INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2008)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1688

RMetS
ROYAL METEOROLOGICAL SOCIETY

Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States

Christopher Daly,[a]* Michael Halbleib,[a] Joseph I. Smith,[a] Wayne P. Gibson,[a] Matthew K. Doggett,[a] George H. Taylor,[a] Jan Curtis[b] and Phillip P. Pasteris[b]

[a] *Oregon State University, Corvallis, OR 97331, USA*

[b] *USDA Natural Resources Conservation Service Water and Climate Center, Portland, OR 97232, USA*

ABSTRACT: Spatial climate data sets of 1971–2000 mean monthly precipitation and minimum and maximum temperature were developed for the conterminous United States. These 30-arcsec (~800-m) grids are the official spatial climate data sets of the U.S. Department of Agriculture. The PRISM (Parameter-elevation Relationships on Independent Slopes Model) interpolation method was used to develop data sets that reflected, as closely as possible, the current state of knowledge of spatial climate patterns in the United States. PRISM calculates a climate–elevation regression for each digital elevation model (DEM) grid cell, and stations entering the regression are assigned weights based primarily on the physiographic similarity of the station to the grid cell. Factors considered are location, elevation, coastal proximity, topographic facet orientation, vertical atmospheric layer, topographic position, and orographic effectiveness of the terrain. Surface stations used in the analysis numbered nearly 13 000 for precipitation and 10 000 for temperature. Station data were spatially quality controlled, and short-period-of-record averages adjusted to better reflect the 1971–2000 period.

PRISM interpolation uncertainties were estimated with cross-validation (C-V) mean absolute error (MAE) and the 70% prediction interval of the climate–elevation regression function. The two measures were not well correlated at the point level, but were similar when averaged over large regions. The PRISM data set was compared with the WorldClim and Daymet spatial climate data sets. The comparison demonstrated that using a relatively dense station data set and the physiographically sensitive PRISM interpolation process resulted in substantially improved climate grids over those of WorldClim and Daymet. The improvement varied, however, depending on the complexity of the region. Mountainous and coastal areas of the western United States, characterized by sparse data coverage, large elevation gradients, rain shadows, inversions, cold air drainage, and coastal effects, showed the greatest improvement. The PRISM data set benefited from a peer review procedure that incorporated local knowledge and data into the development process. Copyright © 2008 Royal Meteorological Society

KEY WORDS physiography; spatial climate; PRISM; interpolation; temperature; precipitation; rain shadow; cold air drainage; temperature inversion; coastal proximity; elevation; Daymet; WorldClim

Received 11 June 2007; Revised 27 December 2007; Accepted 28 December 2007

1. Introduction

Spatial climate data sets in digital form are currently in great demand. The most commonly used spatial climate data sets are gridded estimates of mean daily minimum and maximum temperature and total precipitation on a monthly time step, averaged over a nominal 30-year period. The demand for these data sets has been fueled in part by the linking of geographic information systems (GIS) to a variety of models and decision support tools, such as those used in agriculture, engineering, hydrology, ecology, and natural resource conservation.

Spatial climate data are often key drivers of computer models and statistical analyses, which form the basis for scientific conclusions, management decisions, and other important outcomes. It is therefore imperative that these data sets provide a realistic representation of the major forcing factors that affect spatial climate patterns. To achieve this high level of realism, methods used to create the data sets must explicitly account for these factors. A detailed discussion of these factors is given in Daly (2006), and a brief overview is provided here.

General circulation patterns are largely responsible for large-scale climate variations, and include the positions of storm tracks, prevailing wind directions, monsoonal circulations, and other defining features of a region's climate. It is assumed that most of these patterns occur at scales large enough to be adequately reflected in the station data, and therefore are not explicitly accounted for by interpolation methods. Physiographic features on the earth's surface, namely water bodies and terrain, modulate these large-scale climate patterns. Water bodies provide moisture sources for precipitation, and create

* Correspondence to: Christopher Daly, PRISM Group, Department of Geosciences, 2000 Kelley Engineering Center, Oregon State University, Corvallis, OR 97331, USA. E-mail: daly@nacse.org

complex temperature gradients along coastlines and in adjacent inland areas. Terrain effects include the direct effect of altitude on climate conditions, the blockage and uplift of major flow patterns by terrain barriers, and cold air drainage and pooling in valleys and depressions.

The relationship between elevation and precipitation is highly variable, but precipitation generally increases with elevation (Oke, 1978; Barry and Chorley, 1987). Exceptions are when terrain rises above the height of a moist boundary layer or trade wind inversion (Mendonca and Iwaoka, 1969). Blocking and uplifting of moisture-bearing winds amplifies precipitation on windward slopes, especially those with steep windward inclines, and can sharply decrease it on leeward slopes downwind, producing rain shadows (Smith, 1979; Daly *et al*., 1994, 2002).

Temperature exhibits a strong, predictable decrease with elevation when the atmosphere is well mixed, such as occurs on summer days in inland areas (e.g. Willmott and Matsuura, 1995). The main summer exception is in coastal regions with well-defined marine layers, where maximum temperatures often increase with elevation above the marine inversion. Winter temperatures, and minimum temperatures in all seasons, have a more complex relationship with elevation. In the absence of solar heating or significant winds to mix the atmosphere, temperatures stratify quickly, and cool, dense air drains into local valleys and depressions to form pools that can be hundreds of metres thick (Geiger, 1964; Hocevar and Martsolf, 1971; Bootsma, 1976; Gustavsson *et al*., 1998; Lindkvist *et al*., 2000; Daly *et al*., 2003). This results in temperature inversions, in which temperature sharply increases, rather than decreases, with elevation (Clements *et al*., 2003). In Polar regions, widespread regional inversions hundreds of kilometres in extent can dominate wintertime temperature patterns (Milewska *et al*., 2005; Simpson *et al*., 2005). Terrain can also serve as a barrier between air masses, creating sharply defined horizontal temperature gradients.

Coastal effects on temperature are most noticeable in situations where the water temperature is significantly different from the adjacent land temperature (Haugen and Brown, 1980; Atkinson and Gajewski, 2002). Along the California coastline during summer, the contrast between the cool Pacific Ocean and the adjacent warm land mass can create daytime air temperature gradients of more than 10 °C in just a few kilometres across the coastal strip (Daly *et al*., 2002).

The above factors are most important at scales from less than 1 km to 50 km or more (Daly, 2006). Several additional spatial climate-forcing factors are most important at relatively small scales of less than 1 km, but can have influences at larger scales as well. These factors include slope and aspect (McCutchan and Fox, 1986; Barry, 1992; Bolstad *et al*., 1998; Lookingbill and Urban, 2003; Daly *et al*., 2007), riparian zones (Brosofske *et al*., 1997; Dong *et al*., 1998; Lookingbill and Urban, 2003), and land use/landcover (Davey and Pielke, 2005). Land use/landcover variations are a major consideration in the spatial representativeness of climate stations at much larger scales. For example, stations located near parking lots, buildings, or other heat-absorbing surfaces may have very different temperature regimes than those in open grasslands or heavily vegetated areas (Davey and Pielke, 2005). In data-sparse regions, a single station, and its particular land use/landcover regime, may influence the interpolated climate conditions for tens of kilometres around that station.

This paper describes the development of spatial climate data sets of 1971–2000 mean monthly total precipitation and daily minimum and maximum temperature across the conterminous United States, using methods that strive to account for the major physiographic factors influencing climate patterns at scales of 1 km and greater. These data sets, created at 30-arcsec (∼800-m) grid resolution, were commissioned by the U.S. Department of Agriculture through the Natural Resources Conservation Service (USDA-NRCS) to serve as the official spatial climate data sets of the USDA. They are updates of the 2.5-arcmin (∼4-km) 1961–1990 spatial climate data sets developed in the 1990s (USDA-NRCS, 1998). The new data sets were interpolated using the latest version of the Parameter-elevation Regressions on Independent Slopes Model (PRISM) climate mapping system. Section 2 of this paper describes the study area and the digital elevation model (DEM) used. In Section 3, the preparation of station data is described. Section 4 provides an overview of the PRISM climate mapping system for this application, and summarizes the modelling, review, and revision process. Section 5 presents the resulting gridded data sets, discusses model performance, and compares and contrasts the PRISM data sets to two other spatial climate data sets. Concluding remarks are given in Section 6.

2. Study area and digital elevation model

Climate data sets were developed at 30-arcsec resolution in geographic (lat./long.) coordinates. A 30-arcsec grid cell is approximately 900 × 700 m at 40°N latitude, and is referred to as '800 m' after the discussion of the elevation grid in the next section. The boundaries of the grid were 22 and 50°N and 65 and 125 °W. Memory, CPU, and model parameterization considerations required that interpolation be performed separately in three regions: western, central, and eastern United States, and the resulting grids merged to form a complete conterminous U.S. grid.

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

The DEM was the single most important grid input to the interpolation; it provided the independent variable for the PRISM elevation regression function and served as

DOI: 10.1002/joc

Table I. Methods for preparation of PRISM input grids, part 1.

Model input grid	Description	Climate element/region	Units	Step	Starting grid	Operation
U.S. boundary and actual coastline mask	Mask of conterminous U.S. boundaries and coastlines	All/All	Binary (1 or 0)	–	U.S. boundary vector file	Convert vectors to rasters
Generalizedcoastline and modeling mask	Mask for modelling region, generalized coastlines	All/All	Binary (1 or 0)	–	U.S. boundary and actual coastline mask	Grow mask beyond U.S. boundaries; fill in bays and inlets
Coastal proximity TCE	Distance from coastline, with bays and inlets treated as transition areas	Max and min temp, central and eastern US	Kilometres (0–50)	1	US boundary mask	Distance from coastlines
				2	Modelling mask	Distance from coastlines
				3	Steps 1 and 2 grids	Average the two grids
Coastal proximity TW	Index of marine penetration through complex terrain	Max and min temp, western US	Index (0–1000)	1	US boundary mask and 800-m temperature DEM	Coastal advection model (Daly *et al*., 2003)
				2	Generalized coastline mask	Coastal advection model (Daly *et al*., 2003)
				3	Steps 1 and 2 grids	Step 3 = 0.75(Step 1) + 0.25 (Step 2)
Coastal proximity PCE	Distance from generalized coastline	Precip, central and eastern US	Kilometres (0–90)	–	Modelling mask	Distance from coastlines
Coastal proximity PW	Moisture index	Precip, western US	Index (0–1000)	–	Modelling mask and 800 m precipitation DEM	Coastal trajectory model (Daly *et al*., 2003)

the basis for most of the other physiographic input grids to PRISM (Section 4). The DEM used in this study was derived from the 3-arcsec (∼80-m) National Elevation Database (NED) (NED; http://gisdata.usgs.gov/NED/). The NED elevation data was found to be superior to other DEMs for the conterminous United States, such as GTOPO30, Digital Terrain Elevation Database (DTED), and Space Shuttle Radar Topography Mission (SRTM), for this application. The GTOPO30 and DTED DEMs had noticeable elevation breaks, or 'seams' along some of the original U.S.Geological Survey (USGS) quad boundaries. The SRTM DEM had numerous grid cells with missing data in areas of steep terrain, and possessed significant ground clutter associated with low microwave signal-to-noise ratios.

A 30-arcsec (∼800-m) version of the NED DEM was derived from the 3-arcsec (∼80-m) NED DEM by applying a modified Gaussian filter (Barnes, 1964). This filter provides a truly circular averaging area (as opposed to a typical rectangular block, which causes distortion along the diagonals) and weights the surrounding grid cells in a Gaussian, or normal, distribution with distance, which better preserves local detail at and around the central grid cell than does a uniformly weighted average. This 30-arcsec (∼800-m) DEM was used for interpolating maximum and minimum temperature. An additional filtering step was performed to obtain a DEM suitable for precipitation interpolation. The Gaussian filter was applied to the 30-arcsec (∼800-m) DEM to filter out terrain features up to 3.75 arcmin (∼7 km) in extent, while retaining the 30-arcsec (∼800-m) grid resolution. The direct effects of elevation on precipitation appear to be most important at scales of 5–10 km or greater, owing to a number of mechanisms, including the advective nature of moisture-bearing airflow, the viscosity of the atmosphere, delays between initial uplift and subsequent rainout, and the movement of air around terrain obstacles (Daley, 1991; Daly *et al*., 1994; Funk and Michaelsen, 2004; Sharples *et al*., 2005). In this paper, the two DEMs are hereafter referred to as the temperature and precipitation DEMs (Table II).

3. Station data

3.1. Data sources

Data from surface stations, numbering nearly 13 000 for precipitation and nearly 10 000 for minimum and

DOI: 10.1002/joc

Table II. Methods for preparation of PRISM input grids, part 2.

Model input grid	Description	Climate element	Units	Step	Starting grid	Operation	Radius of influence (km)	Cell weighting profile
800-m Temperature DEM	DEM for temperature interpolation	Min. and max. temperature	m	–	80-m NED DEM	Average, and coarsen resolution to 800 m	1.3	Gaussian (Barnes, 1964)
800-m Precipitation DEM	DEM for precipitation interpolation	Precipitation	m	–	800-m Temperature DEM	Average	7	Gaussian (Barnes, 1964)
Topographic facets	Areas of consistent aspect	Min. and max. temp., precipitation	Direction on eight compass points	1	800-m Precipitation DEM	Average	0.8, 12, 24, 36, 48, 60	Gaussian (Barnes, 1964)
				2	Step 1 grids	Facet (Gibson *et al.*, 1997)	0.8, 12, 24, 36, 48, 60	Uniform
Potential wintertime inversion height	Estimates altitude of boundary between layer 1 (boundary layer) and layer 2 (free atmosphere)	Min. temperature all year; max. temp. in winter (Nov–Mar)	m	1	800-m Temperature DEM	Minimum	19	–
				2	Step 1 grid	Average	44	Uniform
				3	Step 2 grid	Add 250 m	–	–
Moist boundary layer height	Estimates altitude of boundary between layer 1 (moist boundary layer) and layer 2 (drier atmosphere above)	Precip. in western US	m	1	County boundaries	Counties west of the Cascade crest in OR and WA given layer 1 height of 2200 m; all other areas 4000 m	–	–
Topographic index	Estimates how much higher pixel is than surrounding neighbourhood	Min. temp. all year; max. temp. in winter (Nov–Mar)	m	1	800-m Temperature DEM	Minimum	14	–
				2	Step 1 grid	Average	15	Uniform
				3	Step 2 grid	Subtract step 2 grid from DEM	–	–
Effective terrain height	Estimates orographically effective profile of terrain features	Precipitation	m	1	800-m Precipitation DEM	Minimum	22	–
				2	Step 1 grid	Average	11	Uniform
				3	Step 2 grid	Subtract step 2 grid from DEM	–	–
				4	Step 3 grid	Average	15	Uniform

DOI: 10.1002/joc

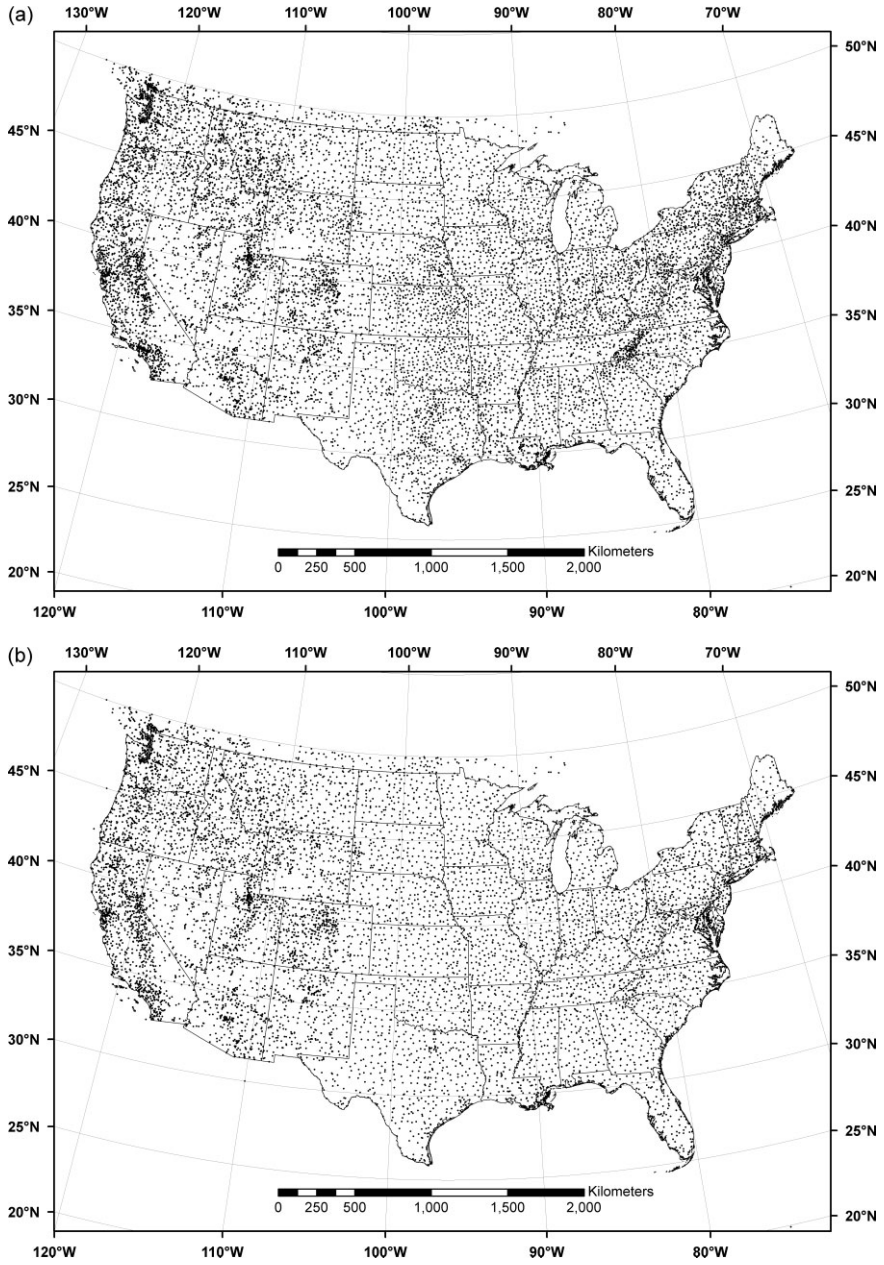

Figure 1. Locations of (a) surface precipitation stations and (b) surface temperature stations used in the interpolation.

maximum temperature, were obtained from a variety of sources with an eye towards creating the most comprehensive data set possible. These sources, summarized in Table III, included National Weather Service Cooperative Observer Program (COOP) and Weather Bureau Army Navy (WBAN) stations (http://cdo.ncdc.noaa.gov/CDO/cdo); USDA NRCS Snow Telemetry (SNOTEL) and snowcourses (http://www.wcc.nrcs.usda.gov/snow/); USDA Forest Service and Bureau of Land Management Remote Automatic Weather Stations (RAWS; http://www.wrcc.dri.edu/); California Data Exchange Center (CDEC) stations from a variety of networks (CDEC; http://cdec.water.ca.gov); Bureau of Reclamation Agrimet sites (http://www.usbr.gov/pn/agrimet); Environment Canada stations south of 50N, (EC; http:// climate.weatheroffice.ec.gc.ca); and miscellaneous stations from Reynolds Creek Experimental Watershed, ID; H.J. Andrews Experimental Forest, OR; Nevada Dept. of Water Resources; USDA Forest Service; USGS; and other local networks (Figures 1(a) and (b)). RAWS precipitation data were used only for the months of May–September because many of these stations are not maintained in winter and their unheated tipping-bucket gauges do not measure snowfall well and are subject to freezing. To better define temperature profiles in complex situations, such as temperature inversions, upper-air temperature grid points for the western and eastern United States were obtained from the Global Upper-Air Climatic Atlas, which consists of European Center for Medium Range Weather Forecasting (ECMWF) model grid point data at 2.5° resolution for the period 1980–1991(http://ols.nndc.noaa.

DOI: 10.1002/joc

Table III. Number of stations used in the mapping process, and native averaging interval. See text for definitions of acronyms.

Source	Native averaging interval	Precipitation	Maximum temperature	Minimum temperature
COOP	Daily	10 784	7663	7697
SNOTEL	Daily	645	623	636
Snowcourse	1961–90 annual mean	324	0	0
RAWS	Hourly	259	787	842
CDEC	Daily, monthly, annual	156	223	238
Agrimet	Daily	51	58	59
EC (Canada)	1961–90 monthly mean	483	326	324
Miscellaneous	Various	142	3	3
Estimated	1971–2000 monthly mean	93	31	3
Upper-air	1980–1995 monthly mean	0	69	69
Total		12 937	9783	9871

gov/plolstore/plsql/olstore.prodspecific?prodnum=C00338-CDR-S0001). The 500 hPa (∼5500 m) level was chosen for the western United States, and the 700 hPa (∼3000 m) for the eastern United States. (Upper-air data were not needed in the central United States because of a lack of elevated terrain.) These levels were sufficiently far above the highest terrain features to minimize errors involved in estimating surface temperatures from free air values. As such, the upper-air values served as distant 'anchor points' for the temperature–elevation regression functions.

3.2. Calculation of monthly values and quality control

As shown in Table III, station data were obtained in a wide variety of time steps. Hourly station data were aggregated to create daily values corresponding to midnight–midnight Local Standard Time (LST). Observations originally in Coordinated Universal Time (UTC, from French translation) were converted to LST. The offset between UTC and LST was approximated by dividing the station longitude by the factor 15°/h and rounding to the nearest hour (e.g. a station at −125° longitude would get an −8 h offset from UTC). Daily maximum and minimum temperatures were the maximum and minimum of the hourly temperature observations for the 24-h local period. Daily precipitation was the sum of the hourly incremental precipitation, calculated from hourly accumulations by subtracting the previous hour's total from the current hour's total. Negative incremental amounts (caused by evaporation, thermal expansion, resetting of accumulations, etc.) were set to missing. Daily values based on hourly observations were subject to the requirement that at least 18 of 24 observations must be non-missing; fewer than 18 non-missing hourly observations resulted in a missing daily value. Daily observations and those based on hourly observations were then aggregated to create monthly precipitation totals and mean maximum and minimum temperatures. A minimum of 85% of non-missing daily values were required for a monthly value to be non-missing. Our 85% data completeness criterion is not dissimilar to those used by the NCDC (NOAA-NCDC, 2003) and the World Meteorological Organization (WMO, 1989) in developing monthly temperature and precipitation statistics.

Monthly precipitation amounts were estimated for annual-only stations by searching for a nearby long-term station with non-missing monthly totals that would be most likely to have similar seasonality. Most of the annual-only stations were located in the mountains of the western United States, where precipitation seasonality can vary dramatically with elevation. Therefore, a three-dimensional (3D) distance measure was calculated, in which 1 km of vertical (elevation) distance was equated with 100 km of horizontal distance. The 100 : 1 ratio was found to adequately consider elevation as well as distance in choosing the optimal station. The annual-only total was multiplied by each of the chosen station's monthly percent-of-annual to generate the corresponding monthly total for the annual-only station.

Stations with data available at different time steps for different periods were combined at the monthly time scale to create the longest possible period of record (POR). Data from some stations were available in more than one time step for periods that overlapped temporally. When overlap occurred, the precedence of data selection was daily, monthly, and hourly. For example, if a station had daily precipitation data from 1972–2000 and monthly precipitation data from 1952–1974, monthly data would have been used from 1952–1971 and daily data from 1972–2000.

Initial range checking was performed for precipitation and maximum and minimum temperatures. Precipitation observations were checked for negative and extreme values, while temperature observations were checked for extreme maximum and minimum values and for maximum temperature less than minimum temperature. Extreme value thresholds were selected on the basis of the state in which the observing station resided and, in the case of temperature, the month in which the observation was made (http://www.ncdc.noaa.gov/oa/climate/severeweather/extremes.html). The threshold for extreme

DOI: 10.1002/joc

precipitation was chosen to be 115% of the state record 24-h precipitation. The threshold for extreme maximum temperature was chosen to be 3 °C above the state monthly record maximum temperature. Likewise, the threshold for extreme minimum temperature was chosen to be 3 °C below the state monthly recorded minimum temperature. The additional allowances were to accommodate potential new records.

Daily precipitation observations and hourly incremental amounts derived from accumulation observations were subject to both negative and extreme value checks, while monthly precipitation observations were checked for negative values only. Monthly and daily maximum and minimum temperatures, as well as hourly air temperatures, were subject to maximum and minimum temperature extreme checks. Observations that failed any of the above tests were set to missing.

Station elevations were checked for consistency against 800-m DEM elevations at the given station locations. Elevation discrepancies of more than 200 m were investigated, and either the station location or elevation corrected as a result.

Several spatial quality control (QC) procedures were conducted on the monthly data. In an initial QC screening step, monthly averages for 1971–2000 were calculated for stations having data during this time period. Stations not having data during 1971–2000 had historical averages calculated from their entire POR. These averages were tested for spatial consistency using the ASSAY QC system, a version of PRISM that estimates data for specific station locations and compares them to the observed values (Daly *et al*., 2000; Gibson *et al*., 2002). Averages failing the ASSAY QC check were immediately omitted from further consideration if (1) they were RAWS or CDEC stations, (2) they had less than three years of historical data, or (3) had three or more years of historical data but had fewer than four consecutive months of non-missing data during those years. These stations were considered to be at highest risk for poor quality.

In the second spatial QC step, all individual monthly values from the remaining stations were tested for spatial consistency using the ASSAY QC system; values failing this test were set to missing. The remaining monthly values were averaged and subjected to adjustment, if needed, as described in Appendix A.

In the third spatial QC step, all SNOTEL temperature data were subjected to a recently developed spatial QC system for temperature data from this network. Details on the operation of this system are available in Daly *et al*. (2005). Temperature data subjected to this QC process were first passed through the aforementioned range checks. SNOTEL temperature data were also tested for extended periods of unchanging values (flatliners); temperatures remaining unchanged (less than 0.1 °C daily difference) for longer than ten consecutive days were set to missing. In addition, temperature values remaining at exactly 0.0 °C for two or more consecutive days (which is a known problem in SNOTEL data) were set to missing.

3.3. Period-of-record adjustment

Monthly station data passing the above QC tests were averaged to create 1971–2000 monthly means. A 1971–2000 monthly mean calculated using data from at least 23 of these 30 years (75% data coverage) was considered to be sufficiently characteristic of the 1971–2000 period, and was termed a 'long-term' station. However, many stations had a POR with fewer than 23 years. It was advantageous to include these 'short-term' stations in the analysis because they often added information to the interpolation process at critical locations. In order to minimize temporal biases from these short-term stations, their POR means were adjusted to the 1971–2000 period. The assessment and adjustment of averages was performed for each month individually; therefore, it was possible for a station to be considered long term for some months and short term for others. A discussion of the adjustment process is given in Appendix A.

4. PRISM overview and modelling

PRISM (Daly *et al*., 1994, 2001, 2002, 2003; Daly, 2006) is a knowledge-based system developed primarily to interpolate climate elements in physiographically complex landscapes. The regression-based PRISM uses point data, a DEM, other spatial data sets, and an encoded spatial climate knowledge base to generate estimates of annual, monthly, daily, and event-based climatic elements. These estimates are interpolated to a regular grid, making them GIS-compatible. Previous mapping efforts with PRISM have included peer-reviewed, official USDA precipitation and temperature maps for all 50 states and Caribbean and Pacific Islands; a new official climate atlas for the United States; a 112-year series of monthly temperature, precipitation, and dew point maps for the conterminous 48 states; detailed precipitation and temperature maps for Canada, China, and Mongolia, and the first comprehensive precipitation maps for the European Alps region (Daly *et al*., 2001; Hannaway *et al*., 2005; Milewska *et al*., 2005; Schwarb *et al*., 2001a,b; Simpson *et al*., 2005). Reports and papers describing PRISM are available from http://prism.oregonstate.edu.

4.1. The climate–elevation regression function

PRISM adopts the assumption that for a localized region, elevation is the most important factor in the distribution of temperature and precipitation (Daly *et al*., 2002). PRISM calculates a linear climate–elevation relationship for each DEM grid cell, but the slope of this line changes locally with elevation (through elevation weighting) as dictated by the data points. Beyond the lowest or highest station, the function can be extrapolated linearly as far as needed. A simple, rather than multiple, regression model was chosen because controlling and interpreting the complex relationships between multiple independent variables and climate is difficult. Instead, weighting of the data points (discussed later) controls for the effects of variables other than elevation.

DOI: 10.1002/joc

The climate–elevation regression is developed from X, Y pairs of elevation and climate observations supplied by station data. A moving-window procedure is used to calculate a unique climate–elevation regression function for each grid cell. The simple linear regression has the form

$$Y = \beta_1 X + \beta_0 \quad (1)$$

where Y is the predicted climate element, β_1 and β_0 are the regression slope and intercept, respectively, and X is the DEM elevation at the target grid cell. The DEM elevation is represented at a spatial scale appropriate for the climate element being mapped (discussed in Section 2).

4.2. Station weighting

Upon entering the regression function, each station is assigned a weight that is based on several factors. The combined weight (W) of a station is given by the following:

$$W = W_c[F_d W_d{}^2 + F_z W_z{}^2]^{1/2} W_p \; W_f \; W_l \; W_t \; W_e \quad (2)$$

where W_c, W_d, W_z, W_p, W_f, W_l, W_t, and W_e are the cluster, distance, elevation, coastal proximity, topographic facet, vertical layer, topographic position, and effective terrain weights, respectively, and F_d and F_z are user-specified distance and elevation weighting importance scalars (Daly *et al*., 2002; Daly *et al*., 2007). All weights and importance factors, individually and combined, are normalized to sum to unity. Table IV summarizes how the PRISM elevation regression and station weighting functions account for the physiographic climate forcing factors described in Section 1, and provide references for further information. The focus in the following discussions is on those station-weighting functions that have not been previously published or have been updated for the current application.

4.2.1. Cluster, distance, and elevation weighting

Cluster weighting seeks to limit the influence of stations that are clustered with other nearby stations, which can lead to over-representation in the regression function. A detailed discussion of cluster weight is provided in Appendix B.

A station is down-weighted when it is relatively distant from the target grid cell. The distance weight is given as

$$W_d = \left\{ \begin{array}{c} 1; d - r_m \le 0 \\ \dfrac{1}{(d - r_m)^a}; d - r_m > 0 \end{array} \right\} \quad (3)$$

where d is the horizontal distance between the station and the target grid cell, a is the distance weighting exponent, and r_m is the minimum radius of influence. In this application, a was set to 2, which is equivalent to an inverse-distance-squared weighting function, and r_m was set to approximately 7 km for precipitation and 10 km for temperature. The 7-km precipitation value matched the estimated minimum scale at which elevation directly affects precipitation patterns (discussed in Section 1). The 10-km temperature value represents the minimum scale of the effects of land use on station representativeness in this application. Implementing this minimum radius reduced 'bull's eyes' and other artifacts created in the interpolated temperature fields caused by variations in station siting and the surrounding land use/land cover. These effects are not yet well understood, but can be highly significant (Mahmood *et al*., 2006).

A station is down-weighted when it is at a much different elevation than the target grid cell. A discussion of elevation weighting is given in Daly *et al*., (2002, Section 4.1).

4.2.2. Coastal proximity

Coastal proximity weighting is used to define gradients in precipitation or temperature that may occur due to proximity to large water bodies (Daly *et al*., 2002, 2003). Stations with coastal proximities similar to that of the target grid cell are assigned relatively high weights in the regression function.

For this application, coastal proximity guidance in the PRISM modelling run was provided through four coastal proximity grids, depending on the region and climate element. In the central and eastern United States, the coastal proximity grid for precipitation (abbreviated PCE) was composed simply of distances from the generalized coastline grid (Table I). A simple distance was adequate because of a lack of terrain features. The generalized coastline is used because bays and inlets were not considered to be as important moisture sources as the open ocean for precipitation. In the central and eastern United States, the coastal proximity grid for temperature (abbreviated TCE) was a weighted average of the distances from the actual and generalized coastlines, the rationale being that bays and inlets experience a temperature environment that is a transition between inland and fully oceanic (Table I).

Complex terrain along the U.S. West Coast required more sophisticated methods of estimating coastal proximity. An advection model was designed to quantify coastal proximity for temperature mapping (abbreviated TW; Table I). The advection model is a cost–benefit algorithm that assesses the optimal path a surface air parcel might take as it moves from the coast to each inland pixel. The basic assumption is that the mean coastal influence experienced at a site will be the result of a flow path from the coast that minimizes two factors: (1) modification of the air by continental influences, which accumulates as the path length over land increases (path length penalty); and (2) loss of momentum caused by flowing over terrain obstacles (terrain penalty). Predominant mesoscale flow patterns, which aid certain flow paths and cause more effective inland penetration, are also accounted for by decreasing the path length penalty for predominant directions and increasing it for infrequent directions. For the West Coast, flow paths from the west and northwest (typical summertime directions) were set to incur substantially

DOI: 10.1002/joc

Table IV. PRISM algorithms and associated physiographic climate forcing factors, inputs to the algorithms, and references on their formulation and use. Methods used to prepare the gridded model inputs are summarized in Tables I and 2.

PRISM algorithm	Description	Physiographic forcing factors	Inputs to algorithm	Reference
Elevation regression function	Develops local relationships between climate and elevation	Direct effects of elevation	Station data DEM	This paper, Section 4.1
Cluster weighting	Downweights stations clustered with others	–	Station locations and elevations	This paper, Appendix B
Distance weighting	Upweights stations that are horizontally close	Spatial coherence of climate regimes	Station locations	This paper, Section 4.2.1
Elevation weighting	Upweights stations that are vertically close	Direct effects of elevation	DEM Station locations and elevations	Daly *et al*. (2002), Section 4.1
Topographic facet weighting	Upweights stations on the same side of a terrain barrier, or on the same exposure, at six scales	Rain shadows Air mass separation	DEM Topographic facet grids Station locations	Daly *et al*. (2002), Section 5
Coastal proximity weighting	Upweights stations having similar exposure to coastal influences	Effects of water bodies on temperature and precipitation	DEM Coastal proximity grid Moisture index grid Station locations	Daly *et al*. (2002), Section 6 Daly *et al*. (2003), Sections 2.3.2–2.3.3
Two-layer atmosphere weighting	If an inversion is present, upweights stations in the same vertical layer (boundary layer or free atmosphere)	Temperature inversions; vertical limit to moist boundary layer (precipitation)	DEM Inversion height grid Station locations	Daly *et al*. (2002), Section 7 Daly *et al*. (2003), Sections 2.3.2–2.3.4
Topographic position weighting	Upweights stations having similar susceptibility to cold air pooling	Cold air drainage	DEM Topographic index grid Station locations	Daly *et al*. (2007), Section 4
Effective terrain height weighting	Modifies regression slope to reflect ability of terrain to affect precipitation patterns; upweights stations on similar terrain	Orographic effectiveness of terrain features	DEM Effective terrain height grid Station locations	This paper, Appendix C

DOI: 10.1002/joc

less path length penalty than paths from the northeast, east, and southeast (unusual directions). Algorithms for, and operation of, the coastal advection model are given in Daly *et al*. (2003). The western U.S. coastal proximity grid for temperature is shown in Figure 2(a).

A simple trajectory model was designed to assess relative moisture availability to guide precipitation mapping (abbreviated PW; Table I). The trajectory model is similar to the coastal advection model in that it accounts for changes in moisture content due to path length and terrain effects. It differs in that the air parcel trajectories are straight line and invariant throughout the simulation, as might be the case within large-scale synoptic circulations that produce significant precipitation. The premise here is that the mean potential for precipitation experienced at a site is the result of (1) loss of moisture through rainout, which accumulates as the path length over land increases, and (2) enhancement and suppression of precipitation caused by adiabatic cooling and warming during flow over terrain obstacles. Convergence is not considered in this simple model. Moisture-bearing air flow along the West Coast is most commonly southwesterly, with some variation around that direction. Therefore, the trajectory model was run for three flow directions: 230, 240, and 250° true, and the results averaged as $0.2(230) + 0.6(240) + 0.2(250)$. Algorithms for, and operation of, the trajectory model are given in Daly *et al*. (2003). The western U.S. coastal proximity grid for precipitation is shown in Figure 2(b).

4.2.3. Topographic facets

In complex terrain, climatic patterns are defined and delineated by topographic slopes and barriers, creating a mosaic of hillslopes, or 'facets', each potentially experiencing a different climatic regime (Daly *et al*., 1994; Gibson *et al*., 1997; Daly *et al*., 2002). Topographic facet weighting effectively groups stations onto these individual facets at a variety of scales, to account for sharp changes in climate regime that can occur across facet boundaries, such as rain shadows and air mass transitions. For example, topographic facet weighting improves the interpolation in the vicinity of rain shadows by reducing the mixing of stations on leeward and windward slopes, which may have very different precipitation–elevation (P–E) relationships.

Topographic facet guidance is provided to PRISM through a series of topographic facet grids (Table II). The current method for delineating topographic facets is described in Gibson *et al*. (1997) and Daly *et al*. (2002). Facet grids are constructed for six DEM smoothing levels, or scales (Daly *et al*., 1994, 2002). The smoothed DEM for each level is prepared by applying the Gaussian filter (Barnes, 1964) to the precipitation DEM. The horizontal filtering cut-off distance (i.e. how far out the

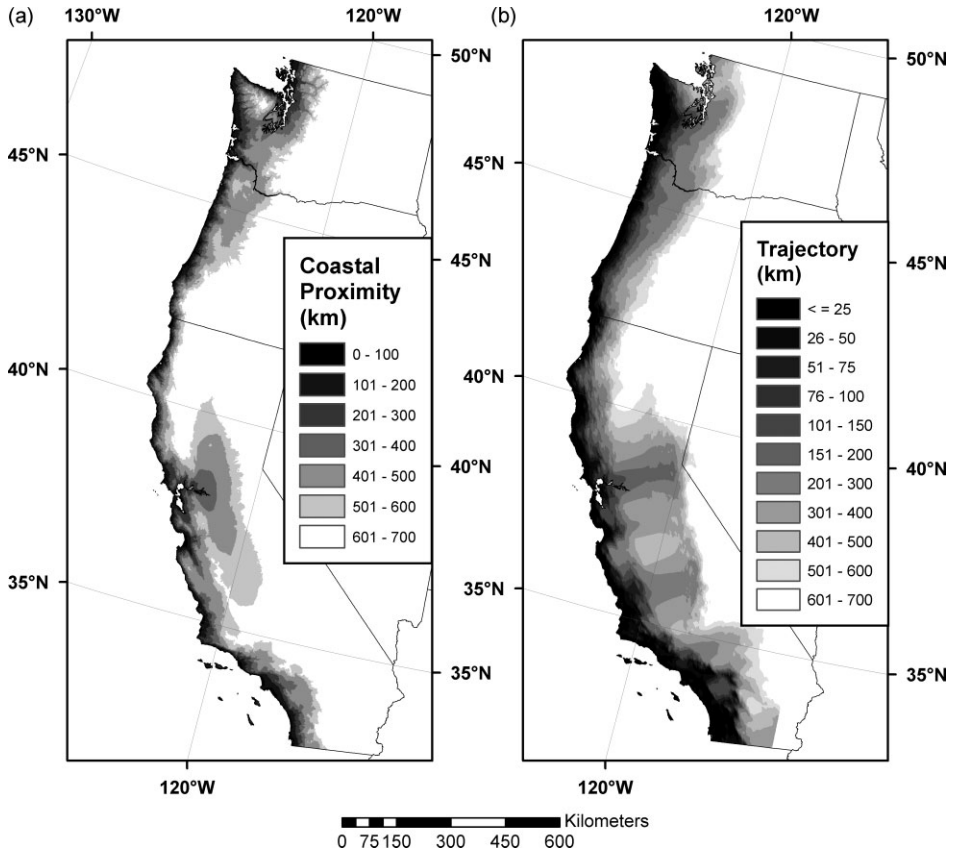

Figure 2. Western U.S. coastal proximity grids for (a) temperature (TW) and (b) precipitation (PW). Coastal proximity describes the optimal path length from the coastline in kilometres, accounting for terrain blockage. Trajectory describes the straight-line path length from the coastline in kilometres, modified by terrain blockage.

DOI: 10.1002/joc

averaging function extends) for each of the six levels is controlled by a user-defined maximum wavelength (Daly *et al*., 2002, Table II). The six levels are spaced evenly from the original cell size of the DEM up to this maximum, which is typically 60 km. In this application, levels at 12, 24, and 36 km were used for precipitation modelling, and levels at 0.8, 24, 36, and 48 km were used for temperature. The topographic facet grid for a filtering distance of 36 km is shown in Figure 3.

4.2.4. Two-layer atmosphere

While climate often varies monotonically with elevation, many cases arise for which a monotonic change is not realistic. Examples are mid-slope precipitation maxima where the moist boundary layer is shallow relative to the terrain height (Giambelluca and Nullet, 1991; Juvik *et al*., 1993), and wintertime and nocturnal temperature inversions in sheltered valleys, where temperature increases of 2.5–3.0 °C/100 m are not uncommon. PRISM divides the atmosphere into two vertical layers to handle these situations. Layer 1 represents the boundary layer, and layer 2 the free atmosphere above. Stations in the same layer as the target grid cell receive full weight, while those in the adjacent layer receive lower weights. In essence, the layer-weighting scheme limits the ability of stations in one layer to influence the regression function of the other. A fuller discussion of the vertical layer weighting function is available from Daly *et al*. (2002, Section 6), and Daly *et al*. (2003, Section 2.3).

A simple method was developed to spatially distribute the approximate height of the top of a climatologically persistent inversion layer, should it exist, for use in mapping temperature. The inversion height represents a candidate dividing line between stations, which may be partially or fully dissolved by PRISM if the relationship between the climate element and elevation is similar above and below it (Daly *et al*., 2002).

A grid of elevations representing the top of the boundary layer (layer 1) for the contiguous United States under persistent temperature inversions was prepared using a 800-m DEM by finding the minimum elevation of all grid cells, spatially averaging these elevations within an appropriate distance to form a 'base elevation' grid, and adding a constant, climatological inversion height to the base elevations (Table II). The combination of search and averaging radii given in Table II was prepared for the present application, and might not be transferable to others. Mountain/valley systems span a wide range of scales, and can therefore be delineated at a variety of search and filtering radii, depending on grid resolution and station density.

Our analyses of rawinsonde data from several cities in the United States with persistent climatological inversions, mostly during winter, indicated that the inversion top typically occurred at roughly 200–300 m above ground level. Therefore, 250 m was added to the base elevation at each grid cell to obtain the potential inversion height above sea level. While this is only a rough approximation of the inversion height, it has been effective for modelling purposes (Table II). The potential inversion height grid for the contiguous United States is illustrated in Figure 4. Valleys, plains, and other low-lying areas tend to fall within the boundary layer, while local ridge tops and other elevated

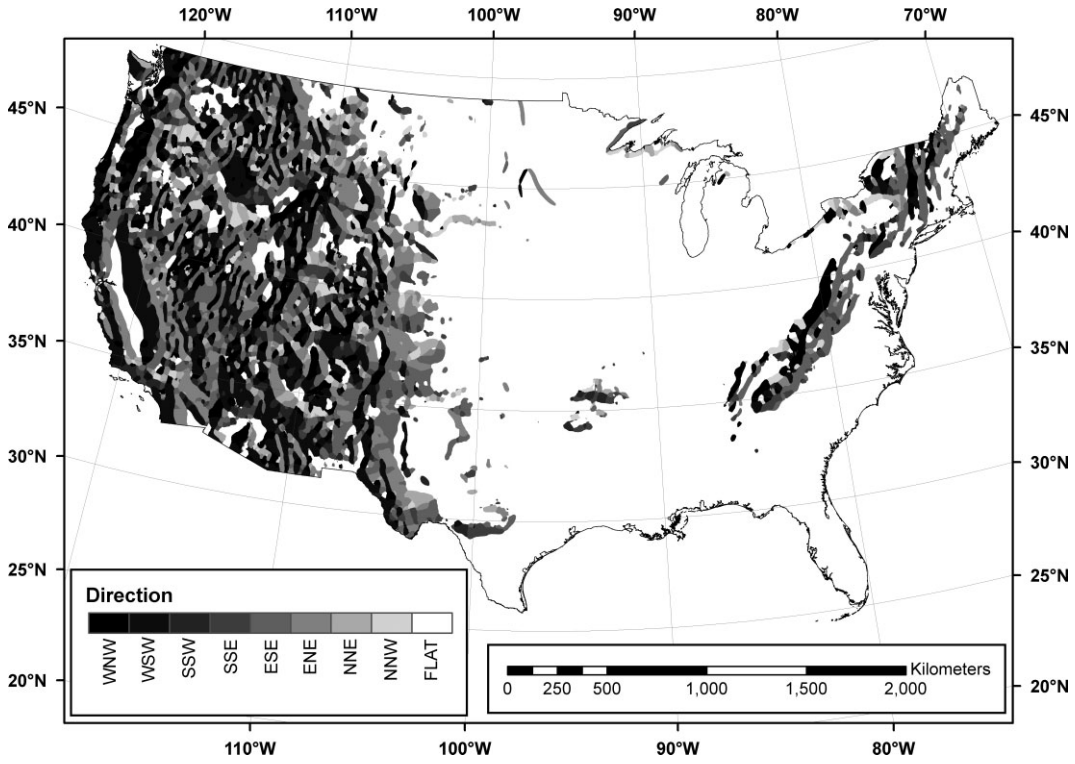

Figure 3. Topographic facet grid for the conterminous United States at a filtering distance of 36 km.

DOI: 10.1002/joc

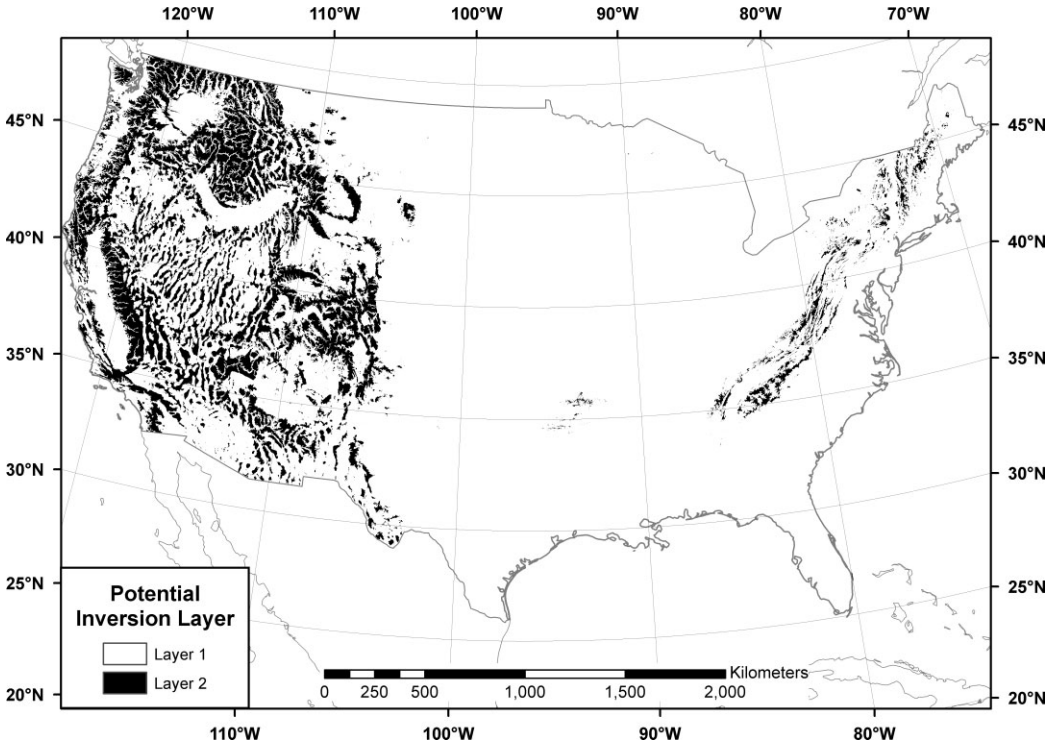

Figure 4. Potential inversion height grid for the conterminous United States, illustrated as the difference between the inversion height and the 800-m DEM elevation. Terrain rising above the inversion height (in layer 2) is shaded, and terrain falling below this height (in layer 1) is unshaded.

terrain emerge into the free atmosphere (Johnson *et al*., 2000).

The inversion height grid was used in a somewhat different manner for precipitation modelling. Here, it represented the estimated height of the moist boundary layer within which precipitation could be expected to increase with elevation; above this level, precipitation would be expected to decrease. Experience working with observational data has suggested that precipitation increases with elevation to the top of the highest available terrain in most areas within the conterminous United States, except within the marine-dominated precipitation regimes along the West Coast. In Washington and Oregon, on the windward slopes of the Cascades and westward, the height of the moist boundary layer was estimated at 2200 m (Table II). In California, a fixed boundary layer height oversimplified what is a very complex P–E relationship in the Sierra Nevada; therefore, it was left unspecified to allow the station data to dictate the local relationships at higher elevations.

4.2.5. Topographic index

A related station weighting function assesses a site's susceptibility to cold air pooling by considering its vertical position relative to local topographic features, such as valley bottom, mid-slope, or ridge top. Stations located in deep valleys and depressions are more susceptible to cold air pooling than those located on slopes or ridge tops.

To simulate this, a 'topographic index' grid was created, which describes the height of a pixel relative to the surrounding terrain height. PRISM used this information to weight stations during temperature interpolation. An 800-m topographic index grid was developed with methods similar to those of the inversion height grid (Table II). The resulting grid represents the local, or relative, elevation variations within a given area (Figure 5). The selection of the search and averaging parameters (Table II) was again somewhat subjective and pragmatic, and depended largely on station data density (Daly *et al*., 2007).

4.2.6. Effective terrain height

Terrain features produce varying P–E gradients, depending partly on their effectiveness in blocking and uplifting moisture-bearing air. Large, steeply rising features that present a significant barrier to air flow can generally be expected to produce steeper P–E gradients than low, gently rising, features. A discussion of how PRISM recognizes and accounts for differences in orographic effectiveness is given in Appendix C.

An 800-m effective terrain height grid for the conterminous United States was created by a method that was again similar to that used in creating the potential wintertime inversion height grid, with different search and averaging parameters (Table II). The resulting grid represented the profiles presented by terrain features above the local base elevation of the terrain. Terrain profiles oriented normal to the moisture-bearing flow would be expected to have greater orographic effects than those oriented parallel to the flow; however, this effective terrain height grid does not account for prevailing

DOI: 10.1002/joc

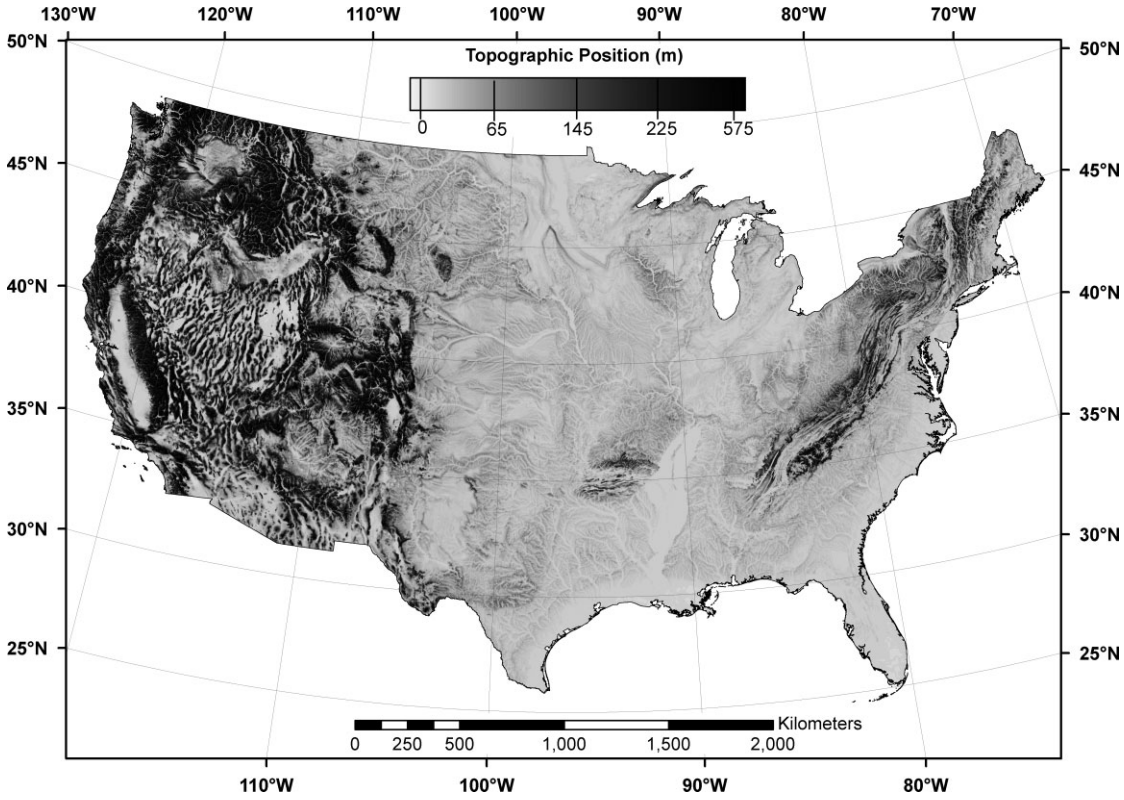

Figure 5. Topographic index grid for the conterminous United States. The scale indicates metres above the local base terrain.

wind directions, and therefore represents omni-directional terrain profiles only. A final radial averaging of the profile grid, not done in the wintertime inversion height estimation, was needed to minimize isolated, single-cell terrain discontinuities that did not appear to represent significant blocking features, such as small escarpments in otherwise gently rolling terrain (Table II).

The PRISM input grid of 'significant' terrain features for orographic precipitation is shown in Figure 6. Effective terrain heights of less than 75 m were treated as having insignificant orographic profiles, those between 75 and 250 m as two-dimensional/three dimensional (2D/3D) transitional profiles, and those exceeding 250 m as fully 3D profiles (Appendix C). A firm threshold for effective terrain height that discerns 2D from 3D features has not yet been established, primarily because of limited observational data, weaknesses in our understanding of how small terrain features impact precipitation patterns, and the likelihood that the threshold varies somewhat from region to region.

4.3. Inter-cell processing

Once PRISM calculated climate estimates for each cell in the domain, the full grid was subjected to a final set of operations designed to ensure spatial consistency among grid cells. A climate–elevation gradient was calculated between each grid cell and its neighbour, and subjected to upper and lower slope bound checks like those used in the target pixel climate–elevation regression functions (Daly *et al*., 2002). If the vertical gradient created by a pixel and its neighbour exceeded or fell below the user-specified upper or lower bounds, respectively, the climate value of each pixel was increased or decreased equally until the vertical gradient no longer violated the bounds check.

Finally, precipitation grids were subjected to a variable filter, designed to ensure a smooth field in low-gradient areas, without affecting high-gradient areas. For each grid cell, the filter performed a distance-weighted average of all surrounding grid cells within 8 km as follows:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i \frac{1}{{d_i}^a}}{\sum_{i=1}^{n} \frac{1}{{d_i}^a}} \tag{4}$$

where $\overline{x}$ is the averaged centre-cell value, x_i is the value at grid cell i, d_i is the distance between the centre grid cell and grid cell i, and a is the distance weighting exponent. Variable filtering was done by varying the exponent a as a function of the complexity of the field (of x values):

$$a = \left\{ \begin{array}{l} a_{\max}; \overline{\Delta x} \geq \Delta x_{\max} \\ a_{\max}(\overline{\Delta x}/\Delta x_{\max}); \overline{\Delta x} < \Delta x_{\max} \end{array} \right\} \tag{5}$$

where $a_{\max}$ is the maximum allowable filtering exponent, $\overline{\Delta x}$ is the mean absolute difference between the centre grid cell and all other grid cells within the specified radius of influence (an equally weighted average was used to characterize the field complexity), and $\Delta x_{\max}$

DOI: 10.1002/joc

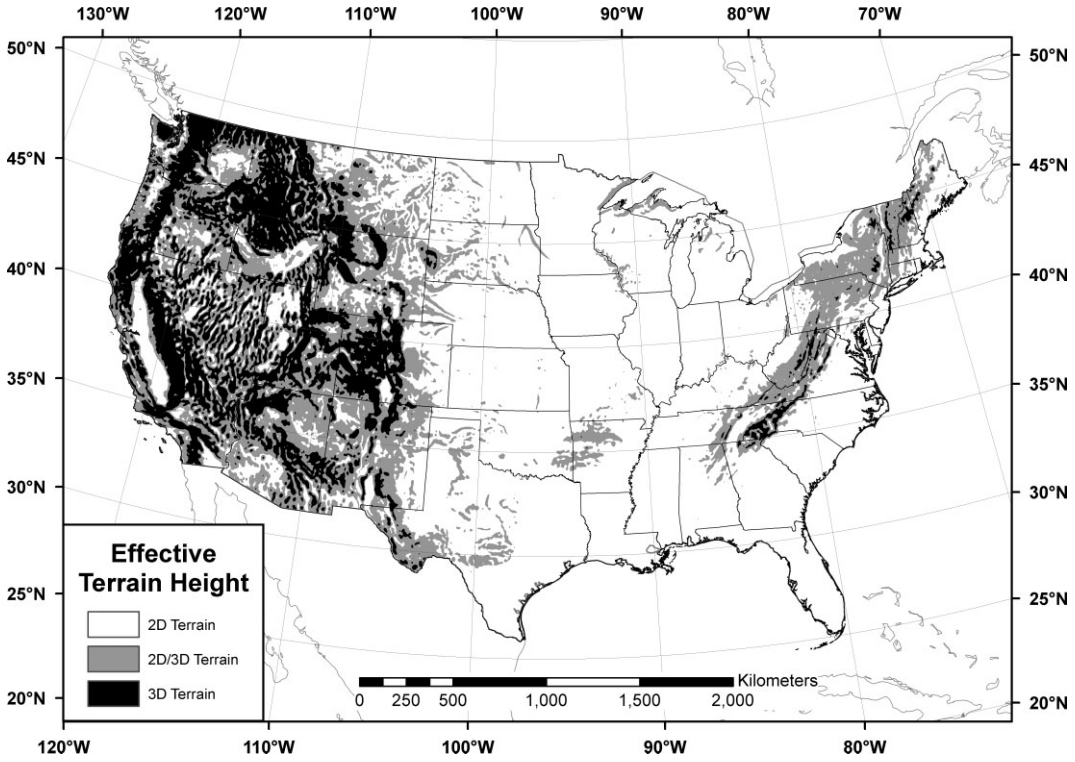

Figure 6. Effective terrain grid for the conterminous United States. Black areas denote terrain features that are expected to produce the strongest terrain-induced (3D) precipitation patterns, and grey areas denote features expected to have some influence on precipitation patterns (2D/3D transition). Unshaded grid cells as far as 100 km away from the shaded areas may also be considered 3D or 2D/3D transition areas. See text and Appendix C for discussion.

is the specified maximum average difference. In this application, $a_{\max}$ was set to 4%, and $\Delta x_{\max}$ was set to 4% of the centre-cell value. Therefore, when $\overline{\Delta x}$ equaled or exceeded 4% of the centre-cell value, the maximum distance weighting exponent of 4 was invoked. Below 4%, a was a linear function of the ratio $\overline{\Delta x}$ to $\Delta x_{\max}$. An a of 4 produces a very steep weighting function that gives an $\overline{x}$ that is nearly unchanged from the original value, thus creating little change on the field when it is complex. An a of zero creates a uniformly weighted average of all grid cells, which gives maximum smoothing within 8 km. The $\Delta x_{\max}$ value of 4% was arrived at by attempting to match the actions of the filter with the spatial coverage of 2D and 3D terrain (Section 4.2.6); the mean inter-pixel variation was typically greater than 4% in 3D terrain, and less than 4% in 2D terrain.

4.4. Application, review, and revision

PRISM was applied as described above to produce draft grids of 1971–2000 mean monthly and annual minimum and maximum temperature and precipitation. Model results for all parts of the country were examined and re-modelled numerous times as modelling and input data issues were discovered and remedied. Annual grids of precipitation were obtained by summing the monthly grids, and annual temperature grids were obtained by averaging the monthly grids.

No one person or group possesses the most accurate and detailed knowledge of spatial climate everywhere within the United States, necessitating the need for external review of spatial climate data sets. The goal of the review process was to ensure that the climate maps reasonably reflected the current state of knowledge regarding the patterns and magnitudes of precipitation and temperature in the conterminous United States (Daly *et al.*, 2002). It was not possible to have all the monthly grids reviewed in a meaningful and time-efficient way, so three grids thought to adequately represent the range of climatic conditions and situations experienced across the conterminous United States were chosen for review: mean annual precipitation, January minimum temperature, and July maximum temperature.

Reviewers were chosen primarily by project managers at the USDA–NRCS Water and Climate Center, and included all state and regional climatologists and hydrologists, as well as other personnel known to have local, regional, or national expertise in spatial climate patterns and analysis. The maps and station data were posted to an internet map server for easy access.

Not surprisingly, reviewers in regions with little terrain variation were concerned primarily with the quality of the station data, while those in mountain states commented mostly on climatic extremes in complex terrain. For example, in the states in the central plains, reviewers generally focussed on how station siting and variations in the period of observational record might affect the resulting maps. On the East Coast, the New Jersey State Climatologist believed that the draft PRISM maps did not show a sufficiently a well-developed coastal strip,

DOI: 10.1002/joc

where temperatures were warmer in winter and cooler in summer than adjacent inland areas (D. Robinson, pers. comm.). PRISM's coastal proximity weighting exponent was increased to resolve this issue. A U.S. Forest Service scientist doing field work came across several U.S. Geologic Survey precipitation gauges on remote mountaintops in Nevada, and was able to locate and provide data from them (D. Westfall, pers. comm.). A Nevada Division of Water Resources scientist was able to provide precipitation data from a local network which improved estimates on the eastern slope of the Sierra Nevada (J. Huntington, pers. comm.). Overall, the reviews raised important issues, expanded the station data sets, and improved the final product.

'Final' grids were produced on the basis of the reviews received. The word 'final' is given in quotes, because there have been, and likely will be, small changes made to these grids as issues are discovered and corrected. Users can download all the current monthly grids and access a log of changes at http://prism.oregonstate.edu.

5. Results and performance evaluation

Maps of mean annual precipitation, January minimum temperature, and July maximum temperature are shown in Figure 7. Average vertical gradients in precipitation and temperature are summarized by region in Table V. A vertical gradient was calculated for each pixel by finding the average climate–elevation gradient between the pixel and the four surrounding pixels. The vertical gradients among pixels reported here, and in the PRISM diagnostic grids shown in subsequent figures, usually do not match the weighted regression slopes calculated by PRISM. This is especially true if adjacent pixels are in different physiographic situations than the target pixel, e.g. a valley bottom pixel and adjacent hillslope pixels, because the PRISM regression slope reflects the relationship between climate and elevation only for stations matching the target cell's physiographic characteristics.

While the maps in Figure 7 do not show much detail at this scale, some features are notable. Rain shadows downwind of the coastal, Olympic, Cascade, and Sierra Nevada ranges are visible, as are increases in precipitation over most elevated terrain, including the Appalachian Mountains. Over orographically significant terrain features (denoted 3D terrain in Figure 6), precipitation increased by about 70–75% of the estimated cell value per kilometre elevation in the western and central United States, and about 50% in the eastern United States (Table V). The mean increase in precipitation over all terrain was much lower than for 3D terrain only, especially in the centre and east, where much of the terrain is not orographically significant.

January minimum temperature exhibited very complex patterns in the western United States, characterized by numerous cold 'ponds' in valleys and depressions: temperatures were relatively mild along the ocean and Great Lakes coastlines. The mean rate of change with elevation was actually positive in layer 1 (white areas in Figure 6) in the western United States, reflecting the dominance of temperature inversions in this region (Table V). The July maximum temperature exhibited a strong decrease with elevation across most of the country. Coastal areas were cooler than inland areas, especially where the water temperature was significantly cooler than that of the land, such as along the West Coast. Regionally averaged vertical gradients were all negative, but positive gradients were common within the marine layer (layer 1) along most coastlines, including those in the central and eastern United States. Some of these features will be discussed in greater detail in subsequent sections.

Table V. Vertical gradients in temperature and precipitation, summarized by effective terrain height (precipitation) and layer (temperature).

Element	Region		
	West	Cent	East
Mean annual precipitation			
3D terrain precipitation gradient (%/km)	75.45	71.02	50.71
Overall precipitation gradient (%/km)	55.42	7.14	11.93
January minimum temperature			
Layer 1 temperature gradient (°C/km)	2.33	−0.82	−2.83
Layer 2 temperature gradient (°C/km)	−2.45	−0.60	−1.65
Overall temperature gradient (°C/km)	1.15	−0.79	−2.80
July maximum temperature			
Layer 1 temperature gradient (°C/km)	−5.74	−4.11	−3.78
Layer 2 temperature gradient (°C/km)	−7.50	−7.68	−7.42
Overall temperature gradient (°C/km)	−6.32	−4.21	−3.89

5.1. Statistical uncertainty analysis

Estimating the true errors associated with spatial climate data sets is difficult and subject to its own set of errors (Daly, 2006). This is because the true climate field is unknown, except at a relatively small number of observed points, and even these are subject to measurement and siting uncertainties.

5.1.1. Cross-validation error

A performance statistic often reported in climate interpolation studies is the cross-validation (C-V) error (Daly *et al.*, 1994; Willmott and Matsuura, 1995; Gyalistras, 2003). C-V error is a measure of the difference between one or more station values and the model's estimates for those stations, when the stations have been removed from the data set. In the common practice of single-deletion jackknife C-V, the process of removal and estimation is performed for each station one at a time, with the station returned to the data set after estimation. Once the process is complete, the overall error statistics, such as mean absolute error (MAE), bias, and others, are calculated (e.g. Willmott *et al.*, 1985; Legates and McCabe, 1999). The obvious disadvantage to C-V error estimation is that no error information is provided for locations where there are no stations. In addition, the single-deletion jackknife method favours interpolation models

DOI: 10.1002/joc

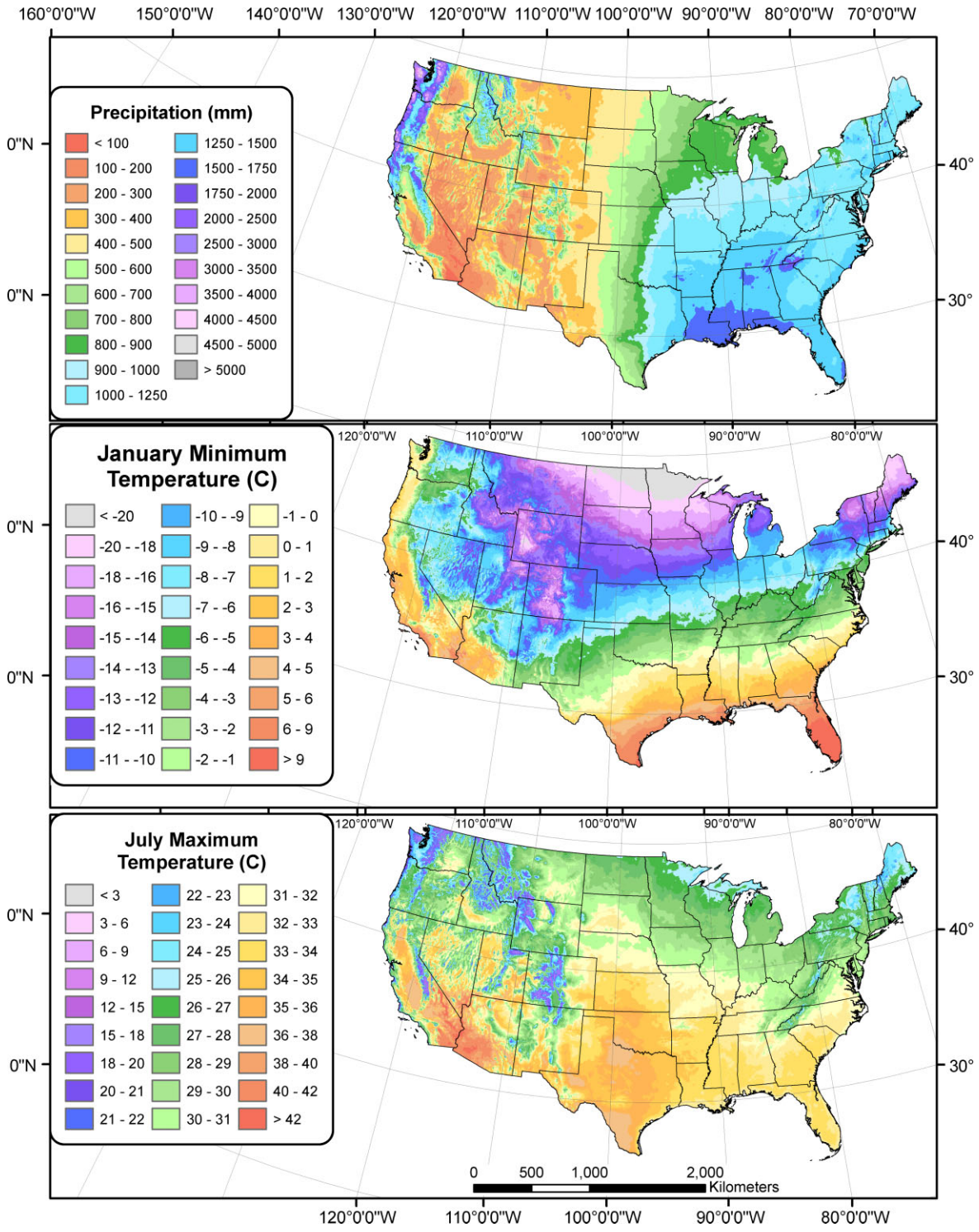

Figure 7. Conterminous U.S. maps of PRISM 1971–2000 mean annual precipitation, January minimum temperature, and July maximum temperature. This figure is available in colour online at www.interscience.wiley.com/ijoc

and parameterizations that heavily smooth the results and reduce local detail, so that deletion of one station is relatively unimportant to the stability of the estimate. Other deletion schemes, such as withholding a group of data points from the analysis, may sometimes be useful in detecting certain weaknesses in the interpolation, but may not give an accurate overall error estimate. For example, withholding high-elevation stations may help determine how well the system can extrapolate beyond the elevation range of the data.

5.1.2. *Prediction interval*

Climate interpolation models themselves typically provide few realistic estimates of error because these estimates rely at least partly on the very same assumptions used in the interpolation process itself and are

DOI: 10.1002/joc

therefore neither independent nor reliable. These error statistics should only be used in a relative sense, with the same model and station data set, and interpreted against the backdrop of model assumptions. However, modelled errors can be useful because they provide information at all grid cells, not just at station locations. The key to assessing the usefulness of modelled errors is to determine how and to what extent they rely on model assumptions, and assess how they compare with C-V errors.

Perhaps the most useful estimate of uncertainty produced by PRISM is the regression prediction interval. Since PRISM uses weighted linear regression to estimate precipitation or temperature as a function of elevation, standard methods for calculating prediction intervals for the dependent variable (Y) are used. Unlike a confidence interval, the prediction interval takes into account both the variation in the possible location of the expected value of Y for a given X (since the regression parameters must be estimated), and variation of individual values of Y around the expected value (Neter *et al*., 1989). The formula used for calculating the variance of Y (temperature or precipitation) for a new value of $X = X_h$ (elevation) is:

$$s^2\{Y_{\mathrm{h}(new)}\} = s^2\{\hat{Y}_{\mathrm{h}}\} + MSE = MSE \times \left[1 + \frac{1}{\sum_{i=1}^{n} w_i} + \frac{(X_{\mathrm{h}} - \overline{X})^2}{\sum_{i=1}^{n} (w_i X_i - \overline{X})^2}\right] \quad (6)$$

where $s^2\{\hat{Y}_{\mathrm{h}}\}$ is the estimated variance of the expected value of Y_{h} at $X = X_{\mathrm{h}}$, MSE is the regression mean square error, $\overline{X}$ is the weighted mean elevation of the regression data set, and X_i and w_i are the elevation and weight for station i, respectively (Neter *et al*., 1989). The prediction interval at significance level α was created as:

$$\hat{Y}_{\mathrm{h}} \pm t_{1-\alpha/2,df}\, s\{\hat{Y}_{\mathrm{h}}\} \quad (7)$$

where $t_{1-\alpha/2,df}$ is the $1 - (\alpha/2)$ percentile value of the t distribution with df degrees of freedom for MSE. A $(1 - \alpha)$ of 0.70 (70%) was chosen for the prediction interval because it approximated one standard deviation (where $(1 - \alpha) = 0.67$) around the model prediction. Many model-based uncertainty estimates assume that the model is perfect, and therefore underestimate the uncertainty. For example, the kriging estimation variance assumes the correctness of a simple semi-variogram model, and does not incorporate any uncertainty related to model goodness of fit (Daly, 2006). The PRISM 70% prediction interval, hereafter referred to as PI70, does include this important source of variability.

As with any error estimate, PI70 has its strengths and weaknesses. A major strength is that PI70 is large when there is a high degree of scatter about the local regression line, indicating a poor relationship between climate and elevation and suggesting a poor prediction. This tends to occur at locations far from stations, in areas within transition zones between two or more climatic regimes (such as coastal temperature boundaries or rain shadows), or at elevations in the vertical transition between the boundary layer and free atmosphere during temperature inversions. PI70 also increases the farther the prediction is extrapolated away from the mean regression elevation. This is seen in high mountain areas that are well above the highest stations in the vicinity, and therefore have uncertain predictions.

Unfortunately, the spatial patterns of PI70 are also sensitive to the way in which PRISM is parameterized to weight stations. For example, the presence of atmospheric-layer weighting produces PI70 patterns that may be highly discontinuous in space, owing to markedly different stations being used above and below the inversion height. Without layer weighting, PI70 patterns become smoother and more generalized in space. In another example, the use of a minimum radius of influence, r_{m} (discussed in Section 4.2.1) affects PI70 patterns. When r_{m} is set to zero, PI70 at pixels containing station locations is reduced to near zero because the co-located stations dominate the weighted regression. However, when r_{m} is set to 7–10 km (as was done for this data set) and there are several nearby stations, PI70 may not be reduced to zero because the weight of the co-located station does not dominate the regression function. In essence, attempts to reduce bull's eyes and other isolated features in the interpolated field will often result in increased PI70 values.

Finally, station weighting can often confound the effect that df (degrees of freedom) has on PI70. Statistically, as df decreases, PI70 increases. Therefore, in areas where there are few stations, the prediction interval should widen, reflecting a more uncertain prediction. However, PRISM is parameterized to successively increase its radius of influence until it retrieves a specified minimum number of stations for the regression function; this minimum number is 15 stations for temperature and 40 for precipitation. Therefore, df is always relatively high. However, station-weighting may favour only a handful of those stations with very large weights, leaving the others with vanishingly small weights, producing a regression function with what is effectively very few stations (and usually limited scatter), but retaining the high df value.

A major weakness of both C-V error estimates and PI70, and which is somewhat counterintuitive, is that these error measures typically increase as the station data set becomes more comprehensive (Daly, 2006). For example, a low-elevation data set that does not resolve high mountain features may be easy to estimate and produce little scatter about the regression lines, but produce very poor estimates in the mountains. In contrast, a station data set that spans a wide range of elevations and samples more of the true variability in the climate field would produce improved estimates, but likely be more challenging for the model to simulate, resulting in

DOI: 10.1002/joc

higher C-V errors and more scatter about the regression lines.

5.1.3. Comparison of cross-validation errors and prediction intervals

C-V error and PI70 statistics for all months are summarized by region in Table VI, and illustrated spatially for mean annual precipitation, January minimum temperature, and July maximum temperature in Figures 8 and 9. C-V biases were very small for all the three climate elements (not shown), less than 1 mm for precipitation and 0.13 °C for temperature; this indicates that values were not systematically over- or under-predicted. In general, regional mean C-V MAE and PI70 were surprisingly similar and were highest in the physiographically complex western United States (Table VI). For precipitation, C-V MAE and PI70 values in the western region were highest in winter and lowest in summer, owing to the predominance of winter-wet and summer-dry conditions over much of the area. C-V MAE for precipitation was somewhat higher in winter than PI70. Annually, both C-V MAE and PI70 were approximately 11% of the predicted grid cell value (not shown). In the central and eastern regions, PI70 was slightly higher than C-V MAE in most months. Annually, both PI70's and C-V MAEs were about 5% of the predicted grid cell value in the central region and about 4% in the east. The spatial distribution of percent C-V MAEs clearly shows a concentration of relatively large errors in the western mountains and in arid regions, with a few larger errors in the Appalachian Mountains (Figure 8).

C-V MAEs and PI70's for minimum temperature were generally higher than those for maximum temperature, likely a result of the increased complexity of the elevation regression function. In the West, minimum temperature C-V MAEs and PI70's exceeding 1 °C were not uncommon, while those for maximum temperature averaged about 0.7 °C (Table VI). Central and eastern errors were similar, and ranged from 0.6 to 0.8 °C for minimum temperature and about 0.5 °C for maximum temperature. Winter errors were slightly higher than summer errors. The spatial distribution of January minimum temperature errors was characterized by large errors in the western United States, but there were a significant number of larger errors in the central and eastern United States, as well (Figures 8(b) and 9(b)). This reflected the high spatial variability of minimum temperature, even in relatively gentle terrain. Relatively few stations showed large C-V MAEs for July maximum temperature (Figures 8(c) and 9(c)). However, large errors were found in mountainous areas of the West, as well as in coastal areas adjacent to the Pacific and Atlantic Oceans and Great Lakes. During summer, water temperatures are significantly lower than land temperatures in these areas, creating large temperature gradients across short horizontal distances (Daly, 2006).

The similarity of the C-V MAE and PI70 error estimates seems to lend support to the idea that the PI70 is a reasonable substitute for C-V MAE between stations, at least at the regional level. To determine how well the two error estimates agree at sub-regional scales, the western United States was subdivided into regular grids with successively smaller cell sizes (500, 400, 300, 200, and 100 km) and the average C-V MAE and PI70 estimates for mean annual precipitation, January minimum temperature, and July maximum temperature compared at each scale. Figure 10 shows the change in R^2 of a linear regression between the mean C-V MAE and PI70 as the averaging scale changed. At small averaging scales, R^2 was relatively low (0.2–0.5), indicating that there was relatively poor agreement between C-V MAE and PI70 within small areas. However, as the size of the area increased, R^2 increased rapidly, reaching maxima of 0.75–0.85.

5.2. Comparison with other data sets

PRISM spatial climate data sets were compared to two other spatial climate data sets of similar resolution for the conterminous United States: Daymet (Thornton *et al*., 1997) and WorldClim (Hijmans *et al*., 2005). The Daymet data set spans the conterminous United States at a resolution of 1 km, and has an averaging period of 1980–1997. Approximately 6000 stations from the NWS COOP and USDA NRCS SNOTEL networks were used in the interpolation. Interpolation was performed with the Daymet model (Thornton *et al*., 1997). Daymet develops local linear relationships between climate and elevation for each grid cell on a DEM, using data from surrounding stations. Each station is weighted only by its distance from the target grid cell. This method takes into account the elevational variation of climate, but does not have the ability to simulate non-monotonic relationships between climate and elevation, such as temperature inversions, and does not explicitly account for terrain-induced climatic transitions or coastal effects (Daly, 2006). Daymet data sets can be accessed online at www.daymet.org.

The WorldClim data set is global in extent, and has a resolution of 30 arcsec. Station data used in the conterminous United States were largely restricted to the USHCN (Williams *et al*., 2004), which is the U.S. portion of the Global Historical Climate Network (GHCN). The USHCN consists of about 1200 COOP stations with long-term records. Averaging period for the data varied, but was typically 1950–2000. The WorldClim data set was interpolated with ANUSPLIN (Hutchinson, 1995), a method that fits thin-plate splines (usually second- or third-order polynomials) through station data in three dimensions: latitude, longitude, and elevation. As with Daymet and PRISM, the relationship between the climate variable and elevation can vary in space, making the method suitable for large domains. Because a spline is by definition smoothly varying, this approach has difficulty simulating sharply varying climate transitions, which are characteristic of temperature inversions, rain shadows, and coastal effects (Daly, 2006). WorldClim data sets can be accessed online at www.worldclim.org.

DOI: 10.1002/joc

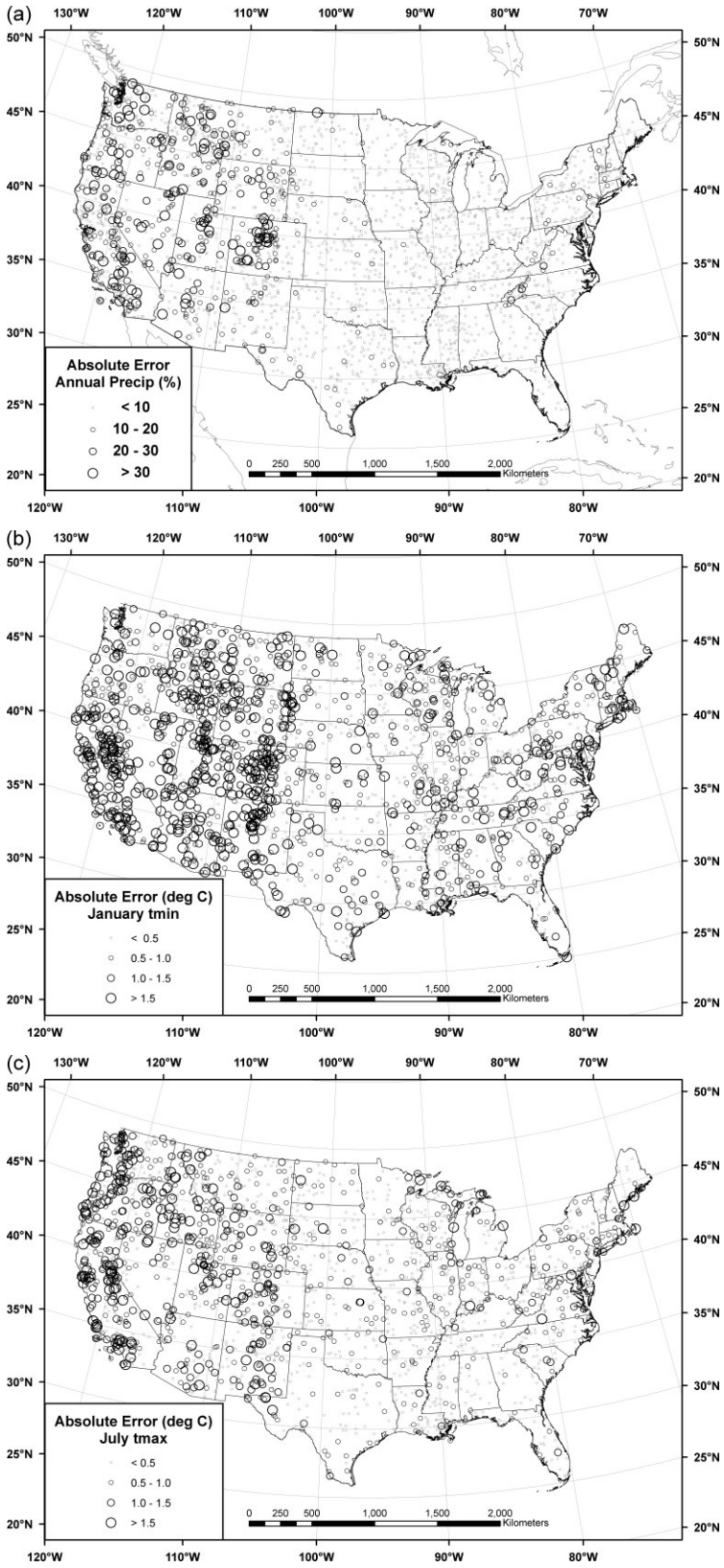

Figure 8. PRISM mean absolute cross-validation errors (MAE) for 1971–2000 mean: (a) annual precipitation; (b) January minimum temperature; and (c) July maximum temperature. Station density was too high to be shown clearly, requiring that 80% of the stations be randomly omitted from the plots.

A conterminous U.S. comparison between PRISM and these two data sets is given below, followed by detailed comparisons at three locations: Washington's Olympic Mountains (annual precipitation); Gunnison Valley, Colorado (January minimum temperature); and the central California coast (July maximum temperature).

DOI: 10.1002/joc

Table VI. PRISM cross-validation and PI70 statistics for precipitation and minimum and maximum temperature for the western, central, and eastern regions. MAE is the mean absolute (unsigned) difference between the prediction and observation during single-deletion jackknife cross-validation, and PI70 is the mean 70% prediction interval of the PRISM climate-elevation regression function.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Precipitation												
West												
MAE (mm)	12.59	11.49	11.20	7.68	6.48	4.68	4.80	4.76	5.23	6.76	10.84	11.66
PI70 (mm)	8.53	7.65	7.70	5.76	5.41	4.29	4.41	4.33	4.29	5.10	7.62	8.08
Central												
MAE (mm)	3.48	3.30	4.38	4.80	5.77	5.97	6.14	5.73	5.06	4.48	4.13	3.68
PI70 (mm)	4.29	3.91	5.08	5.48	6.74	7.32	7.39	6.87	6.34	5.28	4.89	4.47
East												
MAE (mm)	5.65	4.72	5.58	5.06	5.90	6.51	7.24	6.86	5.96	4.87	5.20	5.31
PI70 (mm)	6.57	5.37	6.29	5.79	6.75	7.96	8.80	8.42	7.41	5.75	6.01	6.18
Minimum temperature												
West												
MAE (C)	1.12	1.09	0.99	0.96	1.04	1.19	1.32	1.35	1.35	1.28	1.13	1.16
PI70 (C)	1.34	1.12	1.11	1.11	1.12	1.13	1.14	1.14	1.15	1.14	1.13	1.14
Central												
MAE (C)	0.75	0.72	0.61	0.56	0.57	0.58	0.60	0.62	0.66	0.73	0.66	0.70
PI70 (C)	0.95	0.92	0.78	0.71	0.70	0.67	0.69	0.73	0.77	0.88	0.82	0.89
East												
MAE (C)	0.72	0.69	0.62	0.62	0.63	0.60	0.60	0.63	0.67	0.76	0.65	0.64
PI70 (C)	0.96	0.95	0.85	0.81	0.81	0.76	0.74	0.78	0.84	0.96	0.85	0.86
Maximum temperature												
West												
MAE (C)	0.75	0.74	0.74	0.71	0.71	0.73	0.77	0.75	0.73	0.70	0.68	0.73
PI70 (C)	0.85	0.83	0.79	0.70	0.69	0.71	0.72	0.70	0.69	0.68	0.74	0.82
Central												
MAE (C)	0.56	0.55	0.54	0.50	0.46	0.44	0.43	0.43	0.46	0.46	0.44	0.47
PI70 (C)	0.78	0.77	0.74	0.60	0.55	0.52	0.51	0.51	0.54	0.54	0.60	0.67
East												
MAE (C)	0.49	0.48	0.49	0.49	0.47	0.45	0.43	0.41	0.44	0.44	0.36	0.38
PI70 (C)	0.71	0.70	0.68	0.59	0.54	0.51	0.47	0.46	0.48	0.51	0.52	0.58

5.2.1. Conterminous United States

Differences between PRISM and Daymet and WorldClim for mean annual precipitation, January minimum temperature, and July maximum temperature are shown in Figures 11, 12, and 13, respectively. Differences were found in most areas of the United States, with the largest occurring in the West. For mean annual precipitation, Daymet tended to be drier than PRISM in some locations and wetter in others. In general, Daymet did not resolve rain shadows well, owing to an inability to recognize topographic facets, causing it to be too wet on leeward slopes and sometimes too dry on windward slopes (see Section 5.2.2 for an example). WorldClim was much drier than PRISM, especially at higher elevations. The use of only GHCN stations, which are located mainly at low elevations, rendered the WorldClim analysis far too dry in most mountainous areas. To quantify this dry bias, mean annual precipitation observed at 638 SNOTEL stations was compared to estimated values at the grid cell centres closest to the station locations. The WorldClim dataset had an overall bias of −25%, and underpredicted one-fifth of the SNOTEL stations by 40% or more. In contrast, the Daymet and PRISM data sets, which incorporated SNOTEL stations, had biases of −2% and 0%, respectively.

For January minimum temperature, WorldClim was generally cooler than PRISM, possibly caused by differences in the averaging period (1950–2000 was generally cooler than 1971–2000). Both Daymet and WorldClim were cooler in the mountains and warmer in the valleys than PRISM, sometimes substantially so. This was caused mainly by the inability of the Daymet and WorldClim data sets to recognize inversions, to weight stations by susceptibility to cold air drainage, and to adequately simulate rapid reversals in temperature gradients caused by cold air drainage and inversions. These topographic cold biases are discussed more fully in Section 5.2.3.

Differences between PRISM and Daymet and WorldClim were relatively small for July maximum temperature, except for the Pacific Coast, where a lack of coastal proximity weighting limited the ability of Daymet and WorldClim to recognize the difference between coastal

DOI: 10.1002/joc

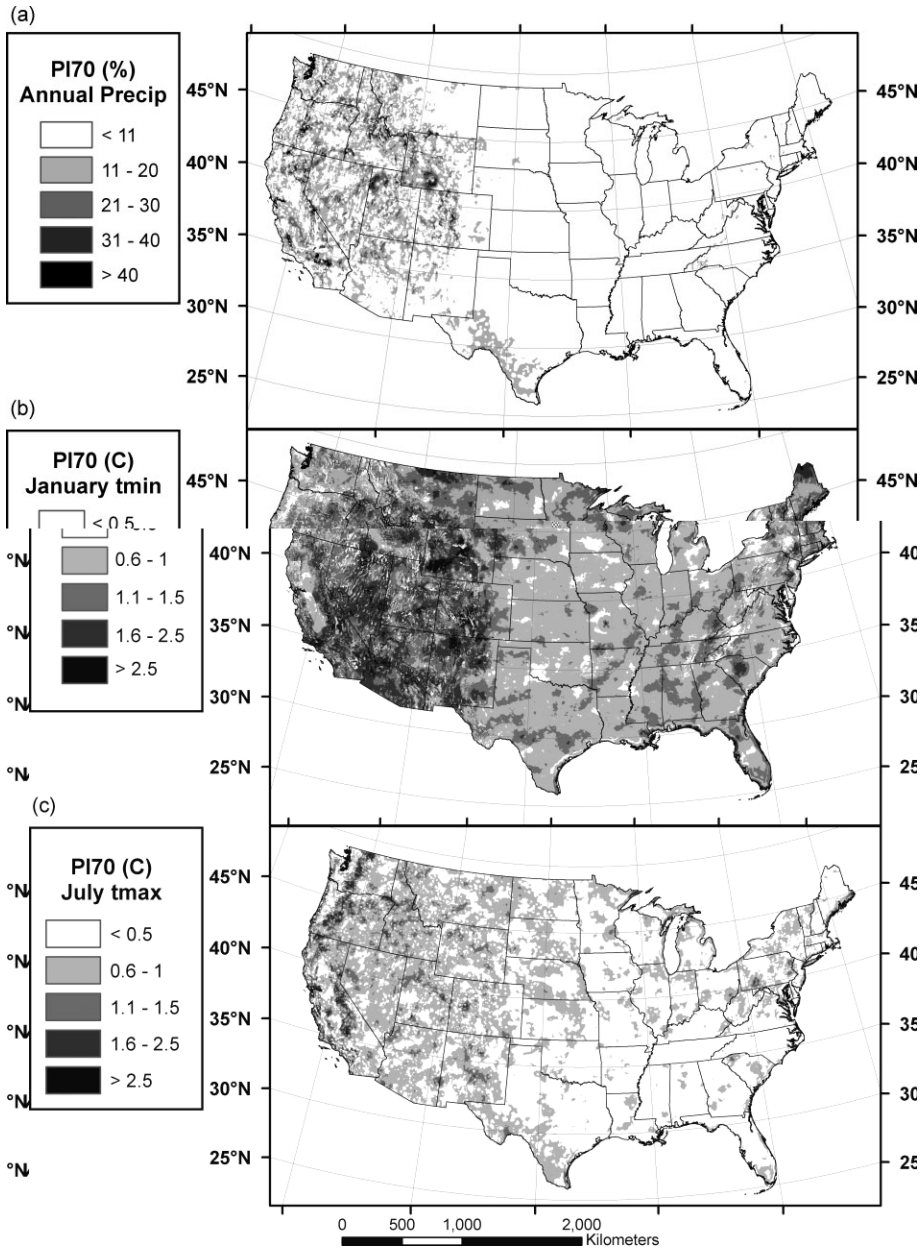

Figure 9. Seventy percent prediction intervals (PI70) for: (a) mean annual precipitation; (b) January mean minimum temperature; and (c) July mean maximum temperature. PI70 for precipitation is expressed as a percentage of the pixel's value.

and inland stations. These issues are discussed in more detail in Section 5.2.4.

The flatter terrain of the eastern United States reduced the differences associated with the interpolation method, and shifted the focus to station data quality as a significant error source. Many of the bull's eyes in the difference maps (Figures 11–13) are the result of differences in stations used, or different values used for the same station. However, differences among the three methods did occur in the eastern United States. For example, a cold bias in January minimum temperature was seen along coastlines, especially in the Great Lakes region. Here, the influence of the relatively warm waters of the Great Lakes resulted in milder winter minimum temperatures than those of inland areas. WorldClim, which did not recognize coastal zones and had limited station data, suffered from the greatest cold bias, averaging a 1.6 °C underestimation of January minimum temperature for 20 coastal stations in the Great Lakes area.

5.2.2. *Olympic Mountains precipitation*

A detailed look at the differences in mean annual precipitation between PRISM and Daymet and WorldClim for Washington's Olympic Mountains is presented in Figures 14 and 15. In western Washington, moisture-bearing winds from the southwest produce precipitation maxima of 6000–7000 mm/yr or greater on the windward slopes near the Olympic crest, and develop a sharp rain shadow on the northeastern slopes, where precipitation drops rapidly to less than 500 mm/yr. There are few precipitation stations in the vicinity, and none in the

DOI: 10.1002/joc

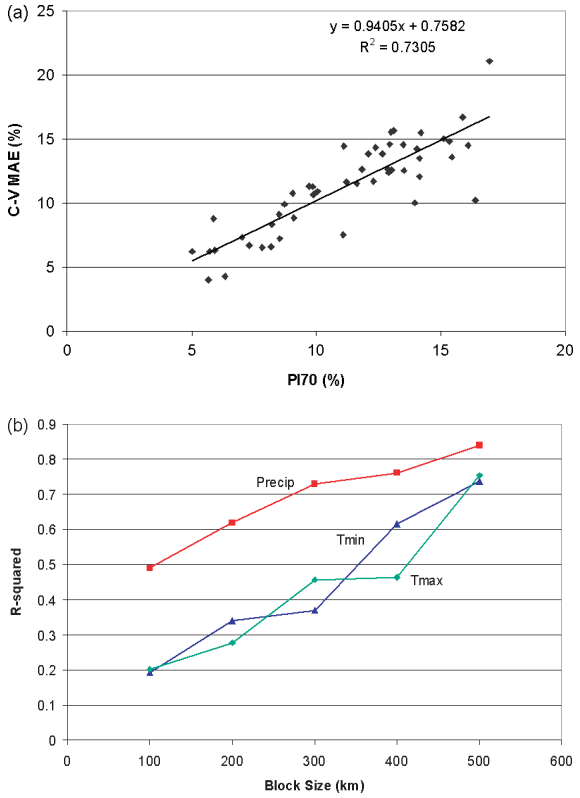

Figure 10. (a) Scatterplot of the relationship between C-V MAE and PI70 for mean annual precipitation (given in percentage of the estimated grid cell value) when averaged over 300 × 300-km blocks in the western United States; and (b) R^2 of the linear relationships between CV-MAE and PI70 for mean annual precipitation, January minimum temperature, and July maximum temperature at block sizes ranging from 100 to 500 km in the western United States. This figure is available in colour online at www.interscience.wiley.com/ijoc

core mountain area, making this a challenging region to interpolate.

As an independent evaluation of the precipitation fields, observed runoff data from USGS gauges (USDOI-USGS, 2007) were obtained for eight basins, four on the windward slope and four on the leeward slope. Runoff periods of record were 1971–2000 for all basins, except for the Queets (1975–2000) and the Quilcene (1995–2000). A certain amount of precipitation is lost to ET before it reaches the gauge outlet. Pan evaporation observations range from 625 to 750 mm annually in this region. A pan coefficient of 0.8 for this area (Farnsworth *et al.*, 1982) yields a range of potential evapotranspiration (PET) values of 500–600 mm/yr. Given that the region's evergreen coniferous forests have the ability to transpire year round, actual ET is likely a large proportion of the PET on an annual basis. Hence, an ET estimate of 500 mm/yr seems reasonable. This value has been adopted as a rough approximation by the National Weather Service's Northwest River Forecast Center for water balance calculations in this area (D. Laurine, pers. comm.), and is thought to be accurate to within about 30%.

Basin RO + ET values in Figure 14 (upper left panel) show the sharp decline in moisture between windward and leeward basins. A comparison of basin-mean precipitation for the three data sets is presented in the remaining three panels of Figure 14, and a profile of the comparisons is shown in Figure 15. Daymet simulated the windward basins well, but at the expense of the leeward side, which was substantially overestimated. WorldClim was quite low on the windward side, and somewhat high on the leeward side. PRISM simulated both the windward

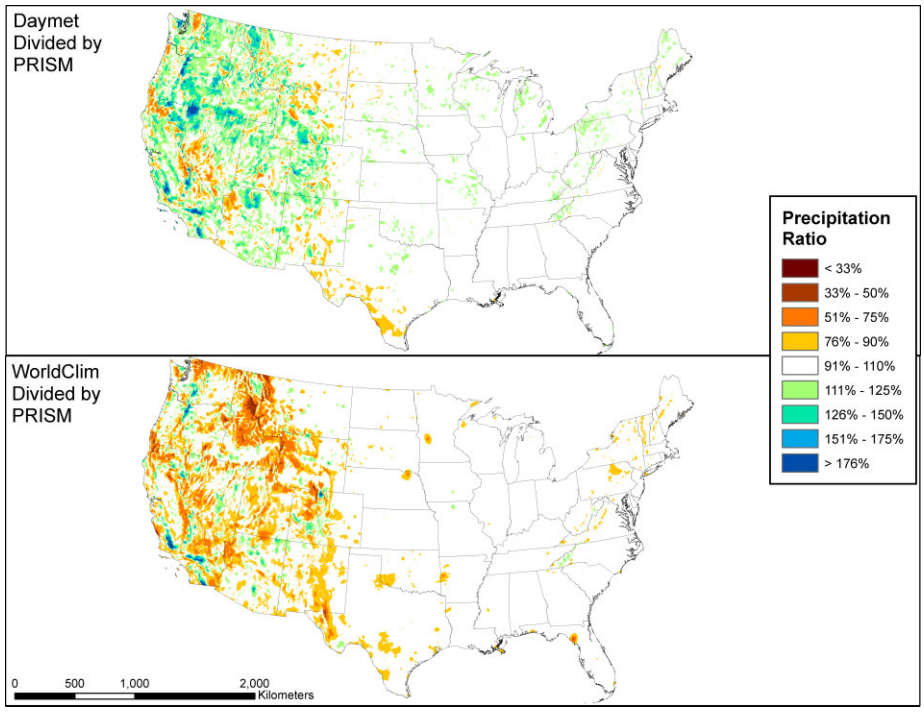

Figure 11. Daymet/PRISM and WorldClim/PRISM ratio maps (expressed as percentage) of mean annual precipitation for the conterminous United States. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

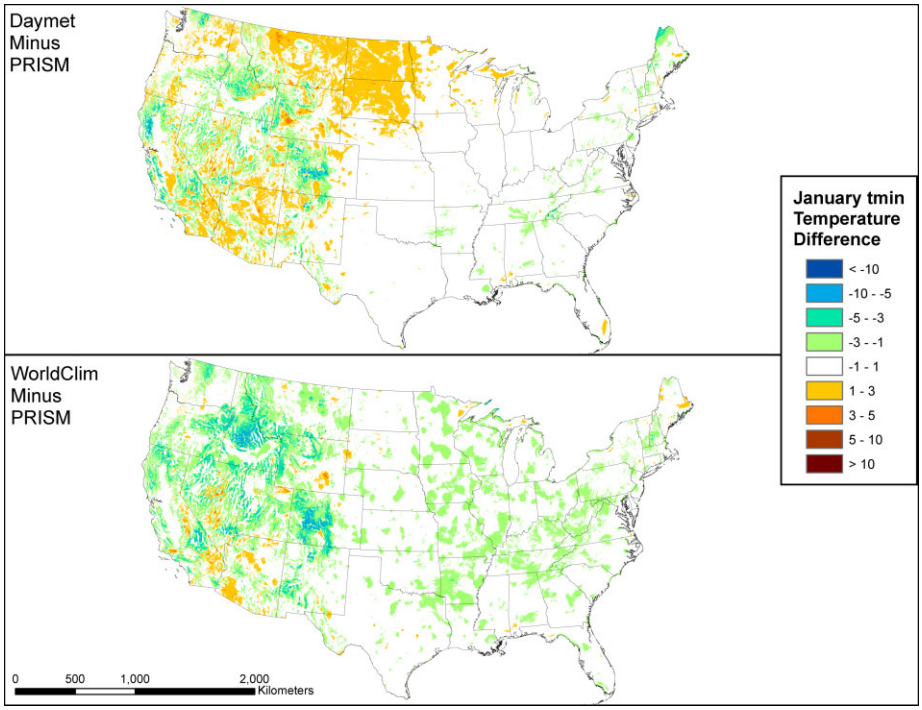

Figure 12. Daymet – PRISM and WorldClim – PRISM difference maps of mean January minimum temperature for the conterminous United States. This figure is available in colour online at www.interscience.wiley.com/ijoc

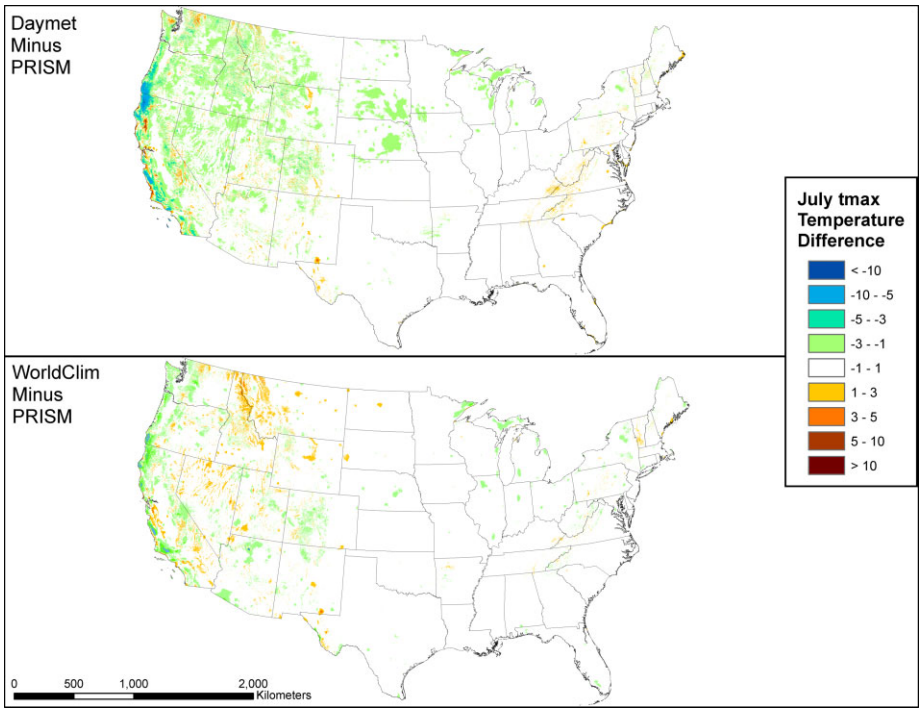

Figure 13. Daymet – PRISM and WorldClim – PRISM difference maps of mean July maximum temperature for the conterminous United States. This figure is available in colour online at www.interscience.wiley.com/ijoc

and leeward basins reasonably well. Only the PRISM data set simulated the sharp decline in precipitation between windward and leeward basins.

PRISM diagnostic grids are shown in Figure 16. By allowing PRISM to separate stations on each exposure in the regression function, topographic facet weighting was largely responsible for PRISM's ability to simulate the sharp divide between windward and leeward precipitation in the Olympics. Further station weighting using the trajectory model output for coastal proximity reinforced the drying trend from southwest to northeast. Vertical precipitation gradients were highest around the fringes of the mountains, often exceeding 150% of the pixel precipitation value per kilometre. In the core mountain

DOI: 10.1002/joc

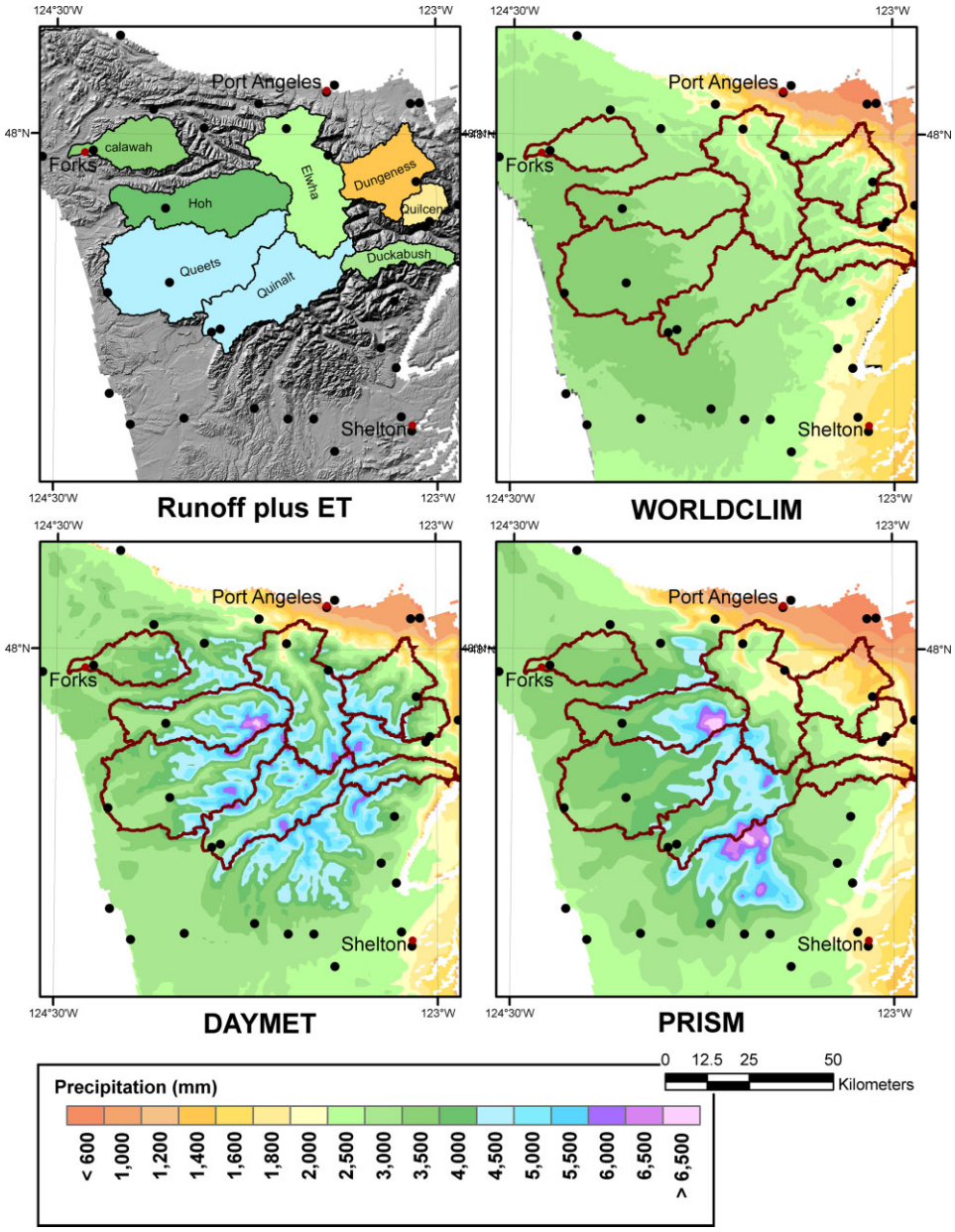

Figure 14. Upper left panel: annual RO + ET (runoff + evapotranspiration), a water balance approximation of total precipitation, for selected basins in Washington's Olympic Mountains, colour-coded following the precipitation legend. ET was roughly estimated at 500 mm/yr for all basins. Remaining panels: WorldClim, Daymet, and PRISM mean annual precipitation grids, with basin overlays, also colour-coded following the precipitation legend. To agree with the basin water balance estimates, the mean grid colour within a basin should match that shown in the upper left panel. Stations used in the PRISM analysis are shown as black dots. This figure is available in colour online at www.interscience.wiley.com/ijoc

region, gradients were typically 50–100% per km, but these numbers are suspect because of poor station coverage. PI70 values were generally less than 30% of the pixel precipitation value, which seems too low for this region. As discussed earlier, PI70 is a measure of regression function scatter; sparse station coverage probably resulted in an underestimate of this scatter by not reflecting the true complexity of the precipitation field.

5.2.3. *Gunnison Valley and vicinity*

A comparison of January minimum temperature for the Gunnison Valley, Colorado, and vicinity is shown in Figures 17–18. This is an area of strong and persistent inversions during winter, caused by cold air settling into protected valleys during frequently dry and calm nights. A typical minimum temperature profile begins with cold temperatures in valley bottoms, a strong increase in temperature to the top of the inversion at approximately mid-slope, followed by a decrease in temperature above the inversion, resulting in temperatures at the highest elevations approaching those in the valley bottoms. Using inversion height, topographic position, and elevation weighting functions, PRISM simulated this complex relationship between temperature and elevation in a realistic manner, with a well-developed inversion layer

DOI: 10.1002/joc

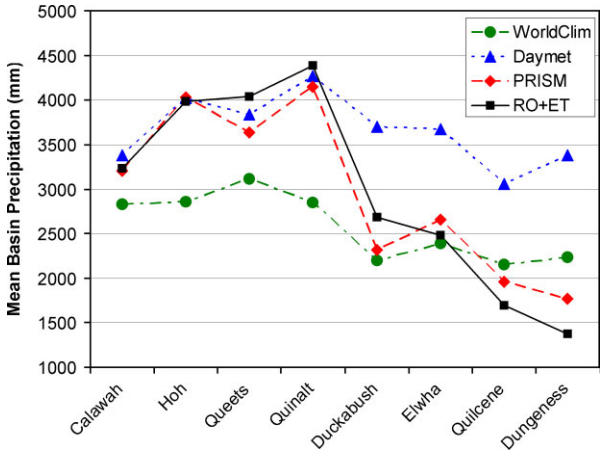

Figure 15. Estimated mean annual precipitation from WorldClim, Daymet, PRISM, and RO + ET (runoff + evapotranspiration), a water balance approximation of total precipitation, for river basins on the windward and leeward slopes of the Olympic Mountains. ET is assumed to be 500 mm for all watersheds. Runoff periods of record are 1971–2000, except for 1975–2000 on the Queets and 1995–2000 on the Quilcene. This figure is available in colour online at www.interscience.wiley.com/ijoc

in the protected Gunnison Valley and adjacent valleys (Figure 17). PRISM diagnostic grids show an estimated inversion height, separating layers 1 and 2, at approximately mid-slope above the valley floors (Figure 18). The topographic index grid depicts valley floors as most susceptible to cold air drainage (white and light grey). The vertical temperature gradient grid shows a distinct sign reversal, from highly positive below the inversion height to negative above. As might be expected, PI70 estimates exceeded 3 °C in the uncertain transition between the two layers where scatter in the regression function was greatest, and fell to 1.5 °C or less in the valleys and in the mountains.

Daymet was not able to simulate inversion conditions at all, despite incorporating a station data set similar to PRISM, because of its assumption of a monotonic, linear temperature–elevation relationship. The result is little spatial variation of temperature, and a cold bias at higher elevations. This high-elevation cold bias occurs when the incorrect assumption is made that temperature decreases with elevation from the cold valley floors to the mountains above. WorldClim, further hampered by sparse station data, also could not simulate the inversion conditions and had an even more pronounced cold bias than Daymet.

Cold biases of the Daymet and WorldClim data sets in western Colorado are illustrated in Figure 19. Here, 369 stations were grouped into four topographic index

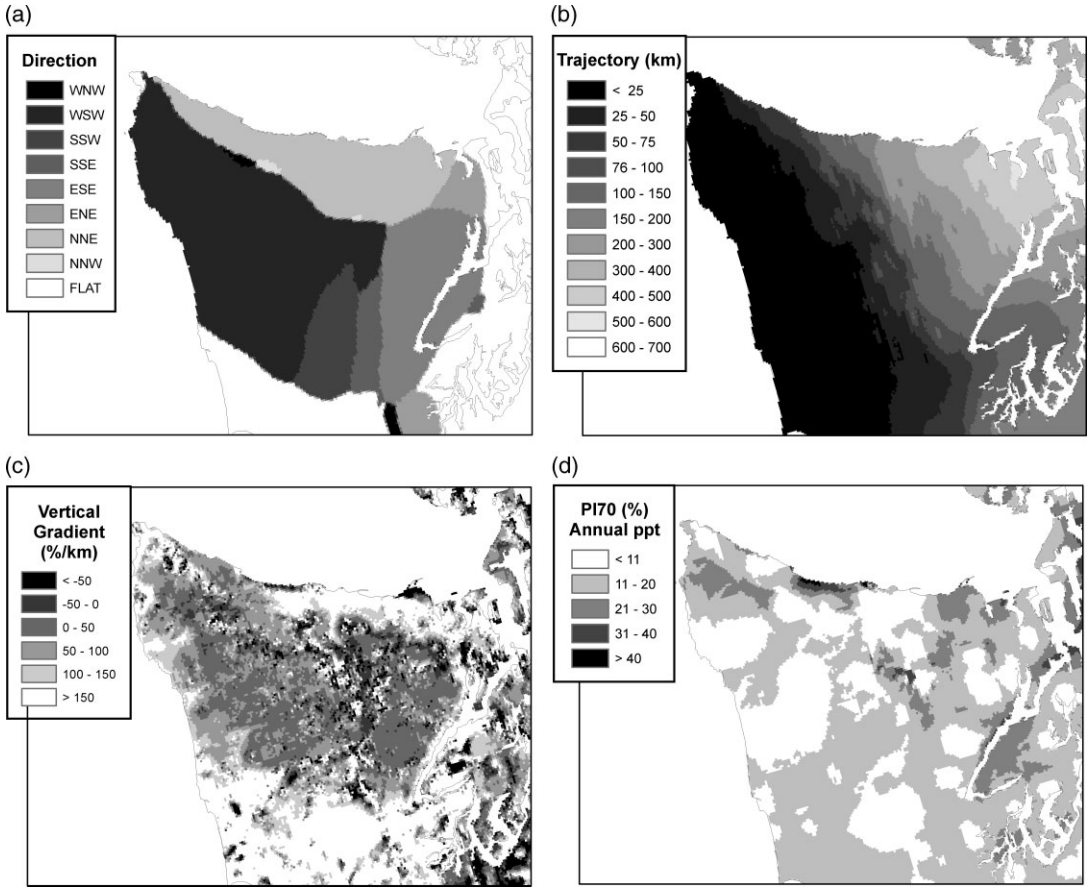

Figure 16. PRISM diagnostic grids associated with mean annual precipitation for the Olympic Mountains, Washington: (a) topographic facet grid at a filtering distance of 36 km; (b) PW coastal proximity, which describes a straight line, approximately southwesterly path length from the coastline in km, modified by terrain blockage; (c) vertical precipitation gradient, expressed as a percentage of the pixel's value per kilometre; and (d) 70% prediction interval (PI70), expressed as a percent of the pixel's value.

DOI: 10.1002/joc

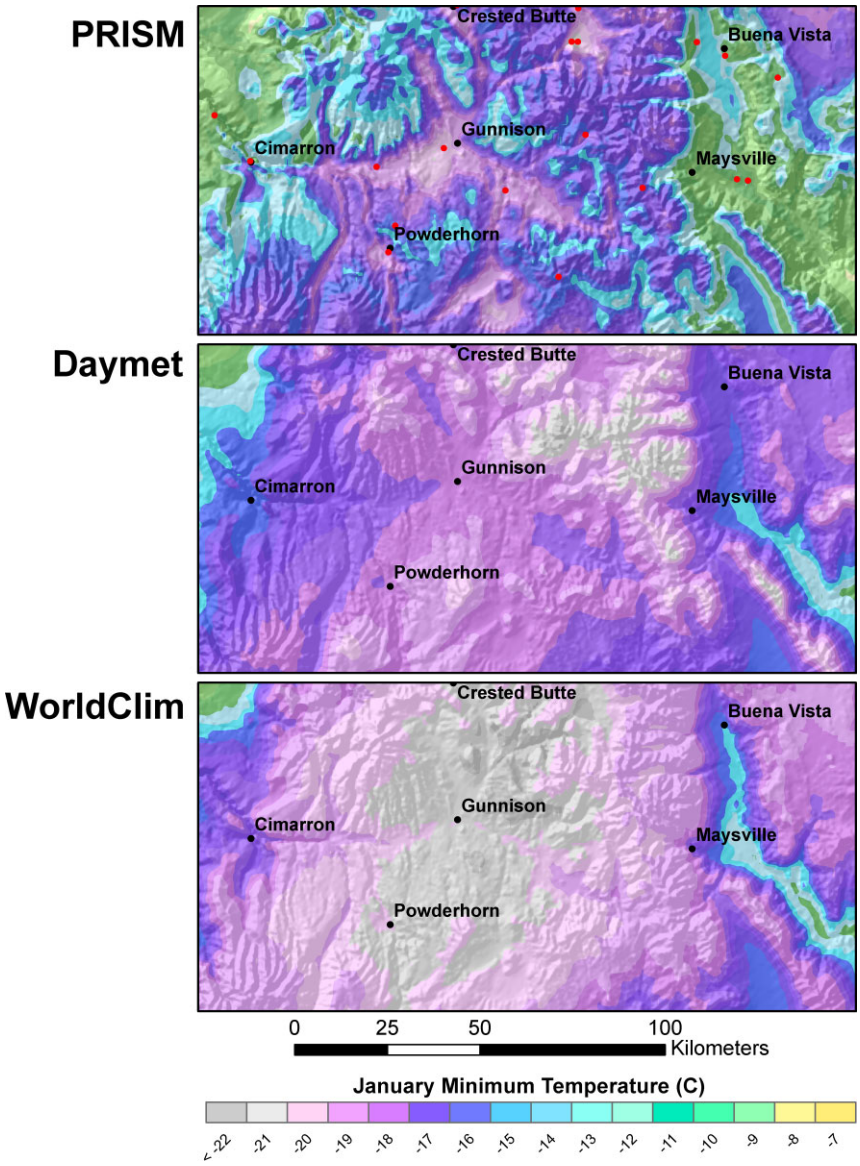

Figure 17. Comparison of PRISM, Daymet, and WorldClim January mean minimum temperature grids for the Gunnison Valley region, CO. Stations used in the PRISM interpolation are shown as red dots in the PRISM panel. This figure is available in colour online at www.interscience.wiley.com/ijoc

bins, and departures of each data set's grid cell value from the accompanying observation averaged over these bins. Stations with high topographic indexes were situated on locally elevated terrain, and those with low indexes were in local valley bottoms. The PRISM data set incorporated all these stations accurately, and showed little bias. Despite incorporating most of these stations, the Daymet data set underestimated temperatures on elevated terrain by an average of about 2 °C. WorldClim, which incorporated selected COOP stations only, substantially underestimated temperatures at all but valley bottom sites. The cold bias averaged nearly 4 °C in the 400–600-m topographic index bin, (with some underestimates as high as 10 °C), and about 3 °C in the >600 m bin. A slight improvement in the estimation of temperatures by WorldClim and Daymet between the 400–600 and >600 m bins may have been due to the fact that stations located near the top of inversions, and often the warmest in the vicinity, were located primarily within the 400–600 m bin. Above this level, temperatures generally decreased with elevation, to reach levels similar those in the valley bottoms where elevations were high enough.

Cold biases in interpolating minimum temperature were not confined to the winter months. For example, WorldClim and Daymet underpredicted July minimum temperature by an average of about 2 °C for stations in the 400–600-m topographic index bin. Nighttime temperature inversions are widespread over the western United States all times of the year, and tend to occur most frequently in seasons dominated by clear, calm weather.

5.2.4. *Central California coast*

A comparison of July mean maximum temperature grids for the central California coast is shown in

DOI: 10.1002/joc

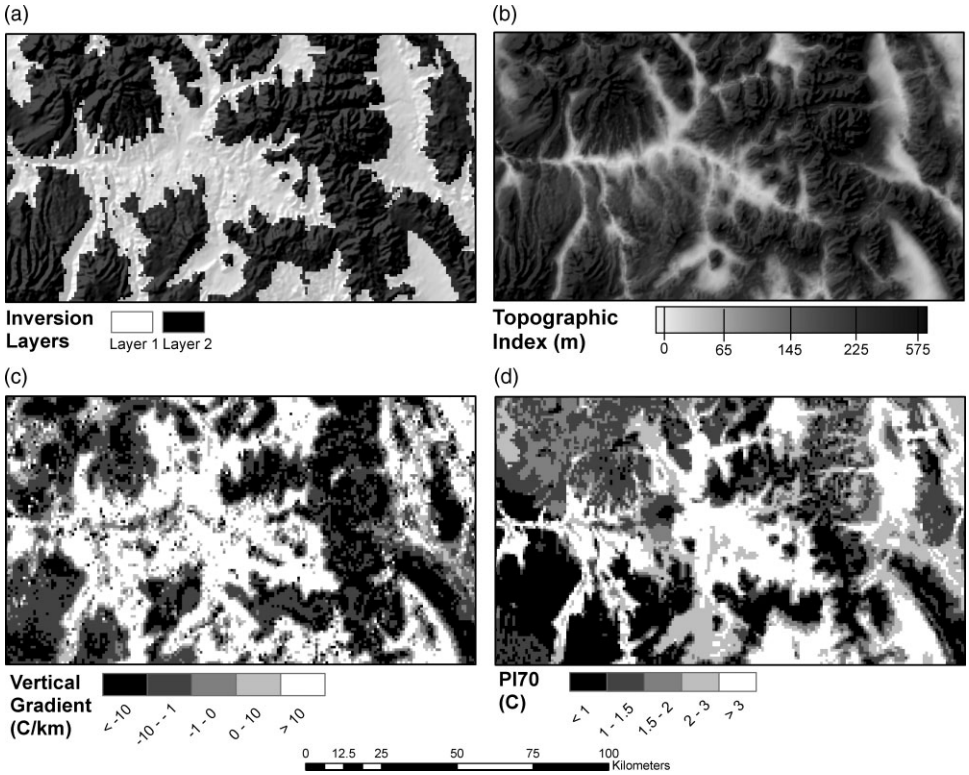

Figure 18. PRISM diagnostic grids associated with January mean minimum temperature interpolation for the Gunnison Valley region, CO: (a) inversion layer; (b) topographic index; (c) vertical temperature gradient; and (d) 70% prediction interval (PI70).

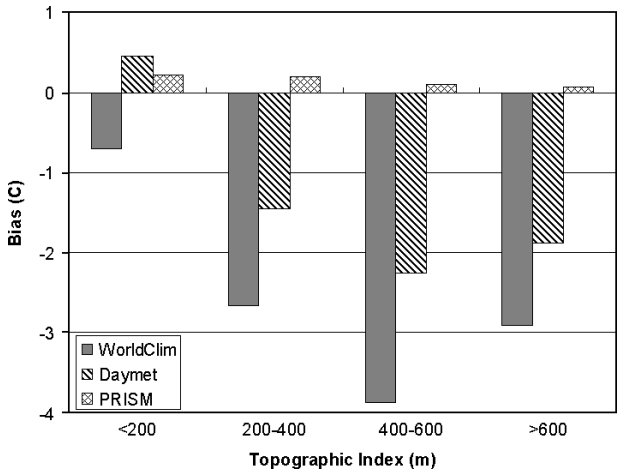

Figure 19. Biases between grid-cell estimates and observed 1971–2000 mean January minimum temperatures (estimate – observation), grouped by topographic index, for 369 stations in western Colorado. Stations with high topographic indexes are on locally elevated terrain, and those with low indexes are in local valley bottoms. Stations with moderate indexes (300–500) are neither in local valley bottoms nor on locally elevated terrain.

Figures 20–22. In summer, Pacific Ocean temperatures are much cooler than over land, resulting in strong temperature gradients from the coast inland. The summer daytime temperature regime of a location is strongly influenced by elevation (more specifically, its vertical position relative to the marine inversion) and exposure to maritime air mass penetration. Cool oceanic air concentrated at the surface undercuts warmer air aloft, forming a marine inversion. Terrain-blocking of maritime air mass penetration inland can be significant, especially if the height of the barrier exceeds the height of the marine inversion. The approximate height of the marine inversion and patterns of maritime air penetration are shown in Figure 21. The PRISM data set simulates the resulting complex temperature pattern in a realistic way. The coolest temperatures are confined to the immediate coast when terrain blocks inland penetration, but are allowed to extend inland where terrain gaps are present. The Big Sur area between King City and the coast is an example of a very thin coastal strip confined by a significant mountain barrier; high temperatures dominate just a few kilometres from the coastline (Figure 20). King City itself is in the Salinas Valley, a well-known conduit for marine air penetration via the Monterey Bay. Vertical temperature gradients were primarily positive along the coast, due to the marine inversion, but reverted to generally negative inland (Figure 21). PI70 values were generally 0.5–1.5 °C, except for sharp marine/non-marine transition areas, such as along the inversion height boundaries and near significant terrain barriers, where values exceeded 2 °C.

Daymet did not simulate July maximum temperature well in this region. The temperature field appeared heavily smoothed, with temperatures overestimated by 3–5 °C along the coast and underestimated by 5–10 °C inland. There was no recognition of coastal proximity or of the marine inversion. WorldClim did somewhat better along the coast, more accurately defining the cooler Salinas Valley, for example. However, a lack of information

DOI: 10.1002/joc

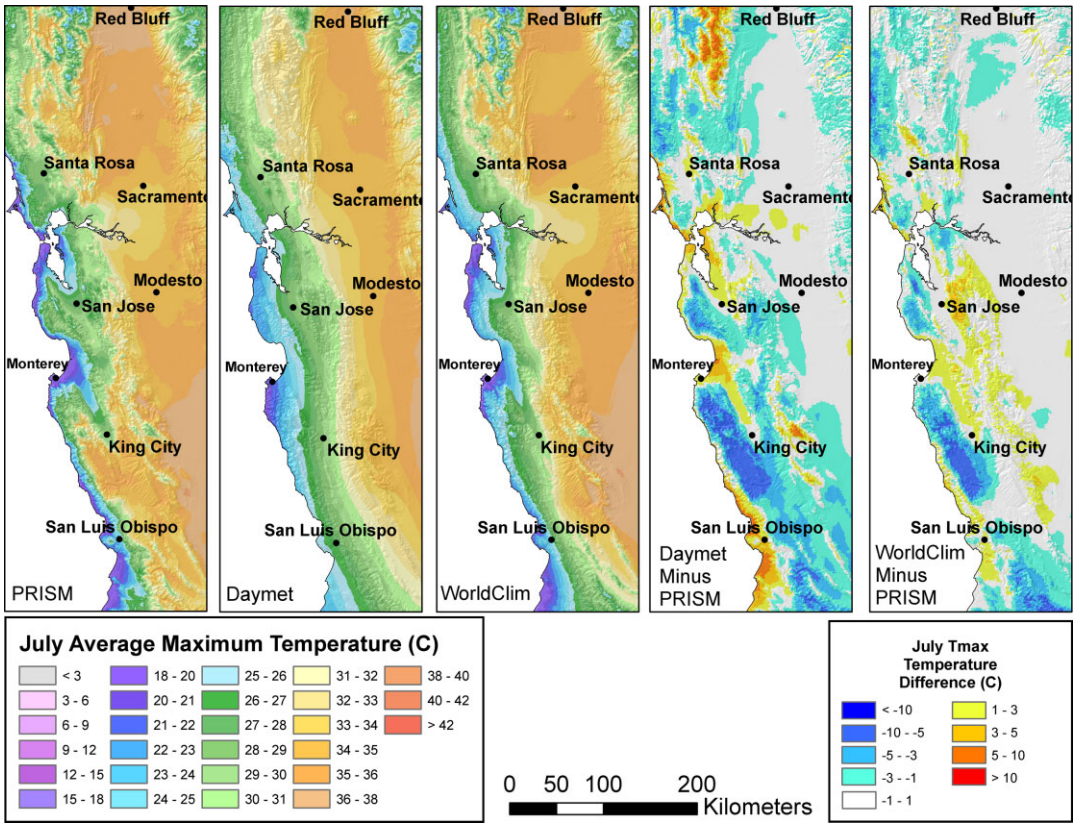

Figure 20. Comparison of PRISM, Daymet, and WorldClim July mean maximum temperature grids for the central coast, CA. This figure is available in colour online at www.interscience.wiley.com/ijoc

about coastal proximity and the marine inversion produced underestimates of inland temperatures by 3–10 °C. Biases in July maximum temperature were quantified by comparing the observed and estimated temperatures at 1201 stations in California, grouped by coastal proximity. The coastal proximity index is not perfect in its depiction of coastal influence, but, in general, stations with low coastal proximity indexes experienced a relatively strong marine temperature influence during summer, and those with high indexes experienced relatively little influence. A strong overestimation of temperatures at immediate coastal sites is apparent for Daymet and WorldClim in Figure 22, with Daymet's warm bias reaching nearly 3 °C along the immediate coastline. In contrast, inland stations that experienced low to moderate coastal influence were underestimated.

6. Conclusions

It is clear from this analysis that accounting for physiographic features in the interpolation process resulted in substantial improvements in the resulting climate grids. The improvement was highly variable, however, depending on the complexity of the region and representativeness of the station data. Regions that do not have major terrain features and are distant from climatically important coastlines, such as the Great Plains and flat areas of the eastern United States, had the simplest long-term spatial climate patterns. Physiographic effects were subdued in this area and station data were plentiful, reducing the need for physiographic weighting in the interpolation process. Near coastlines in flat terrain, coastal proximity was the main physiographic feature to be accounted for. Proximity could be reasonably estimated with simple distance measures, taking into account temperature differences between the open ocean and bays and inlets. Complex terrain in the eastern United States exhibited precipitation and temperature patterns dominated by elevation, requiring interpolation methods that accounted for elevation effects. Rain shadows and temperature inversions were not as pronounced as in the western United States, and these areas were represented by relatively numerous station data.

Most of the western United States, with its complex topography and adjacent Pacific Coast, required much more than basic elevation information for successful climate interpolation. Precipitation interpolation over many mountain ranges benefited from PRISM topographic facet and trajectory-based moisture index station weighting to realistically simulate sharp rain shadows to the lee of the mountain crests. Interpolation of minimum temperature, especially winter minima, benefited from PRISM station weighting functions that accounted for topographic position and the potential for temperature inversions. The reversal of the temperature–elevation relationship due to inversions was a particularly challenging situation to simulate. Maximum temperature, especially during summer, had a generally strong and monotonic relationship with

DOI: 10.1002/joc

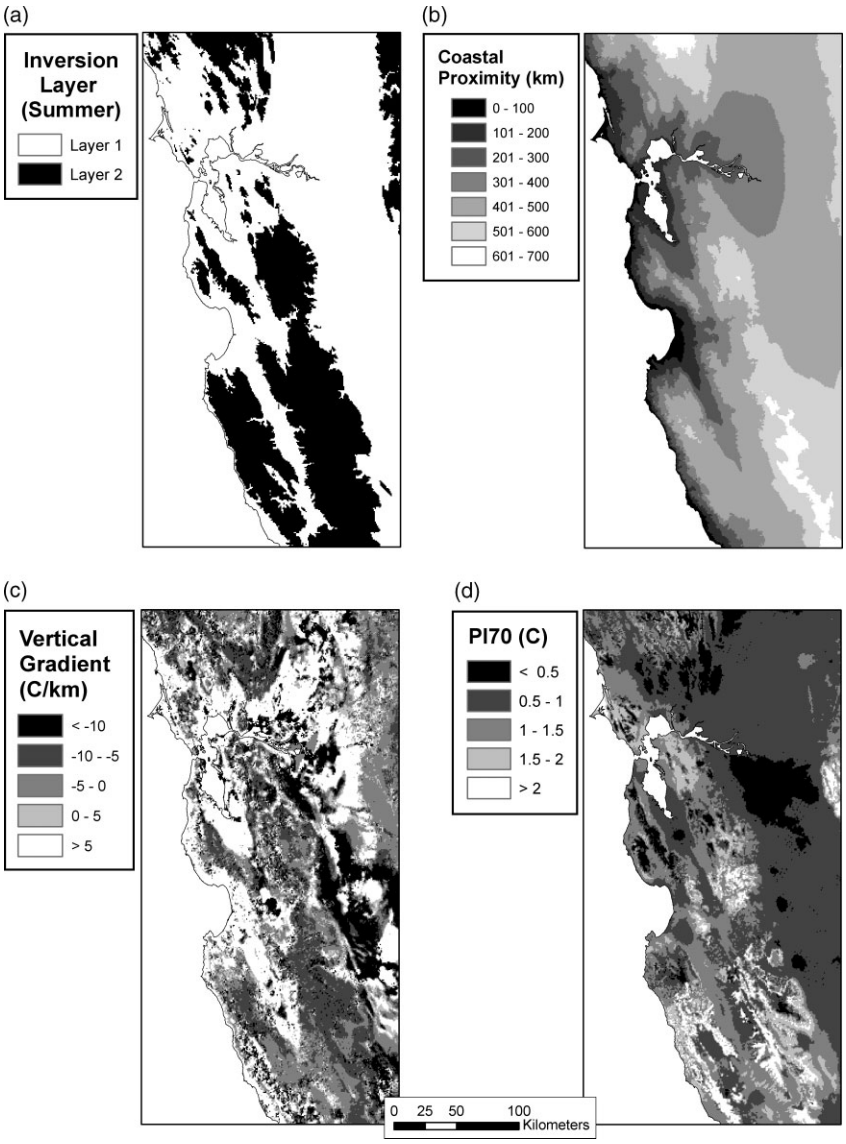

Figure 21. PRISM diagnostic grids associated with July mean maximum temperature interpolation for the central coast, CA: (a) summer inversion layer, given as layer 1 (within inversion) or layer 2 (above inversion); (b) western temperature (TW) coastal proximity, expressed as the optimal path length from the coastline in kilometres, accounting for terrain blockage; (c) vertical temperature gradient; and (d) 70% prediction interval (PI70).

elevation, and therefore was interpolated successfully over inland areas with elevation information only. The most difficult area for interpolating summer maximum temperature was located along, and adjacent to, the Pacific coastline. Simulation of large land–sea temperature differences combined with a marine inversion benefited from station weighting functions that included a trajectory-based coastal proximity index and marine inversion estimates. Again, the reversal of the temperature–elevation relationship near the coast was challenging to reproduce.

An in-depth performance evaluation was presented that included attempts to estimate the interpolation error, and a comparison of PRISM data sets with the Daymet and WorldClim data sets. PRISM interpolation error was estimated in two ways: jackknife C-V and the prediction interval of the climate–elevation regression function at the 70% significance level (PI70). Both methods had advantages and disadvantages. C-V provided information from actual station data, but its purview was limited to those points only, which was quite limited in data-sparse areas. PI70 provided information for all grid cells, but was based on the assumption that the interpolation error was based solely on the scatter of the data points around the climate–elevation regression function. There was an advantage in using both of these methods, in that they estimated error from very different perspectives. Although not well correlated point-by-point, the two measures were found to be reasonably well correlated, and roughly similar in magnitude, when averaged over large domains.

The PRISM data set was shown to be a more accurate representation of the spatial climate patterns in the United States than the WorldClim and Daymet data sets. The

DOI: 10.1002/joc

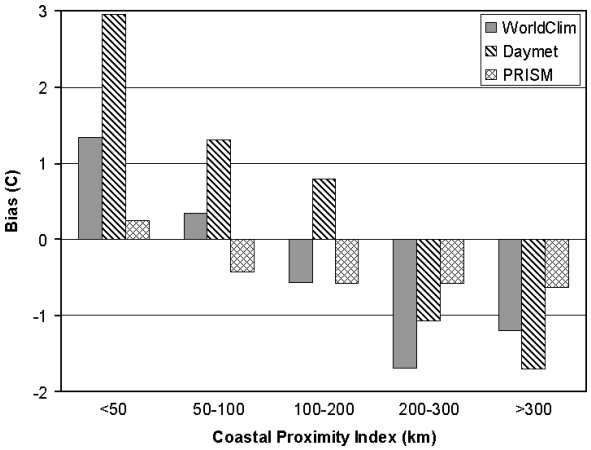

Figure 22. Bias between grid-cell estimates and observed 1971–2000 mean July maximum temperatures (estimate – observation), grouped by coastal proximity (Figure 2(a)), for 1201 stations in California. Stations with low coastal proximity indexes experience a generally strong marine temperature influence during summer, and those with high indexes experience much less of this influence (see Figures 2(a) and 21(b) for spatial context).

WorldClim and Daymet data sets were somewhat similar to PRISM in many areas within the eastern United States, but very different in the western United States. Reasons stemmed from differences in station data density, QC, and interpolation methods. WorldClim used only GHCN stations, which comprised only about 10% of the total number of stations used in PRISM. Most of these stations were in valleys, which resulted in substantial underestimates of both precipitation and minimum temperature in the mountains. WorldClim's ANUSPLIN interpolation method accounted for direct elevation effects only, and had difficulty reproducing sharp changes in the relationship between climate and elevation. Daymet used a much richer station data set that included many COOP and SNOTEL sites. However, the Daymet interpolation method assumed a linear, monotonic relationship between climate and elevation, and accounted for direct elevation effects only, which resulted in poor simulation of minimum temperature in the mountains and maximum temperature near the Pacific coast. WorldClim and Daymet data sets did not account for rain shadows in the lee of most mountain ranges well, owing to a lack of terrain barrier information. An additional advantage to the PRISM data set was a peer review procedure that brought significant knowledge of regional climate patterns and new station data sets into the development process.

Future work to improve the PRISM spatial climate data sets will involve casting an ever-wider net for high-quality meteorological data in remote or complex areas, including data from mesonets. Data from offshore buoys could serve as a useful baseline for along-shore temperature conditions. Additional station-weighting functions under consideration would assess similarities in land use/land cover types, thereby improving stratification of stations in open *versus* forested locations, built *versus* rural settings, etc. However, implementation would require high-quality spatial information on historical and current land use/land cover, highly accurate station coordinates for the POR, and generalizable relationships between variations in land cover and variations in climate. Additional work needs to be done to improve estimates of inversion heights, coastal proximity, and moisture availability, possibly through the use of more sophisticated models.

Acknowledgements

Thanks go to Eileen Kaspar for careful editing of this manuscript. This work was funded in part by Agreement 68-7482-2-36Y with the USDA-NRCS National Water and Climate Center.

Appendix A

A1. Period of record adjustments to short-term stations

Many stations used in this analysis had a POR with fewer than 23 years (75% data completeness threshold). In order to remove temporal biases from these short-term stations, their POR means were adjusted to the 1971–2000 period. Nearby 'anchor' stations with long-term means were used to adjust the short-term 'target' station means. A target station monthly mean was considered usable (and therefore adjustable) if the mean was calculated using at least three years of data from the 1971 to 2000 period. A target station monthly mean was also usable if the target station had at least three years of historical data, and at least one of these years was within 23 years of the beginning or end of the 1971–2000 period, which encompasses the years 1948–2003. A target station monthly mean was not usable if the station had less than three years of historical data, and no data during 1948–2003. Target station monthly means that could not be used were omitted from the data set.

Anchor stations were chosen for each target station using results from the ASSAY QC application using station POR monthly averages as input (initial spatial QC step). For each station for each month, ASSAY QC made a prediction for each target station in the absence of the target station's observation. The weights assigned by ASSAY to surrounding stations in the monthly regression function for the target station served as measures of the physiographic similarity of these stations to the target station. These weights were later used in determining the most suitable anchor stations.

Potential anchor stations were those having data in the same years during 1971–2000, which were used to calculate the target station short-term monthly mean. For example, if a target station had data for the period 1982–1988, the anchor station must have also had data for this period. Secondly, anchor stations were chosen that had data in as many of the additional years outside 1971–2000 where the target station had data, up to a total of 23 years, if possible. The number of additional years that could be provided by the anchor station, divided by the number of years needed to bring the target station

DOI: 10.1002/joc

POR up to 23, was multiplied by the ASSAY weight of the station to produce a ranking score. Anchor stations were then sorted on the basis of the highest score, then highest ASSAY weight, and then the shortest distance to the target station. The result was a list of preferred anchor stations, each of which had a set of up to 23 years for which both the anchor station and target station had data for a given month.

After selecting an anchor station, monthly extended means for the target station and anchor station were calculated using their respective data from the set of common years. These means were called 'extended' because they might have been calculated from a period of years that extended beyond the target period. For precipitation, an adjustment factor was calculated by dividing the anchor station 1971–2000 mean by the anchor station extended mean. The target station extended mean was multiplied by this factor to determine the target station adjusted mean. For maximum and minimum temperature, an adjustment difference was calculated by subtracting the anchor station extended mean from the anchor station 1971–2000 mean. This difference was added to the target station extended mean to determine the target station adjusted mean. Arithmetically, the target adjusted mean ($\overline{X_t}'$) was calculated as follows for precipitation:

$$\overline{X_t}' = \overline{X_{te}}\left(\frac{\overline{X_a}}{X_{ae}}\right) \tag{A1}$$

and for maximum and minimum temperature:

$$\overline{X_t}' = \overline{X_{te}} + (\overline{X_a} - \overline{X_{ae}}) \tag{A2}$$

where $\overline{X_a}$ is the anchor mean, $\overline{X_{ae}}$ is the anchor extended mean, and $\overline{X_{te}}$ is the target extended mean.

$\overline{X_t}'$ was calculated for each target station using the three anchor stations with the highest scores. These three adjusted means were then averaged to obtain the final 1971–2000 mean for the target station. The use of three anchor stations was arrived at by conducting experiments with 84 long-term stations across the United States for which the true 1971–2000 thirty-year means were known. These stations were simulated as short-term sites by reducing their PORs to 5, 11, and 21 years, respectively. Estimates of the 30-year mean were calculated through the above adjustment procedure using between one and five anchor stations. Comparison of the estimated and actual 30-year means indicated that the differences between the two decreased as the number of anchor stations increased from one to three, but leveled off (or actually increased in some cases) as the number of anchors increased to four and five. As was expected, the shorter the POR of the short-term station, the larger the differences, but the mean absolute difference between the estimated and actual 1971–2000 mean was well below 1 °C (Figure A1).

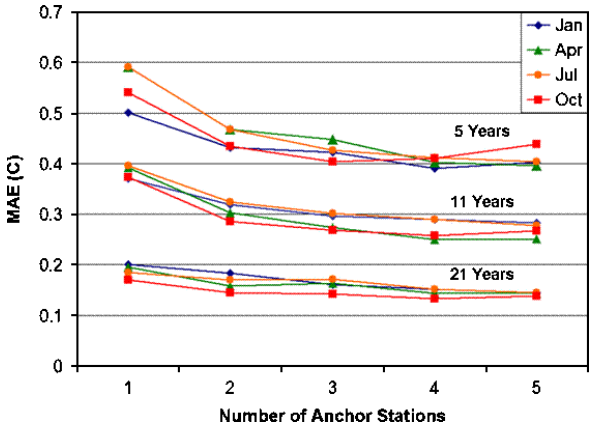

Figure A1. Mean absolute error of the adjusted mean July maximum temperature from the known 30-year mean *versus* number of anchor stations used to compute the adjusted mean. A total of 84 stations were tested using short-term periods of record of 5, 11, and 21 years. This figure is available in colour online at www.interscience.wiley.com/ijoc

Appendix B

B1. PRISM cluster weighting

The PRISM climate–elevation regression function is populated with data from stations surrounding the target grid cell that are within a given radius of influence (described in Daly *et al.*, 2002). Ideally, these stations would be evenly spaced both horizontally and vertically, so that all areas would be represented equally. This is rarely the case, however. Stations often occur in clumps (e.g. populated areas), thereby creating the potential for overrepresentation in the regression function. In an attempt to minimize overrepresentation, the PRISM cluster-weighting algorithm seeks to identify and reduce the regression weights of stations that are clustered together, so that the cluster as a whole has the approximate weight of a single station.

How closely spaced can stations be before they are characterized as representing essentially the same point, and therefore perceived as clustered? Horizontally, how we intuitively define a cluster varies with the scale of the interpolation process, and is best related to the horizontal radius of influence within which stations are drawn for the regression function; the larger the radius of influence, the farther stations can be spaced before they are perceived to be clustered. Over the past decade of modelling with PRISM, it has been found that a horizontal separation distance of less than 20% of the radius of influence suggests clustering. The radius of influence in PRISM varies from pixel to pixel, depending on the local density of stations, but is typically on the order of 30–50 km; this suggests that a clustering influence will be begin to be felt at horizontal separations of 6–10 km.

Vertically, the cluster distance threshold represents how close the elevation stations can be before they are characterized as representing the same elevation. A PRISM input parameter, called the elevation precision, represents the precision to which the station elevations are given for the regression function. This is used to keep stations with very small differences in elevation but

DOI: 10.1002/joc

relatively large differences in climate values from unduly influencing the regression slope. A common value for the elevation precision is 50 m; this means that station elevations within 50 m are assumed to be at the same elevation. This precision was adopted as the vertical cluster threshold.

For a given station, the cluster weighting factor, W_c, is calculated as

$$W_c = \left\{ \begin{array}{l} 1;\ S_c = 0 \\ 1/S_c;\ S_c > 0 \end{array} \right\} \tag{B1}$$

where S_c, the cluster size (number of effective stations in cluster), is given as

$$S_c = \sum_{i=1}^{n} h_i v_i \tag{B2}$$

where n is the number of stations in the regression data set, and h_i and v_i are the horizontal and vertical cluster factors for station i, respectively. These factors are calculated as follows:

$$h_i = \sum_{j=1}^{n-1} \left\{ \begin{array}{l} 0;\ d_{ij} > 0.2\mathrm{r} \\ (0.2r - d_{ij})/0.2\mathrm{r};\ 0 \le d_{ij} \le 0.2\mathrm{r} \end{array} \right\} \tag{B3}$$

where r is the radius of influence and d_{ij} is the horizontal distance between station i and all other stations j in the regression data set; and

$$v_i = \sum_{j=1}^{n-1} \left\{ \begin{array}{l} 0;\ s_{ij} > p \\ (p - s_{ij})/p;\ 0 \le s_{ij} \le p \end{array} \right\} \tag{B4}$$

where p is the elevation precision and s_{ij} is the vertical distance between station i and all other stations j in the regression data set. s_{ij} is calculated as

$$s_{ij} = \left\{ \begin{array}{l} 0;\ |e_i - e_j| < p \\ |e_i - e_j| - p;\ |e_i - e_j| > p \end{array} \right\} \tag{B5}$$

where e_i and e_j are the elevations of station i and station j, respectively. Station pairs with elevation differences within the elevation precision are assumed to have effectively the same elevation: hence the subtraction of p from the elevation difference to obtain the effective separation of the two stations.

Appendix C

C1. PRISM effective terrain height assessment

In the mountainous western United States, terrain is a dominant factor in the spatial pattern of precipitation. In flat or gently rolling areas such as the Great Plains, the role of terrain is more subdued, although variations in precipitation have been documented over low hills in Illinois (Changnon *et al*., 1975) and Sweden (Bergeron, 1968), and in narrow valleys in Canada (Longley, 1975). Conceptually, the effectiveness of a terrain feature in amplifying precipitation depends partially on its ability to block and uplift moisture-bearing air. This ability is determined mainly by the profile the feature presents to the oncoming air flow. Steeply rising features with continuous ridge lines oriented normal to the flow can generally be expected to produce greater P–E regression slopes than low, gently rising features with discontinuous ridge lines oriented parallel to the flow. One might imagine a spectrum of 'effective' terrain heights, ranging from large features that produce highly 3D precipitation patterns, to a nearly flat condition which exhibits 2D patterns only. Between these extremes would be a transition between 2D and 3D patterns, for which P–E slopes would range from zero to values typical of mountainous areas.

Ideally, the effectiveness of the terrain would be reflected in the station data, and therefore in the empirical P–E regression slopes. In reality, the station data are rarely of sufficient density and reliability to provide such a detailed and accurate picture. If spatial estimates of the 2D/3D nature of the terrain were available *a priori*, the range of allowable P–E slopes could be varied to the appropriate degree, providing an independent check and constraint (if necessary) on the empirically-derived P–E slopes.

The effective terrain height for a pixel is estimated by a method similar to that used in estimating the potential wintertime inversion height; the creation of an effective terrain height grid for the conterminous United States is described in Section 4 and Table II of this article.

PRISM uses the effective terrain grid in a multi-step process. A 3D index for the target grid cell (I_{3c}) is determined by comparing the effective terrain height of the target cell with thresholds for 2D and 3D model operation. If the effective terrain height exceeds the 3D threshold, I_{3c} is set to 1.0. If the effective terrain height is less than the 2D threshold, I_{3c} is set to 0.0. An effective terrain height between the two thresholds gives an I_{3c} between zero and 1. The calculation is as follows:

$$I_{3c} = \left\{ \begin{array}{l} 1;\ h_c \ge h_3 \\ \dfrac{h_c - h_2}{h_3 - h_2};\ h_2 < h_c < h_3 \\ 0;\ h_c \le h_2 \end{array} \right\} \tag{C1}$$

where h_c is the effective terrain height for the target grid cell, and h_2 and h_3 are user-defined thresholds for 2D and 3D operation, respectively. At the current time, h_2 and h_3 are empirically defined until more robust methods are developed on the basis of a combination of theoretical and observational studies.

If $I_{3c} < 1$, signalling a 2D or 2D/3D mixed situation, I_{3a}, an areal 3D index, is calculated to assess whether the target grid cell is near a significant 3D terrain feature. Precipitation patterns may be affected by the upstream or downstream effects of mountain barriers, well away from the barriers themselves (Smith, 1979). Indications of this phenomena have been observed during our analysis of observations in many large valleys in the Western United States, and it appears to extend approximately 100 km from the nearest 3D terrain features. The 100-km estimate

DOI: 10.1002/joc

is preliminary, and may be updated as more information is gathered.

I_{3a} is calculated similar to I_{3c}:

$$I_{3a} = \left\{ \begin{array}{l} 1; h_a \geq h_3 \\ \frac{h_a - h_2}{h_3 - h_2}; h_2 < h_a < h_3 \\ 0; h_a \leq h_2 \end{array} \right\} \quad \text{(C2)}$$

where h_a is a distance-weighted effective terrain height, calculated as

$$h_a = \frac{\sum_{i=1}^{n} w_i h_i}{n} \quad \text{(C3)}$$

where h_i is the effective terrain height for grid cell i and n is the number of grid cells within 100 km of the target grid cell. The weight w for a nearby grid cell i is

$$w_i = \frac{1}{d_i} \quad \text{(C4)}$$

where d_i is the horizontal distance between the centres of the target grid cell and nearby grid cell i.

The final 3D index, I_{3d}, is expressed as

$$I_{3d} = \max[I_{3c}, I_{3a}] \quad \text{(C5)}$$

A scalar from 0 to 1, I_{3d} represents the degree of importance terrain should play in the estimation of precipitation. When $I_{3d} = 1$, the PRISM regression function operates in its normal fashion. As I_{3d} approaches zero, the influence of terrain is gradually diminished and values of terrain-related parameters – the minimum, maximum, and default regression slopes (β_{1m}, β_{1x}, and β_{1d}); and the elevation, facet and layer weighting exponents (b, c, and y) – are linearly reduced to zero as a function of I_{3d} (see Daly *et al*., 2002, for definitions of PRISM parameters). When I_{3d} is zero, the slope of the precipitation/elevation regression function is forced to zero and stations are weighted by distance, clustering, and coastal proximity (if enabled) only, resulting in a 2D interpolation.

References

Atkinson DE, Gajewski K. 2002. High-resolution estimation of summer surface air temperature in the Canadian Arctic Archipelago. *Journal of Climate* **15**: 3601–3614.

Barnes SL. 1964. A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology* **3**: 396–409.

Barry RG. 1992. *Mountain Weather and Climate*. Routledge: London.

Barry RG, Chorley RJ. 1987. *Atmosphere, Weather and Climate*, 5th edn. Routledge: London.

Bergeron T. 1968. *Studies of the Oreigenic Effect on the Areal fine Structure of Rainfall Distribution*. Meteorological Institute, Uppsala University, Report No. 6.

Bolstad PV, Swift L, Collins F, Regniere J. 1998. Measured and predicted air temperatures at basin to regional scales in the southern Appalachian Mountains. *Agricultural and Forest Meteorology* **91**: 161–176.

Bootsma A. 1976. Estimating minimum temperature and climatological freeze risk in hilly terrain. *Agricultural Meteorology* **16**: 425–443.

Brosofske KD, Chen J, Naiman RJ, Franklin JF. 1997. Harvesting effects on microclimatic gradients from small streams to uplands in western Washington. *Ecological Applications* **7**: 1188–1200.

Changnon SA Jr, Jones DMA, Huff FA. 1975. Precipitation increases in the low hills of southern Illinois. Part 2. Field investigation of anomaly. *Monthly Weather Review* **103**: 830–836.

Clements CB, Whiteman CD, Horel JD. 2003. Cold-air-pool structure and evolution in a mountain basin: Peter Sinks, Utah. *Journal of Applied Meteorology* **42**: 752–768.

Daley R. 1991. *Atmospheric Data Analysis*. Cambridge University Press: Cambridge; 457.

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

Daly C, Neilson RP, Phillips DL. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology* **33**: 140–158.

Daly C, Helmer EH, Quinones M. 2003. Mapping the climate of Puerto Rico, Vieques, and Culebra. *International Journal of Climatology* **23**: 1359–1381.

Daly C, Smith J, McKane R. 2007. High-resolution spatial modeling of daily weather elements for a catchment in the Oregon Cascade Mountains, USA. *Journal of Applied Meteorology and Climatology* **46**: 1565–1586.

Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P. 2002. A knowledge-based approach to the statistical mapping of climate. *Climate Research* **22**: 99–113.

Daly C, Taylor GH, Gibson WP, Parzybok TW, Johnson GL, Pasteris P. 2001. High-quality spatial climate data sets for the United States and beyond. *Transactions of the American Society of Agricultural Engineers* **43**: 1957–1962.

Daly C, Redmond K, Gibson W, Doggett M, Smith J, Taylor G, Pasteris P, Johnson G. 2005. Opportunities for improvements in the quality control of climate observations. *15th AMS Conference on Applied Climatology*, American Meteorological Society: Savannah, GA, June 20–23, 2005. Paper J3.9. http://ams.confex.com/ams/pdfpapers/94199.pdf.

Daly C, Kittel TGF, McNab A, Gibson WP, Royle JA, Nychka D, Parzybok T, Rosenbloom N, Taylor G. 2000. Development of a 103-year high-resolution climate data set for the conterminous United States. *Proceedings of the 12th AMS Conference on Applied Climatology*. American Meteorological Society: Asheville, NC, May 8–11, 249–252.

Davey CA, Pielke RA Sr. 2005. Microclimate exposures of surface-based weather stations–implications for the assessment of long-term temperature trends. *Bulletin of the American Meteorological Society* **86**: 497–504.

Dong J, Chen J, Brosofske KD, Naiman RJ. 1998. Modelling air temperature gradients across managed small streams in western Washington. *Journal of Environmental Management* **53**: 309–321.

Farnsworth R, Thompson ES, Peck EL. 1982. *Evaporation Atlas for the Contiguous 48 United States*. NOAA Technical Report #33.

Funk C, Michaelsen J. 2004. A simplified diagnostic model of orographic rainfall for enhancing satellite-based rainfall estimates in data-poor regions. *Journal of Applied Meteorology and Climatology* **43**: 1366–1378.

Geiger R. 1964. *The Climate Near the Ground*. Harvard University Press: Cambridge, MA; 611.

Giambelluca TW, Nullet D. 1991. Influence of the trade-wind inversion on the climate of a leeward mountain slope in Hawaii. *Climate Research* **1**: 207–216.

Gibson WP, Daly C, Taylor GH. 1997. Derivation of facet grids for use with the PRISM model. *Proceedings of the 10th AMS Conference on Applied Climatology*. American Meteorological Society: Reno, NV, Oct. 20–23, 208–209. Online at http://prism.oregonstate.edu/pub/prism/docs/appclim97-facets-gibson.html.

Gibson WP, Daly C, Kittel T, Nychka D, Johns C, Rosenbloom N, McNab A, Taylor G. 2002. Development of a 103-year high-resolution climate data set for the conterminous United States. *Proceedings of the 13th AMS Conference on Applied Climatology*. American Meteorological Society: Portland, OR, May 13–16, 181–183.

Gustavsson T, Karlsson M, Bogren J, Lindqvist S. 1998. Development of temperature patterns during clear nights. *Journal of Applied Meteorology* **37**: 559–571.

Gyalistras D. 2003. Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951–2000). *Climate Research* **25**: 55–83.

Hannaway DB, Daly C, Cao W, Luo W, Wei Y, Zhang W, Xu A, Lu C, Shi X, Li X. 2005. Forage species suitability mapping for China using topographic, climatic and soils spatial data and quantitative plant tolerances. *Scientia Agricultura Sinica* **4**: 660–667.

DOI: 10.1002/joc

Haugen RK, Brown J. 1980. Coastal-inland distributions of summer air temperature and precipitation in northern Alaska. *Arctic and Alpine Research* **12**: 403–412.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Hocevar A, Martsolf JD. 1971. Temperature distribution under radiation frost conditions in a central Pennsylvania valley. *Agricultural Meteorology* **8**: 371–383.

Hutchinson MF. 1995. Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Science* **9**: 385–403.

Johnson GL, Daly C, Hanson CL, Lu YY, Taylor GH. 2000. Spatial variability and interpolation of stochastic weather simulation model parameters. *Journal of Applied Meteorology* **39**: 778–796.

Juvik JO, Nullet D, Banko P, Hughes K. 1993. Forest climatology near the tree line in Hawai'i. *Agricultural and Forest Meteorology* **66**: 159–172.

Legates DR, McCabe GJ. 1999. Evaluating the use of "goodness of fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research* **35**: 233–241.

Lindkvist L, Gustavsson T, Bogren J. 2000. A frost assessment method for mountainous areas. *Agricultural and Forest Meteorology* **102**: 51–67.

Longley RW. 1975. Precipitation in valleys. *Weather* **30**: 294–300.

Lookingbill T, Urban D. 2003. Spatial estimation of air temperature differences for landscape-scale studies in montane environments. *Agricultural and Forest Meteorology* **114**: 141–151.

Mahmood R, Foster SA, Logan D. 2006. The GeoProfile metadata, exposure of instruments, and measurement bias in climatic record revisited. *International Journal of Climatology* **26**: 1091–1124.

McCutchan MH, Fox DG. 1986. Effect of elevation and aspect on wind, temperature and humidity. *Journal of Climate and Applied Meteorology* **25**: 1996–2013.

Mendonca BG, Iwaoka WT. 1969. The trade wind inversion at the slopes of Mauna Loa, Hawaii. *Journal of Applied Meteorology* **8**: 213–219.

Milewska EJ, Hopkinson RF, Niitsoo A. 2005. Evaluation of geo-referenced grids of 1961–1990 Canadian temperature and precipitation normals. *Atmosphere-Ocean* **43**: 49–75.

Neter J, Wasserman W, Kutner MH. 1989. *Applied Linear Regression Models*, 2nd edn. Richard D. Irwin, Inc: Boston, MA.

NOAA-NCDC. 2003. *Data Documentation for data set 3220, Summary of the Month Cooperative*. National Climatic Data Center, National Oceanic and Atmospheric Administration. http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3220.pdf.

Oke TR. 1978. *Boundary Layer Climates*. Routledge: New York.

Schwarb M, Daly C, Frei C, Schar C. 2001a. *Mean seasonal precipitation throughout the European Alps, 1971–1990*. Hydrologic Atlas of Switzerland, National Hydrologic Service, Bern, Switzerland, hard copy.

Schwarb M, Daly C, Frei C, Schar C. 2001b. *Mean Annual Precipitation Throughout the European Alps, 1971–1990*. Hydrologic Atlas of Switzerland, National Hydrologic Service, Bern, Switzerland, hard copy.

Sharples JJ, Hutchinson MF, Jellet DR. 2005. On the horizontal scale of elevation dependence of Australian monthly precipitation. *Journal of Applied Meteorology* **44**: 1850–1865.

Simpson JJ, Hufford GL, Daly C, Berg JS, Fleming MD. 2005. Comparing maps of mean monthly surface temperature and precipitation for Alaska and adjacent areas of Canada produced by two different methods. *Arctic* **58**: 137–161.

Smith RB. 1979. The influence of mountains on the atmosphere. *Advances in Geophysics* **21**: 87–230.

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.

USDA-NRCS. 1998. *PRISM Climate Mapping Project–Precipitation*. Mean monthly and annual precipitation digital files for the continental U.S. USDA-NRCS National Cartography and Geospatial Center, Ft. Worth TX. December, CD-ROM and online. http://www.ncgc.nrcs.usda.gov/products/datasets/climate/.

USDOI-USGS. 2007. *Calendar year Streamflow Statistics for Washington*. Water Resources of Washington, US Geologic Survey, Boulder, CO, available online at http://water.usgs.gov/wa/nwis/annual.

Williams CN, Vose RS, Easterling DR, Menne MJ. 2004. *United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data*. ORNL/CDIAC-118, NDP-070. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Willmott CJ, Matsuura K. 1995. Smart Interpolation of annually averaged air temperature in the United States. *Journal of Applied Meteorology* **34**: 2577–2586.

Willmott CJ, Ackleson SG, Davis RE, Feddema JJ, Klink KM, Legates DR, O'Donnell L, Rowe CM. 1985. Statistics for the evaluation and comparison of models. *Journal of Geophical Research* **90**: 8995–9005.

WMO. 1989. *Calculation of Monthly and Annual 30-year Standard Normals*. Prepared by a meeting of experts, Washington, DC, USA, March 1989. World Meteorological Organization, WCDP 10, WMO-TD 341.

DOI: 10.1002/joc