RMetS

Royal Meteorological Society

Development of gridded surface meteorological data for ecological applications and modelling

John T. Abatzoglou*
Department of Geography, University of Idaho, Moscow, ID, USA

ABSTRACT: Landscape-scale ecological modelling has been hindered by suitable high-resolution surface meteorological datasets. To overcome these limitations, desirable spatial attributes of gridded climate data are combined with desirable temporal attributes of regional-scale reanalysis and daily gauge-based precipitation to derive a spatially and temporally complete, high-resolution (4-km) gridded dataset of surface meteorological variables required in ecological modelling for the contiguous United States from 1979 to 2010. Validation of the resulting gridded surface meteorological data, using an extensive network of automated weather stations across the western United States, showed skill comparable to that derived from interpolation using station observations, suggesting it can serve as suitable surrogate for landscape-scale ecological modelling across vast unmonitored areas of the United States. Copyright © 2011 Royal Meteorological Society

🖥 *Additional Supporting information may be found in the online version of this article.*

1. Introduction

Landscape-scale approaches to ecological questions often require spatially and temporally consistent datasets. While ecological studies routinely use *in situ* weather observations, biophysical models run at regional-to-continental scales are often constrained by the lack of suitable meteorological data. The scale mismatch of ecological studies and climatological data results in a host of limitations for ecological modelling (e.g. Root and Schneider, 2003), and prompts the need for such datasets to inform landscape-scale models and the advancement of science (Thornton *et al.*, 1997; Hamlet *et al.*, 2005).

Ecological land-surface models generally require at least air temperature, humidity, precipitation and incident solar radiation as inputs (e.g. Running *et al.*, 1987). Unfortunately, widely available meteorological datasets often are of inadequate spatial and temporal resolution and generally only incorporate a limited set of meteorological parameters (i.e. temperature and precipitation). There is no accepted spatial resolution of climate information needed to inform ecological processes (e.g. Root and Schneider, 2003), rather the spatial resolution of data needs varies across application, variable and physiographic complexity. Although very high-resolution (i.e. 30-m) meteorological surfaces are desired for some applications, our current understanding of physical processes involving coupled mesoscale-topographic drivers is still

evolving (e.g. Holden *et al.*, 2011; Pepin *et al.*, 2011). Likewise, while monthly temporal resolution may suffice for numerous applications, dynamic ecological processes require daily and sub-daily data (e.g. Schwartz and Reiter, 2000; Stockle *et al.*, 2003; Régnière and Bentz, 2007). Finally, spatially explicit meteorological data has generally been limited to temperature and precipitation. Prior studies have developed empirical relationships to derive surface downward shortwave radiation and humidity from landscape position and available temperature and precipitation data (e.g. Running *et al.*, 1987; Kimball *et al.*, 1997); however, such methods may require calibration at local scales, assume fixed relationships that may be overly reliant on a single variable (e.g. Milly and Dunne, 2011), and may be violated under varying synoptic influences (e.g. Daly *et al.*, 2009).

There is a growing need for distributed high-resolution surface meteorological datasets to foster advances in agricultural and ecological modelling (e.g. Stockle *et al.* 2003; Morin and Thuiller, 2009; Meentemeyer *et al.*, 2011). Station-based meteorological observations have historically been used to model ecosystem processes. Unfortunately, station observations suffer from quality control issues and limitations in the observational network. Conversely, standard climate datasets (e.g. National Climate Data Center, United States climate division data and gridded monthly surface air temperature (e.g. Jones *et al.*, 1999)) are often derived from observations that may not adequately represent conditions in mountains (e.g. Pepin and Seidel, 2005), and are often aggregated to resolutions that are inappropriate

---

* Correspondence to: J. T. Abatzoglou, Department of Geography, Moscow, ID 83844-3021, USA. E-mail: jabatzoglou@uidaho.edu

for certain applications. This paper describes and evaluates a methodology of producing high-resolution spatially and temporally consistent surface meteorological datasets across the continental United States from 1979 to 2010 that can be used in landscape-scale ecological modelling and applications.

2. Data and methods

Gridded meteorological datasets have typically been assembled using irregularly spaced observations and a variety of interpolation methods (Daly, 2006, for a review). While the use of station observations has advantages, the observational network suffers from (1) inadequate station density, particularly in mountainous regions, (2) lack of long-term continuous observations, (3) limitations in how spatially representative a single station is to surrounding climate, particularly in regions of heterogeneous physiography and land use (e.g. Pielke *et al.*, 2007), and (4) quality control issues and climate inhomogeneities (e.g. Abatzoglou *et al.*, 2009). Additionally, irregularly spaced observations are problematic for interpolation methods in regions of complex terrain (e.g. Daly *et al.*, 2008). These issues, coupled with the motivation of incorporating a broader suite of meteorological variables used in ecological and hydrological models, necessitates an alternative approach.

2.1. Development of gridded surface meteorological data

This paper employs a hybrid method to create high-resolution gridded surface meteorological data over the continental United States from 1979 to 2010 by combining attributes of two datasets: temporally rich data from the North American Land Data Assimilation System Phase 2 (NLDAS-2, Mitchell *et al.*, 2004), and spatially rich data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM, Daly *et al.*, 2008). Before describing the procedure of merging these two datasets, it is instructive to provide an overview of their sources as well as spatial and temporal characteristics.

The NLDAS-2 dataset features meteorological variables at hourly time scales and 1/8th degree (approximately 12 km) resolution. Data from NLDAS-2 is primarily derived from the North American Regional Reanalysis (NARR, Mesinger *et al.*, 2006), interpolated from the 32-km horizontal resolution NARR grid to the 1/8th degree NLDAS-2 grid, adjusted for elevation differences (e.g. standard lapse rates), and temporally disaggregated from three-hourly to hourly time scales (Cosgrove *et al.*, 2003). Surface downward shortwave radiation from NLDAS-2 is derived by bias correcting NARR output using Geostationary Operational Environmental Satellite (GOES) data (Pinker *et al.*, 2003). Precipitation data from NLDAS-2 is derived by temporally disaggregating Climate Prediction Center daily gauge data to hourly time scales using radar-based estimates, infrared satellite and NARR. While NLDAS-2 has been used both operationally and in research, at its native horizontal and spatial

resolution it (1) contains significant biases (e.g. occasionally exceeding 10 °C on a monthly basis), (2) does not fully capture the diurnal temperature and humidity range, and (3) fails to resolve spatial complexity of meteorological features in diverse terrain.

The PRISM dataset provides high spatial resolution (800 m) climate surfaces of temperature, precipitation and dewpoint temperature at monthly time scales by performing local regressions of station data to physiographic elements using an extensive knowledge base of spatial climate factors (Daly *et al.*, 2008). While PRISM may suffice for monthly and longer time periods, the lack of sub-monthly time scales limits its application to higher-frequency phenomena.

Although NLDAS-2 and PRISM may not individually serve the needs of many ecological applications, both datasets contain desirable and complementary attributes that are blended to create a daily high-resolution meteorological dataset. The approach utilized here preserves the high-spatial resolution PRISM data at monthly and longer time scales, while incorporating the high-temporal resolution NLDAS-2 data at daily and sub-daily time scales.

The process of merging these two datasets involves an initial bilinear interpolation from the 1/8th degree NLDAS-2 grid (approximately 12 km) to a 4-km grid. Although PRISM data exist at 800-m resolution, data are upscaled to a 4-km horizontal resolution using an area-weighted average for computational tractability. Wind velocity and surface downward shortwave radiation are not modified at spatial scales finer than the NLDAS-2 grid; likewise, surface downward shortwave radiation is not further adjusted to correct for local topography (e.g. radiation adjustments for landscape position such as aspect, slope and topographic shading are not accounted for). However, temperature, humidity and precipitation are bias corrected using climatically aided interpolation (Willmott and Robeson, 1995). While climatologically aided interpolation typically superimposes interpolated station anomalies with climatological normals to estimate monthly time series, this study superimposes interpolated daily departures of monthly averages from NLDAS-2 with monthly data from PRISM.

Following Di Luzio *et al.*, (2008), Equation (1) defines daily precipitation as the fraction of monthly precipitation that occurs on a given day from NLDAS-2 (P_N) scaled by the monthly precipitation from PRISM (P_P). Equation (2) defines the daily maximum temperature as the daily maximum temperature from NLDAS-2 (TX_N) plus the difference of the monthly average maximum temperature from PRISM (TX_P) and the monthly average maximum temperature from NLDAS-2 (TX_N). In rare cases where daily minimum temperatures (estimated using Equation (2), performed independently from calculation of daily maximum temperatures) exceed daily maximum temperatures, as may exist in areas with weak diurnal temperature cycles (i.e. near coastal environments), both temperatures are prescribed the daily mean temperature. Following Equation (2), daily mean dewpoint temperature (TD) is defined by bias-correcting daily

mean dewpoint temperature from NLDAS (derived from 2-m specific humidity) using monthly mean dewpoint temperature from PRISM. Hourly temperatures can then be defined using Equation (3) as the daily minimum temperature (TN) plus the ratio of the difference between the hourly temperature (T_N) and daily minimum temperature (TN_N) and the diurnal temperature range from ($TX_N - TN_N$) from NLDAS-2 times the bias-corrected diurnal temperature range ($TX - TN$).

$$P(d, m, y) = \left[\frac{P_N(d, m, y)}{\sum\limits_{d=1}^{nd} P_N(d, m, y)} \right] \times P_P(m, y) \quad (1)$$

$$TX(d, m, y) = [TX_N(d, m, y) - \overline{TX_N(m, y)}] + TX_P(m, y) \quad (2)$$

$$T(h, d, y) = \frac{T_N(h, d, y) - TN_N(h, d, y)}{TX_N(h, d, y) - TN_N(h, d, y)} \times [TX(d, y) - TN(d, y)] + TN(d, y) \quad (3)$$

$$TD(h, d, m, y) = \left(TD_N(h, d, m, y) - \overline{TD_N}(d, m, y) \right) + TD(d, m, y) \quad (4)$$

Derivation of daily maximum and minimum relative humidity (RH) involves the additional steps of (1) identifying the time of day that coincides with the daily RH extremes from NLDAS-2, (2) applying hourly corrections to temperature using Equation (3) and dewpoint temperature using Equation (4) (similar to Equation (3), but uses hourly deviation from daily mean dewpoint (TD)), and (3) calculating RH using standard World Meteorological Organization methods (WMO, 2008). Relative humidity is set to 100% when dewpoint temperature exceeds temperature.

The ability of the gridded datasets to integrate meteorological forcing pertinent to ecological modelling is examined for two metrics: Energy Release Component (ERC) and reference potential evapotranspiration (ET_o). ERC is a component of the United States Forest Service National Fire Danger Rating System (NFDRS) used by fire management to assess large fire potential and tactical operations. ERC integrates daily temperature, precipitation and humidity across a fuel model matrix (here viewed using model G, short needle pine, heavy dead loads) and comprises a hybrid weather-climate buildup type metric that numerically represents the amount of energy at the flaming front of a fireline (Deeming et al., 1978). Ancillary data needed to model equilibrium moisture content in the NFDRS includes surface observations at 1300 local standard time (LST, here 21Z was used following Equations (3) and (4)) and the number of hours in which measureable precipitation ($>0.01''$, 0.25 mm) was recorded over the previous 24 h. Reference ET_o is the atmospheric demand for evaporation and transpiration from a reference static vegetative surface (here a grass reference surface is assumed) in the absence of water limitations.

Reference ET_o is calculated using the Penman-Monteith method (Allen et al., 1998) and requires maximum and minimum temperature, daily average dewpoint temperature (equivalently, vapour pressure or vapour pressure deficit), wind speed and downward shortwave radiation. Daily mean vapour pressure deficit is calculated as the difference between the mean of the saturation vapour pressure concomitant with the daily high and low temperatures and saturation vapour pressure at the daily mean dewpoint temperature following Jensen et al., (1990).

2.2. Validation

Validation efforts are focused over the complex terrain of the western United States. Three sets of observation data covering the 1991–2010 time period are used: (1) over 1500 Remote Automated Weather Stations (RAWS, http://fam.nwcg.gov/fam-web/weatherfirecd, last accessed 13 May 2011), (2) 74 automated weather stations (AgriMet) in irrigated areas of the Pacific Northwest from the US Bureau of Reclamation (http://www.usbr.gov/pn/agrimet/wxdata.html, last accessed 13 May 2011), and (3) 44 AgWeatherNet (AWN) observations located primarily across eastern Washington State from Washington State University (http://weather.wsu.edu, last accessed 13 May 2011). RAWS summaries include daily maximum and minimum temperature and RH, observations of temperature, humidity, wind speed at 1300 LST as well as the accumulated precipitation and precipitation duration over the previous 24 h. Observations from AgriMet and AWN include daily maximum and minimum temperature and RH, daily averaged wind speed and dewpoint temperature and cumulative surface downward shortwave radiation and precipitation ending at local midnight. Time delineations are preserved to pair fields to their intended application, rather than converting all observations to a common reference time (e.g. ending at local midnight). An example of meteorological fields for 14 July 2007 from both station observations and the gridded surface meteorological data are shown in Figure 1. While validation focuses on the western United States, Appendix A provides results for the contiguous United States for a limited set of meteorological variables from the United States Historical Climate Network (USHCN, http://cdiac.ornl.gov/ftp/ushcn_daily/, last accessed 13 May 2011).

Station observations are initially scanned for improbable data following Durre et al., (2010). A screening procedure is then used to identify temporal and spatial outliers. This involves transforming daily observations into standardized anomalies by estimating means and standard deviations across all years using a 31-day moving window to increase sampling robustness. A square root transformation is performed on precipitation and wind speed to reduce skew in the data. Data fail quality control when either the temporal anomaly exceeds 5 standard deviations, or the spatial anomaly, defined as the difference between individual observations and the mean of the nearest 10 stations (e.g. Peterson et al., 1998),
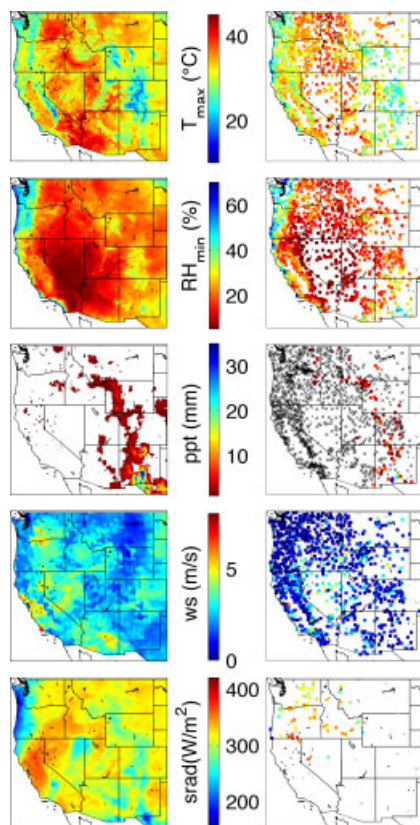
Figure 1. Gridded (left) and observed (right) meteorological data for 14 July 2007. From top to bottom are the daily maximum temperature, daily minimum relative humidity, daily accumulated precipitation (x denotes stations not reporting measureable precipitation; gridded data with less than 0.25 mm are not shown), wind speed (see text for more details), and daily mean downward shortwave radiation. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

exceeds 1.5 standard deviations. It should be noted that this quality control procedure has the potential to eliminate valid observations for stations in uniquely sighted areas under certain synoptic regimes (i.e. a mountain-top station surrounded by stations located in drainages on nights with a strong inversion). The screening process is iterated until all data adhere to the quality control measures and the process is performed independently across variables. All observations for a station on a given day are considered erroneous if one or more variables fail quality control. Collectively, the quality control procedure discarded approximately 1.1% of the data.

Validation is conducted through a direct comparison between station observations and the co-located, closest, grid point for maximum and minimum temperature and RH, precipitation, wind speed, vapour pressure deficit, ERC and ET_o. To evaluate the skill of the gridded dataset, a separate validation is performed using inverse-distance weighting (IDW) interpolation (e.g. Willmott et al., 1985). The IDW approach is sensitive to station density, however, it provides a practical means for estimating observations and comparing the gridded dataset to an approach that only uses surface observations. Validation is conducted separately for the cool season (Oct–Apr) and warm season (May–Sep) due to

the fact that many RAWS are not operational during winter. Statistics of the Pearson correlation coefficient, mean absolute error (MAE, note however, that for precipitation and winds, percent MAE, $\Sigma (|Xi-Oi|)/\Sigma Oi$, is used) and bias (grid minus station) are calculated for each station with at least 1000 valid observations. Additional validation statistics for temperature and precipitation extremes, as well as precipitation frequency are provided in Appendix B. As ERC calculations require continuous data, a limited set of 520 RAWS stations that had no more than 15% missing data are used. Missing data for the limited set of RAWS are estimated using IDW interpolation. Data limitations restrict the validation of ET_o to Agrimet and AWN stations. Finally, a logarithmic adjustment factor is applied to the 10-m gridded wind field to facilitate a direct comparison to the 6.1 m (20 ft) and 2-m observations from RAWS and AgriMet/AWN stations, respectively (Allen et al., 1998).

There has been significant interest in developing long-term high-quality meteorological datasets suited to test ecological hypotheses involving climate variability. To complement validation on daily time scales, two additional validation exercises are performed to evaluate ERC and ET_o on interannual time scales. First, ERC during core fire season, herein defined as 15 July to 15 September, are averaged across 26 stations and co-located gridded data within the perimeter of the Northern Rocky Mountain Forest ecoprovince (Bailey, 1995) from 1991 to 2010. Secondly, cumulative March–October ET_o is averaged over 7 AgriMet stations, and co-located gridded data within the Snake River Plain, Idaho, from 1993 to 2010.

3. Results and discussion

3.1. Validation of meteorological variables

Median correlations of 0.94–0.95 and 0.87–0.90 were found for daily gridded maximum and minimum temperature, respectively, with median MAE between 1.7 and 2.3 °C (Figure 2, Table SI). These results are consistent with previous findings regarding correlation patterns for maximum and minimum temperatures (e.g. Eischeid et al., 2000). Lower correlations of maximum temperatures at coastal sites, defined as sites located less than 20 km from the coastline, result from surface-free air decoupling associated with mesoscale phenomena (e.g. marine boundary layer) that are not adequately resolved by mesoscale reanalysis (i.e. NARR). Observed biases for maximum temperature were significantly correlated to biases in elevation (grid minus station), generally following regional lapse rates (Figure S5). For minimum temperature, lower correlations were found along the coast and in parts of western Montana and Wyoming, particularly during the warm season. Adjacent low ($r < 0.7$) and high ($r > 0.9$) correlations suggest that local (<10 km) factors and station siting are critically important in estimating minimum temperatures in regions of complex terrain, particular under quiescent atmospheric conditions

conducive to decoupling and cold air drainage (e.g. Daly *et al.*, 2009; Holden *et al.*, 2011; Pepin *et al.*, 2011). It is important to note that stations with lower correlations in Montana and Wyoming are RAWS that tend to be positioned on the landscape to maximize fire danger (e.g. south-facing aspects, mid-slope position away from vegetative cover; (Brown *et al.*, 2011)), and are often located in the 'thermal belt' above the nocturnal boundary layer that predisposes such sites to warmer overnight temperatures. The stronger cool bias at RAWS (median bias −0.95 °C) *versus* non-RAWS sites (median bias −0.12 °C) supports this hypothesis. By contrast, supplemental validation performed using daily data from the USHCN found much stronger correlations ($r > 0.9$) with nominal biases across the United States (Appendix A, Figure S3). However, USHCN sites are geographically biased toward valley locations (not necessarily cold-air drainages), and thus do not fully sample conditions across complex terrain as with the primary analysis presented.

Daily maximum and minimum RH exhibited median correlation values between 0.77 and 0.81 and median MAE between 6 and 12% (Figure 3, Table SI). The ability to capture the daily RH extremes is subject to capturing three factors: (1) surface air temperature, (2) surface specific humidity, and (3) the timing of these two variables over the diurnal cycle. Overall, gridded data adequately captured minimum RH ($r > 0.85$, MAE <6%) for a majority of stations, consistent with relationships found for maximum temperature. Biases were generally small, aside from the dry bias seen at coastal stations (Figure S5). Maximum RH generally exhibited slightly weaker correlations than for minimum RH, which may be partially associated with the degradation of measurement accuracy at higher RH values (Interagency Wildland Fire Weather Station Standards and Guidelines, 2009; http://raws.fam.nwcg.gov/nfdrs.html last accessed 13 May 2011) and partially due to the increased error in estimating RH for small dewpoint depressions (Lin and Hubbard, 2004). Broad regions of lower correlation for maximum RH as seen in the Northern Rockies and along the coast, similar to those for minimum temperature, suggest that the ability to simulate maximum RH is dependent on capturing minimum temperature. Similar
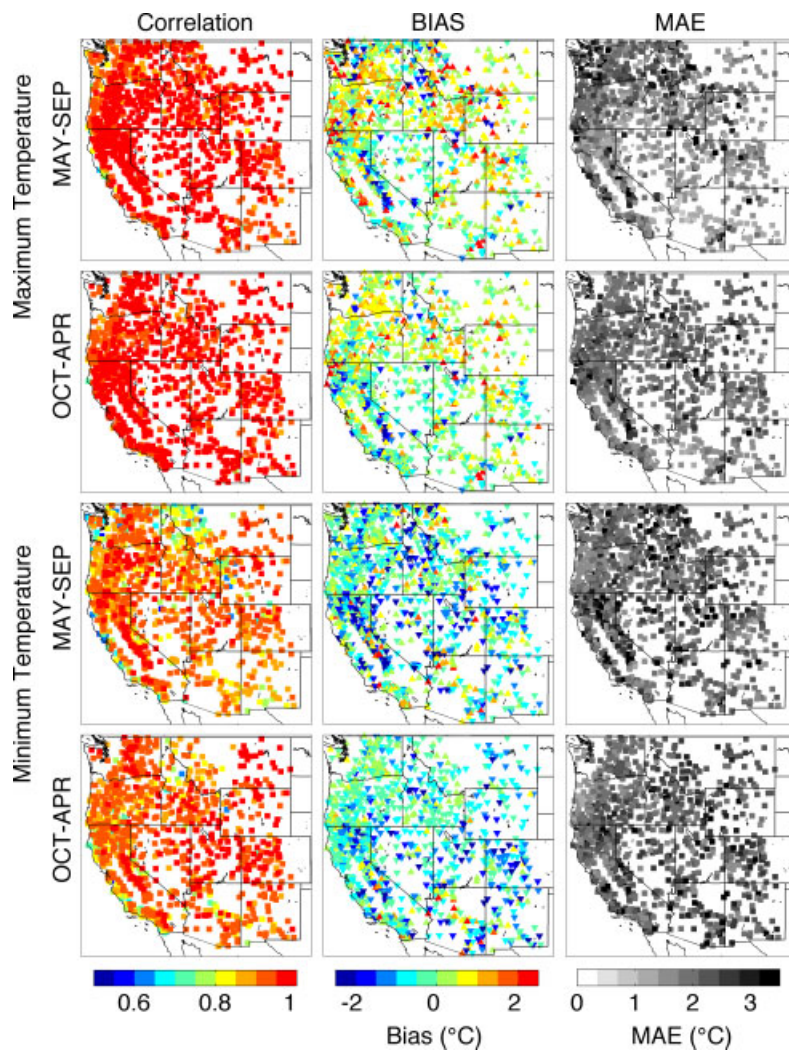

Figure 2. Validation statistics for daily maximum and minimum temperature for all RAWS, AgriMet and AWN data with at least 1000 valid observations. Bias is defined as gridded data minus station observations. This figure is available in colour online at wileyonlinelibrary.com/journal/joc
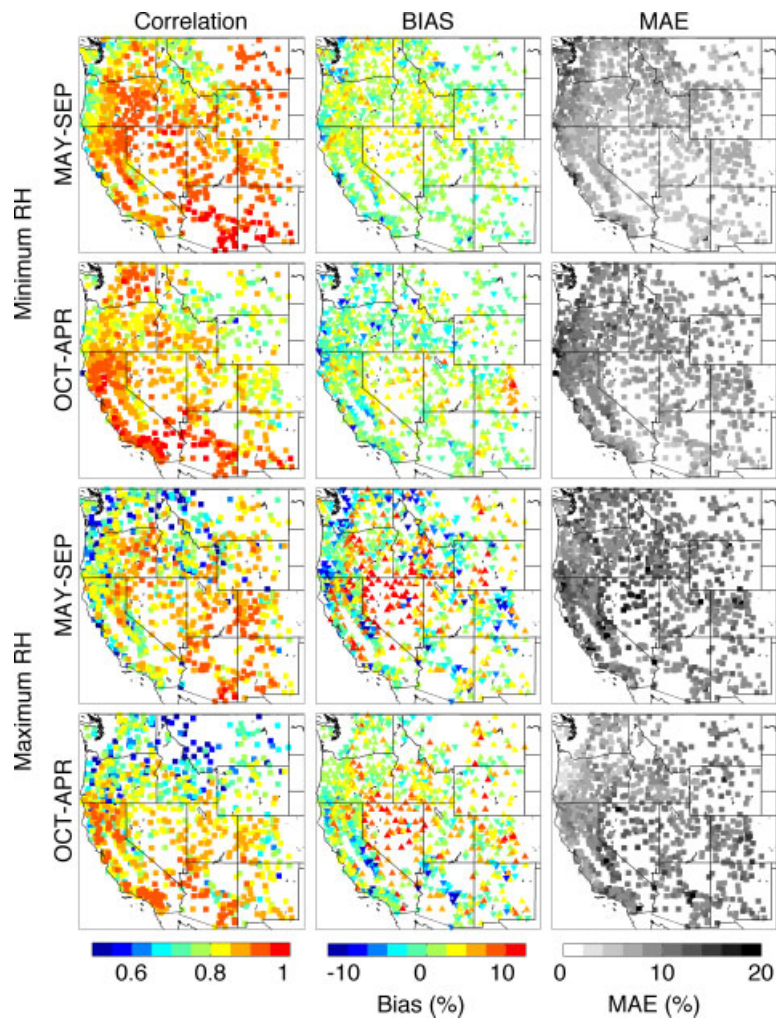
Figure 3. Validation statistics for daily maximum and minimum relative humidity for all RAWS and AgriMet data with at least 1000 valid observations. Bias is defined as gridded data minus station observations. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

to minimum temperatures, the juxtaposition of validation statistics at stations separately by less than 50 km, further suggests that RH is highly sensitive to station siting in varied terrain where local variations in minimum temperatures are terrain-driven (e.g. Weise *et al.*, 2010). By contrast, much stronger relationships were found for the daily average vapour pressure deficit at AWN and AgriMet stations (median correlation of 0.97, Figure S1, Table SII).

Validation statistics for precipitation revealed the strongest correlation ($r > 0.85$) during the cool seasons on the windward side of north-south oriented mountains encompassing the Cascades, Sierra Nevada and southern California's Peninsular-Transverse ranges (Figure 4, Table SI). Most precipitation received in these locales during the cool season is associated with organized synoptic-scale frontal systems, generally resolved by gauge-based observations and precipitation assimilation (e.g. Widmann and Bretherton, 2000). Conversely, areas that receive more of their precipitation via convective processes that often lack a cohesive spatial structure, generally exhibited lower correlations. Precipitation extremes were somewhat underestimated across

most sites (median bias of -5%), similar to results of other daily gridded precipitation datasets (e.g. Haylock *et al.*, 2008; Appendix B). Lower correlations and large positive biases were found during the cool season in the interior western United States. As most of the sites with significant positive biases are located in areas that regularly receive snow during the cool season, it is speculated that automated gauges that are not winterized under-report solid precipitation. More generally, positive biases for automated gauges are likely a manifestation of gauge under-catch that varies with wind speed, precipitation phase and intensity (e.g. Legates and DeLiberty, 1993). While efforts have been made to correct for these issues at first-order climate stations, under-catch in automated precipitation gauges shown here is likely real and non-trivial. By contrast, validation of daily precipitation from USHCN stations revealed similar correlations for the contiguous United States without any significant bias (Appendix A, Figure S4).

A direct comparison between gridded data and RAWS is problematic as the latter represents a 10-min average wind speed ending at 1300 LST, whereas gridded wind data is temporally disaggregated to hourly time scales
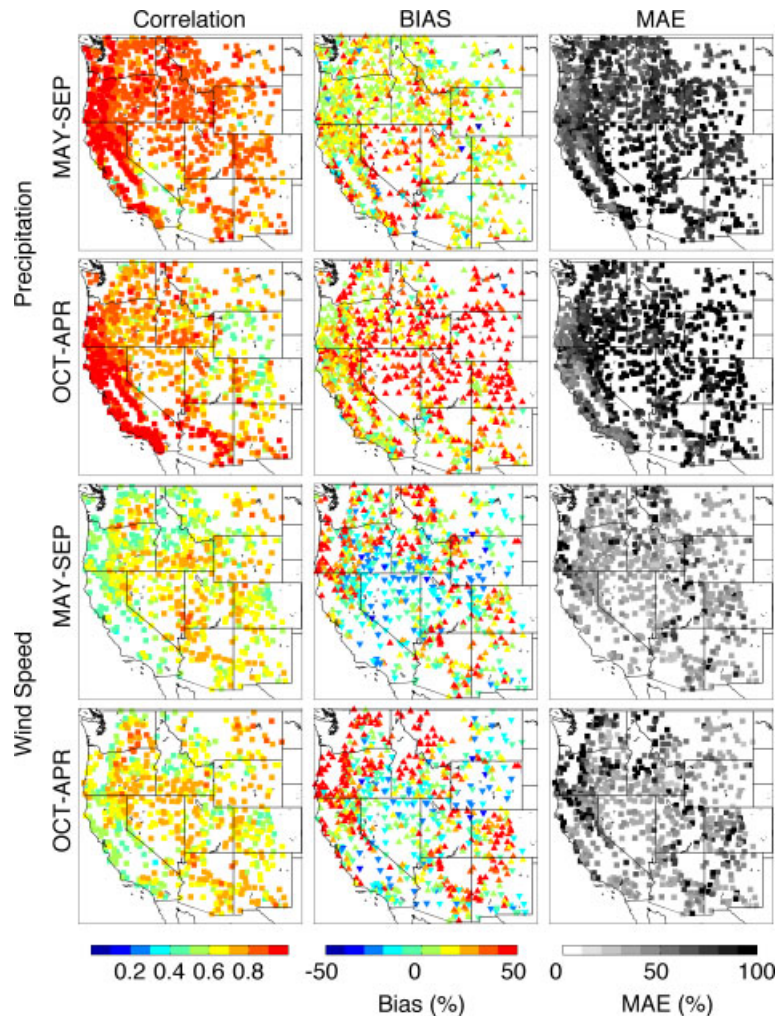
Figure 4. Validation statistics for daily precipitation and wind speed for all RAWS, AgriMet and AWN data with at least 1000 valid observations. Mean absolute error (MAE) is expressed as percent mean absolute error, bias is expressed as absolute percent error where positive values indicate an overprediction compared to station observations. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

from three-hourly average wind speed from NARR. Median correlation values of 0.54 and 0.52 were found for the cold and warm season, respectively (Figure 4, Table SI). Positive biases in wind speeds were found across forested areas, whereas correlation and bias were better captured in non-forested landscapes. Local physiography and surface roughness have been shown to significantly alter surface winds thereby resulting in differences between observed and interpolated wind fields (Wieringa, 1986). A more direct comparison of daily averaged wind speeds for AWN and AgriMet stations resulted in higher correlations (median correlation of 0.68 and 0.62, during the cold and warm season, respectively), although the gridded data generally overestimated wind speeds by 5–30%.

3.2. Validation of energy release component and ET_o

Gridded ERC exhibited strong correlations with correlations at over 90% of stations exceeding 0.9 (Figure S2, Table SII). Spatial patterns of correlation and bias for ERC correspond closely with biases and errors in meteorological variables, particularly RH. For example, positive biases in RH across much of the Great Basin resulted in a positive bias in fuel moisture and subsequently a negative bias in ERC, whereas negative biases in RH across parts of the northern Rockies and the coastal range of California and Oregon resulted in a negative bias in fuel moisture and a positive bias in ERC.

Estimates of ET_o are driven by daily temperature, vapour pressure deficit, wind speed and shortwave radiation. Missing surface downward shortwave radiation observations at several sites required the use of IDW interpolation to estimate ET_o and limit an independent comparison for radiation. Nonetheless, gridded daily ET_o correlated well to daily ET_o from observations (median $r = 0.90$, Figure S1), though showed positive biases across most sites (median bias +0.5 mm, Figure S1). Such biases stem from the aforementioned biases in primary meteorological variables. For example, the positive bias for ET_o in the Snake River Plain was a consequence of the gridded dataset overpredicting vapour pressure deficit, wind speed and maximum temperature. It is

important to point out here that the validation was performed for AgriMet observations in irrigated sites. Land use at these sites is posited to alter the surface energy budget by increasing the latent heat flux and decreasing maximum surface air temperatures, vapour pressure deficits, wind speeds and ET_o (e.g. Adegoke *et al.*, 2003; Jaksa *et al.*, 2011).

3.3. Comparison to inverse distance weighting interpolation

For most variables, the gridded dataset and inverse distance weighting (IDW) interpolation were found to have comparable skill (Table SI). Generally, correlation values of the two methods were not statistically distinguishable from one another; however, IDW interpolation generally outperformed the gridded data in regions of greater station density, and vice versa. The use of climatically aided interpolation with PRISM to develop the gridded dataset resulted in a lower MAE for temperature and precipitation compared to the IDW method. However, IDW interpolation generally outperformed gridded validation for RH. The commonality of RAWS station siting (e.g. southwest slopes) is speculated to favour the use of in-network IDW interpolation to predict site specific information at other RAWS. Nonetheless, lower correlations for maximum RH observed along the coast and in western Wyoming and Montana seen for IDW interpolation, reinforce the notion that maximum RH adheres to localized factors not represented by mesoscale reanalysis or neighbouring observations. For integrated metrics ERC and ET_o, the gridded data exhibited comparable skill to IDW interpolation. These results suggest that while the gridded surface meteorological dataset is prone to errors and biases, it was just as skillful in capturing site-specific data as neighbouring observations.

3.4. Interannual variability

A time series of ERC, averaged seasonally over the core fire season (15 July–15 September) and across the 26 RAWS and co-located grid points within the Northern Rocky Mountain Forest ecoprovince, from 1991 to 2010 is shown in Figure 5(a). The gridded ERC was highly correlated to observed ERC ($r = 0.93$), despite showing a positive bias that arises due to the gridded data under predicting RH. It is unclear whether these biases reflect true biases of the gridded data, or instead, biases associated with station siting; however, the bias appears to be fixed through time. While an intercomparison of trends between observed and gridded data is beyond the scope of this study, and limited by a short record, preliminary results suggest that summer (JJA) maximum temperatures at RAWS across the western US have warmed relative to the gridded data over the 20-year period of record, whereas the opposite is true of minimum temperatures. The origins of these differences are not apparent, but may be associated with station siting of both RAWS and data
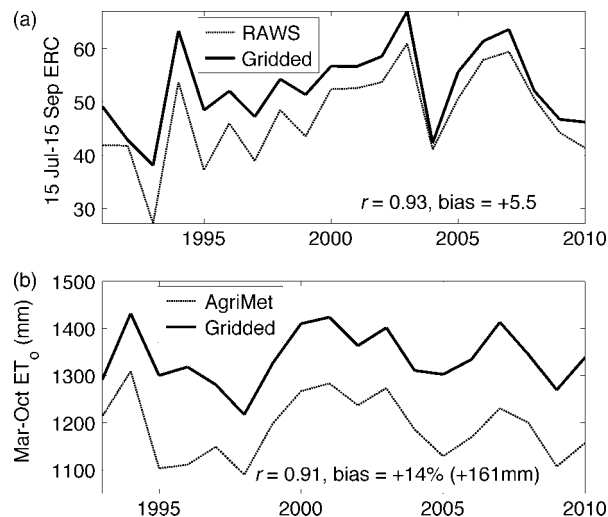

Figure 5. Time series of (a) fire season (15 Jul–15 Sep) Energy Release Component (ERC) averaged for 26 RAWS (dashed) and co-located gridded data (solid) within the Northern Rocky Mountain Forest ecoprovince 1991–2010, and (b) cumulative Mar–Oct reference evapotranspiration (*ETo*) averaged for seven AgriMet stations (dashed) and co-located gridded data (solid) in the Snake River Plain, Idaho, 1993–2010.

included in PRISM (e.g. Pielke *et al.*, 2007), as well as observed differential temperature trends in mountainous areas globally (e.g. Pepin and Lundquist, 2008; Pepin *et al.*, 2011).

Figure 5(b) shows the observed and modelled cumulative ET_o from March to October averaged over 7 AgriMet stations in the Snake River Plain for 1993–2010. The gridded ET_o tracked observed ET_o on an interannual basis (correlation of 0.91), and correlations for individual stations were also strong ($r = 0.84$–0.97). The positive bias in ET_o of about 14% over the March–October period follows the positive bias in vapour pressure deficit and wind speeds, hypothesized to be partially attributed to modifications in the surface energy budget associated with irrigation. While biases were apparent in both examples, the overall strong correlations suggest that the gridded data can be of value for multiyear studies that require long-term and continuous data.

4. Conclusions

The methodology detailed in this paper blends desirable attributes from observations and regional reanalysis to derive a high-resolution gridded surface meteorological dataset. This methodology can be applied in other regions where similar datasets exist, and can be updated as advances in observational networks and regional modelling occur. The validation of the gridded data across the challenging surface meteorological network of the western United States showed comparable skill to IDW interpolation from station observations. However, the gridded dataset is advantageous in that it provides spatially and temporally complete coverage across the continental

United States from 1979 to 2010 across a suite of meteorological variables needed for ecological modelling.

Biophysical models can be sensitive to errors and biases in meteorological forcing data (e.g. Morin and Thuiller, 2009). Four primary caveats of the dataset are provided to address data limitations for ecological modelling. First, high-resolution datasets do not necessarily equate to realism or direct application to point-scale observations. The use of the term 'high-resolution' is relative, as the 4-km gridded data may suffice for certain applications, yet be too coarse for other needs. Gridded data homogenize terrain and land-surface features within a grid cell, whereas station observations are point-based measurements often linked to microclimates that are not capable of being resolved by the methodology presented. This is particularly true for meteorological phenomena such as localized cold air drainage and its subsequent influence on RH and vapour pressure deficit. Additional means to incorporate detailed topoclimate information on terrain and synoptic meteorology may be employed to glean these patterns at richer spatial scales (e.g. Running *et al.*, 1987; Lundquist *et al*, 2008; Holden *et al.*, 2011). Secondly, surface wind speeds were bilinearly interpolated from NARR and do not account for the influence of terrain or the effect of frictional drag of forest canopy. Advanced statistical or dynamical methods that couple prevailing wind velocity, topography and vegetation are needed to facilitate higher-resolution wind fields. Thirdly, the effects of land-surface conditions (e.g. seasonal irrigation in arid and semi-arid areas) may not be reflected in the data and require additional calibration for use in certain applications (e.g. Huntington *et al.*, 2011). Finally, small-scale convective precipitation is not well resolved by either the observational network or atmospheric models, and may be subject to error. Additional high-density observational studies in regions of heterogeneous physiography and vegetation are required to further our understanding of these processes before they are fully and realistically incorporated into gridded meteorological datasets.

The advent of higher-resolution meteorological datasets in regions of complex terrain will likely continue in the foreseeable future. Advances in landscape-scale ecological modelling have been limited to suitable meteorological data. The ability to incorporate data that span space and time may help bridge the gap in interdisciplinary science and facilitate a means to better test larger-scale ecological hypotheses. For example, local-to-regional-scale data similar to that shown in the examples in Figure 5, can provide data for biophysical models, and possibly new insights into the role of weather and climate on wildfire potential and irrigation requirements. Likewise, high-quality gridded datasets are needed to statistically downscale future climate scenarios (e.g. Abatzoglou and Brown, 2011) and assess potential landscape impacts under a changing climate (e.g. Littell *et al.*, 2011).

5. Appendix A

Daily observations for 1218 stations from the United States Historical Climate Network (USHCN) covering the continental United States were used to supplement the validation performed for RAWS, AgriMet and AWN stations. The USHCN data is regarded as a higher-quality dataset and has undergone numerous quality assurance and quality control measures for temperature data. The validation procedure detailed in Section 2 was completed using USHCN data for 1979–2009.

Results show strong agreement across the United States with median correlation coefficients of 0.97 for both maximum and minimum temperature, and 0.74 for daily precipitation (Figures S3–4; Table SIII). Note that these values build in the annual cycle and thus are inflated relative to the primary analysis done for warm and cool seasons separately. Biases were generally quite small with a majority of stations having biases of $\pm 0.5\,°C$ and $\pm 5\%$ for temperature and precipitation, respectively. It should be noted, however, that because PRISM utilizes HCN data at the monthly time step, these results are not completely independent.

6. Appendix B

The ability of the gridded data to capture the magnitude of extremes in the observations was examined by estimating the 1st and 99th percentile for temperature, and the 99th percentile for daily precipitation. To facilitate a direct comparison, percentiles were only considered for days in which station observations were recorded. Maximum and minimum temperature extremes showed a slight warm bias ($+0.5 \pm 1\,°C$) for cool extremes (1st percentile) and cool bias ($-0.7 \pm 0.9\,°C$) for warm extremes (99th percentile). For daily precipitation extremes gridded data exhibited a dry bias (median bias of -5%), suggesting that the heaviest precipitation events were slightly

under-predicted relative to observations. These results are consistent with previous studies that have shown gridded meteorological data tends to reduce the magnitude of extremes when compared to station observations (Haylock *et al.*, 2008).

A contingency table for days with measureable precipitation (>0.25 mm) was used to complement the validation of precipitation. For both the observed and gridded data, days with measurable precipitation were identified and statistics on the hit rate (proportion correct), false alarm rate (precipitation shown in gridded data, but none observed) and false negatives (no precipitation shown in gridded data, not precipitation was observed). The median hit rate was 82%, false alarm rate 8%, and false positive rate 10%. Spatially, the largest regional discrepancies were seen in the central Rockies, with the most agreement in regions where the correlation shown in Figure 3 is high. A contingency table for days with at least 2.5 mm of precipitation exhibited a significant improvement, with a median hit rate of 92%, false alarm rate 3%, and false positive rate 5%. These results suggest that errors are most acute for light precipitation, particularly in regions dominated by convective precipitation regimes.

References

Abatzoglou JT, Brown TJ. 2011. A Comparison of Statistical Downscaling Methods Suited for Wildfire Applications. *International Journal of Climatology*. DOI: 10.1002/joc.2312.

Abatzoglou JT, Redmond KT, Edwards LM, 2009. Classification of Regional Climate Variability in the State of California. *Journal of Applied Meteorology and Climatology* **48**(8): 1527–1541.

Adegoke JO, Pielke RA, Eastman J, Mahmood R, Hubbard KG. 2003. Impact of irrigation on midsummer surface fluxes and temperature under dry synoptic conditions: A regional atmospheric model study of the U.S. high plains. *Monsoon Weather Review* **131**: 556–564.

Allen RG, Pereira LS, Raes D, Smith M. 1998. Crop evapotranspiration, guidelines for computing crop water requirements. FAO Irrig. and Drain. Paper 56, Food and Agricultural Organization of the United Nations. Rome, Italy. 300.

Bailey RG. 1995. *Ecosystem Geography:* New York: Springer-Verlag.

Brown TJ, Horel JD, McCurdy GD, Fearon MG. 2011. What is the Appropriate RAWS Network? CEFA Report 11-01, April 2011, 97.

Cosgrove BA, Lohmann D, Mitchell KE, Houser PR, Wood EF, Schaake J, Robock A, Marshall C, Sheffield J, Luo L, Duan Q, Pinker RT, Tarpley JD, Higgins RW, Meng J. 2003. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *Journal of Geophysical Research* **108**(D22): 8842, DOI: 10.1029/2002JD003118.

Daly C. 2006. Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology* **26**: 707–721.

Daly C, Conklin DR, Unsworth MH. 2009. Local atmospheric decoupling in complex topography alters climate change impacts. *International Journal of Climatology*. DOI: 10.1002/joc.2007.

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PA. 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology* DOI: 10.1002/joc.1688.

Deeming JE, Burgan RE, Cohen JD. 1978. The National Fire-Danger Rating System – 1978. Revisions to the 1978 National Fire-Danger Rating System. General Technical Report INT-39. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station; 63.

Di Luzio M, Johnson GL, Daly C, Eischeid J, Arnold JG. 2008. Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. *Journal of Applied Meteorology and Climatology* **47**: 475–497.

Durre I, Menne MJ, Gleason BE, Houston TG, Vose RS. 2010. Comprehensive Automated Quality Assurance of Daily Surface Observations. *Journal of Applied Meteorology and Climatology* **49**: 1615–1633, DOI: 10.1175/2010JAMC2375.1.

Eischeid JK, Pasteris PA, Diaz HF, Plantico MS, Lott NJ. 2000. Creating a serially complete, national daily time series of temperature and precipitation for the western United States. *Journal of Applied Meteorology* **39**: 1580–1591.

Hamlet AF, Lettenmaier DP. 2005. Production of temporally consistent gridded precipitation and temperature fields for the continental U.S. *Journal of Hydrometeorology* **6**(3): 330–336.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily highresolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research* **113**: D20119, DOI: 10.1029/2008JD010201.

Holden ZA, Abatzoglou JT, Baggett LS, Luce C. 2011. Empirical downscaling of daily minimum air temperature at very fine resolutions in complex terrain. *Agricultural and Forest Meteorology*. DOI: 10.1016/j.agrformet.2011.03.01.

Huntington JL, Szilagyi J, Tyler SW, Pohll GM. 2011. Evaluating the complementary relationship for estimating evapotranspiration from arid shrublands. *Water Resources Research* **47**: W05533, DOI: 10.1029/2010WR009874.

Jaksa WA, Nuss KM, Sridhar V. 2011. *Effects of coupling in understanding the surface energy balance in the Snake River basin, Idaho*. World Environmental and Water Resources Congress, Palm Springs, California, May 22–26, 2011.

Jensen ME, Burman RD, Allen RG. 1990. *Evapotranspiration and irrigation water requirements. ASCE Manuals and Reports on Engineering Practices* No. 70. American Society for Civil Engineers: New York, 360.

Jones PD, New M, Parker DE, Martin S, Rigor IG. 1999. Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics* **37**(2): 173–199, DOI: 10.1029/1999RG900002.

Kimball J, Running SW, Nemani RR. 1997. An improved method for estimating surface humidity from daily minimum temperature. *Agricultural and Forest Meteorology* **85**: 87–98.

Legates DR, DeLiberty TL. 1993. Precipitation measurement biases in the United States. *Water Resources Bulletin* **29**: 855–861.

Lin X, Hubbard KG. 2004. Uncertainties of Derived Dewpoint Temperature and Relative Humidity. *Journal of Applied Meteorology* **43**: 821–825.

Littell JS, Peterson DL, Millar CI, O'Halloran K. 2011. U.S. National Forests adapt to climate change through science-management partnerships. *Climatic Change*. DOI: 10.1007/s10584-011-0066-0.

Lundquist JD, Pepin NC, Rochford C. 2008. Automated algorithm for mapping regions of cold-air pooling in complex terrain, *Journal of Geophysical Research* **113**: D22107, DOI: 10.1029/2008JD009879.

Meentemeyer RK, Cunniffe NJ, Cook AR, Filipe JAN, Hunter RD, Rizzo DM, Gilligan CA. 2011. Epidemiological modeling of invasion in heterogeneous landscapes: Spread of sudden oak death in California (1990-030). *Ecosphere* **2**(2): 17.

Mesinger F, DiMego G, Kalnay E, Mitchell K, Shafran PC, Ebisuzaki W, Jovic D, Woollen J, Rogers E, Berbery EH, Ek MB, Fan Y, Grumbine R, Higgins W, Li H, Lin Y, Manikin G, Parrish D, Shi W. 2006. North American Regional Reanalysis. *Bulletin of the American Meteorological Society* **87**: 343–360.

Milly PCD, Dunne KA. 2011. On the Hydrologic Adjustment of Climate-Model Projections: The Potential Pitfall of Potential Evapotranspiration. *Earth Interactions* **15**: 1–14, DOI: 10.1175/2010EI363.1.

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 JD, 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 GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research* **109**: D07S90, DOI: 10.1029/2003JD003823.

Morin X, Thuiller W. 2009. Comparing niche- and process based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology* **90**: 1301–1313.

Pepin NC, Lundquist JD. 2008. Temperature trends at high elevations: Patterns across the globe. *Geophysical Research Letters* **35**: L14701, DOI: 10.1029/2008GL034026.

Pepin NC, Lundquist JD, Daly C. 2011. The Influence of Surface/Free-Air Decoupling on Temperature Trend Patterns in the Western U.S. *Journal of Geophysical Research* **116**: D10109, DOI: 10.1029/2010JD014769.

Pepin NC, Seidel DJ. 2005. A global comparison of surface and free-air temperatures at high elevations. *Journal of Geophysical Research* **110**: D03104, DOI: 10.1029/2004JD005047.

Peterson TC, Vose R, Schmoyer R, Razuvaev V. 1998. Global Historical Climatology Network (GHCN) quality control of monthly temperature data. *International Journal of Climatology* **18**(11): 1169–1179.

Pielke R, Nielsen-Gammon J, Davey C, Angel J, Bliss O, Doesken N, Cai M, Fall S, Niyogi D, Gallo K, Hale R, Hubbard KG, Lin XM, Li H, Raman S. 2007. Documentation of uncertainties and biases associated with surface temperature measurement sites for climate change assessment. *Bulletin of the American Meteorological Society* **88**: 913–928.

Pinker RT, Tarpley JD, Laszlo I, Mitchell KE, Houser PR, Wood EF, Schaake JC, Robock A, Lohmann D, Cosgrove BA, Sheffield J, Duan Q, Luo L, Higgins RW. 2003. Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) project. *Journal of Geophysical Research* **108**(D22): 8844, DOI: 10.1029/2002JD003301.

Régnière J, Bentz BJ. 2007. Modeling cold tolerance in the mountain pine beetle, Dendroctonus ponderosae. *Insect Physiology* **53**: 559–572.

Root TL, Schneider SH. 2003. Strategic Cyclical Scaling: Bridging Five Orders of Magnitude Scale Gaps in Climatic and Ecological Studies. In *Scaling Issues in Integrated Assessment, Lisse*. Rotmans J, Rothman DS (eds). Swets and Zeitlinger Publishers: The Netherlands, 179–204.

Running SW, Nemani RR, Hungerford RD. 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal of Forest Research* **17**: 472–483.

Schwartz MD, Reiter BE. 2000. Changes in North American spring. *International Journal of Climatology* **20**: 929–932.

Stockle CO, Donatelli M, Nelson R. 2003. CropSyst, a cropping systems simulation model. *European Journal of Agronomy* **18**: 289–307.

Thornton PE, Running SW, White MA. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology* **190**: 214–251.

Weise DR, Stephens SL, Fujioka FM, Moody TJ, Benoit J. 2010. Estimation of fire danger in Hawai'i using limited weather data and simulation. *Pacific Science* **64**(2): 199–220.

Widmann M, Bretherton CS. 2000. Validation of Mesoscale Precipitation in the NCEP Reanalysis Using a New Gridcell Dataset for the Northwestern United States. *Journal of Climate* **13**: 1936–1950.

Wieringa J. 1986. Roughness-dependent geographical interpolation of surface wind speed averages. *Quarterly Journal of the Royal Meteorological Society* **112**: 867–889.

Willmott CJ, Robeson SM. 1995. Climatologically aided interpolation (CAI) of terrestrial air temperature. *International Journal of Climatology* **15**(2): 221–229.

Willmott CJ, Rowe CM, Philpot WD. 1985. Small-Scale Climate Maps: A Sensitivity Analysis of Some Common Assumptions Associated with Grid-point Interpolation and Contouring. *American Cartographer* **12**: 5–16.

WMO. 2008. *Guide to Meteorological Instruments and Methods of Observation*, 7th edn. World Meteorological Organization: 677.