

RESEARCH ARTICLE

# Mapping Atmospheric Moisture Climatologies across the Conterminous United States

Christopher Daly\*, Joseph I. Smith, Keith V. Olson

PRISM Climate Group, Northwest Alliance for Computational Science and Engineering, Oregon State University, Corvallis, Oregon, United States of America

\* [Chris.Daly@oregonstate.edu](mailto:Chris.Daly@oregonstate.edu)



OPEN ACCESS

**Citation:** Daly C, Smith JI, Olson KV (2015) Mapping Atmospheric Moisture Climatologies across the Conterminous United States. PLoS ONE 10(10): e0141140. doi:10.1371/journal.pone.0141140

**Editor:** Robert Guralnick, University of Colorado, UNITED STATES

**Received:** April 17, 2015

**Accepted:** October 4, 2015

**Published:** October 20, 2015

**Copyright:** © 2015 Daly et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All of the PRISM normal grids discussed here are available online at (<http://prism.oregonstate.edu>), and include 800-m and 4-km resolution data, images, metadata, pedigree information, and station inventory files. Raw station data are available without restriction from the URLs listed in the paper.

**Funding:** This study was funded by the USDA Department of Agriculture Risk Management Agency grant number 4500054499. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Spatial climate datasets of 1981–2010 long-term mean monthly average dew point and minimum and maximum vapor pressure deficit were developed for the conterminous United States at 30-arcsec (~800m) resolution. Interpolation of long-term averages (twelve monthly values per variable) was performed using PRISM (Parameter-elevation Relationships on Independent Slopes Model). Surface stations available for analysis numbered only 4,000 for dew point and 3,500 for vapor pressure deficit, compared to 16,000 for previously-developed grids of 1981–2010 long-term mean monthly minimum and maximum temperature. Therefore, a form of Climatologically-Aided Interpolation (CAI) was used, in which the 1981–2010 temperature grids were used as predictor grids. For each grid cell, PRISM calculated a local regression function between the interpolated climate variable and the predictor grid. Nearby stations entering the regression were assigned weights based on the physiographic similarity of the station to the grid cell that included the effects of distance, elevation, coastal proximity, vertical atmospheric layer, and topographic position. Interpolation uncertainties were estimated using cross-validation exercises. Given that CAI interpolation was used, a new method was developed to allow uncertainties in predictor grids to be accounted for in estimating the total interpolation error. Local land use/land cover properties had noticeable effects on the spatial patterns of atmospheric moisture content and deficit. An example of this was relatively high dew points and low vapor pressure deficits at stations located in or near irrigated fields. The new grids, in combination with existing temperature grids, enable the user to derive a full suite of atmospheric moisture variables, such as minimum and maximum relative humidity, vapor pressure, and dew point depression, with accompanying assumptions. All of these grids are available online at <http://prism.oregonstate.edu>, and include 800-m and 4-km resolution data, images, metadata, pedigree information, and station inventory files.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

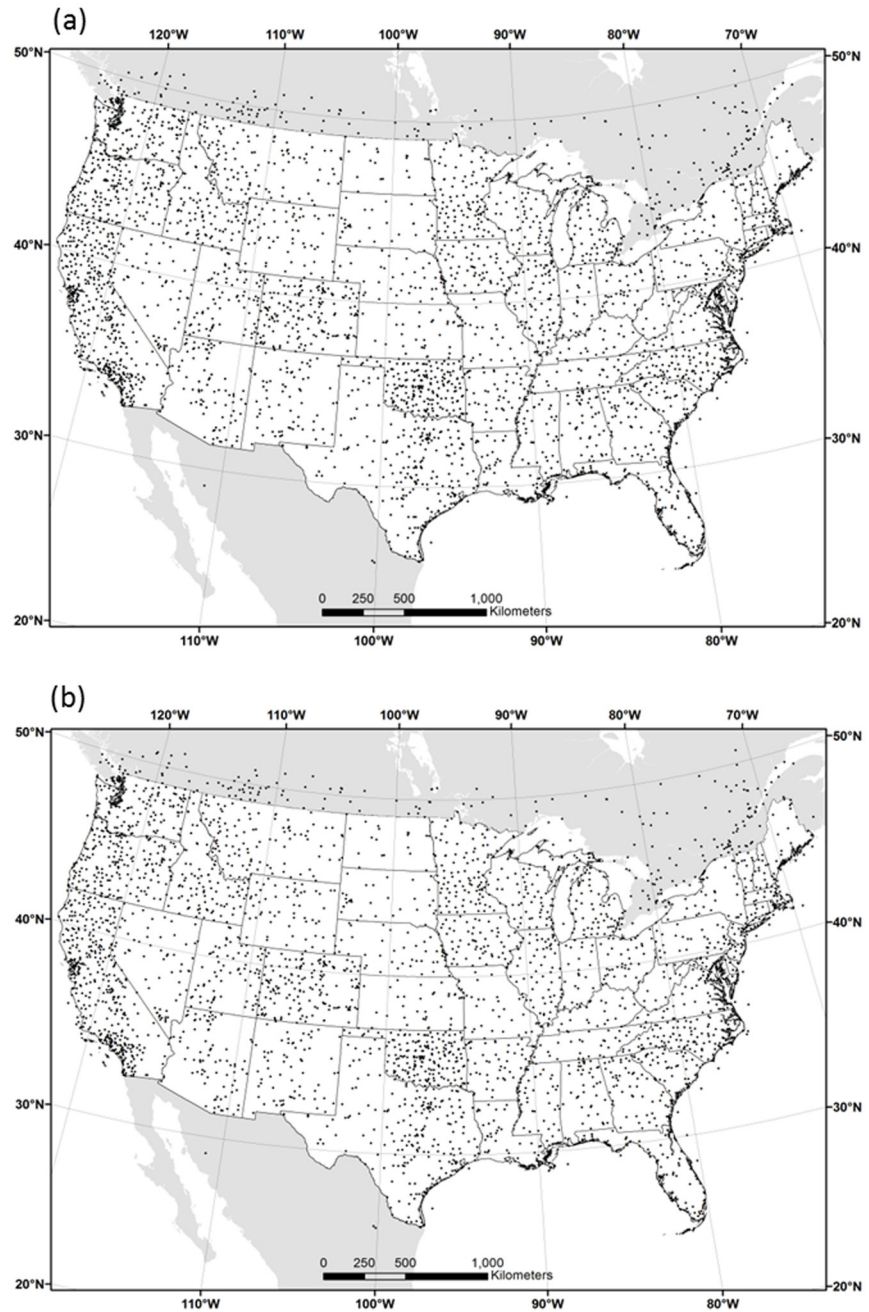
The demand for spatial climate data sets in digital form continues to increase, as more and more climate-driven modeling and analysis activities are performed within spatially-explicit computing environments. Key inputs to these analyses are grids of thirty-year decadal climate averages (e.g., 1971–2000, 1981–2010, etc.), termed “normals,” that describe the values and spatial patterns that can be expected in an average year or month. Through various forms of a technique called Climatologically-Aided Interpolation (CAI; see [Mapping Methods](#) section), climate normal grids also provide the foundation for a number of spatially distributed climate products that cover individual years, months and days [1–4]. Normals produced by the PRISM (Parameter-elevation Relationships on Independent Slopes Model) Climate Group at Oregon State University are often used for this purpose [5–9].

Most spatial climate normals are limited to temperature and precipitation, but measures of atmospheric moisture content and deficit are also important in a variety of processes, including plant function [10–14], drought severity and wildfire behavior [15], [16], and climate change impacts [17], [18]. Relatively little work has been done to produce long-term normal grids of such variables, partly due to a lack of observations. Empirical relationships have been developed to derive humidity measures from available temperature and precipitation data (e.g., [19–21]), but such methods incorporate assumptions that are violated in certain regions (e.g., [22]). Actual observations of mean dew point ( $T_{dmean}$ ) have been included in monthly time series datasets ([9]), which in turn have been incorporated into other datasets (e.g., [23]), but there has been little focus on long-term normals of  $T_{dmean}$ . Further, long-term normals of atmospheric moisture deficit have received even less attention. A useful variable in this regard is vapor pressure deficit ( $VPD$ ), which is the difference between the saturation vapor pressure (dictated by temperature alone) and the actual vapor pressure.  $VPD$  is a direct measure of the deficit, unlike relative humidity ( $RH$ ), for which one must know the temperature at the time of measurement [24]. Ideally, one would want to know the average minimum  $VPD$  ( $VPD_{min}$ ) and maximum  $VPD$  ( $VPD_{max}$ ), which bracket the range of moisture deficit conditions throughout an average day. This is similar to the advantage of knowing the average maximum and minimum temperature as opposed to mean temperature only. An added advantage is that  $T_{dmean}$ ,  $VPD_{min}$  and  $VPD_{max}$  in combination with existing grids of minimum and maximum temperature ( $T_{min}$  and  $T_{max}$ ), allow many other atmospheric moisture variables to be derived, such as minimum and maximum  $RH$ , vapor pressure, and dew point depression, with accompanying assumptions (see Relative Humidity Derivation section, for example).

This paper describes the development of spatial climate normals of 1981–2010 mean monthly  $T_{dmean}$ ,  $VPD_{min}$  and  $VPD_{max}$  across the conterminous United States, using methods that strive to account for the major physiographic factors influencing climate patterns. The  $T_{dmean}$  normals are updates of datasets used in the development of a century-long monthly time series [9]. The  $VPD_{min}$  and  $VPD_{max}$  datasets are, to our knowledge, the first of their kind, although datasets of mean  $RH$  have been developed [25]. Section 2 of this paper describes the study area, preparation of station data, and mapping methods. Section 3 presents the resulting gridded data sets, and discusses model performance. Concluding remarks are given in Section 4.

## Study Area and Work Flow

Climate data sets were developed for the conterminous US at 30 arc-second resolution in geographic (latitude/longitude) coordinates ([Fig 1](#)). A 30-arc-second grid cell is approximately 900 x 700 m at 40N latitude, and is referred to here as “800 m.” The boundaries of the grid are 22°N and 50°N and 65°W and 125°W, which are exactly coincident with other 800-m grids from the



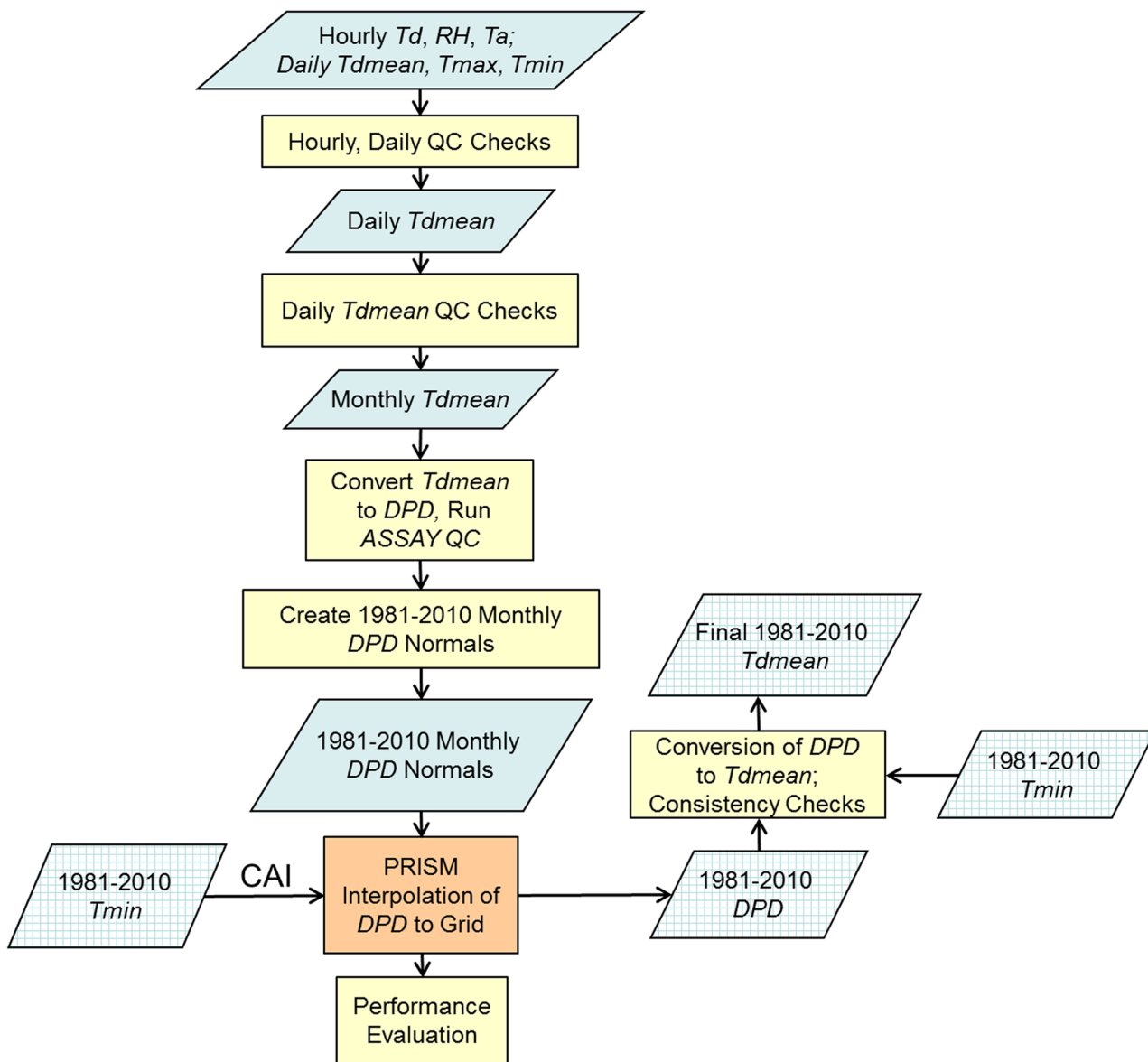
**Fig 1. Map of study area and station locations.** Conterminous US and border area monthly average (a) mean dew point ( $T_{dmean}$ ) stations and (b) minimum and maximum vapor pressure deficit ( $VPD_{min}$  and  $VPD_{max}$ ) stations.

doi:10.1371/journal.pone.0141140.g001

PRISM Climate Group [8]. The interpolation was performed separately in three overlapping regions: western, central and eastern US, and the resulting grids merged to form a complete conterminous US grid. The western region extends from the Pacific Coast to eastern Colorado, central from central Colorado to Lake Michigan, and eastern from eastern Minnesota to the eastern seaboard.

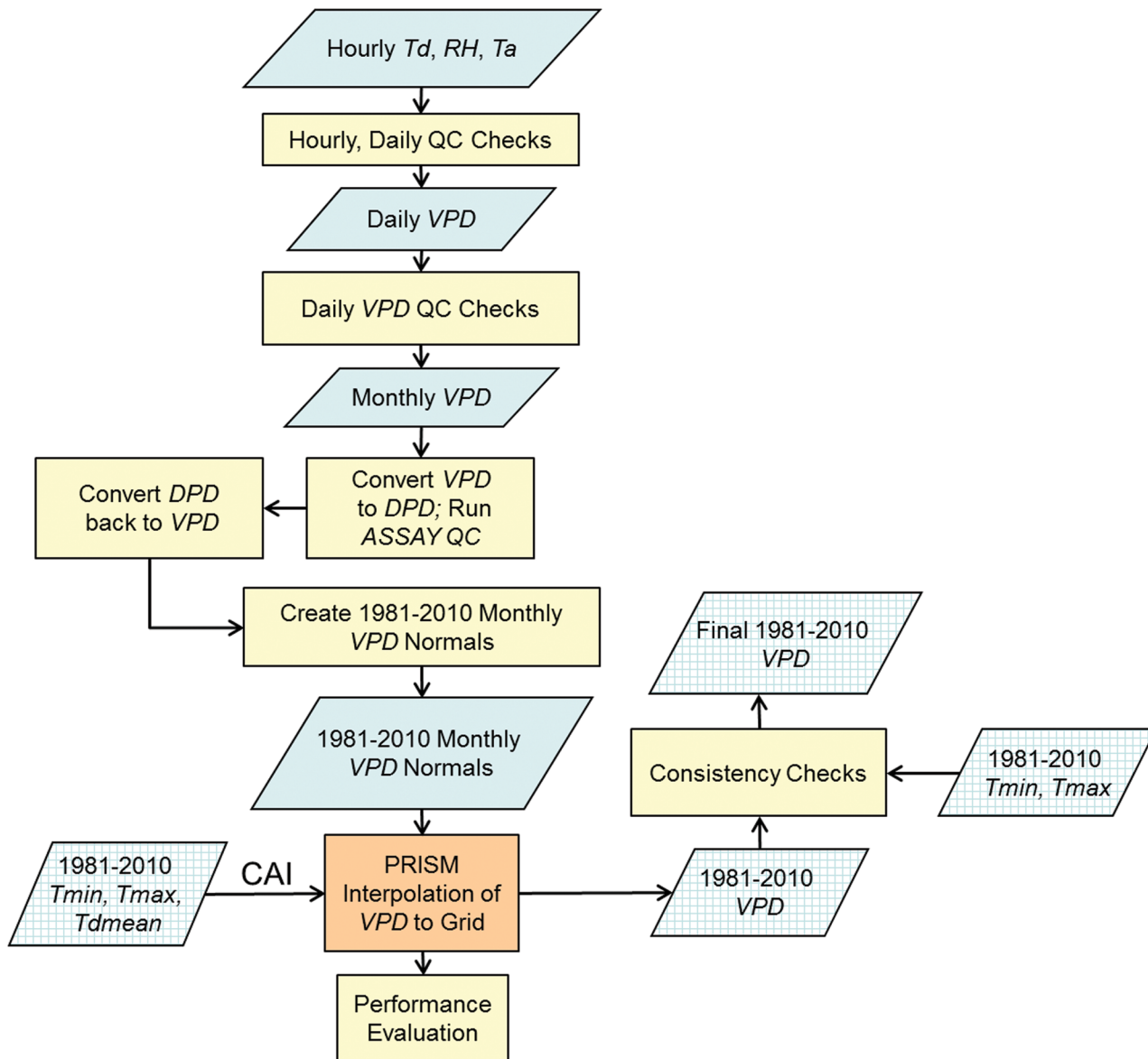
Care was taken to include as many islands offshore the US mainland as possible, but undoubtedly some very small islands were missed. To accommodate GIS shoreline data sets of varying quality and resolution, the modeling region was extended offshore several km and generalized to include bays and inlets. However, the gridded climate estimates are valid over land areas only.

Overviews of the data processing and mapping work flows for  $T_{dmean}$  and for  $VPD_{min}$  and  $VPD_{max}$  are diagrammed in Figs 2 and 3, respectively. The process began with hourly and daily data for  $T_{dmean}$ , and hourly data for  $VPD_{min}$  and  $VPD_{max}$ , which were averaged to daily, and then monthly, time steps, with quality screening done at each stage (see [Station Data and Processing](#) section). At the monthly time step, all three variables were expressed as dew point



**Fig 2. Work flow diagram for data processing and mapping of  $T_{dmean}$ .** Cross-hatched boxes represent gridded data. See text for abbreviation definitions.

doi:10.1371/journal.pone.0141140.g002



**Fig 3. Work flow diagram for data processing and mapping of  $VPD_{min}$  and  $VPD_{max}$ .**  $VPD$  refers to either  $VPD_{min}$  or  $VPD_{max}$ . Cross-hatched boxes represent gridded data. See text for abbreviation definitions.

doi:10.1371/journal.pone.0141140.g003

depression ( $DPD$ ) to take advantage of a spatial consistency quality control (QC) method using a version of PRISM called ASSAY (see Quality Control and Calculation of Monthly Values section).  $T_{dmean}$  was kept in the form of  $DPD$  for interpolation by PRISM, and converted back to  $T_{dmean}$  as a grid post-processing step.  $VPD_{min}$  and  $VPD_{max}$  were converted from  $DPD$  back to their original forms before interpolation by PRISM (see Mapping Methods section). A performance evaluation was conducted on the PRISM interpolation process (see Uncertainty Analysis section). As a final step, output grids from PRISM were checked for consistency with previously-created grids of 1981–2010  $T_{max}$  and  $T_{min}$ .

## Station Data and Processing

### Data Sources

Data from surface stations, numbering about 4,000 for  $T_{dmean}$  (44 km average station spacing) and 3,500 for  $VPD_{max}$  and  $VPD_{min}$  (47 km average station spacing) were obtained from a variety of sources (Table 1; Fig 1). Data from the National Weather Service (NWS) Automated Surface Observing System (ASOS) came via Unidata's Internet Data Distribution system and the National Climatic Data Center (NCDC) Integrated Surface Hourly/Integrated Surface Data archives (<ftp://ftp.ncdc.noaa.gov/pub/data/noaa/>), supplemented by the Solar And Meteorological Observation Network (SAMSON) and Integrated Surface Weather Observations CD-ROMs. NWS Cooperative Observer Program (COOP) and Weather Bureau Army Navy (WBAN) data were obtained from (<http://www.ncdc.noaa.gov/cdo-web/>). USDA Forest Service and Bureau of Land Management Remote Automatic Weather Station (RAWS) data came from archives at the Western Regional Climate Center (<http://www.raws.dri.edu>), the Real-time Observation Monitor and Analysis Network (<http://raws.wrh.noaa.gov>), and from the MesoWest Local Data Manager (LDM) feed and website (<http://mesowest.utah.edu/data>). USDA Natural Resources Conservation Service (NRCS) Soil Climate Analysis Network (SCAN) observations were provided through the National Water and Climate Center web service (<http://www.wcc.nrcs.usda.gov>). Bureau of Reclamation AgriMet (AGRIMET) data were obtained from the Pacific Northwest Region (<http://www.usbr.gov/pn/agrimet/>), Great Plains Region (<http://www.usbr.gov/gp/agrimet/>), Upper Colorado Region in cooperation with the Utah Climate Center (<https://climate.usurf.usu.edu/agweather.php>), and Mid Pacific Region in cooperation with the Desert Research Institute (<http://nicenet.dri.edu/>) and the National Oceanic and Atmospheric Administration NOAA Idaho National Laboratory Weather Center (<http://niwc.noaa.inel.gov/>). Data were also collected from the Oklahoma Climatological Survey Mesonet (OKMESONET; <http://www.mesonet.org/data/public/mesonet/mts>), Colorado State University's Colorado Agricultural Meteorological Network (COAGMET; <http://ccc>.

**Table 1. Source, averaging interval, and number of stations used in the mapping process.**

Source	Averaging Interval	Stations Contributing $\geq 1$ Month to Grids		Stations Contributing All 12 Months to Grids	
		$T_{dmean}$	$VPD_{min}$ $VPD_{max}$	$T_{dmean}$	$VPD_{min}$ $VPD_{max}$
ASOS	Hourly	1960	1947	1789	1619
RAWS	Hourly	1066	1067	885	682
COOP	Daily	388	0	334	0
OKMESONET	Hourly	129	129	123	122
SCAN	Hourly	116	116	87	76
AGRIMET	Hourly	73	73	70	55
NDBC	Hourly	72	72	53	56
COAGMET	Hourly	63	63	54	50
WBAN	Daily	42	0	38	0
Other Networks	Daily	23	0	23	0
<b>Total Surface Stations</b>		<b>4001</b>	<b>3467</b>	<b>3456</b>	<b>2660</b>
NCAR/NCEP (Upper-air) Grid Points	1981–2010 monthly mean	69 <sup>a</sup>	69 <sup>a</sup>	69 <sup>a</sup>	69 <sup>a</sup>

<sup>a</sup> Count for upper air is the number of grid points used.

doi:10.1371/journal.pone.0141140.t001

[atmos.colostate.edu/~coagmet](http://atmos.colostate.edu/~coagmet)), and a small number of miscellaneous stations from state and local networks.

To improve marine representation, data were obtained from coastal stations and offshore buoys operated by the NOAA National Data Buoy Center (NDBC; <http://www.ndbc.noaa.gov/data/historical/stdmet>). To better define humidity profiles at high elevations, mean monthly upper-air temperature, geopotential height, and relative humidity grid points for the western and eastern United States were obtained at 2.5-degree resolution for the period 1981–2010 from the National Center for Environmental Prediction (NCEP) Global Reanalysis (<ftp://ftp.cdc.noaa.gov/Datasets/ncep.reanalysis.derived/pressure>). The 5650-m level (~500 hPa) was chosen for the western US, and the 3050-m level (~700 hPa) for the eastern US. (Upper-air data were not needed in the central US, because of a lack of elevated terrain.) These levels were sufficiently far above the highest terrain features to minimize potential errors involved in estimating surface humidity statistics from free air values.

### Quality Control and Calculation of Monthly Values

As shown in [Table 1](#), station data were available at either an hourly or daily time step. Daily data were sufficient for  $T_{dmean}$ , but hourly data were required to calculate  $VPD_{min}$  and  $VPD_{max}$ . In some cases, hourly data were provided in the form of  $RH$  directly from the instrument. Range checks were performed to screen out values that were either impossible or could cause instabilities in the calculation of related statistics. If  $RH \leq 0$  or  $RH > 105$ , the value was set to missing. Otherwise, if  $RH < 0.5$ , it was set to 0.5, and if  $100 < RH \leq 105$  (as sometimes occurs under saturated conditions), it was set to 100.

Dew point ( $T_d$ ) was derived from hourly  $RH$  (%) and ambient temperature ( $T_a$ , °C) by calculating the saturation vapor pressure at  $T_a$  ( $SATVP_a$ ) ([\[26\]](#), Eq 21):

$$SATVP_a = 6.1094e^{\left(\frac{17.625T_a}{243.04+T_a}\right)} \tag{1}$$

and finding  $T_d$  (°C) [\[27\]](#):

$$T_d = \frac{237.3 \ln \left[ SATVP_a \left( \frac{RH}{611} \right) \right]}{7.5 \ln 10 - \ln \left[ SATVP_a \left( \frac{RH}{611} \right) \right]} \tag{2}$$

A consistency check was done to ensure that the calculated  $T_d$  was less than  $T_a$ ; if not, each was set to missing.

As a range check,  $T_d$  values that fell below -68°C, or exceeded the statewide extreme maximum temperature record (<http://www.ncdc.noaa.gov/extremes/scec/records>) were set to missing. In addition, the hourly  $T_d$  values were subjected to a step check in which each value could not differ from that of the previous hour by more than 10°C; otherwise the current hour's  $T_d$  was set to missing.

Hourly  $VPDs$  were calculated from a combination of  $T_a$  and  $T_d$  or  $T_a$  and  $RH$ , depending on which was available. If  $RH$  was available,  $VPD$  was calculated by: (1) finding  $SATVP_a$  from [Eq 1](#); (2) finding the saturation vapor pressure at the previously calculated  $T_d$ :

$$SATVP_d = 6.1094e^{\left(\frac{17.625T_d}{243.04+T_d}\right)} \tag{3}$$

and; (3) calculating  $VPD$  as

$$VPD = SATVP_a - SATVP_d \tag{4}$$

If  $T_d$  was available,  $VPD$  was obtained by calculating  $SATVP_d$  from Eq 3 and  $VPD$  from Eq 4. Each hourly  $VPD$  was subjected to a range check, where if  $VPD > 200$  hPa or  $< 0$  hPa, or if  $VPD \geq SATVP_a$ , the value was set to missing.

Three persistence tests were done on  $T_a$  and  $T_d$  hourly values for each 24-hour period. If the maximum difference between any pair of hourly observations or the difference between the maximum and minimum value over the 24-hour period was less than  $0.1^\circ\text{C}$ , or the standard deviation of all values in the 24-hour period was less than  $0.1^\circ\text{C}$ , the day was considered a “flat day” and all variables, including the  $VPDs$ , were set to missing.

For networks with hourly data,  $T_{dmean}$  values were calculated by averaging available hourly observations, subject to the requirement that at least 18 of 24 observations be non-missing; fewer than 18 non-missing hourly observations resulted in a missing daily value. Daily  $VPD_{min}$  and  $VPD_{max}$  as well as  $T_{min}$  and  $T_{max}$ , were determined by finding the minimum and maximum hourly value, respectively, subject to the same data completeness requirement.

Daily maximum and minimum  $T_d$  ( $T_{dmax}$ ,  $T_{dmin}$ ) and mean  $T_a$  ( $T_{mean}$ ) values were calculated for diagnostic purposes. Consistency checks were performed on daily combinations of  $T_{dmean}$  and temperature as follows: If  $T_{max} < T_{dmax}$ ,  $T_{min} < T_{dmin}$ , or  $T_{mean} < T_{dmean}$ , all variables, including  $VPD_{min}$  and  $VPD_{max}$  were set to missing.

The daily values were averaged to create monthly mean  $T_{dmean}$ ,  $VPD_{min}$ , and  $VPD_{max}$  for each year of record. A minimum of 85% of non-missing daily values were required for a monthly value to be non-missing [28], [29].

The monthly averages were tested for spatial consistency using the ASSAY QC (quality control) system. ASSAY is a version of PRISM that estimates station values in their absence and compares them to the observed values, a procedure termed cross-validation [8]. The interpolation procedures in ASSAY are exactly the same as in PRISM; the only difference is that ASSAY interpolates to point (station) locations, rather than grid cells. As a rule, a large discrepancy between an observed station value and the interpolated estimate from ASSAY suggests that the station value is unusual compared to nearby stations, and may therefore be erroneous. In previous work, an ASSAY QC analysis was performed for  $T_{dmean}$ , with  $T_{dmean}$  expressed in the form of dew point depression ( $DPD$ ) with  $T_{min}$ ; that is,  $DPD = T_{min} - T_{dmean}$ . (As will be discussed in the next section, the  $DPD$  form of  $T_{dmean}$  was the favored method of expressing  $T_{dmean}$  for interpolation in this study.) In the previous ASSAY analysis of  $DPD$ , QC was done manually by trained personnel, and compared with the ASSAY results. It was found that when the absolute difference between the observation and ASSAY estimate of  $DPD$  exceeded  $4.5^\circ\text{C}$ , the value was considered erroneous by the manual method. Based on these results, in this study ASSAY QC was applied to  $T_{dmean}$  expressed as  $DPD$ , and station values producing absolute differences between the observation and estimate of more than  $4.5^\circ\text{C}$  were flagged as bad and set to missing. Overall, 0.8 percent of the monthly  $T_{dmean}$  values were set to missing using this method.

To take advantage of the ASSAY QC screening method,  $VPD_{min}$  and  $VPD_{max}$  were also expressed as  $DPD$  with  $T_{min}$ , termed  $DPD(VPD_{min})$  and  $DPD(VPD_{max})$ , respectively. This involved: (1) calculating the saturation vapor pressure for the station’s  $T_{min}$  value ( $SATVP_a$ ) using Eq 1; (2) subtracting the  $VPD$  from this value to get  $SATVP_d$ ; (3) obtaining  $T_d$  from Eq 2, substituting  $SATVP_d$  for  $SATVP_a$  and using 100 for the  $RH$  value; and (4) subtracting the resulting  $T_d$  from  $T_{min}$  to obtain the  $DPD$ . Manual QC checks found that the same threshold of  $4.5^\circ\text{C}$  used in QC’ing  $DPD$  was suitable for  $DPD(VPD_{min})$  and  $DPD(VPD_{max})$  as well, and those exceeding this threshold were flagged as bad and set to missing. Overall, 0.7 percent of the monthly  $VPD_{min}$  values and 1.8 percent of the monthly  $VPD_{max}$  values were set to missing using this method.

Monthly values of  $T_{dmean}$ ,  $VPD_{min}$ , and  $VPD_{max}$  passing the ASSAY QC screening were averaged over their period of record (POR), or 1981–2010, if available. A 1981–2010 monthly



mean calculated using data from at least 23 of these 30 years (75% data coverage) was considered to be sufficiently characteristic of the 1981–2010 period, and was termed a “long-term” station. However, many stations had a POR of fewer than 23 years. Averages from stations with short PORs were subjected to adjustment to minimize temporal biases, as described in [8], Appendix A.

## Mapping Methods

Mapping of 1981–2010 mean monthly  $T_{dmean}$ ,  $VPD_{min}$ , and  $VPD_{max}$  was performed using PRISM [6–8], [30]. For each grid cell, PRISM calculated a local linear regression function between the atmospheric moisture variable and a predictor grid (see Climatologically-Aided Interpolation below). Nearby stations entering the regression were assigned weights based primarily on the physiographic similarity of the station to the grid cell. Physiographic factors relevant to this study were distance, elevation, coastal proximity, vertical atmospheric layer, and topographic position (relative to surrounding terrain). Detailed descriptions of the PRISM model algorithms, structure, input grids, and operation are given in [7], [8], and [30]. Details on the specific modeling approach for this study are given below.

### Derive or Interpolate $VPD_{min}$ and $VPD_{max}$ ?

Given that  $T_{dmean}$  is a basic atmospheric moisture variable from which other variables can be derived, early in this study the question was asked: do  $VPD_{min}$  and  $VPD_{max}$  need to be interpolated separately, or can they be derived from a combination of the  $T_{dmean}$  grid and previously-created grids of  $T_{max}$  and  $T_{min}$ ? More specifically, can  $T_{dmean}$  be combined with  $T_{max}$  to estimate  $VPD_{max}$  and  $T_{dmean}$  combined with  $T_{min}$  to estimate  $VPD_{min}$ ? To do this, we must assume that  $T_d = T_{dmean}$  and  $T_a = T_{max}$  at the time of  $VPD_{max}$ , and  $T_d = T_{dmean}$  and  $T_a = T_{min}$  at the time of  $VPD_{min}$ . To test these assumptions, monthly averages of  $VPD_{min}$  and  $VPD_{max}$ , as well as  $VPD_{min}$  and  $VPD_{max}$  estimated as described above, were calculated and averaged for each month over the ten-year period 2003–2012 at 100 randomly selected ASOS stations with hourly data. Results of this exercise are given in Table 2 for January and July, which bracket the range of monthly values observed during the year.

In Table 2, actual and estimated  $VPD_{min}$  and  $VPD_{max}$  are expressed as means and percentiles (5<sup>th</sup> percentile / mean / 95<sup>th</sup> percentile) so that the full distribution of values can be evaluated. Differences between actual and estimated  $VPD_{max}$  were generally within about three hPa or five percent across the distributions in both winter and summer. However, a proportion of the estimated  $VPD_{min}$  values fell below zero, which is not physically possible. This problem was most serious during winter; in January, a  $VPD_{min}$  of less than zero was estimated at 47 of the 100 stations. In these cases,  $T_{dmean}$  exceeded  $T_{min}$ , resulting in a negative VPD. In reality, the opposite was true, resulting in a positive  $VPD_{min}$ . Given this issue, deriving  $VPD_{min}$  and

**Table 2. Monthly averages of actual and estimated vapor pressure deficit.** 2003–2012 January and July averages from 100 randomly-selected ASOS stations. Actual and estimated VPD values are expressed as 5<sup>th</sup> percentile / mean / 95<sup>th</sup> percentile.

VPD (hPa)	January	July
$VPD_{min}$	0.2 / 0.8 / 1.6	0.9 / 3.3 / 9.7
Estimated $VPD_{min}$	-0.6 / 0.7 / 1.6	-0.1 / 4.2 / 11.2
$VPD_{max}$	1.5 / 5.9 / 12.3	15.2 / 26.9 / 47.7
Estimated $VPD_{max}$	1.8 / 6.3 / 12.1	14.7 / 26.3 / 45.4

doi:10.1371/journal.pone.0141140.t002

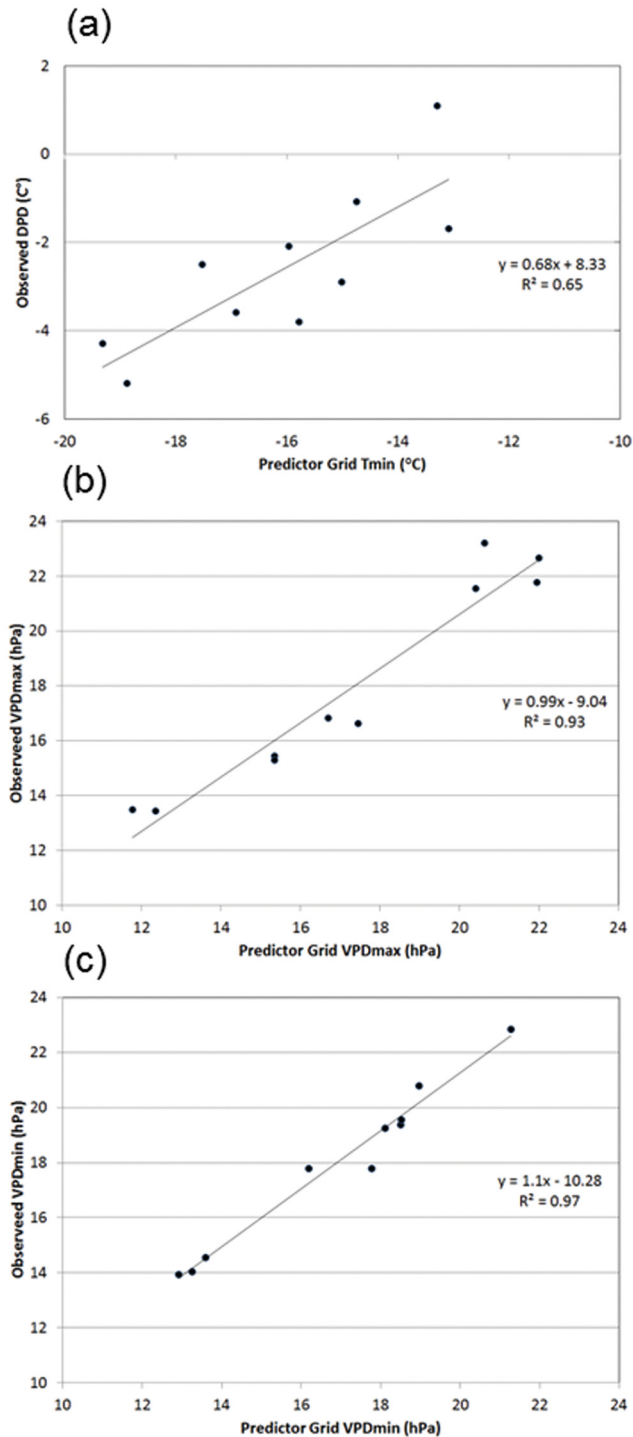
$VPD_{max}$  from  $T_{dmean}$ ,  $T_{min}$ , and  $T_{max}$  was not considered a viable option, leading us to interpolate  $VPD_{min}$  and  $VPD_{max}$  directly from station data.

## Climatologically-Aided Interpolation

In previous work, constructing the 1971–2000 and 1981–2010 monthly normals for  $T_{min}$  and  $T_{max}$  used a DEM as the predictor grid (see [8] for details on methods used for mapping  $T_{min}$  and  $T_{max}$ ). There were nearly 10,000 stations used in the mapping of the 1971–2000  $T_{min}$  and  $T_{max}$  normals, and over 16,000 stations used in the 1981–2010 normals. In contrast, there were only 3,500–4,000 stations available for this study, with poor representation at high elevations. Faced with limited station data, we opted for the CAI method of interpolation. CAI uses an existing climate grid to improve the interpolation of another climate element for which data may be sparse or intermittent in time [31–36]. This method relies on the assumption that local spatial patterns of the climate element being interpolated closely resemble those of the existing climate grid (called the predictor grid). Uses of CAI fall into two main categories: (1) using a long-term mean grid of a climate element to aid the interpolation of the same element over different (usually shorter) averaging periods; and (2) using a grid of a climate variable to aid the interpolation of a different, but related, climate variable, such as interpolating annual extreme minimum temperature using January mean minimum temperature as the predictor grid [37]. A classic example of the first strategy involves mapping a long-term mean climatology carefully with sophisticated methods, then developing time series grids for shorter averaging periods (monthly or daily) using simpler and faster methods such as inverse-distance weighting to interpolate deviations from the mean climatology to a grid. These deviations can then be added to (e.g., temperature) or multiplied by (e.g., precipitation) the mean climatology to obtain the new grid.

Our use of CAI for this study falls into the second strategy, for which we use pre-existing grids of 1981–2010 mean monthly  $T_{min}$  and  $T_{max}$  as predictor grids in the interpolation process. A series of tests was conducted with ASSAY and PRISM to determine which of the 1981–2010 mean monthly  $T_{min}$  and  $T_{max}$  grids were the strongest predictors of the spatial patterns of  $T_{dmean}$ .  $T_{min}$  was found to be a good predictor, which is not surprising given that temperatures often reach the dew point at the time of  $T_{min}$  over much of the US. However, inspection of the interpolated grids revealed that in some areas subject to locally low temperatures, such as cold air pools in mountain valleys, the interpolated  $T_{dmean}$  exceeded the mean temperature, meaning that long-term mean  $RH$  exceeded 100 percent, which was not acceptable. In order to more closely tie the patterns of  $T_{dmean}$  to the patterns of the existing monthly  $T_{min}$  grids and their relatively large supporting station data sets, each  $T_{dmean}$  station value was expressed as the deviation from  $T_{min}$ , or  $DPD$  ( $T_{dmean} - T_{min}$ ). Thus, in the interpolation process, PRISM was run with the 1981–2010 mean monthly  $T_{min}$  grid as the independent variable and  $DPD$  as the dependent variable. The local regression functions were generally not as strong as they were in the case of  $T_{dmean}$  vs.  $T_{min}$ , because  $DPD$  was essentially the residual from the  $T_{dmean}$  vs.  $T_{min}$  relationship. Once the mean monthly  $DPD$  values were mapped in this manner, the final  $T_{dmean}$  grid was obtained by adding the  $DPD$  grid to the 1981–2010  $T_{min}$  grid. The result was an interpolated  $T_{dmean}$  grid that was highly consistent in an absolute, as well as relative sense, with the associated  $T_{min}$  grid. An example of the relationship between 1981–2010 January mean observed  $DPD$  and gridded  $T_{min}$  for a location north of the Wind River Mountains of Wyoming (43.6N, 109.73W) is shown in Fig 4a. This area is characterized by persistent winter-time cold air pools in mountain valleys, where the humidity is high and  $T_{min}$  is less than  $T_{dmean}$  ( $DPD < 0$ ). Above these cold pools,  $T_{min}$  increases and rises above  $T_{dmean}$  ( $DPD > 0$ ).

An additional series of tests was conducted with PRISM and ASSAY to determine whether either of the pre-existing 1981–2010 mean monthly  $T_{min}$  and  $T_{max}$  grids could be used as



**Fig 4. Scatterplots of relationships between predictor grid and interpolated variables.** 1981–2010 mean monthly (a) dew point depression (*DPD*) vs. minimum temperature ( $T_{min}$ ) for January in western Wyoming; (b) minimum vapor pressure deficit ( $VPD_{min}$ ) vs. first-guess  $VPD_{min}$  for June near Las Vegas; and (c) maximum vapor pressure deficit ( $VPD_{max}$ ) vs. first-guess  $VPD_{max}$  for October in San Francisco.

doi:10.1371/journal.pone.0141140.g004

predictors of the spatial patterns of  $VPD_{min}$  and  $VPD_{max}$ .  $T_{max}$  was found to be a good predictor of  $VPD_{max}$ ; and the same was true for  $T_{min}$  and  $VPD_{min}$ . However, once  $T_{dmean}$  was mapped, a superior, second-generation CAI method using  $T_{dmean}$  became available. Termed CAI2, this method involves using the result of a CAI interpolation as the predictor grid in another CAI interpolation. Specifically, a first-guess  $VPD_{min}$  grid was created by calculating the  $VPD$  associated with the combination of the  $T_{dmean}$  and  $T_{min}$  grids using Eqs 1, 3 and 4. Similarly, a first-guess  $VPD_{max}$  grid was found by calculating the  $VPD$  associated with the combination of the  $T_{dmean}$  and  $T_{max}$  grids. These “first-guess” predictor grids represented what the  $VPD_{min}$  and  $VPD_{max}$  spatial patterns would be like if  $T_d$  was held constant throughout the day at  $T_{dmean}$ , and  $VPD_{min}$  occurred at the time of  $T_{min}$  and  $VPD_{max}$  occurred at the time of  $T_{max}$ . While these assumptions are not always correct (see Table 2), they are sufficiently reasonable to produce predictor grids that closely match the relative spatial patterns of  $VPD_{min}$  and  $VPD_{max}$ .

An example of the relationship between 1981–2010 June mean observed  $VPD_{min}$  and first-guess predictor grid  $VPD_{min}$  for a location near Las Vegas, Nevada (35.79N, 115.26W) is shown in Fig 4b. In this desert environment,  $VPD_{min}$  is still relatively large (14–22 hPa), even at its minimum for an average day in June. The lowest  $VPD_{min}$  values are found at higher elevations, where temperatures are cooler. An example of the relationship between 1981–2010 October mean observed  $VPD_{max}$  and the first-guess predictor grid  $VPD_{max}$  for a location in San Francisco, along the California coastline (37.76N, 122.45W), is shown in Fig 4c. In this case, the lowest  $VPD_{max}$  values are found along the immediate coast where temperatures are cooler and moisture is greater, with higher values in warmer and drier inland areas.

### PRISM Weighting Functions

During PRISM interpolation, upon entering the local linear regression function for a pixel, each station was assigned a weight based on several factors. The combined weight ( $W$ ) of a station is given by the following:

$$W = W_c [F_d W_d^2 + F_z W_z^2]^{1/2} W_p W_l W_t \tag{5}$$

where  $W_c$ ,  $W_d$ ,  $W_z$ ,  $W_p$ ,  $W_b$  and  $W_t$  are cluster, distance, elevation, coastal proximity, vertical layer, and topographic position weights, respectively, and  $F_d$  and  $F_z$  are user-specified distance and elevation weighting importance scalars [8], [38]. All weights and importance factors, individually and combined, are normalized to sum to unity. PRISM weighting functions not enabled in this study were topographic facet weighting, which is used primarily to identify rain shadows in precipitation mapping, and effective terrain height weighting, which identifies orographic precipitation regimes based on terrain profiles [8].

Table 3 summarizes how the PRISM climate regression and station weighting functions accounted for physiographic climate forcing factors, and provides citations for further information. Cluster weighting was used to keep clusters of stations that represent similar local conditions from dominating the regression functions; both horizontal and vertical separations were considered [8]. Distance and elevation weighting were used to accommodate the spatial coherence of climatic regimes, both horizontally and vertically [8].

Coastal proximity weighting accounted for sharp gradients in temperature and atmospheric moisture from coastlines to interior regions [7], [8], [30]. Atmospheric layer weighting was useful where temperature inversions occurred, by delineating gradients in the relationship between temperature and atmospheric moisture along the transition from the relatively humid boundary layer near the earth’s surface to the drier free atmosphere above [7], [30]. Topographic position weighting differentiated topographically sheltered locations where cold, moist

**Table 3. PRISM weighting algorithms and associated physiographic climate forcing factors.** Inputs to the algorithms and references on their formulation and use are included. Methods used to prepare the gridded model inputs are summarized in [8], Tables 1 and 2.

PRISM Algorithm	Description	Physiographic Forcing Factors	Inputs to Algorithm	Reference
<b>Climate Regression Function (CAI)</b>	Develops local relationships between a climate predictor grid and the interpolated variable	Incorporates physiographic features implicit in the climate predictor grid	Station data; climate predictor grid	This paper, Mapping Methods Section; [37] Section 4b
<b>Cluster Weighting</b>	Downweights stations clustered with others	–	Station locations and elevations	[8], Appendix B
<b>Distance Weighting</b>	Upweights stations that are horizontally close	Horizontal coherence of climate regimes	Station locations	[8], Section 4.2.1
<b>Elevation Weighting</b>	Upweights stations that are vertically close	Vertical coherence of climate regimes	DEM; station locations and elevations	[7], Section 4.1
<b>Coastal Proximity Weighting</b>	Upweights stations having similar exposure to coastal influences	Effects of water bodies on temperature and moisture	DEM; coastal proximity grid; station locations	[7], Section 6; [30] Section 2.3.2; [8] Section 4.2.2
<b>Two-Layer Atmosphere Weighting</b>	If an inversion is present, upweights stations in the same vertical layer (boundary layer or free atmosphere)	Temperature inversions; vertical limit to moist boundary layer (humidity)	DEM; inversion height grid; station locations	[7], Section 7; [30], Section 2.3.4
<b>Topographic Position Weighting</b>	Upweights stations having similar heights above the local terrain	Cold air drainage and pooling in topographic depressions	DEM; topographic index grid; station locations	[38], Section 4; [8], Section 4.2.5

doi:10.1371/journal.pone.0141140.t003

air may accumulate, from more exposed locations not susceptible to cold air pooling, such as hill slopes and ridge tops [8], [38].

Relevant PRISM control parameters are listed in Table 4. A minimum of 25 stations were required for each pixel’s regression function, and the radius of influence was expanded from a minimum of 20 km to as far as necessary to reach the 25-station threshold.  $T_{dmean}$  slope bounds were expressed as the change in  $DPD$  per unit  $T_{min}$  from the predictor grid, and  $VPD$  slopes expressed as the change in  $VPD$  per unit first-guess  $VPD$  from the predictor grid. The maximum and minimum allowable regression slopes were derived from test runs of PRISM where distributions of slope values were created and outliers examined to determine validity, combined with performance assessments using ASSAY. In the final PRISM interpolation runs, slopes falling outside the designated bounds were set to values that fell halfway between the default slope and the bound that the slope violated (either the maximum or minimum). In the western region, allowable slopes in the relationship between  $VPD_{max}$  and the first-guess  $VPD_{max}$  were constrained to fall between 0.99 and 1.01, so as to avoid a rare circumstance of predicting a  $VPD_{max}$  that might approach the value of  $SATVP_a$  at  $T_{max}$  (hence producing a very low  $RH$ ) in the warmest and driest areas. The exponential relationship between temperature and  $SATVP_a$  steepens at higher temperatures, leaving less room for error in extremely low-humidity situations. Unconstrained, the average slope of the 5.5 million PRISM regression functions (one per pixel) for July (the warmest month) across the western region was 1.017, with the 10th percentile at 0.964 and the 90th at 1.15, so this constraint had a very slight effect.

Station weighting parameters were very similar among the three variables, and were generally set to the same values across the three regions. The exception was the elevation weighting exponent for  $VPD_{max}$ , which was set to 0.5, compared to 1.5 for the other variables.  $T_{max}$  and  $VPD_{max}$  occur primarily during the day, when the atmosphere is more likely to be well-mixed, and vertical gradients in atmospheric moisture more slowly varying. As a rule, the weighting exponents were set to the lowest values possible to achieve the desired effects. Being parsimonious with the weighting exponents ensured that the station data entering the local regression

**Table 4. Values of relevant PRISM interpolation parameters.** Regression slope bounds are sometimes different among regions; values separated by forward slashes are for west, central, and east regions, respectively. See [7] and [8] for details on PRISM parameters and Table 3 for citations to details on weighting function formulations.

Name	Description	$T_{dmean}$	$VPD_{min}$	$VPD_{max}$
<u>Regression Function</u>				
$R$	Radius of influence	20 km*	20 km*	20 km*
$s_t$	Minimum number of total stations desired in regression	25 stations	25 stations	25 stations
$\beta_{1ma}$	Min valid regression slope, layer 1	-0.5 <sup>+</sup>	-10/-1/-1.5 <sup>#</sup>	0.99/0/0 <sup>#</sup>
$\beta_{1xa}$	Max valid regression slope, layer 1	2.5 <sup>+</sup>	10/2/1 <sup>#</sup>	1.01/1.5/1.5 <sup>#</sup>
$\beta_{1da}$	Default regression slope, layer 1	0.5 <sup>+</sup>	1/0.5/0.5 <sup>#</sup>	1/1/1 <sup>#</sup>
$\beta_{1mb}$	Min valid regression slope, layer 2	-0.5 <sup>+</sup>	-10/-1/-1 <sup>#</sup>	0.99/0/0 <sup>#</sup>
$\beta_{1xb}$	Max valid regression slope, layer 2	2.5 <sup>+</sup>	10/2/1 <sup>#</sup>	1.01/1.5/1.5 <sup>#</sup>
$\beta_{1db}$	Default regression slope, layer 2	0.5 <sup>+</sup>	1/1/0.5 <sup>#</sup>	1/1/1 <sup>#</sup>
<u>Distance Weighting</u>				
$A$	Weighting exponent	1.5	1.5	1.5
$F_d$	Importance factor	0.5	0.5	0.5
$D_m$	Minimum allowable distance	0 km	0 km	0 km
<u>Elevation Weighting</u>				
$B$	Weighting exponent	1.5	1.5	0.5
$F_z$	Importance factor	0.5	0.5	0.5
$\Delta z_m$	Station-grid cell elev difference below which weight is maximum	100 m	100 m	100 m
$\Delta z_x$	Station-grid cell elev difference above which weight is zero	5000 m	5000 m	5000 m
<u>Coastal Proximity Weighting</u>				
$v$	Weighting exponent	1.0	1.0	1.0
$\Delta p_m$	Station-grid cell proximity difference below which weight is maximum	100 <sup>§</sup>	100 <sup>§</sup>	100 <sup>§</sup>
$\Delta p_x$	Station-grid cell proximity difference above which weight is zero	300 <sup>§</sup>	300 <sup>§</sup>	300 <sup>§</sup>
<u>Vertical Layer Weighting</u>				
$y$	Weighting exponent	0.5	0.5	0.5
$\Delta l_m$	Station-grid cell distance to adjacent layer below which weight is maximum	100 m	100 m	100 m
$\Delta l_x$	Station-grid cell distance to adjacent layer above which weight is zero	N/A	N/A	N/A
<u>Topographic Position Weighting</u>				
$\mu$	Weighting exponent	1.5	1.5	1.5
$\Delta t_m$	Station-grid cell topo difference below which weight is maximum	50 m	50 m	50 m
$\Delta t_x$	Station-grid cell topo difference above which weight is zero	900 m	900 m	900 m

\* Expands as needed to encompass minimum number of total stations desired in regression ( $s_t$ ).

<sup>+</sup> Slope is expressed as the change in dew point depression (DPD) per unit Tmin from the predictor grid

<sup>#</sup> Slope is expressed as the change in VPD per unit first-guess VPD from the predictor grid

<sup>§</sup> Units are effective distance in grid cells from the coast, incorporating terrain barriers

doi:10.1371/journal.pone.0141140.t004

functions were not down-weighted unnecessarily, which can weaken the statistical results by decreasing the effective number of stations in the regression.

A number of small but noticeable inconsistencies in DPD values between adjacent stations were observed during the summer months in some agricultural areas, most notably in eastern Colorado. Further investigation revealed that relatively low DPD values were coming from two networks: AGRIMET and COAGMET. Stations in these networks are typically sited in or near irrigated fields for use in water management calculations, resulting in more humid conditions than locations away from irrigated areas. Given that the data from these networks were of otherwise high quality, it was unreasonable to simply omit these two networks outright. This issue

raised questions about whether the grids should, or could, represent conditions over irrigated land. Concerns over attempting to do so are summarized as follows:

1. Siting requirements for these station networks state that the station must be located in, or immediately adjacent to, an irrigated field to be representative of an irrigated environment. This suggests that the effects of irrigation on humidity are highly local, and which may not extend beyond one 800-pixel. In practice, however, these stations influenced the interpolated estimates many km away, well beyond the likely limits of irrigation in many instances.
2. A complete picture of the location and extent of irrigated lands across the country was unavailable; further, there was no interpolation mechanism in place to constrain the influence of stations on irrigated fields to just those lands. It was also unclear how to modify  $DPD$  values in the many irrigated areas not represented by station data.
3. We, and likely others, will be using these atmospheric moisture climatologies as the predictor grids for CAI mapping of monthly and daily time series of the same variables. Given that many of these time series will extend back a century or more, when crop patterns and irrigation practices were very different than today, patterns of humidity caused by today's irrigation patterns could be propagated to times when they are not applicable.

Given these concerns, a subjective middle ground was taken, where a few ( $<10$ ) stations causing the most severe spatial discrepancies were omitted from the  $T_{dmean}$  dataset, and the rest retained. For consistency, the same stations were also omitted from the  $VPD_{min}$  and  $VPD_{max}$  datasets.

## Grid Post-processing

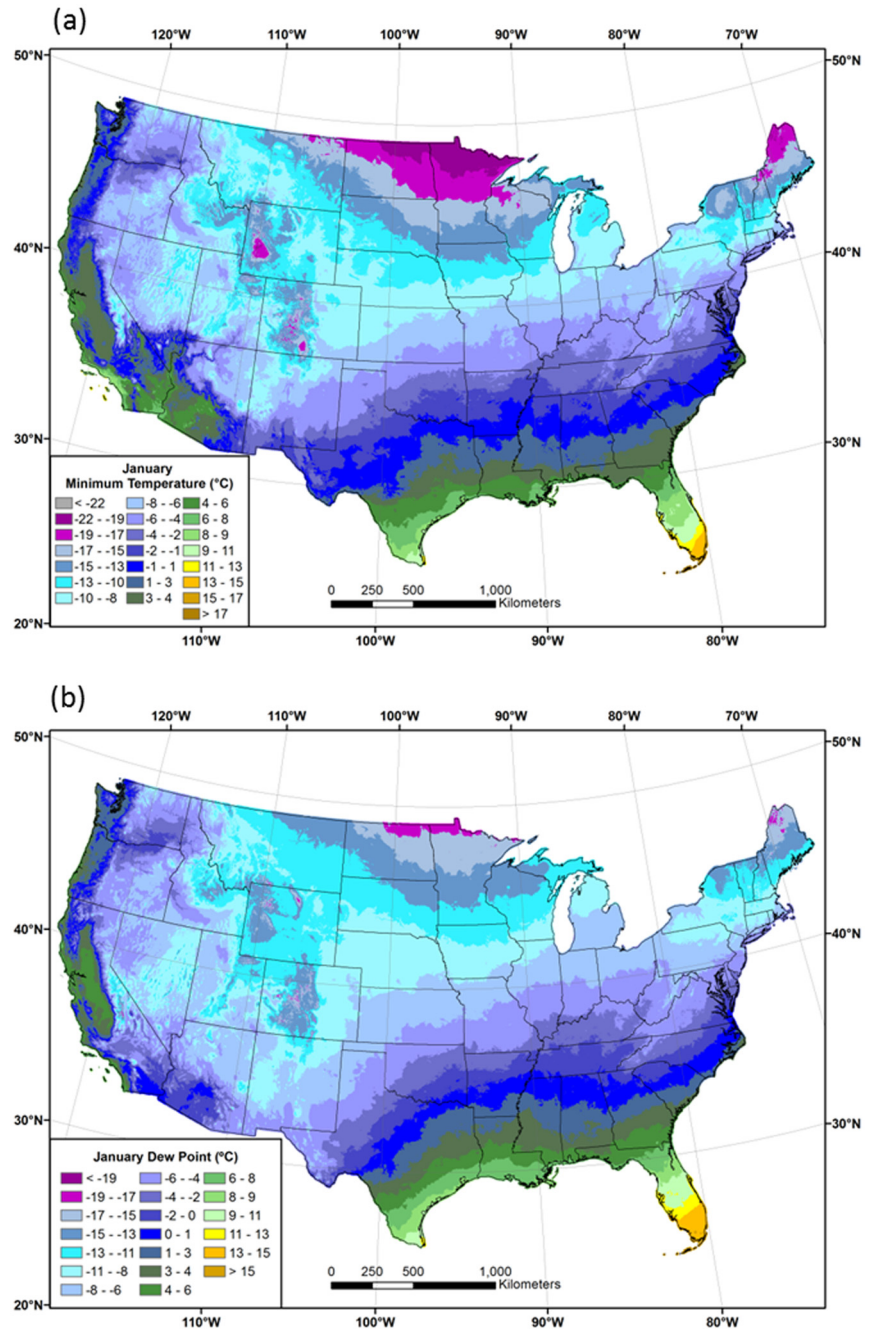
Once monthly grids of  $T_{dmean}$ ,  $VPD_{min}$ , and  $VPD_{max}$  were generated with PRISM, post-processing checks were made to ensure that the interpolated values did not exceed reasonable ranges, and that consistency was maintained among the three variables and with the pre-existing 1981–2010 mean monthly  $T_{min}$  and  $T_{max}$  grids. Given that the 1981–2010 monthly means are by definition made up of yearly values above and below these means, the acceptable ranges were defined to be more restrictive than what would be allowed on, say, a single day. If  $T_{dmean} > (T_{mean} - 0.5^\circ\text{C})$ ,  $T_{dmean}$  mean was set equal to  $T_{mean} - 0.5^\circ\text{C}$ . If  $VPD_{min} < 0.01$  hPa,  $VPD_{min}$  was set equal to 0.01 hPa. If  $VPD_{min} < 0.001 \text{ SATVP}_a$  at  $T_{min}$  (i.e.,  $RH > 99.9\%$ ),  $VPD_{min}$  was set equal to  $0.001 \text{ SATVP}_a$  at  $T_{min}$ . If  $VPD_{max} > 0.95 \text{ SATVP}_a$  at  $T_{max}$  (i.e.,  $RH < 5\%$ ),  $VPD_{max}$  was set equal to  $0.95 \text{ SATVP}_a$  at  $T_{max}$ . The number of grid cells affected by these restrictions was limited to a handful in remote mountainous regions of the western US.

## Results and Discussion

### Spatial Patterns

**Dew Point.** Maps of 1981–2010 mean  $T_{min}$  and  $T_{dmean}$  in January and July are shown in Figs 5 and 6, respectively. The general patterns of  $T_{dmean}$  are similar to those of  $T_{min}$ , especially in winter. In January, values are lowest in the northern tier of states, and in cold pools along valley bottoms in the West (Fig 5).  $T_{dmean}$  is high in the eastern part of the country in July, exceeding  $20^\circ\text{C}$  in much of the Southeast, which is exposed to the transport of moist air from the Gulf of Mexico (Fig 6). Despite relatively warm temperatures, July  $T_{dmean}$  is low in the dry Intermountain West. A notable exception is the Southwest, where the summer monsoon produces locally elevated  $T_{dmean}$  values.

Interpolated  $DPDs$  between  $T_{min}$  and  $T_{dmean}$  ( $T_{min} - T_{dmean}$ ) for January and July are shown in Fig 7.  $DPD$  is the original variable interpolated with PRISM before being added to the  $T_{min}$

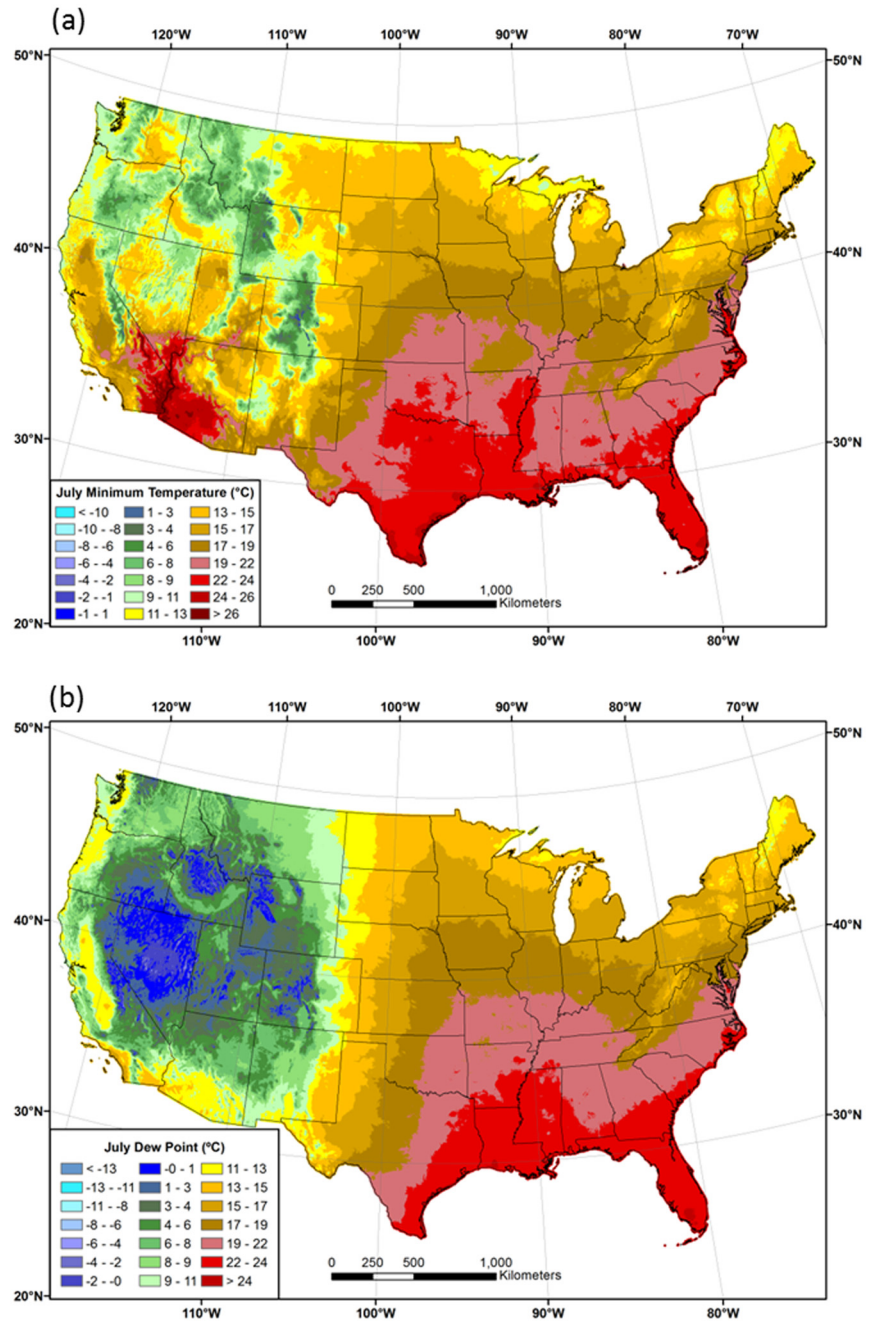


**Fig 5. Maps of January minimum temperature and dew point.** Conterminous US 1981–2010 mean January (a) minimum temperature ( $T_{min}$ ) and (b) dew point temperature ( $T_{dmean}$ ).

doi:10.1371/journal.pone.0141140.g005

grid to obtain  $T_{dmean}$ . In January,  $DPDs$  are mostly near zero or negative, meaning that  $T_{dmean}$  is similar to, or greater than,  $T_{min}$  over much of the country. In the northern US and in persistent cold pools in large inland valleys of the West,  $T_{dmean}$  is as much as 3–5°C greater than  $T_{min}$  (Fig 7a). For example, the 1981–2010 mean January  $T_{dmean}$  at the International Falls, MN ASOS station is -18.1°C and the mean January  $T_{min}$  is -21.3°C. In contrast,  $T_{dmean}$  averages



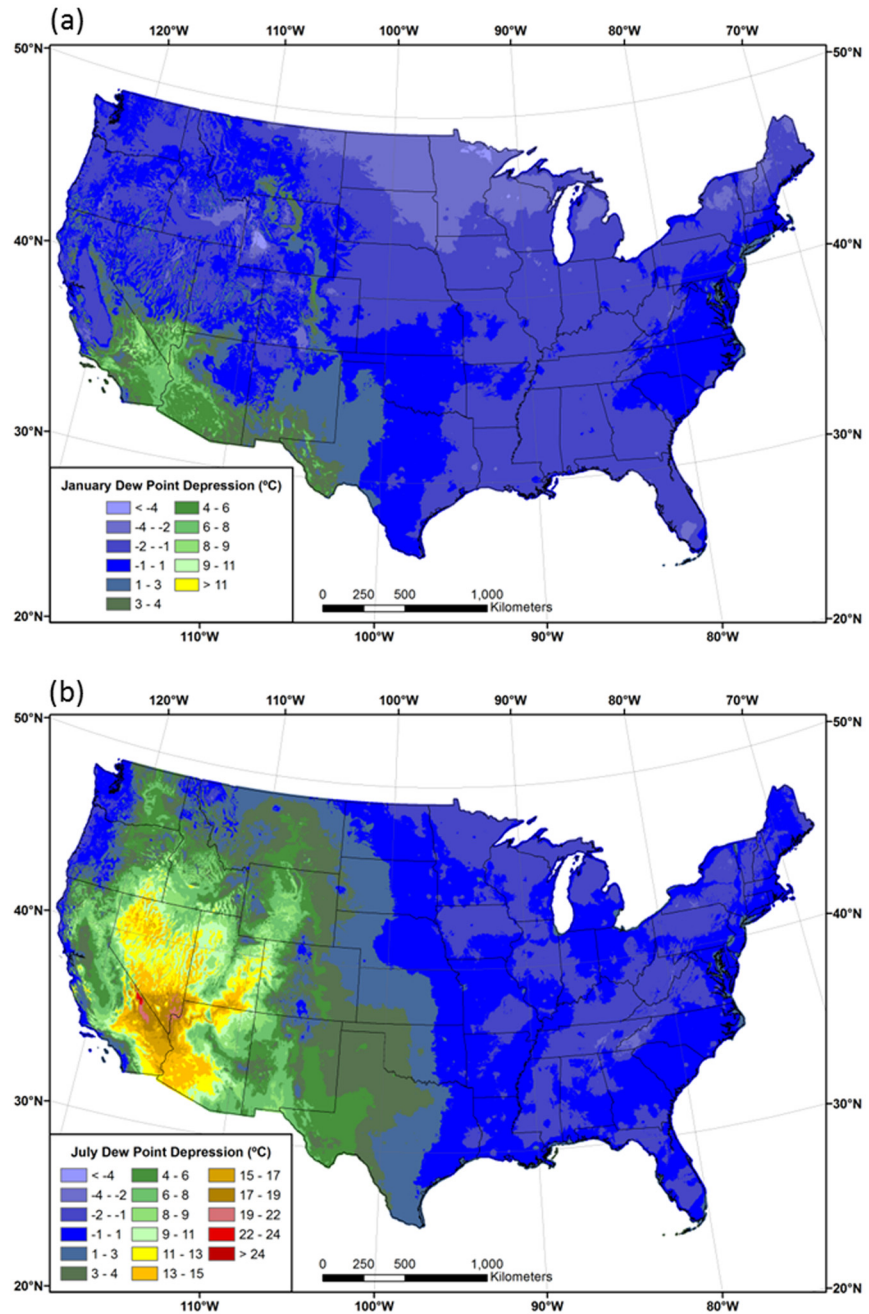


**Fig 6. Maps of July minimum temperature and dew point.** Conterminous US 1981–2010 mean July (a) minimum temperature ( $T_{min}$ ) and (b) dew point temperature ( $T_{dmean}$ ).

doi:10.1371/journal.pone.0141140.g006

6–8°C lower than  $T_{min}$  in the southwestern US, which is a testament to the region’s aridity, even in winter.

In July,  $T_{dmean}$  and  $T_{min}$  have similar values in the eastern half of the country. Interestingly, major metropolitan areas such as St. Louis, Kansas City, Minneapolis, and Detroit, and several eastern seaboard cities, are visible as small areas of relatively warm  $T_{min}$  values, resulting in positive differences between  $T_{min}$  and  $T_{dmean}$  of 1–3°C (dark green spots in Fig 7b). The West is



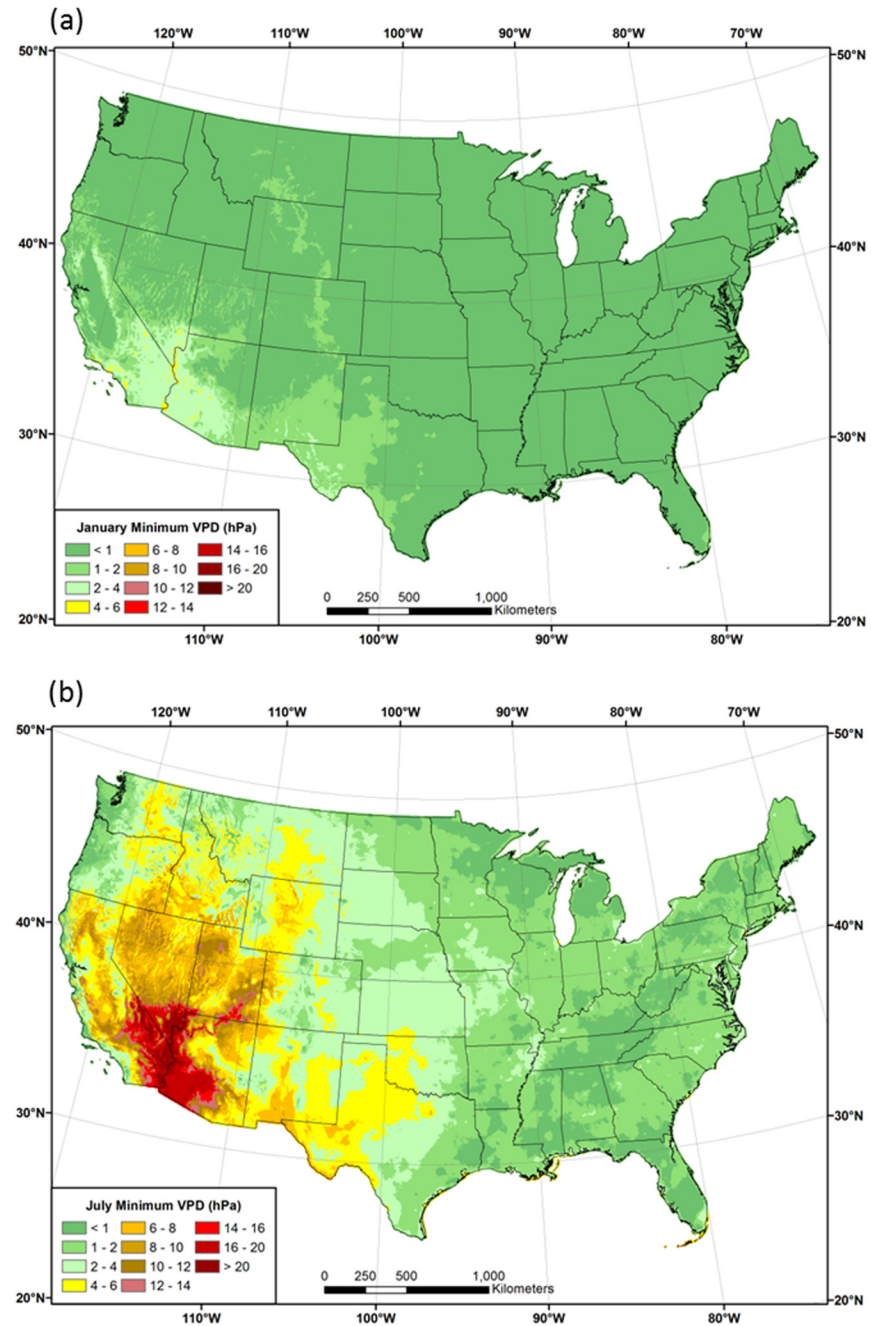
**Fig 7. Maps of dew point depression (DPD).** Conterminous US difference between the 1981–2010 mean minimum temperature and the 1981–2010 mean dew point temperature ( $T_{min} - T_{dmean}$ ) for (a) January and (b) July.

doi:10.1371/journal.pone.0141140.g007

dominated by larger positive differences, where  $T_{min}$  is much higher than  $T_{dmean}$ . The change-over from near-zero to positive differences roughly follows the 100<sup>th</sup> meridian. Notable exceptions are the West Coast, where cool, marine air penetrates inland, and areas of the Southwest affected by the summer monsoon. The core area of maximum DPD is centered on southeastern California, Nevada, western Arizona, and lower elevations of Utah, where  $T_{dmean}$  averages 10–25°C lower than  $T_{min}$ . A combination of low moisture content (low  $T_{dmean}$ ) and limited time

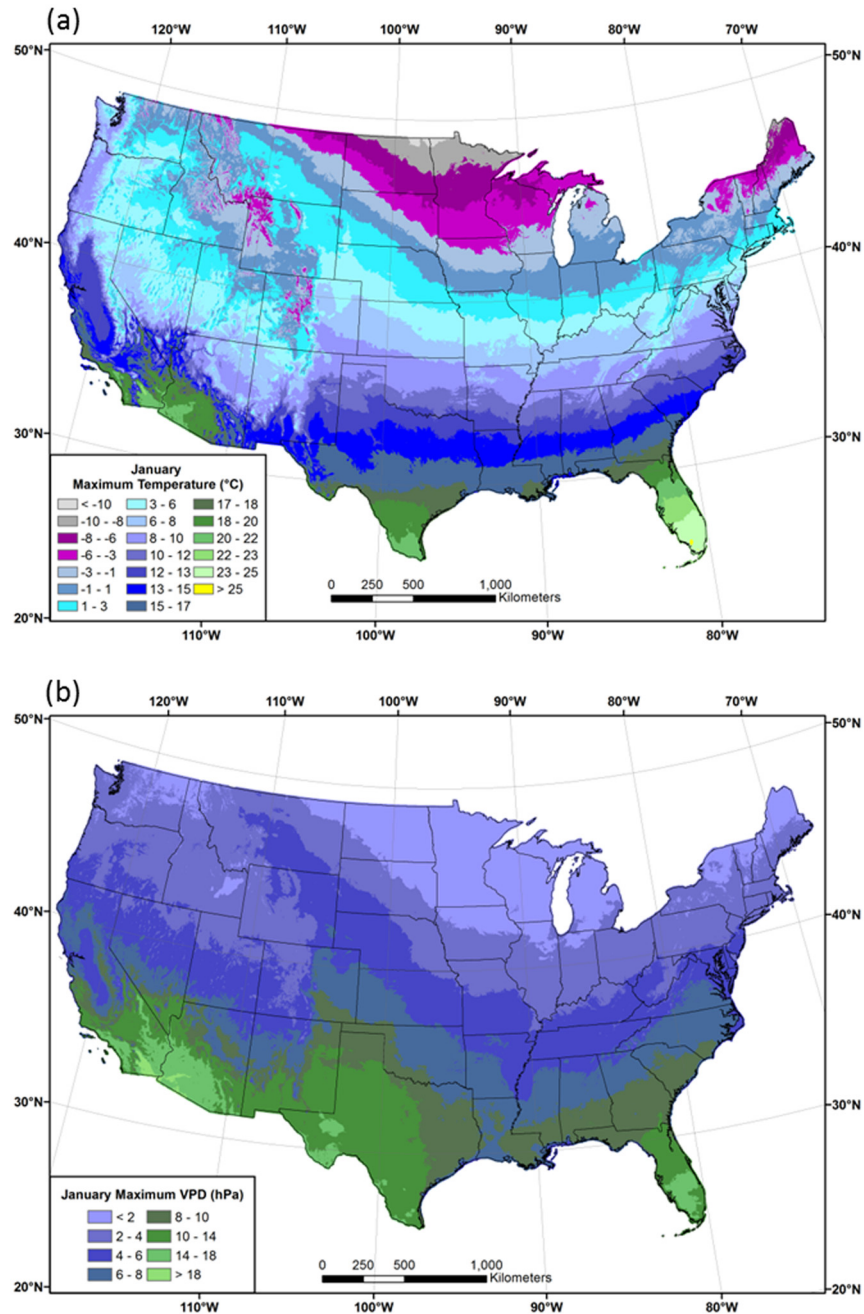
for nighttime cooling (high  $T_{min}$ ) during summer contribute to these high  $DPDs$ . Methods that estimate  $T_{dmean}$  by assuming it is equal to  $T_{min}$  would experience the largest errors in this region [22].

**Vapor Pressure Deficit.** Patterns of 1981–2010 mean  $VPD_{min}$  in January and July are shown in Fig 8. In winter,  $VPD_{min}$  is low (<1 hPa) over most of the country, due to a combination of low temperatures and relatively small differences between  $T_a$  and  $T_d$  during the



**Fig 8. Maps of mean minimum vapor pressure deficit ( $VPD_{min}$ ).** 1981–2010 mean monthly  $VPD_{min}$  for (a) January and (b) July.

doi:10.1371/journal.pone.0141140.g008

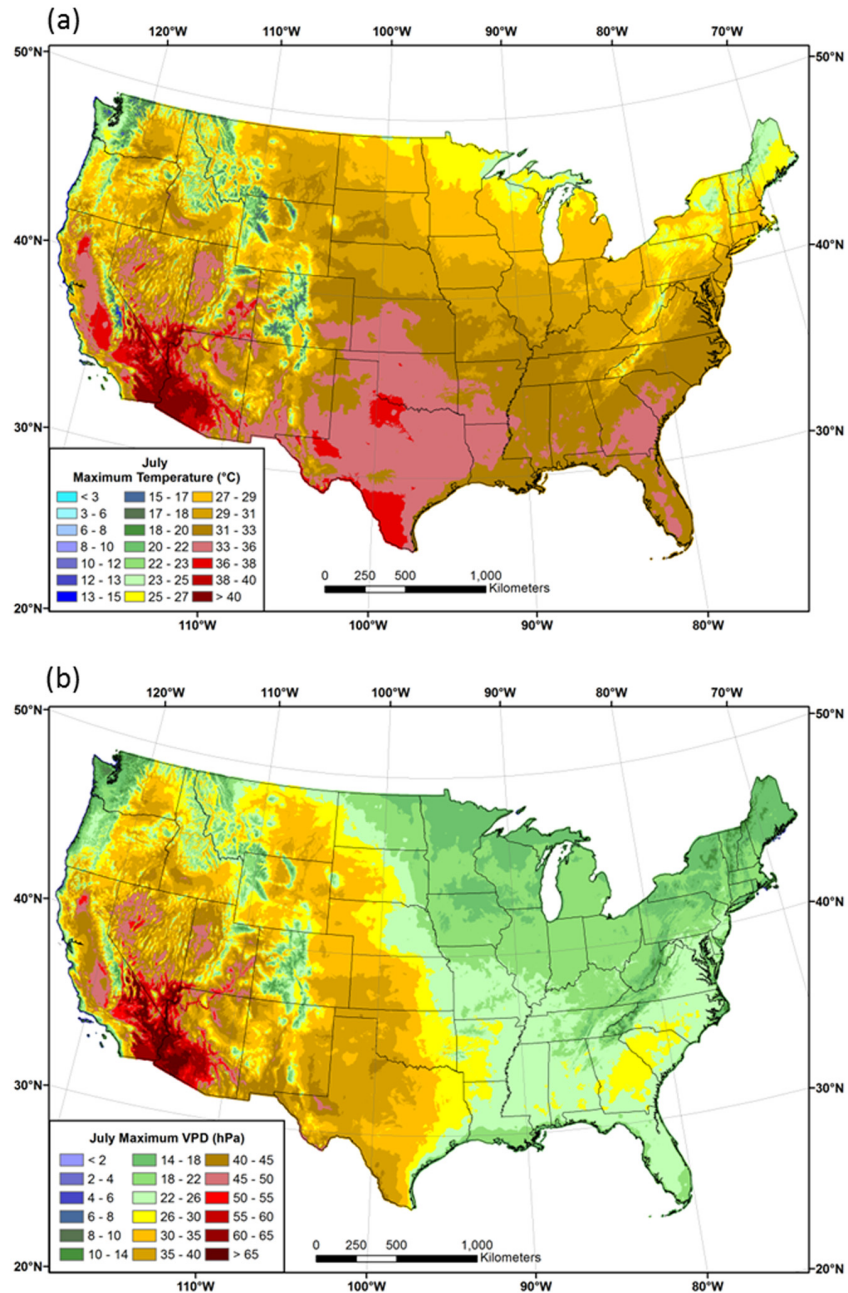


**Fig 9. Maps of January maximum temperature and vapor pressure deficit.** Conterminous US 1981–2010 mean January (a) maximum temperature ( $T_{max}$ ) and (b) maximum vapor pressure deficit ( $VPD_{max}$ ).

doi:10.1371/journal.pone.0141140.g009

morning hours (Fig 8a). Vapor pressure varies exponentially with temperature; for example, a difference between  $T_a$  and  $T_d$  of  $1^{\circ}\text{C}$  amounts to a VPD of 0.15 hPa at  $20^{\circ}\text{C}$ , but less than 0.05 hPa at  $0^{\circ}\text{C}$ . In July,  $VPD_{min}$  is also low in the eastern US, but exceeds 5 hPa over much of the West, reaching maxima of over 20 hPa in the desert southwest (Fig 8b).

Spatial patterns of 1981–2010 mean  $T_{max}$  and  $VPD_{max}$  in January and July are shown in Figs 9 and 10, respectively. Patterns of  $VPD_{max}$  roughly follow those of  $T_{max}$  in January, with the lowest values in the northern tier and western mountains, and the highest in the southern states



**Fig 10. Maps of July maximum temperature and vapor pressure deficit.** Conterminous US 1981–2010 mean July (a) maximum temperature ( $T_{max}$ ) and (b) maximum vapor pressure deficit ( $VPD_{max}$ ).

doi:10.1371/journal.pone.0141140.g010

(Fig 9). In July, the area of high  $VPD_{max}$  expands considerably, exceeding 25 hPa over much of the western US, and reaching a maximum of over 60 hPa in the desert southwest (Fig 10). Coastal regions of the West exhibit relatively low  $VPD_{max}$  values, as they did for  $DPD$ . Similarly low values are found at higher elevations, illustrating the strong spatial relationship between  $VPD_{max}$  and  $T_{max}$ . In the east,  $VPD_{max}$  values in the Midwest and northeast are mostly less than 20 hPa, while those in the southeast can range above 25 hPa; a notable maximum occurs in the piedmont region of Georgia and the Carolinas.

**Table 5. Results of temperature and vapor pressure deficit substitution in the derivation of minimum and maximum relative humidity.** 2003–2012 January and July averages from 100 randomly-selected ASOS stations. Actual and estimated *RH* values are expressed as a distribution: 5<sup>th</sup> percentile / mean / 95<sup>th</sup> percentile.

Variable	January	July
<i>T<sub>a</sub></i> at the hour of <i>RH<sub>min</sub></i> (°C)	4.7	29.8
<i>T<sub>max</sub></i> (°C)	6.2	30.2
<i>VPD</i> at the hour of <i>RH<sub>min</sub></i> (hPa)	5.7	26.4
<i>VPD<sub>max</sub></i> (hPa)	5.9	26.9
<i>RH<sub>min</sub></i> (%)	19 / 44 / 64	14 / 39 / 55
Estimated <i>RH<sub>min</sub></i> (%)	20 / 48 / 67	15 / 40 / 55
<i>T<sub>a</sub></i> at the hour of <i>RH<sub>max</sub></i> (°C)	-1.1	19.2
<i>T<sub>min</sub></i> (°C)	-4.2	18.4
<i>VPD</i> at the hour of <i>RH<sub>max</sub></i> (hPa)	0.8	3.4
<i>VPD<sub>min</sub></i> (hPa)	0.8	3.3
<i>RH<sub>max</sub></i> (%)	51 / 86 / 94	51 / 85 / 96
Estimated <i>RH<sub>max</sub></i> (%)	50 / 85 / 93	50 / 85 / 96

doi:10.1371/journal.pone.0141140.t005

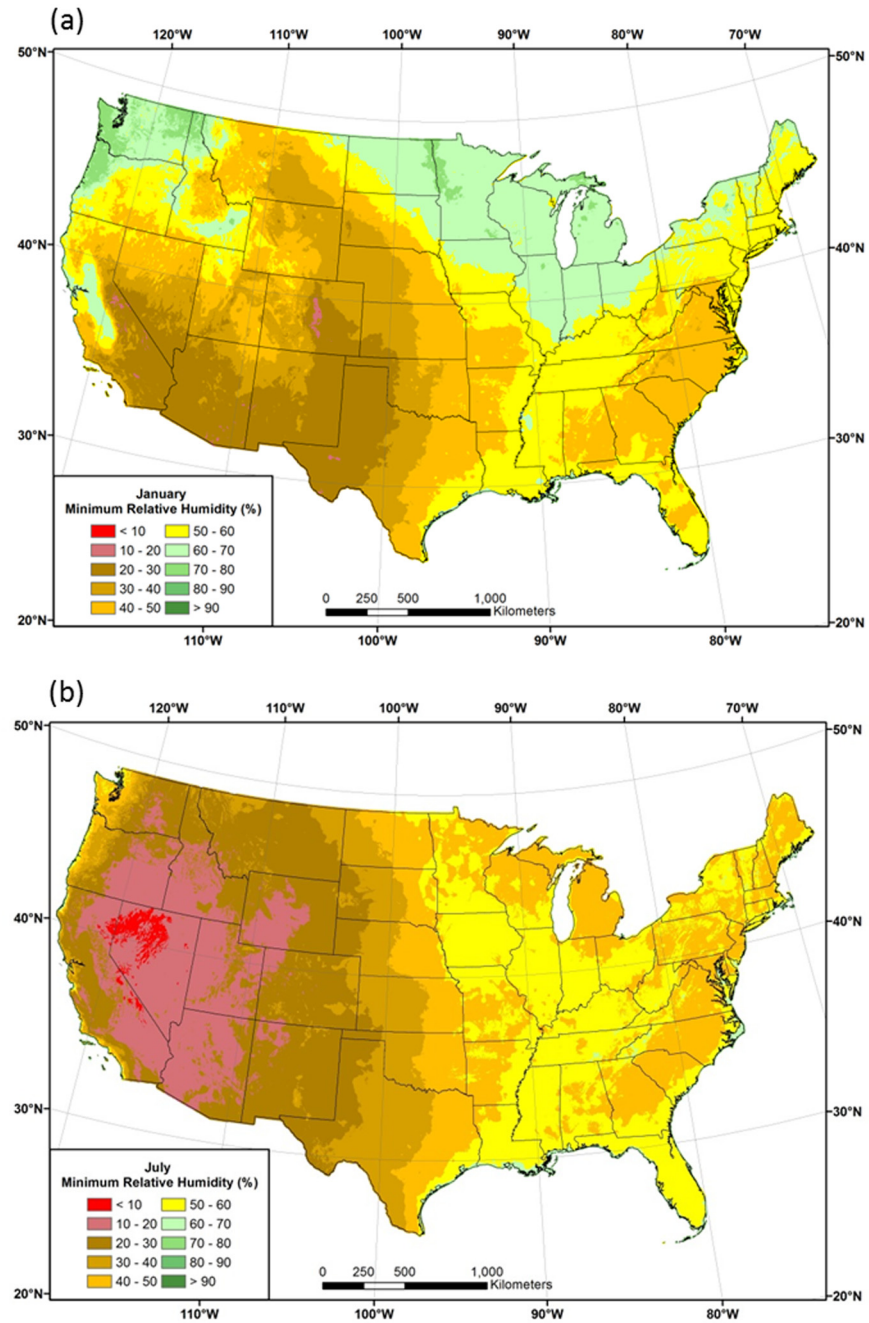
**Relative Humidity Derivation.** *RH* can be derived from *VPD* and *T<sub>a</sub>* by calculating  $SATVP_a$  at the desired temperature from Eq 1, and factoring in the appropriate *VPD*:

$$RH = 100(SATVP_a - VPD)/SATVP_a \tag{6}$$

On a gridded basis, if only *T<sub>max</sub>* and *T<sub>min</sub>* are available to estimate *T<sub>a</sub>*, and only *VPD<sub>min</sub>* and *VPD<sub>max</sub>* available to estimate *VPD*, grids of minimum *RH* (*RH<sub>min</sub>*) can be approximated by substituting *VPD<sub>max</sub>* for *VPD* and *T<sub>max</sub>* for *T<sub>a</sub>*, and maximum *RH* (*RH<sub>max</sub>*) can be similarly derived using *VPD<sub>min</sub>* and *T<sub>min</sub>*. In performing these derivations, we make the assumption that *T<sub>a</sub>* = *T<sub>min</sub>* and *VPD* = *VPD<sub>min</sub>* at the time of day when *RH<sub>max</sub>* occurs, and *T<sub>a</sub>* = *T<sub>max</sub>* and *VPD* = *VPD<sub>max</sub>* at the time of day when *RH<sub>min</sub>* occurs. To evaluate the error introduced by these assumptions, we calculated 2003–2012 mean monthly *RH<sub>min</sub>* and *RH<sub>max</sub>* for the same 100 ASOS stations used in Table 2 and compared them with the estimated *RH<sub>min</sub>* that would result from substituting *VPD<sub>max</sub>* for *VPD* and *T<sub>max</sub>* for *T<sub>a</sub>*, and *RH<sub>max</sub>* resulting from using *VPD<sub>min</sub>* and *T<sub>min</sub>*. As expected, *T<sub>max</sub>* and *VPD<sub>max</sub>* were slightly higher than *T<sub>a</sub>* and *VPD* at the hour of *RH<sub>max</sub>*, and *T<sub>min</sub>* and *VPD<sub>min</sub>* were slightly lower than *T<sub>a</sub>* and *VPD* at the hour of *RH<sub>min</sub>* (Table 5). As a result, the substitutions resulted in an overestimation of *RH<sub>min</sub>* and an underestimation of *RH<sub>max</sub>*. The differences averaged less than five percent in all months, suggesting that the derived grids are reasonable, but not exact, measures of the true *RH*.

Maps of estimated 1981–2010 mean January and July *RH<sub>min</sub>*, estimated from grid values of *VPD<sub>max</sub>* and *T<sub>max</sub>*, reveal some interesting features not easily seen in the maps of *VPD<sub>max</sub>* (Fig 11). In January, *RH<sub>min</sub>* values are as low as 10–15% on the lee side of the Rocky Mountains, associated with dry, Chinook (downslope) winds produced by a strong westerly jet stream. In contrast, *RH<sub>min</sub>* values exceed 70% in the winter-wet Pacific Northwest. *RH<sub>min</sub>* values are also above 70% in California’s Central Valley and Idaho’s Snake Plain, where extended periods of high pressure and calm winds promote temperature inversions that often result in persistent fog and low clouds (Fig 11a). Higher *RH<sub>min</sub>* values in the upper Midwest reflect frequent cloudy weather during winter in this region.

*RH<sub>min</sub>* patterns in July show a divided country, with a dry west and moist east (Fig 11b). The exception is the West Coast, where cool, moist onshore flow maintains relatively high



**Fig 11. Maps of minimum relative humidity.** Conterminous US 1981–2010 mean minimum relative humidity ( $RH_{min}$ ) estimated from mean maximum temperature ( $T_{max}$ ) and mean maximum vapor pressure deficit ( $VPD_{max}$ ) for (a) January and (b) July.

doi:10.1371/journal.pone.0141140.g011

humidity values throughout the day. The lowest  $RH_{min}$  (<10%) values are centered in Nevada, and more generally in the Great Basin. In the east, the highest values are found in regions that receive substantial moisture from the Gulf of Mexico, such as the southern Appalachians and Mississippi Valley.

## Uncertainty Analysis

Estimating the true errors associated with spatial climate data sets is difficult, and subject to its own set of errors [36]. This is because the true climate field is unknown, except at a relatively small number of observed points, and even these are subject to measurement and siting uncertainties (as has already been noted in the case of irrigated land, for example). Leave-one-out cross validation (C-V) approach is the most common evaluation method where each station is omitted from the dataset one at a time, the station value estimated in its absence, and the estimate and observation compared. The mean absolute error (MAE) and bias are typically calculated once the process is complete. While this approach is commonly used to assess error in interpolation studies, and is reported here, there are several disadvantages. An obvious disadvantage is that no error information is provided for locations where there are no stations. The single-deletion method favors interpolation models that heavily smooth the results, so that deletion of one station is relatively unimportant to the stability of the estimate. Randomly withholding a larger percentage of the station data at once can help to minimize this issue, as well as provide more robust error statistics. Withholding a stratified sample from the analysis is useful in detecting specific weaknesses or issues in the interpolation, as was done to investigate irrigated stations (also, see [36] for an example).

In this study, uncertainties in the mapped estimates were initially estimated by performing a C-V exercise with ASSAY, and the results compiled for each month for each of the three modeling regions (west, central, and east).

Even accounting for weaknesses in the C-V methodology, these errors are underestimates of the actual C-V uncertainty. One overlooked aspect of CAI is that it relies on predictor grids which have their own interpolation errors. These errors accumulate from one CAI generation to the next. For  $T_{dmean}$ , uncertainties in the interpolated  $T_{min}$  predictor grids must be accounted for. In turn,  $VPD_{min}$  relies on interpolated  $T_{min}$  and  $T_{dmean}$ , and  $VPD_{max}$  relies on  $T_{max}$ , as well as  $T_{min}$  and  $T_{dmean}$ .

To quantify the effects of error propagation on the CAI MAEs, the predictor grid interpolation error was introduced at each CAI step by using ASSAY interpolated estimates at station locations in their absence instead of the predictor grid values, which already have those station values built in. For  $T_{dmean}$ ,  $T_{min}$  values for all stations used in the mapping of the  $T_{min}$  predictor grid were estimated in their absence using ASSAY, and estimates common to both the  $T_{min}$  and  $T_{dmean}$  station datasets used as the values of  $T_{min}$  in an ASSAY C-V exercise for  $T_{dmean}$ .  $T_{min}$  values for stations used in the interpolation of  $T_{dmean}$  but not  $T_{min}$ , such as those from the SCAN and COAGMET networks, could not be estimated by ASSAY, because they were not used to create the predictor grid. However, they could still be included in the  $T_{dmean}$  C-V error estimation because, by definition, the estimates at their location on the  $T_{min}$  predictor grid were made in their absence.

For  $VPD_{min}$ , ASSAY estimates of station  $T_{dmean}$  values, predicted using ASSAY estimates of  $T_{min}$  as described above, were used in combination with ASSAY estimates of  $T_{min}$  to form station values of first-guess  $VPD_{min}$ , used as the predictor in the interpolation of  $VPD_{min}$ . The result was station values of first-guess  $VPD_{min}$  that accounted for errors in the interpolation of both  $T_{min}$  and  $T_{dmean}$ . These were used as the values of first-guess  $VPD_{min}$  in a C-V assessment of  $VPD_{min}$  with ASSAY.  $VPD_{max}$  involved the same steps as  $VPD_{min}$ , except that first-guess  $VPD_{max}$  values were formed from a combination of ASSAY estimates of  $T_{max}$  and  $T_{dmean}$ . Again, a C-V exercise for  $VPD_{max}$  was performed with ASSAY.

Table 6 reports monthly ASSAY cross-validation MAEs for each of the modeling regions. MAEs that account for CAI interpolation error propagation are denoted with a CAI or CAI2. Figs 12, 13 and 14 show the distribution of cross-validation absolute errors at stations across the country for  $T_{dmean}$ ,  $VPD_{min}$ , and  $VPD_{max}$ , respectively.

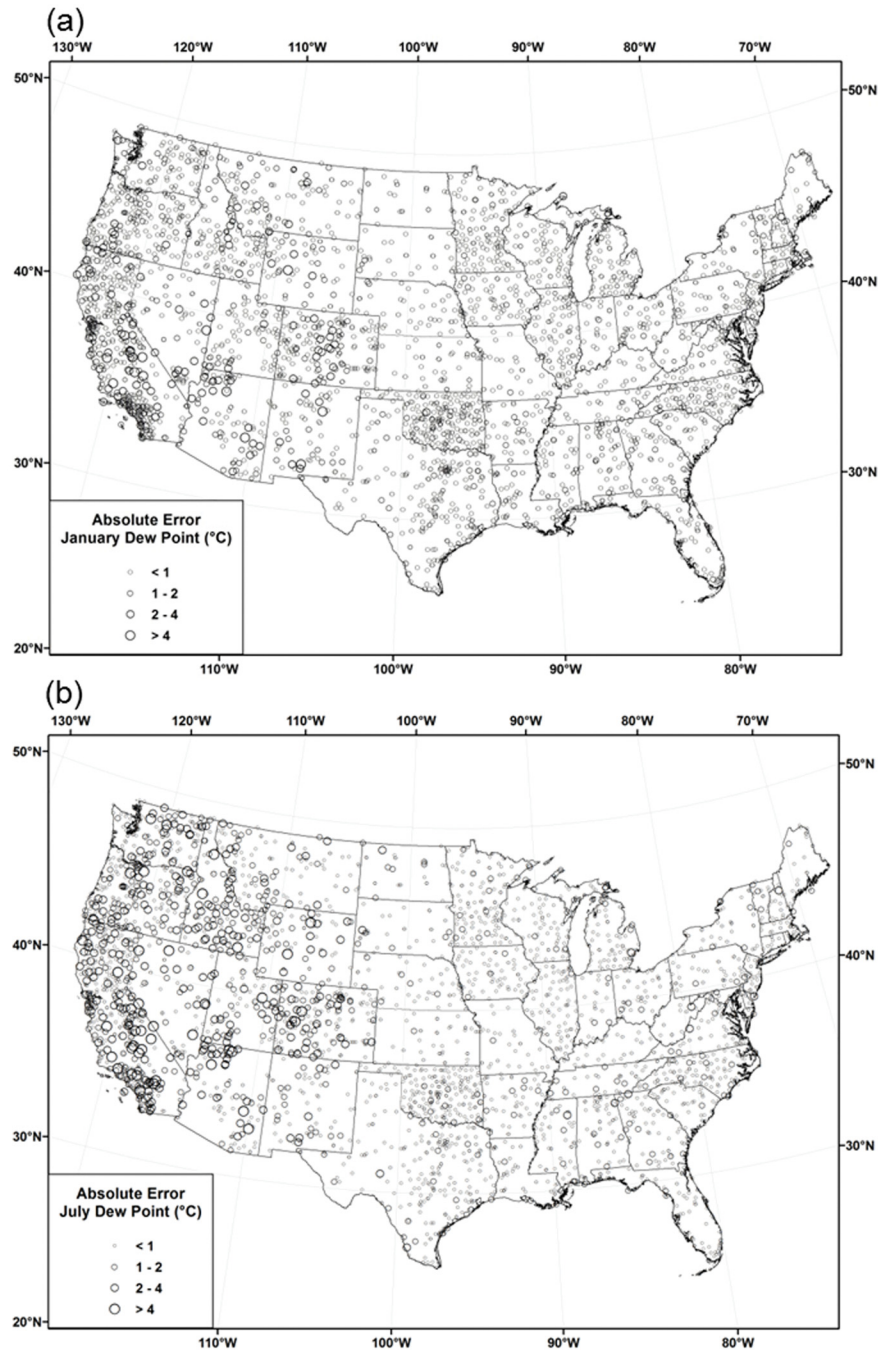


**Table 6. PRISM monthly cross-validation mean absolute errors (MAEs).** Both interpolation-only MAEs, and MAEs accounting for uncertainties in the predictor grids (denoted by CAI or CAI2), are given below. See text for details.

Region/Variable	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<b>West</b>												
$T_{max}$ (°C)	0.65	0.63	0.63	0.64	0.64	0.66	0.69	0.68	0.66	0.62	0.57	0.61
$T_{min}$ (°C)	0.98	0.93	0.85	0.82	0.87	0.97	1.08	1.09	1.11	1.04	0.95	0.96
$T_{dmean}$ (°C)	0.69	0.65	0.66	0.69	0.76	0.84	0.92	0.91	0.87	0.75	0.70	0.69
$T_{dmean}$ CAI (°C)	0.76	0.75	0.77	0.79	0.88	1.00	1.12	1.10	1.03	0.86	0.80	0.79
$VPD_{min}$ (hPa)	0.24	0.19	0.23	0.29	0.41	0.60	0.79	0.76	0.61	0.41	0.27	0.19
$VPD_{min}$ CAI2 (hPa)	0.31	0.28	0.36	0.45	0.65	0.89	1.25	1.17	0.99	0.64	0.41	0.28
$VPD_{min}$ (%)	41.9	31.2	32.2	31.0	28.5	30.1	30.5	32.1	32.2	29.2	34.0	33.7
$VPD_{min}$ CAI2 (%)	58.5	45.2	50.0	45.8	42.9	45.8	48.3	48.7	52.0	44.7	51.3	49.7
$VPD_{max}$ (hPa)	0.50	0.49	0.59	0.69	0.89	1.21	1.47	1.45	1.17	0.76	0.52	0.43
$VPD_{max}$ CAI2 (hPa)	0.57	0.58	0.70	0.82	1.10	1.46	1.86	1.81	1.43	0.97	0.63	0.52
$VPD_{max}$ (%)	11.1	8.7	7.4	6.4	6.1	6.3	6.0	5.9	5.8	5.8	7.6	10.1
$VPD_{max}$ CAI2 (%)	11.8	10.2	8.4	7.0	6.8	6.8	6.9	6.9	6.5	6.7	8.5	11.4
<b>Central</b>												
$T_{max}$ (°C)	0.44	0.46	0.48	0.48	0.46	0.44	0.44	0.43	0.42	0.41	0.38	0.40
$T_{min}$ (°C)	0.65	0.63	0.55	0.55	0.53	0.53	0.55	0.58	0.63	0.67	0.64	0.63
$T_{dmean}$ (°C)	0.45	0.42	0.40	0.42	0.42	0.45	0.45	0.44	0.43	0.41	0.41	0.42
$T_{dmean}$ CAI (°C)	0.47	0.44	0.43	0.45	0.45	0.48	0.48	0.48	0.46	0.43	0.43	0.44
$VPD_{min}$ (hPa)	0.11	0.12	0.15	0.20	0.28	0.37	0.45	0.43	0.34	0.23	0.16	0.11
$VPD_{min}$ CAI2 (hPa)	0.13	0.15	0.19	0.27	0.37	0.49	0.60	0.56	0.46	0.32	0.21	0.15
$VPD_{min}$ (%)	22.4	22.0	20.3	19.3	22.8	23.6	29.6	33.0	31.3	28.4	23.3	23.2
$VPD_{min}$ CAI2 (%)	28.5	28.8	26.4	26.5	31.9	33.4	42.0	47.0	44.0	38.7	31.2	29.9
$VPD_{max}$ (hPa)	0.24	0.28	0.37	0.50	0.64	0.75	0.90	0.86	0.67	0.42	0.30	0.24
$VPD_{max}$ CAI2 (hPa)	0.29	0.34	0.43	0.62	0.79	0.96	1.13	1.07	0.84	0.55	0.37	0.30
$VPD_{max}$ (%)	5.3	5.1	4.2	3.8	3.9	3.6	3.8	3.7	3.5	3.2	3.7	5.1
$VPD_{max}$ CAI2 (%)	6.3	6.1	5.0	4.7	4.7	4.6	4.7	4.5	4.4	4.1	4.5	6.1
<b>East</b>												
$T_{max}$ (°C)	0.40	0.46	0.49	0.51	0.50	0.49	0.48	0.47	0.43	0.42	0.37	0.37
$T_{min}$ (°C)	0.62	0.64	0.61	0.65	0.65	0.60	0.60	0.63	0.69	0.71	0.64	0.61
$T_{dmean}$ (°C)	0.46	0.42	0.40	0.40	0.40	0.40	0.38	0.38	0.40	0.39	0.39	0.39
$T_{dmean}$ CAI (°C)	0.46	0.43	0.42	0.43	0.42	0.42	0.41	0.41	0.44	0.42	0.41	0.40
$VPD_{min}$ (hPa)	0.10	0.12	0.19	0.25	0.31	0.36	0.38	0.38	0.33	0.24	0.18	0.13
$VPD_{min}$ CAI2 (hPa)	0.13	0.15	0.23	0.32	0.39	0.48	0.50	0.48	0.41	0.30	0.22	0.15
$VPD_{min}$ (%)	25.0	26.5	28.7	29.0	35.1	40.1	42.5	47.5	54.0	43.4	32.3	29.3
$VPD_{min}$ CAI2 (%)	33.2	33.9	36.7	40.1	51.9	63.7	77.8	78.9	78.9	64.4	42.8	38.1
$VPD_{max}$ (hPa)	0.25	0.28	0.50	0.59	0.75	0.93	1.01	0.91	0.73	0.48	0.34	0.26
$VPD_{max}$ CAI2 (hPa)	0.28	0.32	0.45	0.69	0.89	1.07	1.23	1.12	0.89	0.59	0.40	0.30
$VPD_{max}$ (%)	6.1	5.4	5.1	4.9	5.1	5.2	5.4	5.1	4.7	4.2	4.6	5.8
$VPD_{max}$ CAI2 (%)	6.6	5.7	5.5	5.6	5.9	6.1	6.6	6.3	5.9	5.1	5.2	6.3

doi:10.1371/journal.pone.0141140.t006

**Dew Point.** In general,  $T_{dmean}$  interpolation errors in the west were greater than those in the central and eastern US, due primarily to terrain-induced complexities in the vertical distribution of moisture and temperature. This is evidenced by higher absolute errors in the Rocky Mountains, Cascades, and Sierra Nevada (Fig 12). Regionally, MAEs for  $T_{dmean}$  interpolated as



**Fig 12. Maps of  $T_{dmean}$  absolute cross-validation errors.** Errors for 1981–2010 mean monthly  $T_{dmean}$  in (a) January and (b) July.

doi:10.1371/journal.pone.0141140.g012

$DPD (T_{min} - T_{dmean})$ , were less than  $1^{\circ}\text{C}$  in the west and less than  $0.5^{\circ}\text{C}$  in the central and east (Table 6). CAI MAEs that accounted for error propagation were  $0.1\text{--}0.2^{\circ}\text{C}$  larger than those that did not in the west, but less than  $0.05^{\circ}\text{C}$  larger in the central and east, owing to larger  $T_{min}$  interpolation errors in the west.

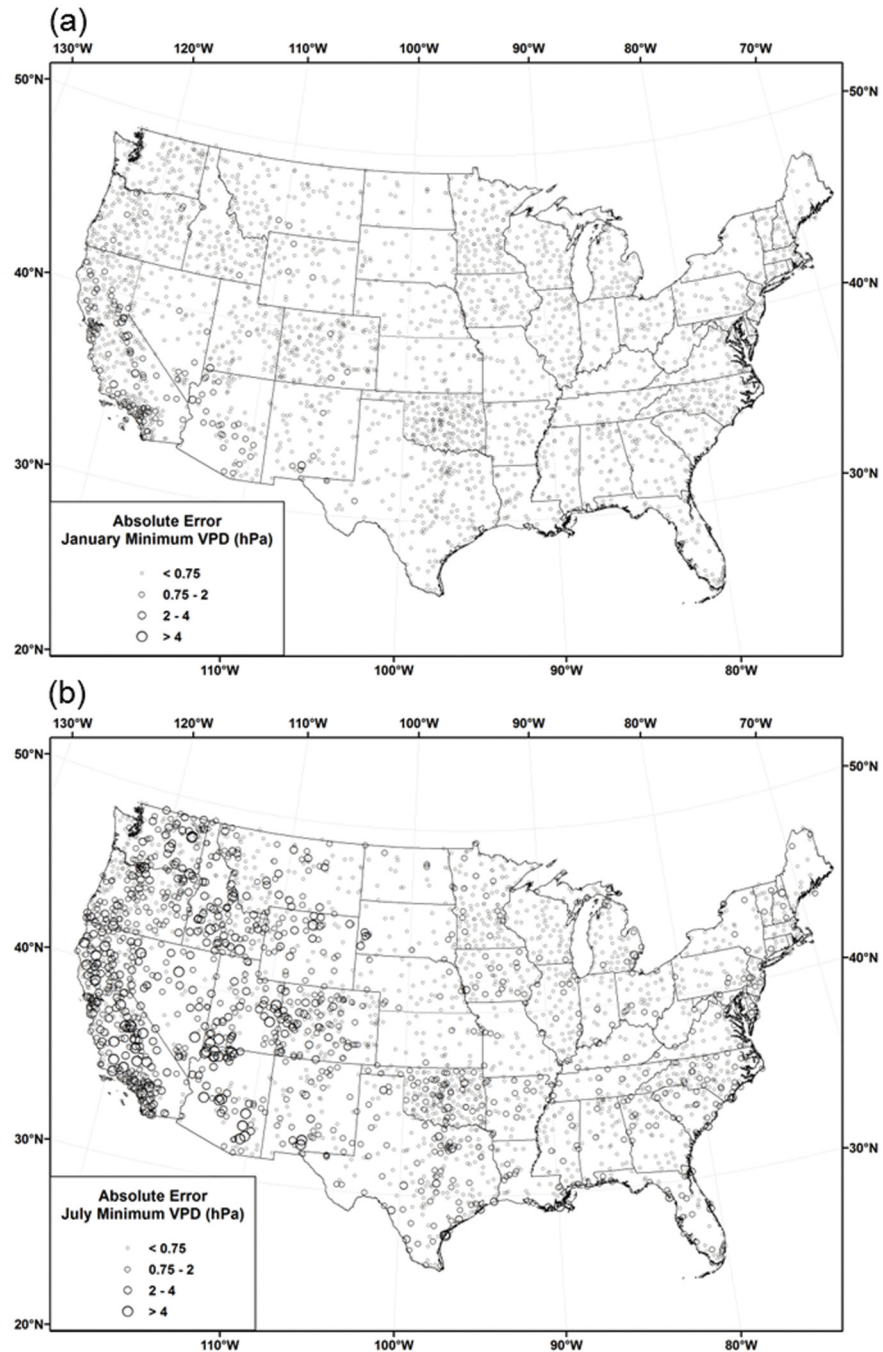
The  $T_{dmean}$  C-V analysis showed that the MAEs were inflated by systematic positive biases (lower observed  $DPD$  and higher  $T_{dmean}$  than predicted) in the COAGMET and AGRIMET

networks. In fact, bias accounted for nearly all of the MAE. As discussed previously, stations in these networks were typically sited in or near irrigated fields for use in water management calculations. Despite subjective omission of stations producing the largest spatial inconsistencies, network-wide biases were still noticeable. Fig 15a and 15b show monthly  $T_{dmean}$  MAE and bias for each of the major station networks when each network is entirely eliminated from the dataset at one time. Plots are for the central region, to better control for the effects of complex physiographic features on interpolation bias; the exception is AGRIMET, which operates in the western region only, but stations are in flat agricultural areas, so interpolated predictions should be otherwise relatively unbiased. The peak MAE and bias in mid-summer correspond to the maximum difference in atmospheric moisture content one would expect between irrigated and non-irrigated land. At their summer peaks, the COAGMET MAE and bias were 1.7 and 1.6°C, respectively. Peak AGRIMET MAE and bias were 1.5 and 1.4°C, respectively. These are in contrast to RAWS, for which the peak MAE was 1.4°C but the bias was only 0.1°C in the central region. The MAEs for these networks were much larger than the overall MAE for the central region of less than 0.5°C (Table 6).

A second source of systematic bias was found regionally in RAWS stations (Fig 15c). In the western region, biases were negative (higher observed  $DPD$  and lower  $T_{dmean}$  than predicted). In this region, the culprits may have been both local land cover and location. RAWS stations in the West are used primarily for fire weather applications, and are typically sited in open, ventilated areas away from transpiring vegetation. In addition, many RAWS stations are located on exposed terrain above locally humid cold air pools, and in foothill locations, which can be above larger-scale valley inversions, where a dry, free atmosphere aloft overlies relatively moist air below. Spring is the time of year when good vertical mixing causes a reduction in the incidence of inversions and cold air pools, which may explain the minimal bias at this time of year. The case for siting and location as causes for bias appears to be supported in the central region, where RAWS biases are near zero (Fig 15c). It is not clear why biases become somewhat positive (lower observed  $DPD$  and higher  $T_{dmean}$  than predicted) in the eastern region. One explanation may be that RAWS stations, often located in heavily forested areas in the east, are sampling more locally humid environments than other networks located at airports and in developed areas. The seasonal maximum bias in summer would correspond with the time of maximum transpiration from vegetation.

**Vapor Pressure Deficit.** MAEs for  $VPD_{min}$  and  $VPD_{max}$  were calculated regionally as percentages and in absolute hPa units (Table 6) and absolute errors mapped across the country (Figs 13 and 14). MAEs in absolute units were typically larger in summer than in winter, in keeping with characteristically larger vapor pressure deficits in summer due to higher temperatures (Table 6). Regionally,  $VPD_{min}$  MAEs were less than 1 hPa in all months and regions, owing to relatively small absolute values. In January, the highest absolute  $VPD_{min}$  errors were concentrated in the desert southwest, where moisture deficits are greatest (Fig 13). These higher errors spread to most of the west in July, when deficits are seasonally high. Regional MAEs ranged from roughly 30–40 percent in the west, 20–30 percent in central, and 25–50 percent in the east. CAI2 MAEs that accounted for error propagation were 0.1–0.5 hPa and 15–20% larger in the west, 0.05–0.15 hPa and 5–15% larger in the central, and 0.05–0.15 hPa and 10–25% larger in the east.

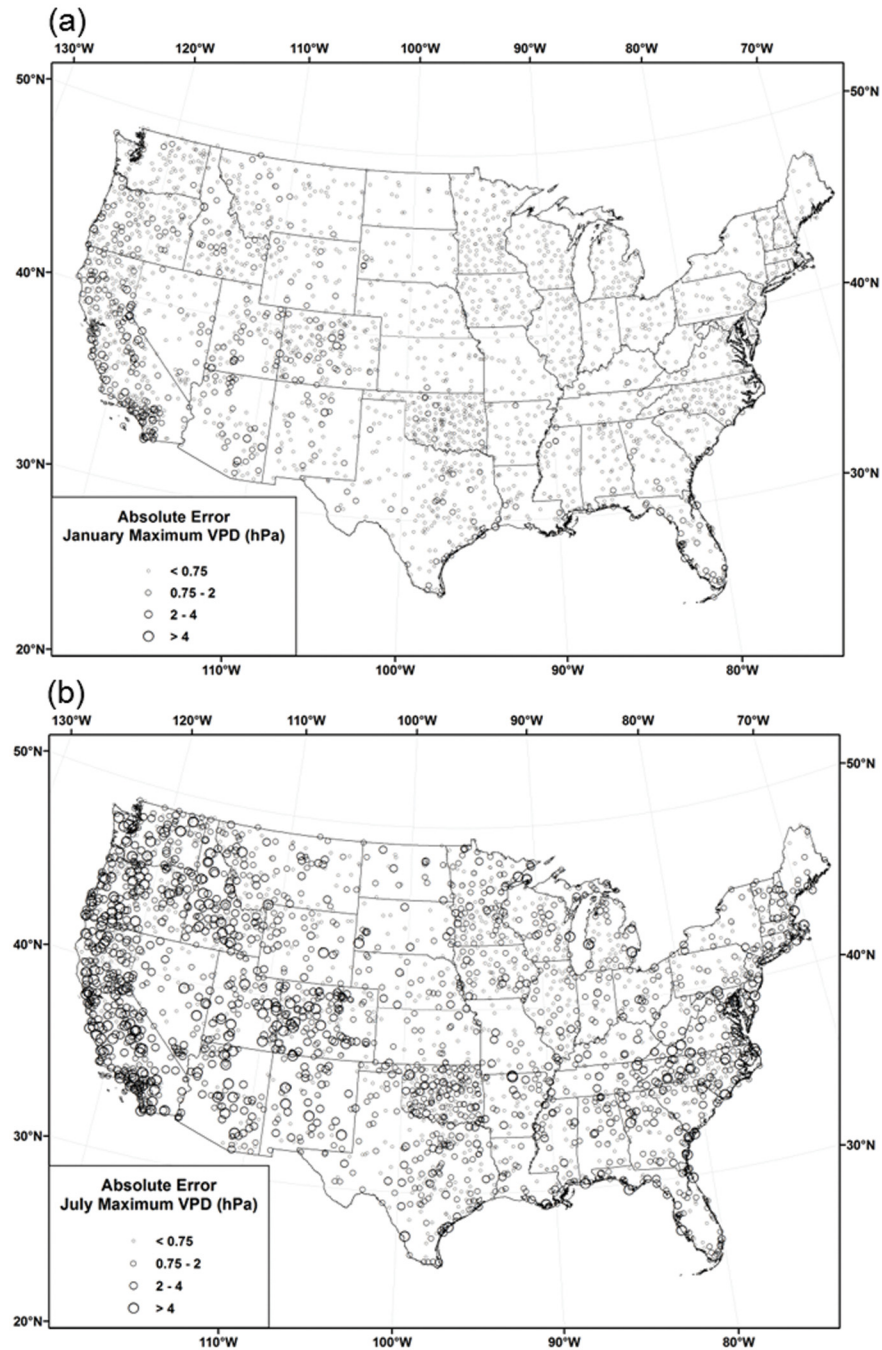
Regional  $VPD_{max}$  absolute MAEs were approximately double those for  $VPD_{min}$  in absolute units. Percentage MAEs were accordingly much lower, ranging from roughly 6–11, 3–5, and 4–6 percent in the west, central, and east, respectively. CAI2 MAEs that accounted for error propagation were 0.1–0.5 hPa and 0.5–1.5% larger in the west, and 0.05–0.25 hPa and 0.5–1.5% larger in the central and east. MAEs were again inflated by systematic positive biases (lower observed  $VPD$  observations than predicted) from COAGMET and AGRIMET. The



**Fig 13. Maps of  $VPD_{min}$  absolute cross-validation errors.** Errors for 1981–2010 mean monthly  $VPD_{min}$  in (a) January and (b) July.

doi:10.1371/journal.pone.0141140.g013

spatial and temporal distributions of absolute errors were similar to that of  $VPD_{min}$ , with the highest January errors in the west, spreading to the east in July (Fig 14). July  $VPD_{max}$  errors were higher in the eastern US than those of  $VPD_{min}$ , because warm maximum temperatures in this region produce significant daytime vapor pressure deficits, despite relatively high  $RH$  values.

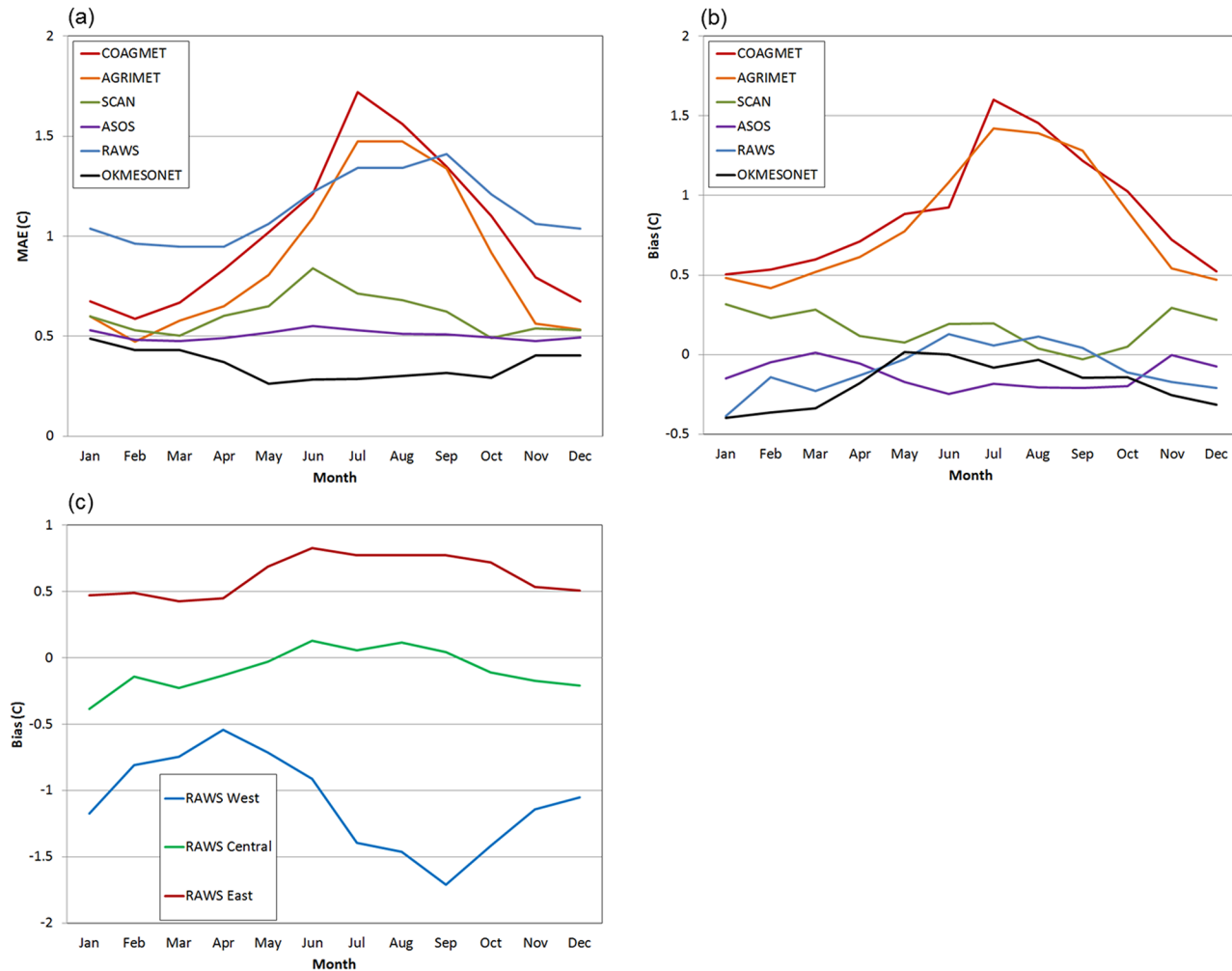


**Fig 14. Maps of  $VPD_{max}$  absolute cross-validation errors.** Errors for 1981–2010 mean monthly  $VPD_{max}$  in (a) January and (b) July.

doi:10.1371/journal.pone.0141140.g014

## Summary and Conclusions

Long-term normal grids of 1981–2010 mean monthly average daily  $T_{dmean}$ ,  $VPD_{min}$  and  $VPD_{max}$  were developed for the conterminous United States. The  $T_{dmean}$  grids update previous unpublished  $T_{dmean}$  normals used as CAI predictor grids in PRISM monthly time series



**Fig 15. Network cross-validation MAE and bias of mean dew point depression (DPD).** By month, station network, and region, when each network is entirely eliminated from the dataset at one time: (a) MAE for each network in the central region, except for AGRIMET, which is in the western region; (b), same as panel a, except showing bias; and (c) RAWS network bias for each region. The number of stations in each network is given in Table 1.

doi:10.1371/journal.pone.0141140.g015

datasets, and to our knowledge, the  $VPD_{min}$  and  $VPD_{max}$  normal grids are the first of their kind. Interpolation of the long-term monthly averages was performed using PRISM. Nearby stations entering the PRISM local regression functions (one per pixel) were assigned weights based on the physiographic similarity of the station to the grid cell that included the effects of distance, elevation, coastal proximity, vertical atmospheric layer, and topographic position. Relatively few stations were available for these variables, prompting us to use CAI to improve interpolation accuracy. In the CAI process, 1981–2010 monthly  $T_{min}$  grids served as predictor grids for the interpolation of  $T_{dmean}$ , expressed as the dew point depression ( $DPD = T_{min} - T_{dmean}$ ). Second-generation CAI (CAI2) involving  $T_{dmean}$  and  $T_{min}$ , and  $T_{dmean}$  and  $T_{max}$  predictor grids were used to interpolate  $VPD_{min}$  and  $VPD_{max}$ , respectively.

The general patterns of  $T_{dmean}$  were similar to those of  $T_{min}$  in both winter and summer. However, the assumption that  $T_{dmean}$  is equal to  $T_{min}$  did not hold over large parts of the country. In winter,  $T_{dmean}$  was several degrees higher than  $T_{min}$  in the northern US and in cold valley locations of the West. In summer,  $T_{dmean}$  averaged 10–25°C lower than  $T_{min}$  in the desert

southwest.  $VPD_{min}$  was very low over most of the country, due to a combination of low temperatures and relatively small differences between ambient and dew point temperatures in the morning hours. Patterns of  $VPD_{max}$  roughly followed those of  $T_{max}$  in winter, with the lowest values in the northern tier and western mountains, and the highest in the southern states. In summer, the area of high  $VPD_{max}$  expanded considerably, reaching a maximum of over 60 hPa in the desert southwest. Coastal regions of the West exhibited relatively low  $VPD_{max}$  values, where cool, marine air penetrates inland.

A PRISM interpolation uncertainty analysis was performed using C-V exercises. Since CAI relies on predictor grids which have their own interpolation errors, these errors were included in the overall error estimates. To quantify the effects of error propagation on the CAI MAEs, the predictor grid interpolation error was introduced at each step by using ASSAY predictions at station locations instead of the interpolated predictor grid values. When accounting for error propagation, MAEs for  $T_{dmean}$  were 0.8–1.1°C in the west modeling region and less than 0.5°C in the central and east, with the greatest errors typically occurring in summer.  $VPD_{min}$  MAEs were generally less than 1 hPa in all months and regions, and  $VPD_{max}$  MAEs about double those values.

Overall, accounting for CAI error propagation increased C-V errors, but not as much as expected, especially when compared to the initial MAEs associated with the  $T_{max}$  and  $T_{min}$  predictor values. One reason for this appears to be that the errors were not systematically biased in one direction; that is, overestimates and underestimates partly canceled each other, producing lower net error increases.

Some stations in the AGRIMET and COAGMET networks were found to have consistently high  $T_{dmean}$  and low  $VPD$  values during the summer months. Stations in these networks were typically sited in or near irrigated fields for use in water management calculations, resulting in more humid conditions than locations distant from irrigated areas. This raised questions about whether the grids should represent conditions over irrigated land. A middle ground was taken, where a few stations causing the most severe spatial discrepancies were omitted from the datasets. In addition, the siting of RAWS stations was posited as the reason for lower than expected humidity in the west and higher than expected in the east. These issues highlight the effect that local topographic position and land use/land cover properties can have on the spatial patterns of atmospheric moisture content and deficit.

In combination with existing PRISM grids of  $T_{min}$  and  $T_{max}$ , grids of  $T_{dmean}$ ,  $VPD_{min}$  and  $VPD_{max}$  allow the user to derive many other atmospheric moisture variables, such as minimum and maximum  $RH$ , vapor pressure, and  $DPD$ . Accompanying assumptions may need to be made, however, such as those outlined in the derivation of  $RH_{min}$  and  $RH_{max}$ . These new grids will serve as the predictor grids in second-generation CAI to produce an updated version of the PRISM  $T_{dmean}$  monthly time series dataset that covers the years 1895–present [9], and initial versions of monthly  $VPD_{min}$  and  $VPD_{max}$  time series grids for the same period using third-generation CAI. Similar methods will be used to produce initial versions of daily  $T_{dmean}$ ,  $VPD_{min}$  and  $VPD_{max}$  time series datasets, spanning 1981–present. All of the PRISM normal grids discussed here are available online at <http://prism.oregonstate.edu>, and include 800-m and 4-km resolution data, images, metadata, pedigree information, and station inventory files. Links to the papers cited in Table 3 and elsewhere are also available from this website.

## Acknowledgments

We thank Mike Halbleib for creating the cartographic maps, and Eileen Kaspar for editing support. We also thank two anonymous reviewers whose comments resulted in a substantially

improved manuscript. We are grateful to all of the station data providers, without whom this mapping effort would not have been possible.

## Author Contributions

Conceived and designed the experiments: CD. Performed the experiments: CD JIS KVO. Analyzed the data: CD JIS KVO. Contributed reagents/materials/analysis tools: CD JIS KVO. Wrote the paper: CD JIS KVO.

## References

1. Maurer EP, Wood AW, Adam JC, Lettenmaier DP, Nijssen B. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *J. Climate*. 2002; 15: 3237–3251.
2. Daly C, Gibson WP, Doggett M, Smith J, Taylor G. Up-to-date monthly climate maps for the conterminous United States. Proc., 14th AMS Conf. on Applied Climatology, 84th AMS Annual Meeting Combined Preprints, Amer. Meteorological Soc., Seattle, WA, January 13–16, 2004, Paper P5.1. Available: <http://ams.confex.com/ams/pdfpapers/71444.pdf>
3. Hamlet AF, Lettenmaier DP. Production of temporally consistent gridded precipitation and temperature fields for the continental U.S. *Journal of Hydrometeorology*. 2005; 6: 330–336.
4. Livneh B, Rosenburg EA, Lin C, Nijssen B, Mishra V, Andreadis KM, et al. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *J. Climate*. 2013; 26, 9384–9392. doi: [10.1175/JCLI-D-12-00508.1](https://doi.org/10.1175/JCLI-D-12-00508.1)
5. Lundquist JD, Hughes M, Henn B, Gutmann ED, Livneh B, Dozier J, et al. High-elevation precipitation patterns: using snow measurements to assess daily gridded datasets across the Sierra Nevada, California. *Journal of Hydrometeorology*. 2015. doi: [10.1175/JHM-D-15-0019.1](https://doi.org/10.1175/JHM-D-15-0019.1)
6. Daly C, Neilson RP, Phillips DL. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.* 1994; 33: 140–158.
7. Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P. A knowledge-based approach to the statistical mapping of climate. *Climate Research*. 2002; 22: 99–113.
8. Daly C, Halbleib MD, Smith JI, Gibson WP, Doggett MK, Taylor GH, et al. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology*. 2008; 28: 2031–2064.
9. PRISM Climate Group, Oregon State University [internet]. Descriptions of PRISM Spatial Climate Datasets for the Conterminous United States. Available: [http://www.prism.oregonstate.edu/documents/PRISM\\_datasets.pdf](http://www.prism.oregonstate.edu/documents/PRISM_datasets.pdf). Accessed 25 September 2015.
10. Hinckley TM, Lassoie JP, Running SW. Temporal and spatial variations in the water status of trees. *Forest Science Monographs*. 1978; 20: 72 pp.
11. Running SW. Environmental and physiological control of water flux through *Pinus contorta*. *Can. J. For. Res.* 1980; 10: 82–91.
12. Johnson JD, Ferrell WK. Stomatal response to vapour pressure deficit and the effect of water stress. *Plant, Cell and Environ.* 1983; 6: 451–456.
13. Bunce JA. Does transpiration control stomatal responses to water vapour pressure deficit? *Plant, Cell and Environ.* 1997; 20:131–135.
14. Zeppel MJB, Lewis JD, Phillips NG, Tissue DT. Consequences of nocturnal water loss: a synthesis of regulating factors and implications for capacitance, embolism and use in models. *Tree Physiol.* 2014; 34: 1047–1055. doi: [10.1093/treephys/tpu089](https://doi.org/10.1093/treephys/tpu089) PMID: [25413023](https://pubmed.ncbi.nlm.nih.gov/25413023/)
15. Littell J.S., McKenzie D, Peterson DL, Westerling AL. Climate and wildfire area burned in Western US ecoprovinces, 1916–2003. *Ecological Applications*. 2009; 19:1003–1021.
16. Williams A, Seager R, Berkelhammer M, Macalady A, Crimmins M, Swetnam T, et al. Causes and implications of extreme atmospheric moisture demand during the record-breaking 2011 wildfire season in the southwest United States. *J. Appl. Meteor. Climatol.* 2009; doi: [10.1175/JAMC-D-14-0053.1](https://doi.org/10.1175/JAMC-D-14-0053.1), in press.
17. Littell JS, Oneil EE, McKenzie D, Hicke JA, Lutz JA, Norheim RA, et al. Forest ecosystems, disturbance, and climate changes in Washington state, USA. *Climatic Change*. 2010; 102: 120–158.
18. Ashcroft BA, Gollan JR. Moisture, thermal inertia, and the spatial distributions of near-surface soil and air temperatures: Understanding factors that promote microrefugia. *Agricultural and Forest Meteorology*. 2013; 176: 77–89.



19. Running SW, Nemani RR, Hungerford RD. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal of Forest Research*. 1987; 17: 472–483.
20. Kimball J, Running SW, Nemani RR. An improved method for estimating surface humidity from daily minimum temperature. *Agricultural and Forest Meteorology*. 1997; 85: 87–98.
21. Thornton PE, Running SW, White MA. Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*. 1997; 190: 214–251.
22. McEvoy DJ, Meija JF, Huntington JL. Use of an observation network in the Great Basin to evaluate gridded climate data. *J. Hydrometeorology*. 2014; 15: 1913–1931.
23. Abatzoglou JT. Development of gridded surface meteorological data for ecological applications and modeling. *International Journal of Climatology*. 2013; 33: 121–131. doi: [10.1002/joc.3413](https://doi.org/10.1002/joc.3413)
24. Anderson D B. Relative humidity or vapor pressure deficit. *Ecology*. 1936; 17: 277–282.
25. New M, Hulme M, Jones PD. Representing twentieth century space-time climate variability. Part II: Development of 1901–1996 monthly grids of terrestrial surface climate. *J Climate*. 2000; 13: 2217–2238.
26. Alduchov OA, Eskridge RE. Improved Magnus form approximation of saturation vapor pressure. *J. Appl. Meteor*. 1996; 35: 601–609.
27. National Oceanic and Atmospheric Administration [Internet]. Wet-bulb temperature and dewpoint temperature from air temperature and relative humidity. Available: <http://www.srh.noaa.gov/images/epz/wxcalc/wetBulbTdFromRh.pdf>. Cited 25 September 2015.
28. World Meteorological Organization. Calculation of monthly and annual 30-year standard normals. Prepared by a meeting of experts, Washington, DC, USA, March 1989. WCDP 10, WMO-TD 341.
29. NOAA-National Climatic Data Center [Internet]. Data documentation for data set 3220, summary of the month cooperative. 2003. Available: <http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3220.pdf>. Cited 25 September 2015.
30. Daly C, Helmer EH, Quinones M. Mapping the climate of Puerto Rico, Vieques, and Culebra. *International Journal of Climatology*. 2003; 23: 1359–1381.
31. Willmott CJ, Robeson SM. Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatology*. 1995; 15: 221–229.
32. Plantico MS, Goss LA, Daly C, Taylor G. A new U.S. climate atlas. In: Proc., 12th AMS Conf. on Applied Climatology, Amer. Meteorological Soc., Asheville, NC, May 8–11, 2000, 247–248. Available: <http://www.ncdc.noaa.gov/oa/about/cdrom/climatls2/info/atlasad.html>.
33. Funk C, Michaelsen J, Verdin J, Artan G, Husak G, Senay G, et al. The collaborative historical African rainfall model: description and evaluation. *Int J Climatology*. 2000; 1: 47–66.
34. Daly C, Gibson WP, Doggett M, Smith J, Taylor G. Up-to-date monthly climate maps for the conterminous United States. Proc., 14th AMS Conf. on Applied Climatology, 84th AMS Annual Meeting Combined Preprints, Amer. Meteorological Soc., Seattle, WA, January 13–16, 2004, Paper P5.1, Available: <http://ams.confex.com/ams/pdfpapers/71444.pdf>.
35. Hamlet AF, Lettenmaier DP. Production of temporally consistent gridded precipitation and temperature fields for the continental U.S. *Journal of Hydrometeorology*. 2005; 6: 330–336.
36. Daly C. Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*. 2006; 26: 707–72.
37. Daly C, Widrechner MP, Halbleib MD, Smith JI, Gibson WP. Development of a new USDA Plant Hardiness Zone Map for the United States. *Journal of Applied Meteorology and Climatology*. 2012; 51: 242–264.
38. Daly C, Smith JI, McKane R. High-resolution spatial modeling of daily weather elements for a catchment in the Oregon Cascade Mountains, United States. *Journal of Applied Meteorology and Climatology*. 2007; 46: 1565–1586.