

# Development of high-resolution (250 m) historical daily gridded air temperature data using reanalysis and distributed sensor networks for the US Northern Rocky Mountains

Zachary A. Holden,<sup>a,b\*</sup> Alan Swanson,<sup>b</sup> Anna E. Klene,<sup>b</sup> John T. Abatzoglou,<sup>c</sup>  
Solomon Z. Dobrowski,<sup>d</sup> Samuel A. Cushman,<sup>e</sup> John Squires,<sup>f</sup> Gretchen G. Moisen<sup>g</sup>  
and Jared W. Oyler<sup>h</sup>

<sup>a</sup> USDA Forest Service Region 1, Missoula, MT, USA

<sup>b</sup> Department of Geography, University of Montana, Missoula, MT, USA

<sup>c</sup> Department of Geography, University of Idaho, Moscow, ID, USA

<sup>d</sup> College of Forestry, University of Montana, Missoula, MT, USA

<sup>e</sup> U.S. Forest Service, Rocky Mountain Research Station, Flagstaff, AZ, USA

<sup>f</sup> U.S. Forest Service, Rocky Mountain Research Station, Missoula, MT, USA

<sup>g</sup> U.S. Forest Service, Rocky Mountain Research Station, Ogden, UT, USA

<sup>h</sup> Division of Biological Sciences, University of Montana, Missoula, MT, USA

**ABSTRACT:** Gridded temperature data sets are typically produced at spatial resolutions that cannot fully resolve fine-scale variation in surface air temperature in regions of complex topography. These data limitations have become increasingly important as scientists and managers attempt to understand and plan for potential climate change impacts. Here, we describe the development of a high-resolution (250 m) daily historical (1979–2012) temperature data set for the US Northern Rocky Mountains using observations from both long-term weather stations and a dense network of low-cost temperature sensors. Empirically based models for daily minimum and maximum temperature incorporate lapse rates from regional reanalysis data, modelled daily solar insolation and soil moisture, along with time invariant canopy cover and topographic factors. Daily model predictions demonstrate excellent agreement with independent observations, with mean absolute errors of <1.4 °C for both minimum and maximum temperature. Topographically resolved temperature data may prove useful in a range of applications related to hydrology, fire regimes and fire behaviour, and habitat suitability modelling. The form of the models may provide a means for downscaling future temperature scenarios that account for potential fine-scale topographically mediated changes in near-surface temperature.

**KEY WORDS** topoclimate; air temperature; cold air drainage; solar radiation; sensor networks; reanalysis

Received 30 June 2015; Revised 25 September 2015; Accepted 2 November 2015

## 1. Introduction

The heterogeneity of climate in complex mountainous topography has emerged as an important challenge in hydrology and ecology. However, a paucity of long-term high-quality climate observations in mountainous regions limits a thorough understanding of spatiotemporal temperature variability in such regions (Minder *et al.*, 2010) including differences in temperature trends by elevation (Mountain Research Initiative EDW Working Group, 2015). Gridded temperature products have been used to estimate temperature variability in mountainous regions, albeit at spatial scales inconsistent with the scale of many physical and biological processes in mountainous terrain (Millar *et al.*, 2007). These data limitations challenge our ability to model hydrologic and ecological processes at

appropriate scales, and have become increasingly important as scientists and land managers attempt to understand and manage for potential climate change impacts at actionable scales.

Mountains create steep gradients of moisture, energy and light that vary at a range of scales, resulting in spatial variation in near-surface air temperature over short distances. Much of the spatial variability in temperature in such environments can be attributed to variations in the atmospheric temperature profile, variability in shortwave and longwave radiation associated with slope geometry and incident solar radiation (Thorntwaite, 1961; Bristow and Campbell, 1984), as well as the influence of over-story vegetation in attenuating shortwave radiation and limiting longwave cooling at night (Geiger, 1966). The influence of shortwave radiation on daytime temperatures can be further mediated by variations in surface properties including ground moisture conditions via differences in latent and sensible heating (Bowen, 1926). Nocturnal cold air drainage (CAD) and pooling during favourable

\* Correspondence to: Z. A. Holden, USDA Forest Service Region 1, 200 East Broadway Street, Missoula, MT 59807, USA. E-mail: zaholden@fs.fed.us

synoptic conditions can produce very pronounced gradients in minimum temperatures in regions of incised terrain (Dobrowski *et al.*, 2009; Daly *et al.*, 2010; Holden *et al.*, 2011a).

A variety of methods have been used to create gridded temperature data sets at spatial resolutions fine enough to account for mountain climates. These include statistical methods, dynamical methods and hybrid dynamical-statistical methods. Statistical methods typically use the existing observations from long-term weather stations and empirical models that rely heavily on interpolation approaches to estimate temperatures at unsampled locations. Some correct for elevation using constant lapse rates of  $-6.5^{\circ}\text{C km}^{-1}$  (Willmott and Matsuura, 1995; Maurer *et al.*, 2002), while others use thin plate spline models that account for latitude, longitude and elevation (Hijmans *et al.*, 2005). The daily 1-km Daymet product uses geographically weighted regression to model daily varying linear lapse rates (Thornton *et al.*, 1997; Thornton and Running, 1999). The Parameter Regression on Independent Slopes Model (PRISM; Daly *et al.*, 2008) product uses local geographically weighted regression models to estimate temperature at 30 arcsec ( $\sim 800$  m) and 2.5 min ( $\sim 4$  km) spatial resolution, but weights observations by physiographic similarity to the prediction point as well. The ‘Topographic-Weather’ data set (TopoWx; Oyler *et al.*, 2014) uses moving window regression Kriging and incorporates both elevation and land surface temperature derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) to estimate temperature at 800-m spatial resolution. Simple interpolation methods such as ClimateWNA (Wang *et al.*, 2006, 2012) can be used to further smooth gridded 2.5-min PRISM data using finer resolution elevation data.

Dynamical models for obtaining high-spatial resolution meteorological surfaces rely on a regional numerical weather model (e.g. the Weather Research and Forecasting model). These models resolve physical processes governing mesoscale atmospheric dynamics and atmosphere–land surface feedbacks and are theoretically a superior means to obtain such information. However, the computational demands of dynamical models typically limit the output resolution to relatively coarse grids (e.g.  $>4$  km).

Hybrid models integrate outputs from dynamical models or reanalyse products with observational or statistically derived data sets. Such hybrid models aim to correct for biases in dynamical modelling and refine their output to more localized scales. Examples of hybrid models that cover the continental US include the North American Land Data Assimilation System (NLDAS; Mitchell *et al.*, 2004) that combined output from the North American Regional Reanalysis (NARR, Mesinger *et al.*, 2006) and other sources to produce hourly output at a 12-km resolution, Abatzoglou (2013) who biased the corrected output from NLDAS-2 with monthly PRISM data to create daily output at a 4-km resolution and the Real-Time Mesoscale Analysis data (De Pondeca *et al.*, 2011) that generated output at 2.5–5 km grid cells.

Distributed networks of low-cost sensors have emerged as a promising tool for supplementing the limited observations of surface air temperature in mountains and have been used extensively to map spatiotemporal variability in temperature in montane regions (Lundquist and Cayan, 2007; Lundquist *et al.*, 2008; Fridley, 2009; Holden and Jolly, 2011; Holden *et al.*, 2011a, 2011b; Ashcroft and Gollan, 2012). While low-cost sensor networks are useful for short-term field studies and validating the existing data sets, development of long-term, high-quality data sets that integrate these observations may be challenging. First, data sets from sensor networks are likely to be temporally short (1–3 years) due to the battery life, limited memory and expense of maintaining these networks over time. Second, sensors are often distributed using various siting and instrumentation protocols that differ from standards for long-term reference weather stations, making it problematic to compare observations across networks.

Here, we describe the development of models for producing high-resolution historical gridded daily surface air temperature using observations from distributed networks of inexpensive sensors, historical weather station observations and reanalysis data for the Northern Rocky Mountains of the USA. Our approach considers a set of dynamic and static covariates that have established physical links to surface air temperature including solar insolation, soil moisture, local topography, canopy cover, geopotential height and humidity.

## 2. Materials and methods

### 2.1. Overview

Models for daily minimum ( $T_{\min}$ ) and maximum ( $T_{\max}$ ) temperature were developed for the US Northern Rocky Mountains ( $42^{\circ}$ – $49^{\circ}\text{N}$ ,  $105^{\circ}$ – $118^{\circ}\text{W}$ ) covering Idaho, Montana and Northwest Wyoming. Models for  $T_{\min}$  and  $T_{\max}$  were developed separately, but each followed the same general three-stage approach described in more detail in the following sections. First, temperature was estimated by interpolating pressure-level free-air temperature and geopotential height from NARR to a digital elevation model (DEM) derived from the 30-m National Elevation Data Set (Gesch *et al.*, 2002). Differences between surface observations and initial NARR estimates were then modelled using a suite of physically based spatial predictors extracted at the location of each weather station and sensor. These models were then applied to the study domain, and the residual errors were interpolated to create a daily offset that spatially corrects the initial temperature estimate from NARR. Model fitting was done using subsampled data from 1979 to 2012 and evaluated using an independent set of observations from permanent weather stations.

### 2.2. Temperature observations

#### 2.2.1. Weather station observations

We used a data set of quality-assured and homogenized temperature observations from most major observational

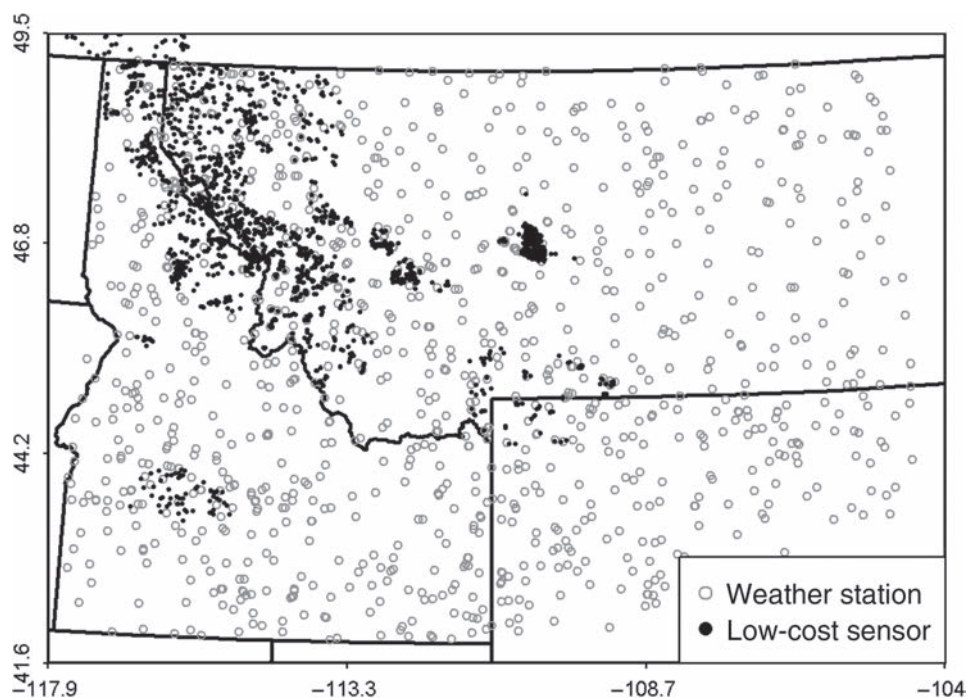


Figure 1. Study area figure with location of low-cost sensors and permanent weather stations used for modelling.

networks in the US Northern Rocky Mountains (Figure 1) that includes the Global Historical Climatology Network stations (GHCN-D; Menne *et al.*, 2012), Snowpack Telemetry (SNOTEL) stations and Remote Automated Weather stations (RAWS). These data were subjected to quality assurance (Durre *et al.*, 2010), homogenization procedures described by Menne *et al.* (2009), and infilling methods described by Oyler *et al.* (2014). Homogenization was noted by Oyler *et al.* (2015) as being particularly important for temperature trends at SNOTEL stations because of the changes in instrumentation over time. Observations from the GHCN-D were excluded from the model fitting procedure for both  $T_{\min}$  and  $T_{\max}$  and were used only for model validation. This decision was made because we did not have confidence in the coordinate precision of some of these stations, which we felt could have impacted the accuracy of extracted estimates of fine-scale topographic predictors used for modelling.

### 2.2.2. Distributed sensor network data

Additional temperature observations were obtained from a large network of low-cost distributed temperature sensors deployed by the authors across the US Northern Rocky Mountains and Canada (Figure 1). In 2009, 535 ThermoChron iButton temperature dataloggers (model 1922B) housed in two inverted plastic funnels (Hubbart *et al.*, 2005; Hubbart *et al.*, 2007) were distributed across western Montana and northern Idaho. In 2010, additional 1100 ThermoWorks Logtag<sup>®</sup> temperature dataloggers (model TRIX-8) were deployed across a larger domain, extending from the Boise Basin, Idaho to southern British Columbia. Logtags were housed in inexpensive solar

radiation shields with performance characteristics comparable to commercially available non-aspirated Gill shields (Holden *et al.*, 2013). In 2011, all iButton and Logtag dataloggers were retrieved, and each Logtag was replaced with a model TRIX-16 Logtag sensor. Owing to the larger storage capacity of the TRIX-16 model, these sensors remained in the field until summer 2013. Both Logtags and iButtons were distributed in a random stratified design within three classes of elevation and four classes of aspect, and at sites within 300 m of roads to facilitate rapid deployment. Sensors were installed directly on the north side of a tree at a height of 2 m to minimize the influence of direct solar heating. Owing to memory limitations of the sensors, both Logtags and iButtons were programmed to record temperature every 90 min, beginning at 0600 LST (1500 UTC). The location of each sensor was recorded using a field-grade Garmin GPS, with a minimum of 100 points collected and averaged at each site. All sensors were removed from the field between July and October 2013. Because of radiation biases associated with the funnel shield noted by Holden *et al.* (2013), iButton data were collected from summer 2009 to summer 2010 and only used for estimating nighttime minimum temperature. Additional information on methods used to screen sensor network data can be found in Supporting Information.

### 2.3. Spatial predictors

A suite of time-varying and time-invariant gridded spatial covariates were used as predictors in the models. These included the daily free-air lapse rate, daily solar insolation, daily soil moisture, CAD potential (CAD-P) and canopy cover (Table 1). Each variable was produced at 8 arcsec ( $\sim 250$  m) resolution, unless otherwise specified. Details about the development of each variable are given below.

Table 1. Spatial and temporal predictors used for modelling minimum and maximum temperature.

| Variable name            | Description                                               | Units             | Source      |
|--------------------------|-----------------------------------------------------------|-------------------|-------------|
| NARR $T_{\min}/T_{\max}$ | Elevation-adjusted free-air $T_{\min}/T_{\max}$ from NARR | °C                | NARR        |
| smoist                   | FASST modelled 0–10 cm moisture                           | %                 | FASST       |
| Hgt_z                    | Normalized geopotential height                            | unitless          | NARR        |
| VPD                      | Vapour pressure deficit                                   | Kpa               | NARR        |
| RH                       | Relative humidity                                         | %                 | NARR        |
| CC                       | MODIS canopy fraction                                     | %                 | MODIS       |
| Longwave                 | Downward longwave radiation at the surface                | W m <sup>-2</sup> | NARR        |
| Solar insolation         | Shade/cloud/aspect corrected net radiation                | W m <sup>-2</sup> | GRASS/r.sun |

### 2.3.1. Daily lapse rate modelling

Models for  $T_{\min}$  and  $T_{\max}$  rely on data from NARR, which is a regional atmospheric model that assimilates available data such as twice-daily radiosonde data and provides 3-h data at a 32-km resolution for a large number of atmospheric and surface hydrologic variables, including air temperature and geopotential height at 29 pressure levels. Owing to computational and storage constraints, we developed a method to reduce the temperature/height pairs from 32-km resolution NARR to a linear approximation while attempting to preserve the nonlinear features of the NARR free-air temperature profiles. Figure 2 shows how daily lapse rate adjusted temperatures were derived from 3-h NARR air temperature and geopotential height at the first 16 pressure levels (1000–550 mb). Figure 2A shows the extent of a single NARR cell. For each 32-km grid cell, the 3-h air temperature/height paired observations were interpolated using a local regression smoother (the R function loess) to a set of fixed elevations corresponding to the mean elevation of the different pressure levels (150 to 4950 m). Daily  $T_{\min}$  and  $T_{\max}$  free-air temperatures were then extracted at each elevation (Figure 2B). These values were interpolated to a 5-min (~8 km) resolution grid. Then, linear regression was used to estimate the lapse rate for  $T_{\min}$  and  $T_{\max}$  as a function of elevation derived from a DEM (Figure 2C). Data for the linear regression were limited to the set of fixed elevations falling within the elevation range of the 5-min grid cell and up to 2500 m. The goal of this step was to preserve nonlinear features in the NARR temperature profiles to better represent free-air temperatures across the wide range of elevations which can occur within a single 32-km NARR cell. This is evident in Figure 2, where nonlinear free-air profiles (Figure 2A) translate into spatial variation within the NARR grid cell domain (Figure 2B). Finally, the 5-min linear regression estimates were resampled with a cubic spline to the 8-arcsec grid and applied to the DEM to produce 250-m resolution gridded lapse rate-adjusted free-air temperatures (Figure 2D). Although the lapse rate estimation method attempts to preserve local variation in the reanalysis lapse rates, its linear form ignores local inversions. This is a potential limitation in the model, and one that could be particularly important for estimating  $T_{\max}$  in winter and spring, when persistent daytime inversions are common. Although we examined NARR temperature profiles, we found little evidence for inverted lapse rates in these relatively coarse data.

### 2.3.2. Solar radiation

Topographically adjusted daily incoming shortwave radiation grids were created at an 8-arcsec (~250 m) spatial resolution for the study domain by downscaling 32-km resolution NARR shortwave radiation. We followed the methods of Suri and Hofierka (2004), described in detail in Dobrowski *et al.* (2013), but with modifications to account for daily variation associated with cloud cover and bias in NARR downward shortwave radiation. Additional details describing the development of daily solar insolation grids are provided in Supporting Information.

### 2.3.3. Soil moisture model

To represent the influence of soil moisture on latent and sensible heat fluxes, we developed mean daily near-surface (0–10 cm depth) soil moisture grids using the Fast All-Season Soil Strength model (FASST; Frankenstein and König, 2004). FASST is a point-based energy balance model designed to estimate ground surface properties, including soil moisture, strength and temperature. FASST has a single-layer snow model that has been shown to perform well when compared with a well-established multi-layer snow model (Frankenstein *et al.*, 2008). In addition, the model is quite flexible, and can be run at sub-hourly to daily time steps with outputs that can be retrieved at a range of user-defined depths. Addition details on the physical schemes implemented in the model can be found in Frankenstein and König (2004). Because FASST is a point-based model and computationally intensive to run daily at fine scale over the entire study domain, we ran the model daily across a systematic grid of 14 830 points and then interpolated predictions using a regression approach. FASST was run using daily minimum and maximum temperature, relative humidity, precipitation and wind speed inputs derived from Abatzoglou (2013). Daily fractional cloud cover was extracted from NARR for each point, and solar radiation was then calculated internally by the model using slope, aspect and elevation retrieved from a DEM. To produce fine-scale daily gridded soil moisture estimates, we performed a regression of FASST soil moisture at each model point using 7-day mean radiation and 7-day cumulative precipitation as predictors. Daily estimates were predicted to the 8-arcsec grid and the residual errors at the gridded FASST points were then interpolated using a two-dimensional thin plate spline model and added

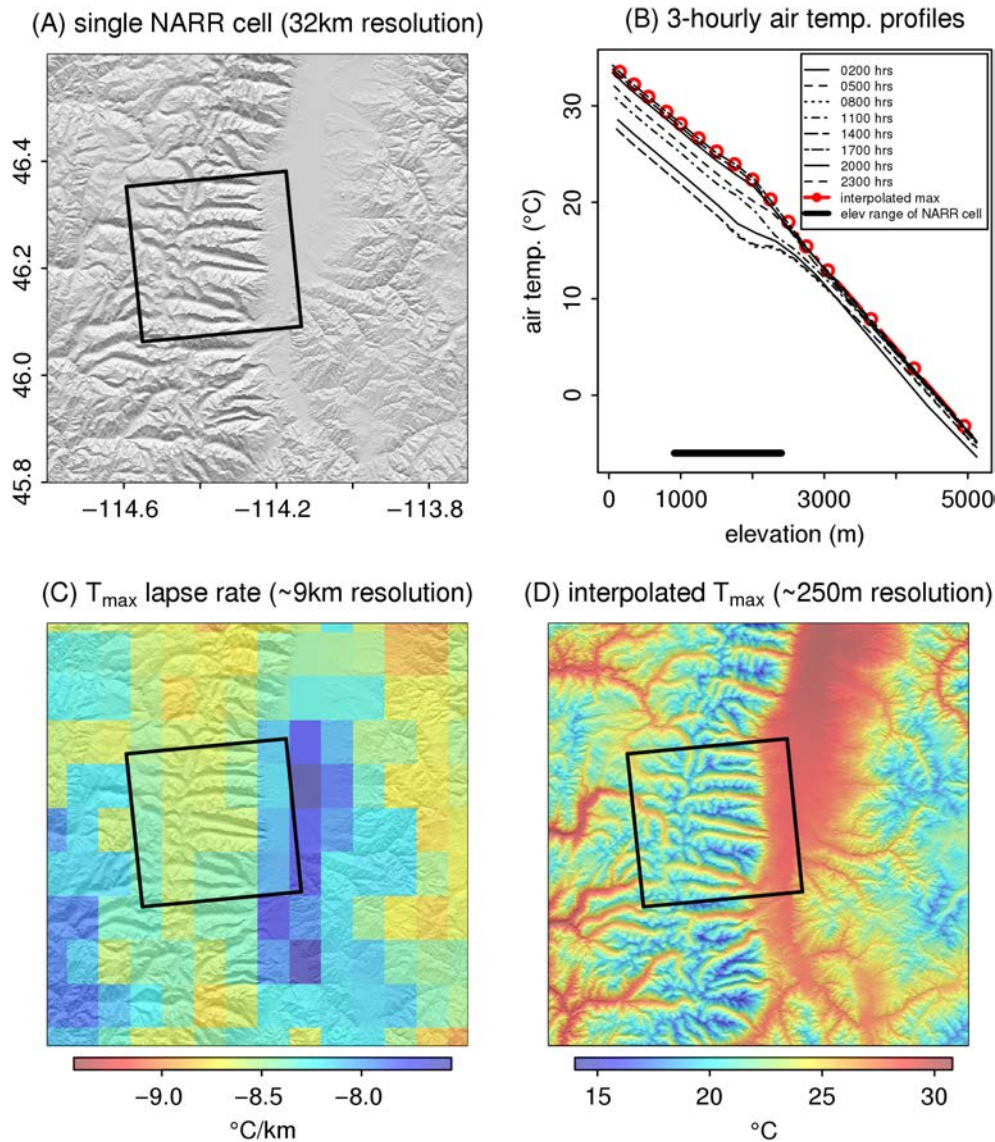


Figure 2. Derivation of  $T_{\max}$  raster from 3-h NARR air temperature and geopotential height for a single NARR cell with considerable topographic relief. (A) Black box shows the extent of a single NARR grid cell draped on a 8 arcsec (~250 m) hillshade. (B) Temperature *versus* geopotential height for the eight readings in a single day, 13 June 2010. The maximum temperatures at a set of fixed elevations are retained. In this example, the lapse rate varies across the cell domain. (C) The  $T_{\max}$  profile from (B) is applied to the DEM to get a lapse rate tuned to the range of elevations contained within each coarse (5 min) grid cell. (D) Gridded lapse rates from (C) are then resampled with a cubic spline and then applied to the final DEM to get NARR estimated  $T_{\max}$  at 8-s (~250 m) resolution.

back to the regression estimates to produce the daily soil moisture grids.

#### 2.3.4. Canopy cover

Canopy cover at each site was estimated using canopy fraction from the MODIS Vegetation Continuous Fields product (VCF; Hansen *et al.*, 2003). The VCF data set provides 250-m resolution estimates of percent canopy cover globally. The maximum VCF value from 2009 to 2012 was extracted for each sensor and weather station location. The maximum (rather than mean or median) was selected because of variability in the number of uncontaminated pixels across the study domain. Canopy cover was thus assumed to be constant through time in the model.

#### 2.4. Air temperature model fitting

##### 2.4.1. Minimum temperature model

Daily  $T_{\min}$  was modelled using a three-stage approach with the goal of estimating the spatial (topographic) and temporal (synoptic climate) influences on minimum temperature, and their interaction. Our approach was to first model CAD-P under standardized synoptic conditions using local physiography, and then model daily NARR lapse adjusted  $T_{\min}$  anomalies ( $\delta T_{\min}$ ) as a function of CAD-P and its interaction with daily synoptic weather conditions. Finally, a daily offset was applied using relationships between predicted temperatures and withheld observations to correct for errors in the NARR. A graphical illustration of the modelling process is shown in

Figure 3. The daily offset correction procedure is described in Section 2.5.

A static map of CAD-P was developed following Lundquist *et al.* (2008) and Holden *et al.* (2011b) whereby temperature observations from high-density observations are correlated with physiographic indices to produce maps estimating the spatial pattern of CAD and pooling. Our approach was to calculate the mean difference ( $\delta T_{\min}$ ) between  $T_{\min}$  of surface observations and of NARR free-air temperature on a selected group of days with stable conditions favourable for CAD, and then model that temperature difference using a suite of physiographic indices derived from a DEM. To identify conditions associated with strong CAD, an initial stochastic gradient boosting model (GBM; Friedman, 2001, 2002) was used to model  $\delta T_{\min}$  on a sample of data, with geopotential height, soil moisture and specific humidity as predictors. This model was used to identify a set of observations from 2010 to 2012 with predicted  $\delta T_{\min}$  greater than  $-8^{\circ}\text{C}$ , resulting in an average of 100 observations per station. Mean  $\delta T_{\min}$  was calculated at each station and used as an estimate of CAD-P under standardized conditions (Figure S1). Next, a model was developed to predict the spatial pattern of CAD-P. Four terrain indices calculated using 15 window sizes (60 indices total) were evaluated as predictors of CAD-P. Similar terrain indices have been used for estimating terrain effects on minimum temperature (Holden *et al.*, 2011a, 2011b; Pepin *et al.*, 2011). The indices evaluated here included a topographic divergence index and local minimum maximum and standard deviation of the focal cell within a variable radius window, and are listed in the Table S1. GBMs were used to model CAD-P as a function of the terrain indices. Initially, all 60 terrain indices were evaluated as candidate predictors. A cross-validation using spatially independent groups was used to identify an optimal predictor set by iterative removal of variables with the lowest contributed explanatory power. The final CAD model included 12 predictor variables. Additional details of the model selection and cross-validation procedure are provided in Supporting Information. A map of the final CAD-P model can be found in Figure S6.

At the second stage of modelling, the time-varying intensity of CAD was characterized as a function of local synoptic climatic conditions, which we call the CAD intensity (CAD-I) model. The CAD-I model was fit as a linear model with  $\delta T_{\min}$  as the response, and interactions between CAD-P and a set of daily climate/weather covariates as the explanatory variables. The daily covariates considered were relative humidity, specific humidity, vapour pressure deficit, longwave radiation, 700-mb geopotential height and modelled soil moisture (Table 1). Geopotential height is strongly correlated with air temperature, such that high pressure days tend to occur during warmer summer periods, potentially resulting in seasonal biases in modelled  $T_{\min}$ . We therefore also considered a normalized 700 mb height value calculated by subtracting the long-term (1979–2012) average geopotential height

for each calendar day and dividing by the long-term standard deviation.

We used a two-stage procedure for selecting the final model for  $T_{\min}$  from a list of candidate models. The candidate models included all possible combinations of the covariates considered, excluding those with more than one humidity or pressure metric, yielding a list of 95 possible models. In the first stage, we performed a 1000-fold repeated random subsampling validation, using spatially and temporally independent training and testing data. To generate each of the 1000 train/test data sets, we selected a random point within the study domain and a random year between 1979 and 2012. The nearest 25% of stations to the random point were selected, and a sample of data within a 3-year period centred on the test year was used as the test set, while a sample of data from the remaining stations and three other randomly selected years were used as the training data set. Each of the candidate models was fit to the 1000 training data sets and tested on the corresponding test sets without inclusion of the daily offset, and the accuracy statistics were collected and evaluated.

The 30 models with the lowest mean absolute error (MAE) in stage 1 were subjected to a second stage of testing. In the second stage, the top 30 models were fit using a simple random sample of 75% of stations and tested over the other 25% on a simple random sample of 1000 dates with inclusion of the daily offset. MAE for each date was calculated, and its mean value over the 1000 days was used to select the top model.

To predict actual daily  $\delta T_{\min}$ , daily CAD-I was multiplied by the static CAD-P to generate a fine-scale daily estimate of CAD. This was added to the free-air temperatures from NARR to obtain the predicted  $T_{\min}$ , and in a third stage of modelling (described below in Section 2.5) a daily offset was applied to each daily  $T_{\min}$  whereby errors from observed and predicted  $T_{\min}$  were projected over a 5-km grid. This daily offset was added to the NARR + CAD prediction to obtain a final gridded estimate of daily  $T_{\min}$ . The full modelling procedure is illustrated graphically in Figure 3.

In an attempt to more evenly weight the contribution of each unique location to the fit of the model, stations with more than 709 (mean number of observations per Logtag station) were randomly subsampled to include only 709 observations. This yielded a data set of approximately 1 million observations for model fitting.

#### 2.4.2. Maximum temperature model

Daily  $T_{\max}$  model form was similar to that of the  $T_{\min}$  model. First,  $T_{\max}$  was estimated using a lapse rate derived from NARR pressure-level data following the procedure described in Section 2.3.1. The residual temperature difference between surface observations and the lapse-estimated temperature ( $\delta T_{\max}$ ) was then modelled as a function of daily solar insolation, soil moisture and canopy cover. A model selection procedure identical to that used for  $T_{\min}$  was applied to select the form of the final model. Additional details of the model fitting results for  $T_{\max}$  and a

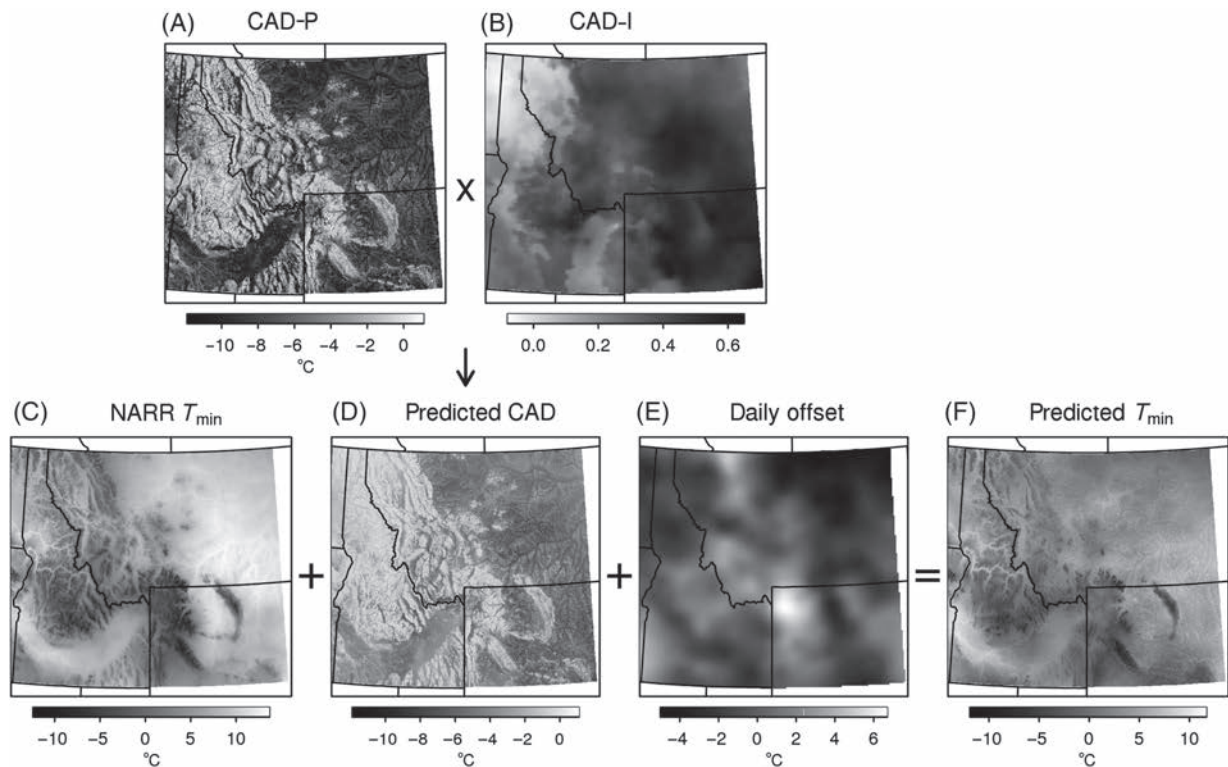


Figure 3. Diagram illustrating the modelling procedure for minimum temperature model for a single day (12 April 1982). The CAD-P map is multiplied by the daily estimate of CAD-I (B) to give a prediction of actual CAD (D).  $T_{\min}$  from NARR (C) is combined with the CAD prediction (D) and the daily offset (E) to get a final prediction of  $T_{\min}$  (F).

simplified schematic diagram illustrating the overall model procedure are provided in Appendix S4.

### 2.5. Daily offset error correction

After estimating the lapse-adjusted temperature and correcting for local terrain effects (radiation, canopy cover and soil moisture), temperature predictions were adjusted to account for daily errors in the model using a subset of independent withheld surface weather station observations. Predicted temperatures for each day were compared with station observations. These residual temperature errors were assumed to be primarily a product of errors in the reanalysis including lapse rate estimates, although it is possible that they could result from observation errors or unmeasured conditions. The residual error was then interpolated to a 5-km grid using a thin plate spline model with  $X$  and  $Y$  as predictors and then resampled to 8 arcsec resolution. The final temperature prediction for each day at each grid cell was then adjusted using the daily offset interpolated error surface. Maps of the daily offset grids averaged by month for the 1979–2012 period are provided in Appendix S4.

### 2.6. Model validation

Models for  $T_{\min}$  and  $T_{\max}$  were validated using a fivefold cross-validation to estimate error rates at each station. For each iteration, we selected 80% of the SNOTEL, RAWS and Logtag stations as training data, and reserved the remaining stations (including all GHCN-D) as testing data.

A linear model was fit to the training data and applied daily (including estimation of the daily offset using the training stations only) over the full time period to collect a full error history at each of the withheld test stations. This was repeated five times, and errors were summarized by station. Model accuracy is reported as MAE between model predictions and withheld data at each station.

An additional validation was performed for the purpose of comparing our modelling results to those from TopoWx (Oyler *et al.*, 2014). To arrive at a valid comparison, we fit a single model using all data, but withheld the Logtag and iButton data when fitting the daily offset in order to avoid the unfair advantage of including additional data. Our objective here was primarily not only to evaluate whether fine-scale topographic variation in the predicted temperature grids was capturing real terrain-scale effects, but also to better understand potential bias associated with the use of low-cost sensors.

## 3. Results and discussion

### 3.1. Model selection results and error statistics

The overall cross-validated MAE for the daily maximum temperature model is 1.02 °C. Error distributions for  $T_{\max}$  at withheld stations are shown in Figure 4 and subset by data source. The model for  $T_{\max}$  includes coefficients for solar radiation, soil moisture, canopy cover and the interaction between soil moisture and radiation. Partial response plots for predictors in the  $T_{\max}$  model developed from

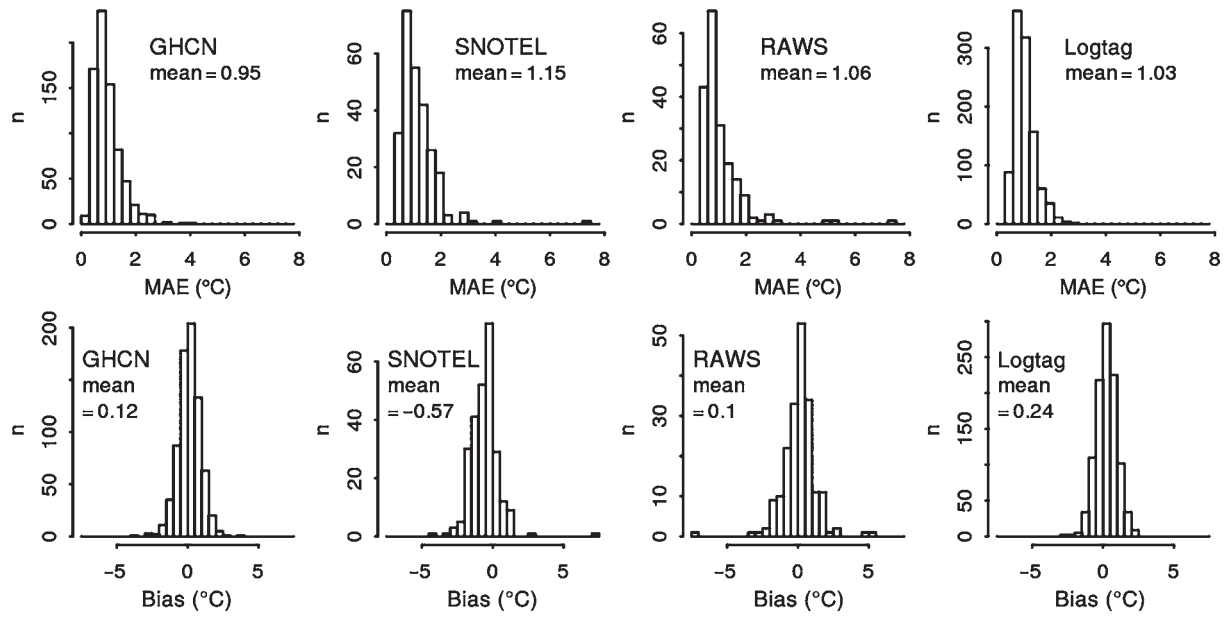


Figure 4. MAE and bias for maximum temperature at independent withheld stations separated by station type. Mean error rates are driven higher by a small number of stations with very large MAE, which we suspect is because of poor locational information.

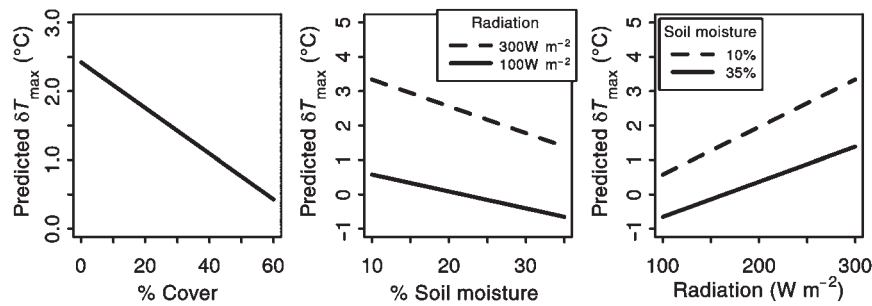


Figure 5. Partial response plots for predictor variables used in the model for  $\Delta T_{\max}$ , the difference between observed and NARR  $T_{\max}$ .

the linear regression model are shown in Figure 5. The response curves for each variable reflect expectations of the physical responses of surface heating to interactions between topography and atmospheric variables. Temperatures were warmer at sites with higher solar insolation with a dampened response at higher (wetter) levels of surface soil moisture. As expected,  $T_{\max}$  is lower (cooler) under higher canopy cover (Figure 5). Because NARR-derived lapse rates are used as an initial estimate of temperature, implicitly maximum temperature decreases with increased elevation.

Cross-validated error summaries for the daily  $T_{\min}$  estimates are shown in Figure 6. The model for minimum daily temperature includes coefficients for CAD-P, standardized 700 mb heights, relative humidity, long-wave radiation and soil moisture. Partial response plots developed from the regression model for minimum temperature are shown in Figure 7. CAD and pooling effects are stronger under stable, high pressure conditions and with decreased atmospheric moisture. The influence of humidity and longwave radiation is small, but significant ( $\sim 0.8^\circ\text{C}$  for humidity and  $0.4^\circ\text{C}$  for longwave radiation). The use of height and humidity fields from

NARR provides a means of accounting for spatiotemporal variation in the magnitude of CAD (Holden *et al.*, 2011a). The soil moisture term in the  $T_{\min}$  model likely reflects the influence of moisture content in the upper soil layers on the magnitude of daytime ground heating and the resulting energy exchange between the ground and atmosphere at night (Whiteman, 1982).

### 3.2. Comparison with the existing data sets

Comparisons of model predictions with the existing gridded data sets provided by PRISM (Daly *et al.*, 2008) and TopoWx (Oyler *et al.*, 2014) resampled to an 8-arcsec grid reveal fine-scale variation of near-surface air temperature that is not captured at the 800-m resolution of these products. Difference maps for both  $T_{\max}$  and  $T_{\min}$  for the Bitterroot Mountains, Montana, are shown in Figure 8. Strong insolation effects are evident, with more than  $2^\circ\text{C}$  differences on north and south facing slopes. Analysis of withheld Logtag data revealed that predictions from our model show less bias with respect to solar insolation (Figure 9). Difference maps for  $T_{\min}$  for the August 1981–2010 normal period reveal cold air pools in narrow valleys and narrow thermal belts above the valley floor

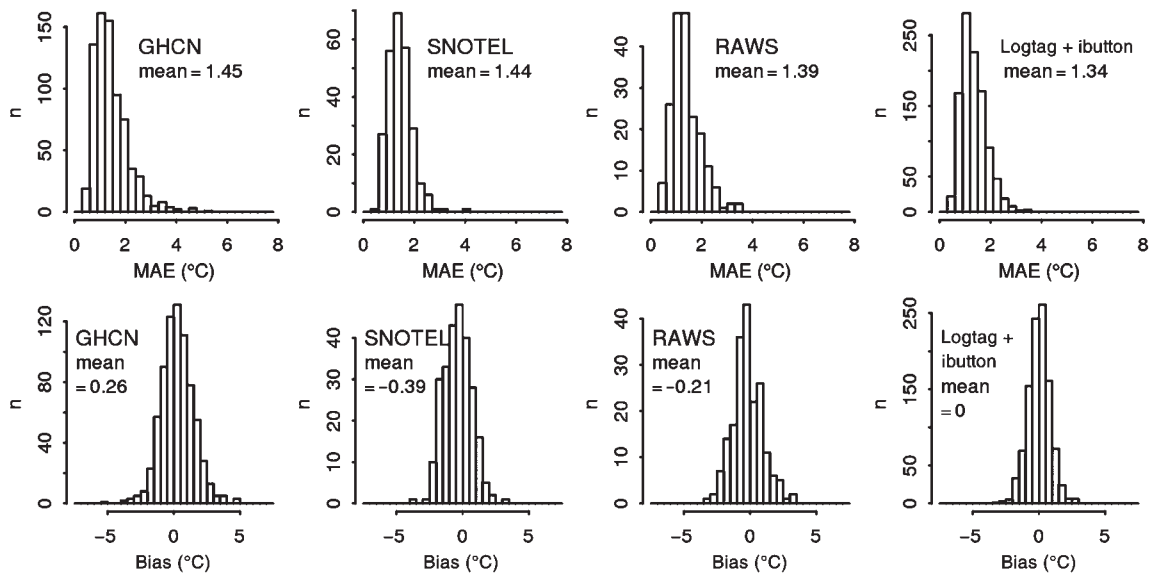
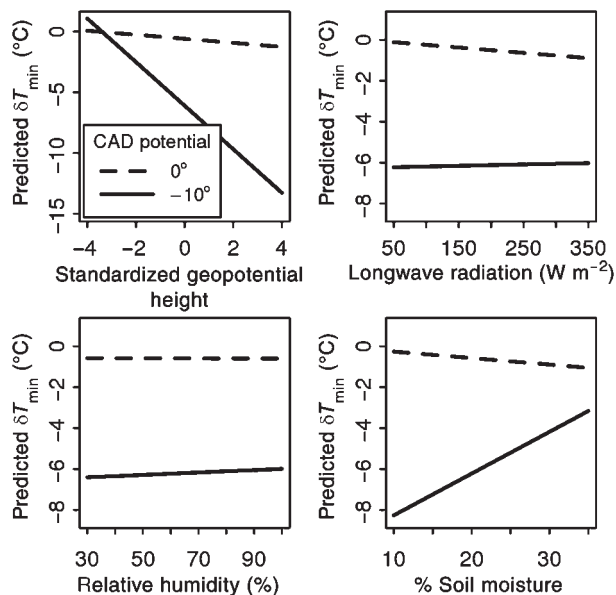


Figure 6. MAE and bias for minimum temperature at withheld stations.

Figure 7. Partial response plots for predictor variables used in the model for  $\delta T_{\min}$ , the difference between observed and NARR  $T_{\min}$ .

(Figure 8). These differences are partly a function of increased spatial resolution, but may also reflect improved characterization of the physical processes governing the fine-scale spatial variation in mountain air temperature. Evidence from withheld Logtag data suggests the latter may be true, as our predictions showed less bias with respect to terrain position (Figure 9). Figure S11 shows model  $T_{\max}$  and  $T_{\min}$  resampled to 30-arcsec resolution and differenced with PRISM and TopoWx. At the coarser spatial resolution of these products, topographic effects are still evident, suggesting that the form of the models in addition to the higher spatial resolution may help to better characterize the effects of topography on surface air temperature.

The interaction between soil moisture and solar insolation reflects an important but previously undescribed source of fine-scale (<1 km) spatiotemporal variation in daytime air temperature. Solar insolation varies daily and seasonally as a function of cloud cover, aspect and solar position, and also drives aspect-scale variation in soil water balance. Soil moisture varies as a function of precipitation amount and timing, as well as slope position and temperature. These factors influence the timing of snowmelt (Lundquist and Flint, 2006; Holden and Jolly, 2011), which in turn influences the onset of soil drying (Harpold *et al.*, 2014). A number of factors, including lower solar insolation and delayed snowmelt timing, decrease evapotranspiration on shaded slope positions and at higher elevations. Moisture retention on more mesic slopes, combined with lower incident radiation, could result in seasonally varying temperature differentials on cool *versus* warm aspect positions. FASST is a point-based model, and although it does capture differences in solar radiation associated with slope geometry, the internal radiation calculations do not consider shading by adjacent terrain, which would contribute further to local radiation influences on modelled soil moisture. Thus, our model likely underestimates the magnitude of aspect-scale variation in near-surface soil moisture. The daily mean insolation maps developed as a predictor for  $T_{\max}$  would have been difficult to incorporate sensibly into the FASST model, which was run at a twice daily time step. Improved modelling of fine-scale moisture patterns could ultimately result in better characterization of the surface air temperature response to interactions between insolation and ground moisture.

The inclusion of canopy cover as a predictor of surface air temperature represents a potentially useful advance for some ecological or hydrologic applications. For example, canopy-adjusted temperatures could be useful for understanding near-surface, sub-canopy processes

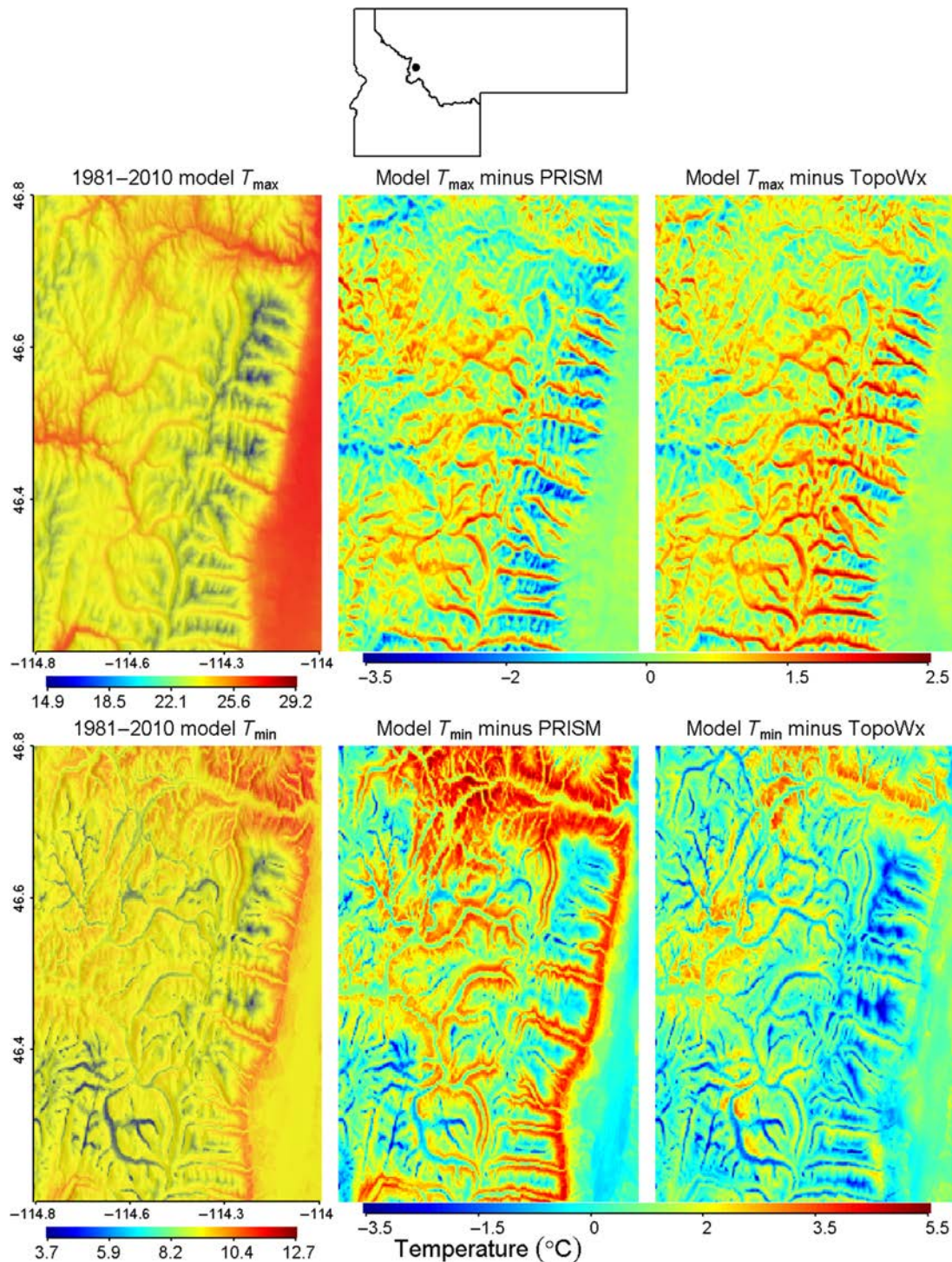


Figure 8. 30-Year average (1981–2010) August maximum and minimum temperature in the Bitterroot Mountains near Missoula, Montana, and differences with PRISM and TopoWx 30-year normal minimum August temperatures. Elevation in the figure domain ranges from 955 to 2874 m. The model predicts significantly warmer maximum temperature on higher insolation (south and west) facing slopes (upper panel). For minimum temperature, the model appears to capture thermal belts and cold air pooling in valley bottoms not fully resolved by PRISM or TopoWx (lower panel).

such as fuel moisture and fire danger modelling (Holden and Jolly, 2011), predicting climatic controls on tree recruitment (Dobrowski *et al.*, 2015) or modelling snow-pack dynamics beneath a forest canopy. However, use of canopy cover estimated from satellite imagery as a predictor is also problematic, particularly at the spatial

resolution of this analysis, as many of the RAWS and SNOTEL stations used in our analysis sit in partial forest openings that may not be reflective of the relatively coarse satellite-derived estimate. The use of canopy cover extracted at this resolution is further complicated by the fact that each low-cost sensor was installed on a live

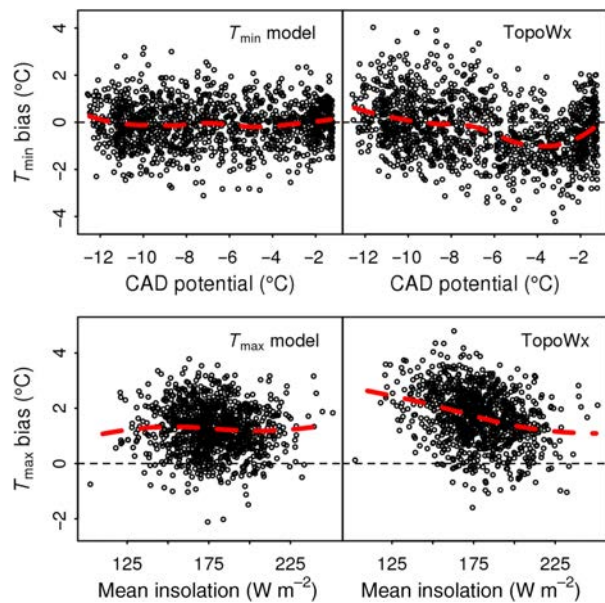


Figure 9. Mean annual model bias for  $T_{\max}$  and  $T_{\min}$  at independent Logtag temperature sensors. Model  $T_{\min}$  shows very little bias with respect to CAD-P, while TopoWx underpredicts  $T_{\min}$  at sites with mild CAD-P, which correspond to the mid-elevation thermal belts evident in Figure 7. Both models overestimate  $T_{\max}$ , which is likely is an artefact of sensor placement next to trees. TopoWx appears to be overestimating  $T_{\max}$  on shaded slopes with lower solar insolation.

tree and is thus sited by default under high local canopy cover. Additionally, with canopy cover fixed in the model through time, the model will be unable to resolve any historical variations in canopy cover associated with past disturbance events, timber harvest or natural vegetation changes during the 1979–2012 period. Further examination of different vegetation data sets such as indices derived from Landsat Thematic Mapper or the National Land Cover Data and the effect of spatial resolution on the effect of canopy cover is warranted in future work. Despite these potential confounding effects, our results are generally consistent with past studies (Rambo and North, 2009) describing the effects of forest cover on surface air temperature.

The use of data from low-cost temperature sensors to parameterize historical temperature models raises a number of potential issues. Chief among these are the sampling rate of the instruments (i.e. whether a 90-min sampling interval can accurately capture diurnal extremes) and siting the sensors in close proximity to vegetation. In addition, the use of observations from a short 3-year time period assumes stationarity in the factors observed during the study period. Our model and TopoWx both overpredict maximum temperature at Logtag sites (Figure 8). This bias is likely driven by an underestimate of highly localized canopy cover due to their location at the base of a tree relative to the co-located 250-m value, but could also reflect the relatively infrequent sampling interval of the sensors. We did not rigorously quantify the differences in model quality with and without distributed sensor observations. However, despite the issues noted above,

during exploratory analysis and model development, we did find that inclusion of low-cost sensors improved the accuracy of predictions to withheld data regardless of time period or sensor type. The utility of these data may arise in part because these sensors were intentionally sited to capture topographic variation in temperature (i.e. on north and south-facing slopes at a range of elevations) presently not well represented by long-term stations. Furthermore, while RAWS and SNOTEL data are increasingly treated as high-quality weather stations and are included in several gridded climate data sets (PRISM, Daymet and TopoWx), they are also often sited near or within forest canopies using passively aspirated radiation shields and using sensors and siting characteristics that changed over time (Oyler *et al.*, 2015). Additionally, RAWS and SNOTEL sensors were not explicitly sited to capture terrain-scale influences on temperature. On the contrary, RAWS were deployed primarily on low-elevation, south-facing sites to capture worst-case fire conditions, while many SNOTEL stations are located at relatively flat, high-elevation saddles.

The models presented here offer several important advances over existing gridded air temperature products. Their spatial resolution is relatively fine compared with the existing data sets, and thus captures information about insolation effects on daytime temperature and CAD patterns in narrow valleys at night. Short duration local studies have noted the potential for using variable lapse rates and/or solar insolation to estimate daytime temperature (Daly *et al.*, 2007; Huang *et al.*, 2009; Frei, 2014). Our models extend this concept by incorporating a suite of variables including lapse rates, atmospheric pressure, shortwave radiation and humidity from reanalysis data. The form of the models, while still empirical, separately estimates the primary physical characteristics of terrain-climate interactions and their variation over time. Statistical models that interpolate air temperature using existing observations presumably integrate these sources of spatial variation which are implicit in the observations. However, they will have a limited ability to transfer those effects forward or backward in time to periods where no observations are available. In contrast, these hybrid models offer a potential empirical approach for downscaling outputs from global or regional climate models with empirical functions that resolve topographically mediated interactions between the land surface and atmosphere.

#### 4. Conclusions

There is a growing awareness among researchers and resource managers of the need for data sets that capture topographic influences on surface air temperature. In regions of complex topography, the available gridded climate data sets may be too coarse to capture the effects of insolation and CAD and pooling. This study describes the development of daily gridded minimum and maximum temperature data sets at a resolution of  $\sim 250\text{m}$  for the

US Northern Rocky Mountains. Model outputs are of the highest spatial resolution currently available for data sets covering regional scales or larger. While the accuracy of this data set is difficult to compare directly with others, error rates were shown to be relatively low in a climatologically complex region. The form of the models account for many of the physically based sources of spatial and temporal variations in surface air temperature and suggest the potential for developing future temperature predictions from coarse resolution general circulation models that account for fine-scale topographically mediated changes in surface air temperature.

### Acknowledgements

This work was funded by USFS Region 1 Fire and Aviation Management through a Challenge Cost-Share agreement between the US Forest Service and the University of Montana (agreement no. 10-CS-11015600-007) and through a NASA Applied Science Program - Wildland Fire award (agreement number NNH11ZDA001N-FIRES). Additional funding was provided by JS. Support was also provided by the Interior West Forest Inventory and Analysis program. Additional support for SZD and ZAH was provided by the National Science Foundation (DEB; 1145985). We gratefully acknowledge the many ecologists and field technicians who distributed and retrieved temperature sensors used in this study including Jessica Page, Jeff DiBenedetto, John Tsroka of the Clark Fork Coalition, and technicians at the Idaho Bird Observatory. We thank Stephan Pracht and Brian Holden for managing field operations and distributing and collecting sensors. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper (<http://www.tacc.utexas.edu>). The authors declare no conflict of interest.

### Supporting Information

The following supporting information is available as part of the online article:

**Appendix S1.** Distributed sensor network data screening.

**Appendix S2.** Daily solar radiation modelling.

**Appendix S3.** Modelling of CAD-P.

**Appendix S4.** Maximum daily air temperature model details.

**Table S1.** Predictors used to model CAD-P.

**Table S2.** Model summary and coefficients for the model of daily maximum temperature. Smoist, modelled soil moisture; radiation, daily mean solar insolation; cover, MODIS percent canopy fraction.

**Figure S1.** Station mean  $\delta T_{\min}$  under stable and well-mixed conditions. The former was used as an estimate of CAD-P, the response for the second stage of modelling. Insets show the Bitterroot valley, Montana and transect of iButtons in the West Fork drainage of the Bitterroot traversing the valley floor to a ridgetop with a maximum sensor elevation of 1670 m.

**Figure S2.** Model error rate *versus* number of covariates for two forms of cross-validation. The y-axis denotes the change in cross-validated MSE with changes in the number of predictors included in the model. The two red lines indicate the MSE of the full model, and the variability of that statistic between the cross-validation folds of the full model. The first method, shown in the upper plot, uses a simple random sample for cross-validation groups while the lower plot shows CV error using spatially independent groups. In the top plot, spatial autocorrelation of the response and predictors results in an artificially low error rate, while the lower plot shows a more realistic error rate. The use of spatially independent validation data appears to give an implicit penalty for overfitting.

**Figure S3.** Partial response curves for the final model. Predictors are described in Table S1.

**Figure S4.** The top four two-way interactions in the final model.

**Figure S5.** Observed *versus* predicted CAD-P from the final model.

**Figure S6.** Raster prediction of CAD-P from the final model.

**Figure S7.** The relationship between modelled CAD-P and the difference between station mean  $\delta T_{\min}$  under well-mixed conditions for the Bitterroot Valley, Montana. The inset shows the same West Fork Bitterroot transect as in Figure S1.

**Figure S8.** Diagram illustrating the procedure for estimating  $T_{\max}$  using reanalysis and observations. Dark grey boxes indicate intermediate or final gridded outputs. Light grey boxes denote statistical models.

**Figure S9.** Monthly mean daily offset error grids applied for  $T_{\max}$  during 1979–2012.

**Figure S10.** Monthly mean daily offset grids for  $T_{\min}$  during 1979–2012.

**Figure S11.** As for Figure 8 in the main article but with model  $T_{\max}$  and  $T_{\min}$  resampled to the 30 arcsec resolution of PRISM and TopoWx.

### References

- Abatzoglou JT. 2013. Development of gridded surface meteorological data for ecological applications and modeling. *Int. J. Climatol.* **33**: 121–131, doi: 10.1002/joc.3413.
- Ashcroft MB, Gollan JR. 2012. Fine-resolution (25 m) topoclimatic grids of near-surface (5 cm) extreme temperatures and humidities across various habitats in a large (200 × 300 km) and diverse region. *Int. J. Climatol.* **32**: 2134–2148.
- Ashcroft MB, Gollan JR, Warton DI, Ramp D. 2012. A novel approach to quantify and locate potential microrefugia using topoclimate, climate stability, and isolation from the matrix. *Glob. Change Biol.* **18**: 1866–1879.
- Bowen IS. 1926. The ratio of heat losses from conduction and by evaporation from any water surface. *Phys. Rev.* **76**: 779–787.
- Bristow KL, Campbell GS. 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric. For. Meteorol.* **31**: 159–166.
- Daly C, Smith JI, McKane R. 2007. High-resolution spatial modeling of daily weather elements for a catchment in the Oregon Cascade Mountains, United States. *J. Appl. Meteorol. Climatol.* **46**: 1565–1586.
- Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.* **28**: 2031–2064, doi: 10.1002/joc.1688.

- Daly C, Conklin DR, Unsworth MH. 2010. Local atmospheric decoupling in complex topography alters climate change impacts. *Int. J. Climatol.* **30**: 1857–1864, doi: 10.1002/joc.2007.
- De Pondaca MSFV, Manikin GS, DiMego G, Benjamin SG, Parrish DF, Purser RJ, Wu W-S, Horel JD, Myrick DT, Lin Y, Aune RM, Keyser D, Colman B, Mann G, Vavra J. 2011. The real-time mesoscale analysis at NOAA's National Centers for Environmental Prediction: current status and development. *Weather Forecast* **26**: 593–612.
- Dobrowski SZ, Abatzoglou JT, Greenberg JA, Schladow S. 2009. How much influence does landscape-scale physiography have on air temperature in a mountain environment? *Agric. For. Meteorol.* **149**: 1751–1758, doi: 10.1016/j.agrformet.2009.06.006.
- Dobrowski SZ, Abatzoglou JT, Swanson AS, Greenburg JA, Myenseberg AR, Holden ZA, Schwartz MK. 2013. The climate velocity of the contiguous United States during the 20<sup>th</sup> century. *Glob. Change Biol.* **19**: 241–251, doi: 10.1111/gcb.12026.
- Dobrowski SZ, Swanson AK, Abatzoglou JT, Holden ZA, Safford HD, Schwartz MK, Gavin DG. 2015. Forest structure and species traits mediate projected recruitment declines in western US tree species. *Global Ecol. and Biogeog.* **24**: 917–927.
- Durre I, Menne MJ, Gleason BE, Houston TG, Vose RS. 2010. Comprehensive automated quality assurance of daily surface observations. *J. Appl. Meteorol. Climatol.* **49**: 1615–1633, doi: 10.1175/2010JAMC2375.1.
- Frankenstein S, König G. 2004. Fast all-season soil strength (FASST). ERDC/CRREL Special Rep. SR-04-1, 107 pp.
- Frankenstein S, Sawyer A, Koeberle J. 2008. Comparison of FASST and SNTherm in three snow accumulation regimes. *J. Hydrometeorol.* **9**: 1443–1463, doi: 10.1175/2008JHM865.1.
- Frei C. 2014. Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. *Int. J. Climatol.* **34**: 1585–1605.
- Fridley J. 2009. Downscaling climate over complex terrain: high finescale (<1000 m) spatial variation of near-ground temperatures in a montane forested landscape (Great Smoky Mountains). *J. Appl. Meteorol. Climatol.* **48**: 1033–1049.
- Friedman JH. 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **29**: 1189–1232.
- Friedman JH. 2002. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **38**: 367–378.
- Geiger R. 1966. *The Climate near the Ground*. Harvard University Press, 610 pp.
- Gesch D, Oimoen M, Greenlee S, Nelson C, Steuck M, Tyler D. 2002. The National Elevation Dataset. *Photogramm Eng Rem S* **68**: 5–11.
- Hansen M, DeFries RS, Townshend JRG, Carroll M, Dimiceli C, Sohlberg RA. 2003. Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS vegetation continuous fields algorithm. *Earth Interact.* **7**: 1–15.
- Harpold AA, Molotch NP, Musselman KN, Bales RC, Kirchner PB, Litvak M, Brooks PD. 2014. Soil moisture response to snowmelt timing in mixed-conifer subalpine forests. *Hydrol. Processes* **29**: 2782–2798, doi: 10.1002/hyp.10400.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**: 1965–1978.
- Holden ZA, Jolly WM. 2011. Modeling topographic influences on fuel moisture and fire danger in complex terrain to improve wildland fire management decision support. *For. Ecol. Manage.* **262**: 2133–2141.
- Holden ZA, Crimmins M, Cushman SA, Littell J. 2011a. Empirical modeling of spatial and temporal variability in warm season nocturnal air temperatures across two North Idaho mountain ranges, USA. *Agric. For. Meteorol.* **151**: 261–269.
- Holden ZA, Abatzoglou JT, Luce CH, Baggett LS. 2011b. Empirical downscaling of daily minimum air temperature at very fine resolutions in complex terrain. *Agric. For. Meteorol.* **151**: 1066–1073, doi: 10.1016/j.agrformet.2011.03.011.
- Holden ZA, Klene A, Keefe R, Moisen GG. 2013. Design and evaluation of an inexpensive radiation shield for monitoring surface air temperatures. *Agric. For. Meteorol.* **180**: 281–286.
- Huang S, Connaughton Z, Potter CS, Genovese V, Crabtree RL, Fu P. 2009. Modeling near-surface air temperature from solar radiation and lapse rate: new development on short-term monthly and daily approach. *Phys. Geogr.* **30**: 517–527.
- Hubbart J, Link T, Campbell C, Cobos D. 2005. Evaluation of a low-cost temperature measurement system for environmental applications. *Hydrol. Processes* **19**: 1517–1523.
- Hubbart JA, Kavanagh K, Pangle R, Link T, Schotzko A. 2007. Cold air drainage and modeled nocturnal leaf water potential in complex forested terrain. *Tree Physiol.* **27**: 631–639.
- Lundquist JE, Cayan D. 2007. Surface temperature patterns in complex terrain: daily variations and long-term change in the central Sierra Nevada, California. *J. Geophys. Res.* **112**: D11124.
- Lundquist JE, Flint A. 2006. Onset of snowmelt and streamflow in 2004 in the western United States: how shading may affect spring streamflow timing in a warmer world. *J. Hydrometeorol.* **7**: 1199–1217.
- Lundquist JE, Pepin N, Rochford C. 2008. Automated algorithm for mapping regions of cold-air pooling in complex terrain. *J. Geophys. Res.* **113**: D22107.
- Maurer EPA, Wood W, Adam JC, Lettenmaier DP, Nijssen B. 2002. A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States. *J. Clim.* **15**: 3237–3251.
- Menne MJ, Williams CN, Vose RS. 2009. The U.S. historical climatology network monthly temperature data, version 2. *Bull. Am. Meteorol. Soc.* **90**: 993–1007, doi: 10.1175/2008BAMS2613.1.
- Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG. 2012. An overview of the Global Historical Climatology Network-Daily Database. *J. Atmos. Ocean. Tech.* **29**: 897–910.
- Mesinger F, DiMego G, Kalnay E, Mitchell K, Shafran PC, Ebisuzaki W, Jovic D, Woollen J, Rogers E, Berbery E, Ek MB, Fan Y, Grumbine R, Higgins W, Li H, Lin Y, Mankin G, Parrish D, Shi W. 2006. North American regional reanalysis. *Bull. Am. Meteorol. Soc.* **87**: 343–360.
- Millar CI, Stephenson NL, Stephens SL. 2007. Climate change and forests of the future: managing in the face of uncertainty. *Ecol. Appl.* **17**: 2145–2151.
- Minder J, Mote PW, Lundquist JD. 2010. Surface temperature lapse rates over complex terrain: lessons from the Cascade Mountains. *J. Geophys. Res.* **115**: 1–13, doi: 10.1029/2009JD013493.
- Mitchell KE, Lohmann D, Houser PR, Wood EF, Schaake JC, Robock A, Cosgrove BA, Sheffield J, Duan Q, Luo L, Higgins RW, Pinker RT, Tarpley D, Lettenmaier DP, Marshall CH, Entin JK, Pan M, Shi W, Koren V, Meng J, Ramsay BH, Bailey AA. 2004. The multi-institution North American Land Data Assimilation System (NLDAS): utilizing multiple GCM products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.* **109**: 1–32, doi: 10.1029/2003JD003823.
- Mountain Research Initiative EDW Working Group. 2015. Elevation-dependent warming in mountain regions of the world. *Nat. Clim. Change* **5**: 424–430.
- Oyler JW, Ballantyne AP, Jencso K, Sweet M, Running SW. 2014. Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature. *Int. J. Climatol.* **35**: 2258–2279, doi: 10.1002/joc.4127.
- Oyler JW, Dobrowski SZ, Ballantyne AP, Klene AE, Running SW. 2015. Artificial amplification of warming trends across the mountains of the western United States. *Geophys. Res. Lett.* **42**: 153–161.
- Pepin N, Daly C, Lundquist J. 2011. The influence of surface/free-air decoupling on temperature trend patterns in the western U.S. *J. Geophys. Res. - Atmos.* **116**: D10109, doi: 10.1029/2010JD014769.
- Rambo TR, North M. 2009. Canopy microclimate response to pattern and density of thinning in a Sierra Nevada forest. *For. Ecol. Manage.* **257**: 435–442.
- Suri M, Hofierka J. 2004. A New GIS-based solar radiation model and its application to photovoltaic assessments. *Trans. GIS* **8**: 172–190.
- Thorntwaite CW. 1961. The task ahead. *Ann. Assoc. Am. Geogr.* **51**: 345–356.
- Thornton PE, Running SW. 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agric. For. Meteorol.* **93**: 211–228.
- Thornton PE, Running SW, White MA. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. *J. Hydrol.* **190**: 214–251, doi: 10.1016/S0022-1694(96)03128-9.
- Wang T, Hamann A, Spittlehouse D, Aitken SN. 2006. Development of scale-free climate data for western Canada for use in resource management. *Int. J. Climatol.* **26**: 383–397.
- Wang T, Hamann A, Spittlehouse D, Murdoch TN. 2012. ClimateWNA - high-resolution spatial climate data for western North America. *J. Appl. Meteorol. Climatol.* **51**: 16–29.
- Whiteman CD. 1982. Breakup of temperature inversions in deep mountain valleys: Part I. Observations. *J. Appl. Meteorol.* **21**: 270–289.
- Whiteman CD, Allwine KJ, Fritschen LJ, Orgill MM, Simpson MM. 1989. Deep valley radiation and surface energy budget microclimates, Part II: energy budget. *J. Appl. Meteorol.* **28**: 427–437.
- Willmott CJ, Matsuura K. 1995. Smart interpolation of annually averaged air temperature in the United States. *J. Appl. Meteorol.* **34**: 2577–2586.