 3(c), with some slight improvement. Except for some parts of western US and Mexico, the timing of Daymet V4 daily precipitation is later than midnight and peak at t = 7. Since the daily timing is now pushed closer to midnight, stronger correlation is showed in Fig. 3(b), as expected. This analysis corroborates that the whole-day shifting approach has improved the timing bias, moving the Daymet results closer to a local midnight-to-midnight measurement, albeit shifting from a large early bias to a smaller late bias. The fact that this compromise is an overall improvement in timing is supported by the generally higher correlations shown for V4 than for V3, comparing Fig. 3(b) with Fig. 3(a).

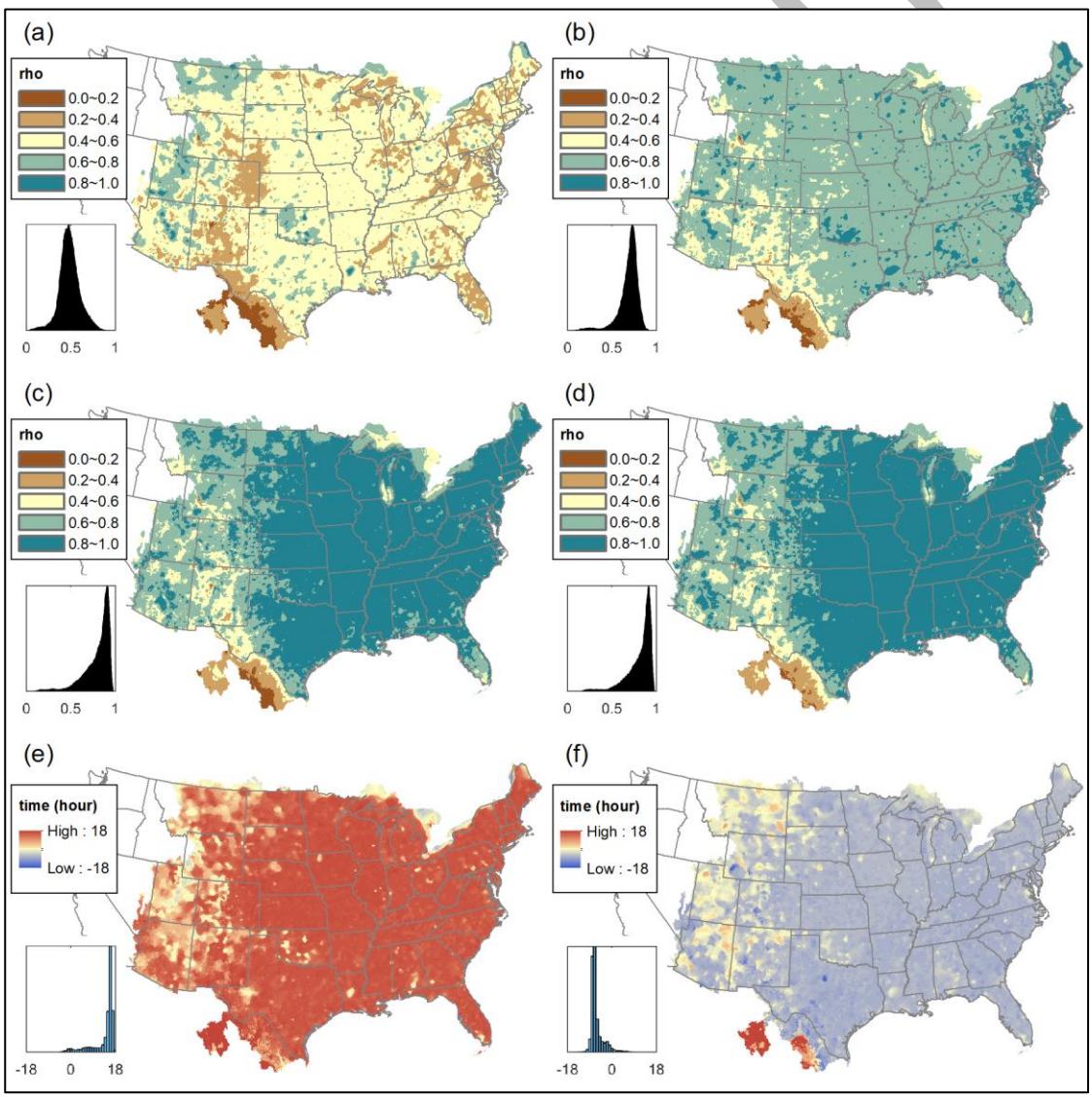

Figure 3. Analysis of Daymet precipitation timing by 2010–2019 hourly ST4 data. Panel (a) shows the correlation coefficient between Daymet V3 and 24-hour ST4 aggregated from local midnight to midnight. By testing different ST4 aggregation timing, panel (c) shows the maximum correlation coefficient between Daymet V3 and 24-hour ST4, and panel (e) shows the corresponding timing. Panels (b), (d), and (f) are similar to (a), (c), and (e) but calculated by Daymet V4.

II. Assessment of SNOTEL temperature sensor correction

We used the Daymet cross-validation framework to quantify the relationship between observed temperatures before and after sensor replacement in the SNOTEL network of high-elevation stations, and to estimate any empirical relationship between those temperature differences and observed temperature. We examined the difference in cross-validation prediction bias at each SNOTEL station before and after sensor replacement, as a function of observed temperature. The theory of this analysis depends on the existence of a large number of stations outside the SNOTEL network (mainly the US COOP stations) but within the same geographic region. We assume that the instrumentation within the non-SNOTEL networks is consistent over time, and that the cross-validation estimates at each SNOTEL station location will be significantly influenced by the presence of the non-SNOTEL network stations. Comparing estimates so derived with the SNOTEL station observations (as per the cross-validation protocol) will result in a characteristic pattern of bias at the SNOTEL measurement site. If the change in SNOTEL instrumentation causes a systematic shift in the observed temperature at the station, then this pattern of cross-validation bias should shift accordingly when examined before and after the sensor change. Furthermore, since one sensor is used for measurement of minimum and maximum temperature at the SNOTEL station, we expect that the differences in cross-validation bias patterns, such as exist, should overlap for the part of the temperature range with significant overlap in the measurement temperatures.

We aggregated cross-validation results from all SNOTEL stations and binned them by measurement temperature and found that there are distinguishable patterns of cross-validation bias for both Tmax and Tmin observations. Examining the differences in bias before and after sensor replacement at each station, we found a coherent and nearly linear pattern of bias differences in both Tmax and Tmin cross-validation results over the range of observed temperature from about -5 °C to 30 °C (Figure 4). The same linear pattern extended down to observed temperatures of -15 °C for Tmax, but not for Tmin (Figure 4). Based on these results we estimated bias corrections using one piece-wise regression fit for observed values from 30 °C down to a break-point at -5.3 °C, and a second fit for values from -5.3 °C to -20 °C (Figure 4). For values above 30 °C and below -20 °C, we held the bias correction constant as given by the relevant piece-wise fit at those values.

We used a synthetic analysis of station data from a cold region outside the SNOTEL domain to assess whether our approach can provide a bias correction for a known (synthetic) imposed bias in a subset of the station data. We found that the cross-validation bias method gave a reliable estimate of the known synthetic bias pattern (results not shown).

SUBMITTED

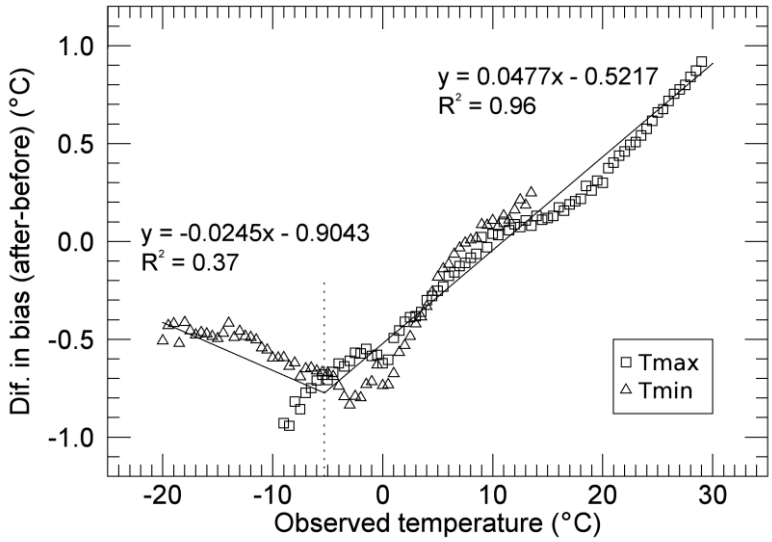

Figure 4. Piece-wise linear regressions based on combined Tmax and Tmin bias differences. Each symbol represents the difference between daily cross-validation bias values after vs. before sensor replacement, for a single variable (Tmax or Tmin) and for a binned sample of observed temperatures taken across all SNOTEL stations. Observed temperature bins are 0.5 °C wide. Solid lines show the piece-wise regression fits, and dashed line shows the piece-wise break-point at -5.3 °C.

III. Summary of Daymet V4 cross-validation results

Complete cross-validation results for Tmax, Tmin, and Prcp are provided with the Daymet V4 data release, for every input weather station and every estimation day over the 40-year period. Users are encouraged to consult the cross-validation data to assess suitability of the results for the region, period, and variables of interest for different data applications. Here we provide high-level summaries of some of the key cross-validation metrics, focusing on daily MAE for Temperature and precipitation. Many other metrics can be calculated from the raw cross-validation results provided in the dataset, according to user needs.

The mean daily MAE of Tmin averaged over the 40-year period is about 1.78 °C, and does not show strong trends over time. This is very similar to the same metric for the Daymet V3 database, which is expected since Tmin processing is unchanged from V3 except for differences in the input weather stations. The same daily MAE metric for Tmax is significantly improved compared to V3 (1.52 °C for V4 compared to 1.75 °C for V3), and shows greater improvement in recent years (Figure 5a). These results are consistent with improvements related to TOO bias correction for Tmax.

SUBMITTED

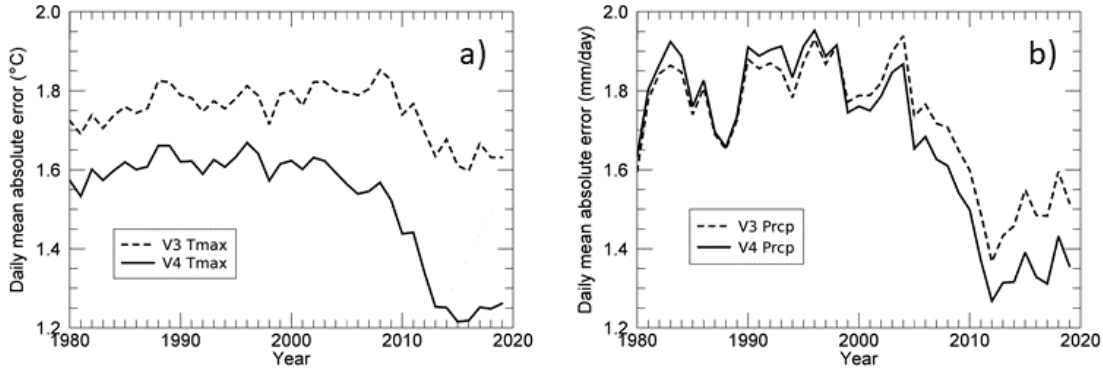

Figure 5. Annual timeseries of daily cross-validation mean absolute error (MAE), averaged over all station days for each year, comparing results for Daymet V4 and Daymet V3. a) MAE for estimation of daily maximum temperature. b) MAE for estimation of daily total precipitation.

Mean daily MAE of Prcp is significantly lower in recent years compared to the early part of the record, consistent with the expansion of the CoCoRaHS network. V4 daily MAE of Prcp is similar to or slightly higher than V3 through about year 2000, after which V4 shows some improvement (lower MAE) compared to V3 (Figure 5b). This pattern is consistent with the TOO bias correction analysis presented above.

Variation in station density, terrain, and large-scale atmospheric patterns all contribute to spatial heterogeneity in cross-validation statistics. Introduction of TOO bias correction reduces cross-validation error in regions where different TOO protocols occur in close proximity, such as across the US-Canada border (Figure 6a and b). Improvements in Prcp cross-validation error are more localized, and are most noticeable in the Great Lakes region and across parts of the eastern seaboard and southeastern US (Figure 6c and d).

SUBMITTED

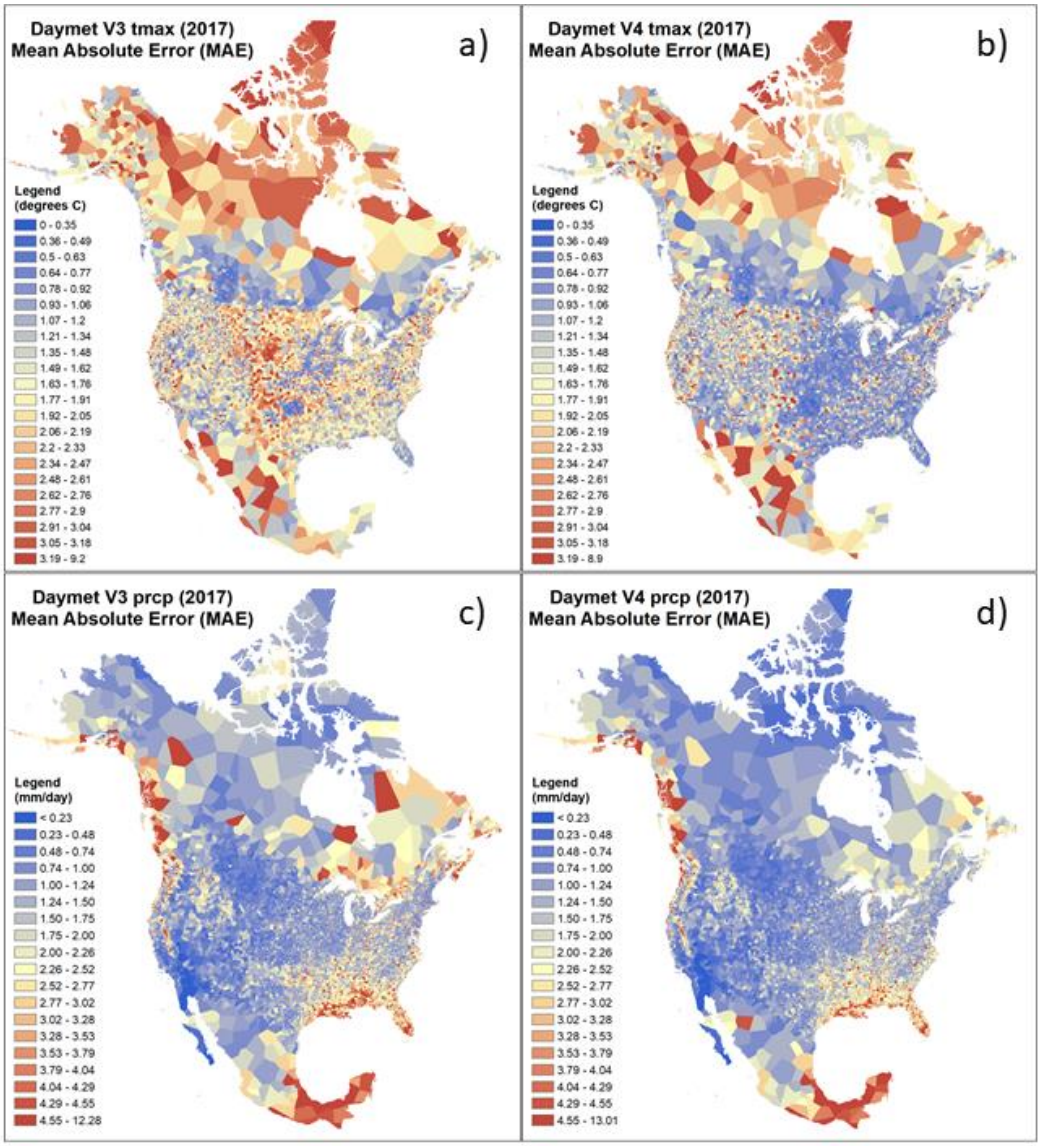

Figure 6. Cross-validation results for all stations, showing daily mean absolute error (MAE) averaged over all days in a single year (2017) and mapped using a nearest-neighbor interpolation to objectively display the influence of varying station density and other factors on data accuracy. Results from the previous version of the Daymet dataset (V3) are shown alongside the latest (V4) results to highlight spatial patterns of improvement based on the new methods described here. a) Annual mean of daily MAE for Tmax, V3. b) Annual mean of daily MAE for Tmax, V4. c) Annual mean of daily MAE for Prcp, V3. d) Annual mean of daily MAE for Prcp, V4.

To help evaluate the dataset for the presence of obvious spatial discontinuities or other similar anomalies we provide a sample of Daymet V4 daily data aggregated to annual climatologies and mapped for 2019, for Tmax, Tmin, Prcp, and the secondary output variable vapor pressure (Figure 7). These climatology include the separately generated subregions for Hawaii and Puerto Rico. The climatologies show both large and small-scale features, and visual inspection indicates that there are not obvious spurious patterns or discontinuities.

One consequence of changes in the Daymet algorithm for V4 is that hard upper limits on daily precipitation amounts have been removed. This is important in resolving extreme events where more than 200 mm of precipitation might be observed in a day (200 mm/day being the V3 hard limit).

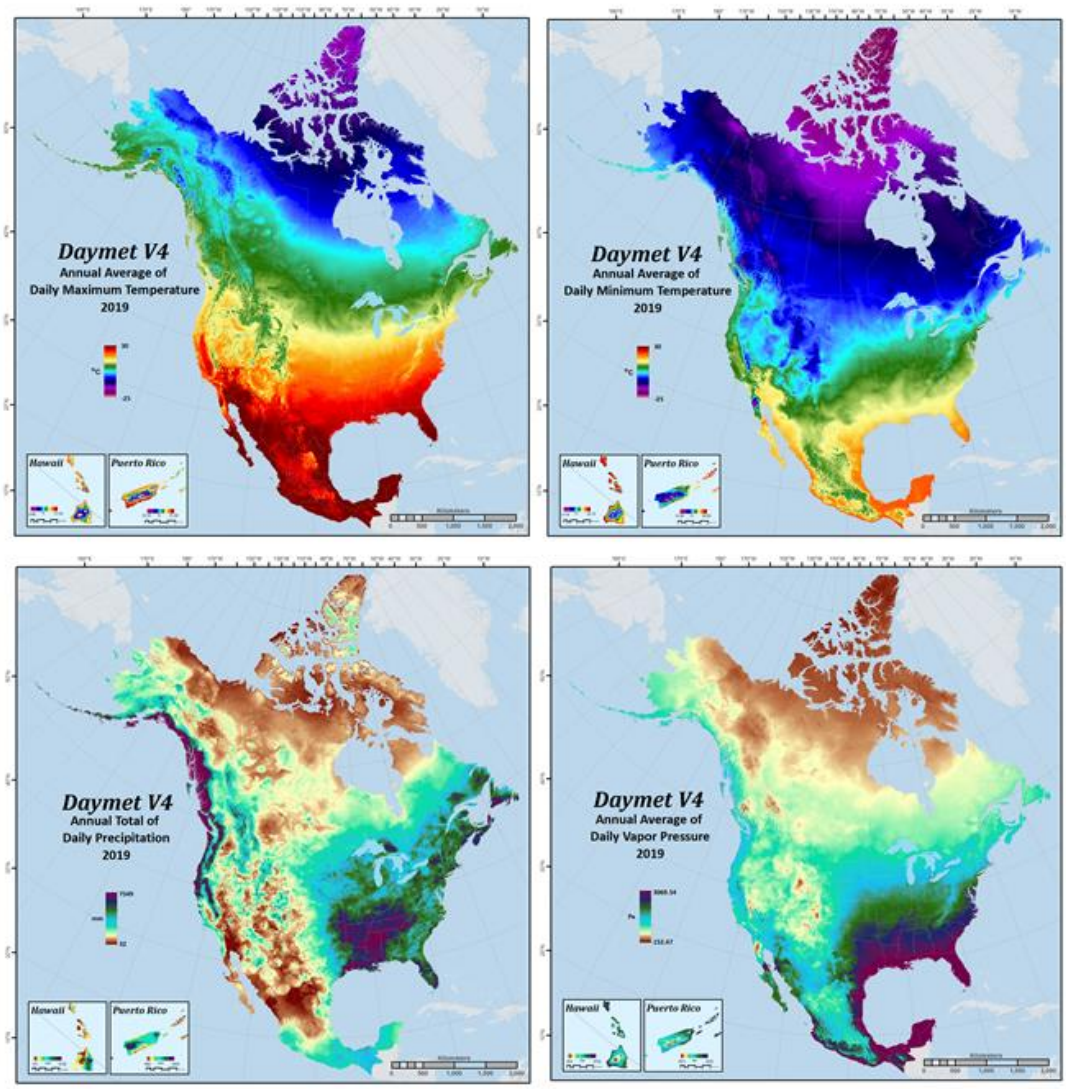

Figure 7. Annual Daymet V4 climatologies for 2019, for Tmax (upper left), Tmin (upper right), Prcp (lower left), and vapor pressure (lower right).

Usage Notes

The Daymet project website (https://daymet.ornl.gov) provides links to all of the methods for obtaining Daymet data through the ORNL DAAC. The data are available for direct download through the dataset landing pages pointed to by the dataset DOIs https://daac.ornl.gov/cgi-bin/dataset_lister.pl?p=32. The data are also available through different services, including:

- Thematic Real-time Environmental Distributed Data Services (THREDDS) Data Server (TDS) at https://thredds.daac.ornl.gov/thredds/catalogs/ornldaac/Regional_and_Global_Data/DAYMET_COLLECTIONS/DAYMET_COLLECTIONS.html provides the ability to obtain subsets of Daymet daily surface weather data, monthly climatology, and annual climatology data sets. Instructions for using the THREDDS netCDF subsetting web service can be found at https://daymet.ornl.gov/web_services.html.
- The ORNL DAAC's Spatial Data Access Tool (SDAT) at https://webmap.ornl.gov/ogc provides visualization and Open Geospatial

Consortium (OGC)-compliant web services for the monthly and annual climatologies.

- The Daymet Single Pixel Extraction tool at https://daymet.ornl.gov/single-pixel/ provides a graphical user interface and web services for extracting a time series of Daymet daily surface weather data at a specific location.
- The Daymet Tile Selection Tool at https://daymet.ornl.gov/gridded/ provides a method for getting the daily surface weather data in gridded 2 degree by 2 degree tiles.
- The ORNL DAAC's Fixed Sites Subsetting Tool at https://modis.ornl.gov/sites/ provides pre-computed time-series visualizations and data downloads (csv and JSON) for approximately 1400 sites in North America which are part of one or more ecological research networks.

All Daymet data are publicly available, without restriction, according to the NASA Earth Observing System Data and Information System (EOSDIS) Data Use Policy at https://earthdata.nasa.gov/earth-observation-data/data-use-policy. A NASA Earthdata Login account is generally needed to download data, available through https://urs.earthdata.nasa.gov. Any person can get an Earthdata Login account through an automated process.

Code Availability

The source code implementing the core Daymet algorithms (Equations 2 through 17 and associated text) is available for examination by editors and reviewers of this manuscript upon request. Upon acceptance and publication, the core algorithm code will be made publicly available through a suitable archive.

Acknowledgements

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the US Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research (BER). The extended temporal record of Puerto Rico and surrounding islands was supported by the NGEE Tropics project, funded by the US DOE, BER. Development of previous versions of Daymet as well as its curation and distribution have been supported by funding from NASA through the Earth Science Data and Information System (ESDIS) and the Terrestrial Ecology Program.

Author Contributions

Peter Thornton: Responsible for overall scientific direction, contributed to all aspects of algorithm and processing improvements, analysis of results, and writing of this manuscript.

Rupesh Shrestha: Contributed to all aspects of algorithm and processing improvements, particularly the corrections at high latitudes, authored multiple aspects of the data processing codes, and contributed to writing of this manuscript.

Michele Thornton: Contributed to all aspects of algorithm and processing improvements, led the data generation and writing of this manuscript.

Shih-Chieh Kao: Contributed to all aspects of algorithm and processing improvements, particularly the precipitation corrections, and writing of this manuscript.

Yaxing Wei: Contributed to all aspects of algorithm and processing improvements, and writing of this manuscript.

Bruce Wilson: Contributed to the temperature and precipitation correction improvements and writing of this manuscript.

SUBMITTED

Competing Interests

The authors declare that they have no conflicts of interest.

References

1 Thornton, P. E., Running, S. W. & White, M. A. Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology* **190**, 214-251, doi:https://doi.org/10.1016/S0022-1694(96)03128-9 (1997).

2 Thornton, P. E. & Running, S. W. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agricultural and Forest Meteorology* **93**, 211-228, doi:https://doi.org/10.1016/S0168-1923(98)00126-9 (1999).

3 Thornton, P. E., Hasenauer, H. & White, M. A. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. *Agricultural and Forest Meteorology* **104**, 255-271, doi:https://doi.org/10.1016/S0168-1923(00)00170-2 (2000).

4 Thornton, P. E. *et al.* Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2. *ORNL Distributed Active Archive Center* https://doi.org/10.3334/ORNLDAAC/1219 (2014).

5 Thornton, P. E. *et al.* Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. *ORNL Distributed Active Archive Center*, https://doi.org/10.3334/ORNLDAAC/1328 (2016).

6 Thornton, M. M., Shrestha, R., Wei, Y., Thornton, P. E. & Wilson, B. E. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4. *ORNL Distributed Active Archive Center*, https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1840 (2020).

7 Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. & Houston, T. G. 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 (2012).

8 Durre, I., Menne, M. J., Gleason, B. E., Houston, T. G. & Vose, R. S. Comprehensive Automated Quality Assurance of Daily Surface Observations. *Journal of Applied Meteorology and Climatology* **49**, 1615-1633, doi:10.1175/2010JAMC2375.1 (2010).

9 Vose, R. S., Williams Jr, C. N., Peterson, T. C., Karl, T. R. & Easterling, D. R. An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network. *Geophysical Research Letters* **30**, doi:10.1029/2003GL018111 (2003).

10 Hopkinson, R. F. *et al.* Impact of Aligning Climatological Day on Gridding Daily Maximum–Minimum Temperature and Precipitation over Canada. *Journal of Applied Meteorology and Climatology* **50**, 1654-1665, doi:10.1175/2011JAMC2684.1 (2011).

11 DeGaetano, A. T. A Serially Complete Simulated Observation Time Metadata File for U.S. Daily Historical Climatology Network Stations. *Bulletin of the American Meteorological Society* **81**, 49-68, doi:10.1175/1520-0477(2000)081<0049:ASCSOT>2.3.CO;2 (2000).

12 Mekis, E. *et al.* An Overview of Surface-Based Precipitation Observations at Environment and Climate Change Canada. *Atmosphere-Ocean* **56**, 71-95, doi:10.1080/07055900.2018.1433627 (2018).

13 CoCoRaHS. *CoCoRaHS: List Daily Precipitation Reports*, <(https://www.cocorahs.org/ViewData/ListDailyPrecipReports.aspx)> (2020).

14 USDA/NRCS. Chapter 8 Standards and Specifications. *Part 622 Snow Survey and Water Supply Forecasting National Engineering Handbook* (2015).

15 Harms, D. S., Seybold, C. A., Weeks, J., & Lea, J. Comparing a SNOTEL extended-range air temperature sensor and equations to a NIST-certified sensor in an environmental chamber. . *84th Annual Western Snow Conference (pp. 123–128). Seattle,*

SUBMITTED

Washington., doi:Harms, D. S., Seybold, C. A., Weeks, J., & Lea, J. (2016). Comparing a SNOTEL extended-range air temperature sensor and equations to a NIST-certified sensor in an environmental chamber. In 84th Annual Western Snow Conference (pp. 123–128). Seattle, Washington. (2016).

16 Ma, C., Fassnacht, S. R. & Kampf, S. K. How Temperature Sensor Change Affects Warming Trends and Modeling: An Evaluation Across the State of Colorado. *Water Resources Research* **55**, 9748-9764, doi:10.1029/2019WR025921 (2019).

17 Oyler, J. W., Dobrowski, S. Z., Ballantyne, A. P., Klene, A. E. & Running, S. W. Artificial amplification of warming trends across the mountains of the western United States. *Geophysical Research Letters* **42**, 153-161, doi:10.1002/2014GL062803 (2015).

18 Menne, M. J. & Williams, C. N., Jr. Homogenization of Temperature Series via Pairwise Comparisons. *Journal of Climate* **22**, 1700-1717, doi:10.1175/2008JCLI2263.1 (2009).

19 GRASS Developement Team. *Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.6.1. Open Source Geospatial Foundation. (http://grass.osgeo.org).* (2017).

20 NASA LP DAAC. (2016).

21 Eng, K., Chen, Yin-Yu, and Kiang, J.E. *User's guide to the weighted-multiple-linear-regression program (WREG version 1.0): U.S. Geological Survey Techniques and Methods, book 4, chap. A8, 21p (Also available at https://pubs.usgs.gov/tm/tm4a8)*. (Also available at https://pubs.usgs.gov/tm/tm4a8) edn, Vol. book 4, chap. A8, 21p 21 p (2009).

22 Hungerford, R. D. N., Ramakrishna R.; Running, Steven W.; Coughlan, Joseph C. MTCLIM: a mountain microclimate simulation model. *Res. Pap. INT-RP-414. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 52 p.* (1989).

23 Bristow, K. L. & Campbell, G. S. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agricultural and Forest Meteorology* **31**, 159-166, doi:https://doi.org/10.1016/0168-1923(84)90017-0 (1984).

24 Running, S. W., Nemani, R. R. & Hungerford, R. D. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal of Forest Research* **17**, 472-483, doi:10.1139/x87-081 (1987).

25 Kimball, J. S., Running, S. W. & Nemani, R. An improved method for estimating surface humidity from daily minimum temperature. *Agricultural and Forest Meteorology* **85**, 87-98, doi:https://doi.org/10.1016/S0168-1923(96)02366-0 (1997).

26 Thornton, M. M. *et al.* Daymet: Station-Level Inputs and Cross-Validation Result for North America, Version 4. *ORNL Distributed Active Archive Center*, https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1850 (2020).

27 Du, J. GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data. Version 1.0. UCAR/NCAR - Earth Observing Laboratory. . doi: https://doi.org/10.5065/D6PG1QDD (2011).

28 Gourley, J. J., Hong, Y., Flamig, Z. L., Li, L. & Wang, J. Intercomparison of Rainfall Estimates from Radar, Satellite, Gauge, and Combinations for a Season of Record Rainfall. *Journal of Applied Meteorology and Climatology* **49**, 437-452, doi:10.1175/2009JAMC2302.1 (2010).

SUBMITTED