INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **26**: 707–721 (2006)
Published online 27 March 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.1322

GUIDELINES FOR ASSESSING THE SUITABILITY OF SPATIAL CLIMATE DATA SETS

CHRISTOPHER DALY*
The PRISM Group, Department of Geosciences, College of Science and College of Oceanographic and Atmospheric Sciences, 326 Strand Agricultural Hall, Oregon State University, Corvallis, Oregon USA 97331, USA

Received 19 October 2005
Revised 25 January 2006
Accepted 30 January 2006

ABSTRACT

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. The recent availability of very high-resolution climate data sets raises important questions about the tendency to equate resolution with realism. This paper discusses the relationship between scale and spatial climate-forcing factors, and provides background and advice on assessing the suitability of data sets. Spatial climate patterns are most affected by terrain and water bodies, primarily through the direct effects of elevation, terrain-induced climate transitions, cold air drainage and inversions, and coastal effects. The importance of these factors is generally lowest at scales of 100 km and greater, and becomes greatest at less than 10 km. Except in densely populated regions of developed countries, typical station spacing is on the order of 100 km. Regions without major terrain features and which are at least 100 km from climatically important coastlines can be handled adequately by most interpolation techniques. Situations characterized by significant terrain features, but with no climatically important coastlines, no rain shadows, and a well-mixed atmosphere can be reasonably handled by methods that explicitly account for elevation effects. Regions having significant terrain features, and also significant coastal effects, rain shadows, or cold air drainage and inversions are best handled by sophisticated systems that are configured and evaluated by experienced climatologists. There is no one satisfactory method for quantitatively estimating errors in spatial climate data sets, because the field that is being estimated is unknown between data points. Perhaps the best overall way to assess errors is to use a combination of approaches, involve data that are as independent from those used in the analysis as possible, and use common sense in the interpretation of results. Data set developers are encouraged to conduct expert reviews of their draft data sets, which is probably the single most effective way to improve data set quality. Copyright © 2006 Royal Meteorological Society.

KEY WORDS: spatial climate data; temperature; precipitation; climate mapping; climate interpolation; cross-validation; ANUSPLIN; kriging; Daymet; PRISM

1. INTRODUCTION

There is currently a great demand for spatial climate data sets in digital form. Basic climate elements provided by these gridded data sets typically include minimum and maximum temperature and precipitation, given over a monthly (or smaller) time step, and averaged over a nominal 30-year period. The demand for these data sets has been fueled in part by the widespread adoption of computer technology that enables a variety of hydrologic, ecological, natural resource, and other models and decision support tools to be linked to geographic information systems (GIS). 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. However, despite the great importance of climate inputs to their analyses, many users do not have a substantial background in geospatial climatology (the study of the spatial patterns of climate), and are,

* Correspondence to: Christopher Daly, The PRISM Group, Department of Geosciences, College of Science and College of Oceanographic and Atmospheric Sciences, 326 Strand Agricultural Hall, Oregon State University, Corvallis, OR 97331, USA; e-mail: daly@coas.oregonstate.edu

therefore, not in a position to critically assess the suitability of spatial climate data sets for a particular application. Errors associated with spatial climate data are often, by default, assumed to be negligible, or at least smaller than those associated with other spatial data sets. This assumption can lead to serious errors in the interpretation of results, and conclusions and decisions made from those results.

The recent availability of very high-resolution climate data sets at the continental to global scale affords a good opportunity for discussion of the suitability of spatial climate data sets. These data sets are seeing heavy usage, but also raise important questions about the tendency to equate resolution with realism. In the past, gridded, computer-generated data sets designed to cover large areas were developed mainly at coarse resolution (~50 km). They were not held to the same standards of accuracy as data sets that covered small areas in great detail, or the traditional hand-drawn climate maps that engaged human knowledge in the process. Coarse-grid climate data sets could not possibly incorporate many of the spatial climate features we know to be present, such as elevational gradients, coastal effects, temperature inversions, and rain shadows, because these basic features occurred primarily at scales below those of the coarse grid. Elevation guidance (used heavily in the preparation of hand-drawn climate maps since the 1950s) began to be used in these data sets only a few years ago, with the implicit understanding that this variable was being specified at a coarse scale, thus allowing interpolation techniques to avoid close scrutiny. Instead, attention was turned to issues other than the detailed climatic realism of the gridded data, such as the inclusion of station data, quality control (QC), data processing steps, and other activities required to produce data sets covering large areas.

Today's computing capabilities now allow very fine-resolution climate grids (~1 km) to be created over most of the globe. To many users, an increase in resolution from tens of km to 1 km carries expectations of increased realism and accuracy. This may or may not be true, depending on the circumstances. For example, it is likely that less error and bias would be involved when estimating a climate value for a given location from a 1-km grid cell average than that from a 50-km grid (Willmott and Johnson, 2005). However, a potential problem exists: while the interpolation methods used to develop these fine grids are not any more sophisticated than those used to develop the coarse grids of the past, known spatial climate features previously assumed to be unimportant at the 50-km scale are now very important at the 1-km scale. In addition, a number of poorly understood climate-forcing factors become important at or just below this scale. Understanding what a spatial climate data set can and cannot deliver requires the user to have a working knowledge of what the basic spatial climate-forcing factors are; how they affect climatic patterns; where, when, and at what spatial scale they occur; and how they are handled by the major interpolation techniques.

The goal of this paper, then, is to provide users of spatial climate data, climate specialists as well as nonclimatologists, with background and advice for assessing the suitability of a climate data set for a given application, thus leading to better interpretations, conclusions, and decisions. Space limitations do not allow more than the briefest overview of some important points, with many equally important issues being left out, but perhaps the discussion will stimulate the reader to seek out more detailed literature. In Section 2, well-known physiographic factors that produce observed patterns of climate are introduced, and the spatial scales at which these factors operate are discussed. In Section 3, popular interpolation techniques are briefly introduced, with a focus on how they accommodate the major climate-forcing factors. Section 4 discusses the issues and difficulties with assessing error in spatial climate data sets. The paper concludes with a set of general guidelines for assessing the suitability of a spatial climate data set for a given application.

2. OVERVIEW OF MAJOR SPATIAL CLIMATE-FORCING FACTORS

Long-term climate patterns observed across the globe are a result of a combination of many different processes that manifest themselves at many spatial scales. General circulation patterns provide a large-scale backdrop, including 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 thus are not explicitly accounted for by the major interpolation methods. This assumption may not always be true over data-sparse parts of the globe, however.

Physiographic features on the earth's surface modulate large-scale climate patterns produced by the general circulation, sometimes creating steep gradients over short distances. Table I provides an overview of some of the more well-known physiographic forcing factors on precipitation and temperature patterns, and suggests questions users can ask to assess their importance in a given region. Table II summarizes the importance of well-known spatial climate-forcing factors at various spatial scales, compared to the densities of routinely available climate stations.

The main physiographic features affecting spatial patterns of climate are terrain and water bodies. The influence of terrain, often referred to as orography, is a dominant factor in all but the largest spatial scales.

Table I. Overview of the effects of well-known spatial climate-forcing factors on precipitation and temperature patterns, and suggested questions for assessing their importance in a given region

Factor	Assessment question	Precipitation	Summer maximum temperature	Summer minimum and winter maximum and minimum temperature
Elevation	Are there valleys, hills or mountains present?	Complex patterns; mainly local increase with elevation, except above maritime layer or tradewind inversion	Strong, predictable decrease with elevation, except maritime inversions	General decrease with elevation, but complicated or even reversed by inversions and cold air drainage
Terrain-induced climate transitions	Is there evidence of sharply- defined climate regimes defined by terrain features?	Rain shadows – precipitation maxima on windward slopes, sharply transitioning to minima on leeward slopes	Blockage of marine airflow penetration, producing sharp contrast between coastal and inland temperatures	Divides air masses, such as continental and maritime
Cold air drainage	Are there valleys, hills or mountains present?	Limited effects	Limited effects	Temperature inversions in protected valleys, even in tropics; very extensive at higher latitudes
Coastal zones	Are there oceans or large lakes present?	Wetter coastal areas, if water body is significant source of moisture	If water–land temps differ, large gradients between coastal and inland temperatures	If water–land temps differ, large gradients between coastal and inland temperatures

Table II. Importance of well-known spatial climate-forcing factors at various spatial scales, compared to the densities of routinely available climate stations

	Scale (km)				
	1	10	50	100	>100
Station spacing	Intensive study areas only	Densely populated areas in developed countries	Moderately populated areas in developed countries	Sparsely populated areas in developed countries	Most other areas of the globe
Factor					
Elevation	High	High	Moderate	Low/Moderate	Low/Moderate
Terrain-induced climate transitions	High	High	Moderate	Low	Low
Cold air drainage	High	Moderate	Low	Low	Low
Coastal zones	High	High	Moderate	Low	Low

Terrain affects 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. Large water bodies, such as oceans and large lakes, provide moisture sources for precipitation; precipitation patterns are dictated by how this moisture is transported and deposited inland. The thermodynamic properties of water are very different than those of land, creating complex temperature gradients along coastlines and at adjacent inland areas.

The relationship between elevation and precipitation is complex and highly variable in space, but in general, precipitation generally increases with elevation, owing to forced uplift and cooling of moisture-bearing winds by terrain barriers (Oke, 1978; Barry and Chorley, 1987). Exceptions are when the terrain rises above the height of a moist boundary layer or trade wind inversion, resulting in an increase in precipitation with elevation on lower slopes, and a rapid drying and a decrease of precipitation with elevation on the upper slopes (Mendonca and Iwaoka, 1969). The elevation of maximum precipitation in such situations is variable, and depends on factors such as the depth of the moist boundary layer, wind speed and direction, terrain profile, and others. The process of blocking and uplifting of moisture-bearing winds amplifies precipitation on windward slopes, and can sharply decrease it on leeward slopes downwind (Smith, 1979; Daly *et al*., 1994, 2002). The resulting 'rain shadow' effect is common to many mountain ranges worldwide, and is especially noticeable near major moisture sources, such as oceans. These effects are most important at grid cell resolutions of less than 100 km, the scale below which major terrain features can be resolved. (However, large areas of elevated terrain, such as the Tibetan Plateau and western United States, have significant climatic effects at scales approaching 1000 km.) The direct effects of elevation on precipitation do not appear to increase further below scales of about 5–10 km, 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; Sharples *et al*., 2005).

Temperature exhibits a strong, predictable decrease with elevation where the atmosphere is well mixed, such as on summer days in inland areas. The main summer exception is in coastal regions with well-defined marine layers, where maximum temperatures often increase with elevation above the marine inversion, caused by cool coastal air near the surface undercutting warmer air aloft. Winter temperatures, and daily 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 meters thick (Geiger, 1965; Hocevar and Martsolf, 1971; Bootsma, 1976; Gustavsson *et al*., 1998; Lindkvist *et al*., 2000). This results in temperature inversions, in which temperature sharply increases, rather than decreases, with elevation (Clements *et al*., 2003). The scales at which inversions are important varies with valley size and terrain configuration, but are typically less than 50 km. In polar regions, however, widespread regional inversions hundreds of kilometers in extent can dominate wintertime temperature patterns (Milewska *et al*., 2005; Simpson *et al*., 2005). Even in the subtropics, cold air drainage within sheltered valleys can have noticeable effects on temperature (Daly *et al*., 2003). Terrain can also serve as a barrier between air masses, creating sharply defined horizontal temperature gradients. For example, the Cascade and Sierra Nevada Mountains in the United States effectively block arctic outbreaks from reaching the Pacific Coast during winter, resulting in a sharp separation of continental and maritime temperature regimes along the mountain crests.

Coastal effects on temperature are most noticeable in situations where the water temperature is significantly different than the adjacent land temperature. Such effects are well known in arctic coastal regions (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 kilometers across the coastal strip (Daly *et al*., 2002). In subtropical climates such as Puerto Rico, relatively warm sea-surface temperatures increase minimum temperatures along the coast, compared to inland areas (Daly *et al*., 2003). Climate gradients associated with coastal proximity usually occur within 50–100 km of the coastline.

Several additional spatial climate-forcing factors are most important at scales of less than 1 km, but may also have effects at larger scales. These factors, which include slope and aspect, riparian zones, and land use/landcover, are typically not accounted for in spatial climate interpolation. Slope and aspect at relatively small scales may play a role in determining the local orographic enhancement of precipitation. Slope and aspect also modulate near-surface temperatures on the basis of exposure to solar radiation and wind (McCutchan and Fox, 1986; Barry, 1992; Bolstad *et al*., 1998; Lookingbill and Urban, 2003). The effects of rivers and streams on the temperature environment are similar to those of coastal effects, except that they are at a smaller scale, and can be pronounced within a few hundred meters of the water (Brosofske *et al*., 1997; Dong *et al*., 1998; Lookingbill and Urban, 2003). Land use/landcover variations, while most important below 1 km, are a major consideration in the spatial representativeness of climate stations at much larger scales. For example, stations located near buildings, roads, 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 many tens of kilometers around that station. A further complication is that land use/landcover near a station is often nonstationary over time, making it difficult to track and account for these effects. Some aspects of land use/landcover, such as snow cover and irrigation, are discontinuous and intermittent in space and time. On a related subject, studies have linked changes in land use/landcover to temporal variations in the mesoscale climate, through changes in surface roughness, albedo, heat and moisture fluxes, and other factors (Lawton *et al*., 2001; Pielke, 2001; Pitman *et al*., 2004).

As shown in Table II, except in the most densely populated regions of developed countries, station spacing is likely to be insufficient to directly represent the major climate-forcing factors discussed above. Over most areas of the globe, station spacing is typically 100 km or greater (Willmott *et al*., 1994). This spacing is larger than the scales at which elevation, terrain-induced climate transitions, cold air drainage, and coastal effects are most important. This means that climate patterns caused by these factors will likely be incorrectly located, inaccurately represented, or not represented at all, if interpolated with simple methods. As discussed in the next section, some interpolation techniques attempt to improve the representation of climate patterns by developing relationships between existing station data and explanatory physiographic variables such as elevation, aspect, coastal proximity, etc. for which much higher-density information is available.

3. OVERVIEW OF COMMON INTERPOLATION TECHNIQUES

Today's most commonly used spatial climate data sets have been created through the process of statistically interpolating data values from irregularly spaced station locations to a regular grid. Some data sets have been derived from remotely sensed data or numerical model output, but we will focus on statistically interpolated data sets here. It is worth mentioning, however, that future climate interpolation methods are likely to employ hybrid approaches that combine remote sensing, numerical models, and station data interpolation (Pandey *et al*., 1999; Daly *et al*., 2003; Smith *et al*., 2005).

Most major data sets in use today have been developed using one of six interpolation techniques: inverse-distance weighting (IDW), various forms of kriging, ANUSPLIN tri-variate splines, local regression models Daymet and PRISM, and regional regression models. These six techniques represent a mixture of general numerical methods and specific models. IDW and kriging are related formulations that have been implemented in a variety of models too numerous to discuss here, so they have been grouped together into one category. ANUSPLIN employs general cubic spline theory, but because this specific model is extensively used in the development of popular spatial climate data sets, it is the focus of discussion here. Daymet and PRISM are specific models that both use local regression techniques and have been used extensively to develop popular spatial climate data sets, but are otherwise very different in their formulation; they are therefore discussed separately. Finally, the regional regression technique was treated as a general category, because there is no one specific model used for data set development. Table III summarizes the strengths and weaknesses of these interpolation techniques, and includes the URL of an example data set or data set description that employed each technique. Table IV reports how the techniques account for well-known climate-forcing factors. A popular hybrid method not included in these tables, but worth discussing in this paper, uses an existing spatial climate data set to improve the interpolation of another data set. This is termed climatologically aided interpolation (CAI).

Table III. Summary of strengths and weaknesses of major interpolation techniques used to produce today's popular spatial climate data sets, and Web location of a major associated data set. If an entry is a specific model, the general interpolation approach it employs is given in parenthesis after the name

Interpolation Technique	Description	Strengths	Weaknesses	Example data set URL
IDW/2D Kriging	General method that uses horizontal distance to a station to determine weight in averaging function	Readily available; IDW very easy to apply; both account for distance relationships	Very simple; accounts for distance effects only; kriging requires domain-wide semivariogram, which limits size and heterogeneity of domain	http://climate.geog.udel.edu/~climate/
ANUSPLIN (thin plate splines)	Specific model that fits smoothing splines to the station data in three dimensions	Readily available; relatively easy to apply; accounts for spatially varying elevation relationships	Simulates elevation relationship only; difficulty handling sharp spatial gradients in relationship	http://biogeo.berkeley.edu/worldclim/worldclim.htm
Daymet (local regression)	Specific model that fits local linear regressions of climate *versus* elevation	Local regression accounts for spatially varying elevation relationships	Not readily available; simulates elevation relationship only; cannot handle nonlinear and nonmonotonic elevation relationships	http://www.daymet.org
PRISM (local regression)	Specific model that fits local linear regressions of climate *versus* elevation, with slopes that vary with elevation	Local regression accounts for spatially varying elevation relationships; also accounts for effectiveness of terrain as barriers, terrain-induced climate transitions, cold air drainage and inversions, and coastal effects	Not readily available; requires significant effort to take advantage of full capability	http://www.ocs.oregonstate.edu/prism/
Regional regression	General method that develops domain-wide, multivariate functions	Accounts for effects of multiple variables (usually latitude, longitude, and elevation) on climate patterns; stable statistical relationship	A single, domain-wide relationship limits size and heterogeneity of modeling domain; may not reproduce station values	http://geog.arizona.edu/~comrie/climas/anim.htm

IDW is the simplest of the six interpolation techniques. It takes advantage of the well-known principle that as the distance between a station and a location to be interpolated increases, the influence of that station on the interpolated value is lowered (Renka, 1984; Dodson and Marks, 1997). IDW is a two-dimensional method, and alone does not explicitly account for any of the climatic forcing factors listed in Table IV. However, even sophisticated interpolation schemes incorporate some form of IDW algorithm. IDW is also used quite often to interpolate deviations from a long-term mean in CAI interpolation schemes (see following text). Through their research with IDW-based algorithms, Willmott *et al*. (1985b) demonstrated that distances across the earth's surface should be calculated in spherical, rather than planar coordinates, to avoid errors and distortions in the

Table IV. Summary of how major interpolation techniques used to produce today's most popular spatial climate data sets account for well-known climate-forcing factors. Techniques that do not account for a factor explicitly (denoted by 'station data only') rely on the density and placement of station data to implicitly reflect that factor to varying degrees, rather than the actual physiographic features and barriers that cause it

Interpolation technique	Elevation	Terrain-induced climate transitions	Cold air drainage	Coastal zones
IDW/2D Kriging	Station data only	Station data only	Station data only	Station data only
ANUSPLIN	Station data, polynomial fit	Station data only	Station data only	Station data only
Daymet	Station data, linear regression	Station data only	Station data only	Station data only
PRISM	Station data, variable linear regression	Station data, weighted by moisture index, effective terrain height, and topographic facets	Station data, weighted by two-layer atmosphere (inversions), and topographic index (cold air drainage)	Station data, weighted by coastal proximity and topographic facets
Regional regression	Station data, multivariate regression	Station data only	Station data only	Station data only

resulting fields. Examples of data sets that used a form of IDW as the primary interpolation method include Legates and Willmott (1990a,b) and Dai *et al*. (1997).

Kriging and its variants also rely on the distance-weighting approach, but involve the development of a semivariogram function that describes the characteristics of the relationship between weight and distance that is specific to the data set (Matheron, 1971; Isaaks and Srivastava, 1989). Several kriging variants, such as elevation-detrending kriging and cokriging (Phillips *et al*., 1992), have been developed in an attempt to bring elevation into the process as an explanatory variable. The main drawback to these approaches thus far is that the relationship between climate and elevation is often assumed to be constant, which usually makes it unsuitable for large domains. Continued development of kriging algorithms may overcome this weakness in the future, however. Examples of data sets developed using kriging include Dingman *et al*. (1988), Hevesi *et al*. (1992), Phillips *et al*. (1992) and Garen *et al*. (1994).

Thin-plate smoothing splines is a related statistical technique (Wahba and Wendelberger, 1980; Cressie, 2003). The software package ANUSPLIN (Hutchinson, 1995) fits thin-plate splines (usually second- or third-order polynomials) through station data in three dimensions: latitude, longitude, and elevation. The main advantages of ANUSPLIN over kriging are that a semivariogram need not be developed (instead, a smoothing term is automatically tuned to minimize the cross-validation error), and 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. Examples of data sets developed using tri-variate thin-plate splines include New *et al*. (2002) and Hijmans *et al*. (2006).

Daymet focuses on the effects of elevation on climate (Thornton *et al*., 1997). Daymet develops local linear regressions between climate and elevation for each grid cell on a digital elevation model, using data from surrounding stations. Each station is weighted in the regression function by its distance from the target grid cell. This method takes into account the elevational variation of climate, and its simple station distance-weighting algorithm is computationally efficient, allowing its use in the development of nearly 20 years of daily climate grids for the conterminous United States at 1-km resolution (http://www.daymet.org). Daymet does not have the ability to simulate nonmonotonic relationships between climate and elevation, such as temperature inversions, and does not explicitly account for terrain-induced climatic transitions or coastal effects.

PRISM (Parameter-elevation Regressions on Independent Slopes Model) also develops local climate-elevation regression functions for each DEM grid cell (Daly *et al*., 1994), but calculates station weights on the basis of an extensive spatial climate knowledge base that assesses each station's physiographic similarity to the target grid cell (Daly *et al*., 2002, 2003). The knowledge base and resulting station weighting functions currently account for spatial variations in climate caused by elevation, terrain orientation, effectiveness of terrain as a barrier to flow, coastal proximity, moisture availability, a two-layer atmosphere (to handle inversions), and topographic position (valley, midslope, ridge) (Table III). While PRISM explicitly accounts for more spatial climate factors than other methods, it also requires more effort, expertise, and supporting data sets to take advantage of its full capability. Examples of PRISM climate data sets can be found at www.ocs.oregonstate.edu/prism/ and www.climatesource.com.

While Daymet and PRISM develop a local regression function between elevation and climate at each DEM grid cell, regional regression techniques develop a single, domain-wide, multivariate regression function between climate and latitude, longitude, and elevation (and sometimes other variables such as wind direction, distance from the coast, etc.). This approach has the advantage of being statistically stable, and often explains a large proportion of the climate variability within the domain. However, a single regression function has difficulty handling spatially varying relationships between one or more explanatory variables and climate across a region, and is therefore usually confined to regional, rather than continental domains that are relatively homogeneous. Examples of the regional regression approach include Ollinger *et al*. (1995), Goodale *et al*. (1998), Brown and Comrie (2002), and Johansson and Chen (2005).

CAI is a hybrid approach that uses an existing spatial climate data set to improve the interpolation of another data set (Willmott and Robeson, 1995b). This method relies on the assumption that local spatial patterns of the element being interpolated closely resemble those of the existing climate grid (sometimes called the background or predictor grid). This method is useful for interpolating climate variables and time periods for which station data may be relatively sparse or intermittent. Uses of CAI fall into two broad categories: (i) using a long-term mean grid of a climate element to aid the interpolation of the same element over a different (usually shorter) averaging period; and (ii) using a grid of a climate element to aid the interpolation of a different, but related, climate element (e.g. interpolating growing degree days using mean temperature as the predictor grid). In one popular CAI strategy, a mean 30-year climatology is mapped carefully with sophisticated methods, then time series grids for shorter averaging periods (monthly or daily) are developed using simpler and faster methods such as IDW to interpolate deviations from the mean climatology to a grid. These deviations can then be added to (e.g. temperature) or multiplied by (e.g. precipitation) the mean climatology to obtain the new grid. A key assumption of this method is that the spatial patterns of the deviations from climatology are relatively smooth and simple, and can be interpolated with relatively quick and simple methods. This is not always the case, such as when temperature inversions occur in an area where they do not occur climatologically. Examples of data sets developing using CAI approaches include Willmott and Robeson (1995b), New *et al*. (2000), Plantico *et al*. (2000), Funk *et al*. (2002), Daly *et al*. (2004), and Hamlet and Lettenmaier (2005).

4. ASSESSING ERRORS IN SPATIAL CLIMATE DATA SETS

The obvious question that could be asked at this point is: 'Rather than talk about pros and cons of various interpolation methodologies, why can't each data set simply come with accuracy information, like that provided for DEMs developed by the US Geological Survey(US Geological Survey, 2005)?' It would seem a reasonable thing to do, until one realizes that to do so requires 'ground truth' sampled at extremely high resolutions over a large array of regions – feasible for a DEM, but not for climate, which requires nontrivial instrumentation over many years, and is subject to siting and data collection errors and biases. In essence, the climate field is unknown, except at a relatively small number of observed points. Climate interpolation models themselves provide few useful internal estimates of error, because these estimates rely on the very same assumptions used in the interpolation process itself, and are therefore, not independent or reliable. (In a sense, this is akin to asking students to grade their own examinations!) These error statistics should only be used in a relative sense, with the same model and data set, and interpreted against the backdrop of model assumptions.

A somewhat more useful error statistic often reported in climate interpolation studies is the cross-validation error (Willmott and Matsuura, 1995a; Gyalistras, 2003). This is a measure of the difference between a station's value and the model's estimate for that station, when the station has been removed from the data set. In the common practice of jackknife cross-validation, 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, overall error statistics, such as mean absolute error (MAE), bias, and others are calculated (e.g. Willmott *et al*., 1985a; Legates and McCabe, 1999). The obvious disadvantage to cross-validation error estimation is that no error information is provided for places where there are no stations. In addition, the single-deletion jackknife method favors interpolation model 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. A related problem is the tendency for jackknife cross-validation to underestimate errors when stations are located in pairs or clusters; this type of station configuration increases the likelihood that a nearby station will be present to produce a good estimate for one that has been omitted from the data set during the jackknife deletion process. Other deletion schemes, such as withholding a stratified sample of data points from the analysis, may sometimes be useful in detecting certain weaknesses in the interpolation. For example, withholding high-elevation stations may help determine how well the system can extrapolate beyond the elevation range of the data. Unlike internally generated model error estimates, cross-validation errors can be compared among interpolation techniques. However, the comparison is valid only when all of the parameters of the interpolation – the domain, input data, grid resolution, etc. – are identical (e.g. Phillips *et al*., 1992). As shown in the example below, cross-validation errors can be very misleading when this stipulation is violated.

As an example, let us evaluate several mean annual precipitation grids for the Spring Mountains and its vicinity, outside Las Vegas in southern Nevada, USA. The Spring Mountains is an isolated range that rises above the desert floor at 600–900 m to a maximum elevation of 3633 m at Mount Charleston (Figure 1). Precipitation stations in and around the Spring Mountains indicate that precipitation increases dramatically from less than 100 mm per year at the lowest elevations, to more than 800 mm per year in the vicinity of Mount Charleston. A vegetation summary of the Spring Mountains by the Biological Resources Research Center at the University of Nevada, Reno (www.brrc.unr.edu/mtn/html/springr.html) describes a gradient of cover types from desert shrub vegetation at the lowest elevations to mixed coniferous forest above about 2300 m. Mount Charleston receives enough snow to support a ski area on its upper slopes, adding additional independent evidence that the higher elevations receive significantly more precipitation than lower elevations.

Mean annual precipitation for the averaging period 1971–2000 was interpolated to a 30-arc-s (~800-m) grid. A fairly comprehensive station data set was available, including both high- and low-elevation stations that represent the range of precipitation in the region (Figure 1). Interpolation was performed under four different scenarios. Scenario 1 used the full PRISM modeling system to interpolate precipitation using all available station data; it was deemed the 'control' scenario, since it represented the best available interpolated grid for this example. Scenario 2 also used all available station data, but PRISM was configured to operate as an IDW interpolator, with only the distance-weighting algorithm operating. Scenario 3 used the full PRISM configuration, but all stations above 1500 m elevation were omitted from the data set before interpolation. Scenario 4 also used stations below 1500 m only, but applied the IDW-configured PRISM.

Interpolated grids for the four scenarios are shown in Figure 1. Jackknife cross-validation statistics are given in Table V. Statistics include bias, which is the mean of the signed difference between the interpolated prediction and the observation; and MAE, the mean of the unsigned differences between the interpolated prediction and the observation. Bias indicates whether the interpolator tends to over-predict (positive bias) or under-predict (negative bias) the observations in their absence. MAE is an indicator of the overall performance of the interpolator; a high MAE suggests that the interpolator does a poor job in predicting the observations, while a low MAE suggests that the interpolator is generally reproducing the observations well. In Table V,

Table V. Error analysis for four interpolation scenarios of 1971–2000 mean annual precipitation for Spring Mountains/Las Vegas, Nevada and its vicinity. Error statistics are given for jackknife cross-validation, consisting of single-station deletion with replacement; and a stratified sample cross-validation, for which the nine high-elevation stations (>1500 m) are omitted as a block, with no replacement. Scenarios are ranked by visual similarity to the control grid (Scenario 1) and mean absolute cross-validation error (MAE). See Figure 1 for plots of precipitation grids

Error statistic	Interpolation scenario			
	1. Control – full PRISM, all stations used ($N = 23$)	2. IDW, all stations used ($N = 23$)	3. Full PRISM, only low-elevation stations used ($N = 14$)	4. IDW, only low-elevation stations used ($N = 14$)
Visual similarity to control grid	1 (control)	3	2	4
Jackknife cross-validation				
Rank based on MAE	3	4	1	2
Bias (mm)	1.6	41.0	5.3	−2.0
Bias (%)	3.8	40.8	5.2	2.2
MAE (mm)	49.7	78.3	17.2	19.9
MAE (%)	17.6	47.0	12.7	14.1
Stratified sample cross-validation, high-elevation stations only ($N = 9$)	–	–	–	–
Bias (mm)	–	–	−49.2	−359.8
Bias (%)	–	–	−5.2	−66.8
MAE (mm)	–	–	90.4	359.8
MAE (%)	–	–	17.7	66.8

the quality of each interpolated grid is ranked, on the basis of two different criteria: (1) visual similarity to the control scenario grid; and (2) jackknife cross-validation MAE. As shown in Figure 1 and Table I, Scenario 3 is visually most like the control, but the maximum value on Mount Charleston is somewhat low. Despite using low-elevation data only, PRISM's ability to extrapolate precipitation-elevation relationships in a climatically reasonable way results in high-elevation precipitation patterns that are similar to those of the control scenario. Scenario 2 is the next most similar. There is no elevation extrapolation in IDW, but the presence of station data in the mountains provide some semblance of the precipitation patterns there. Scenario 4, IDW with low-elevation station data only, does not reproduce the mountain precipitation pattern at all, because there are no high-elevation stations or elevation extrapolation.

At a glance, the cross-validation statistics appear to tell a very different story. Scenario 3 has the lowest MAE, followed closely by Scenario 4, which ranked a distant last in its visual similarity to the control scenario. However, adhering to the stipulation that all of the parameters of the interpolation must be identical, we cannot compare results from Scenarios 1 and 2 with those of 3 and 4, because they used very different station data sets. Comparing Scenarios 1 and 2, which did use the same station data sets, Scenario 2 is correctly shown to have a much larger MAE than that of Scenario 1. The comparison of Scenarios 3 and 4 reveals another difficulty, however. Because cross-validation errors do not give any information for locations where there are no stations, the lack of high-elevation stations in Scenarios 3 and 4 made interpolation performance in the mountains 'invisible' to the error analysis. Thus, similar MAE values were reported for radically different interpolation grids. In this instance, more useful error information was obtained by conducting a stratified sample cross-validation exercise, in which the interpolation schemes in scenarios 3 and 4 were run to estimate precipitation at the nine high-elevation stations (already omitted from the data set). As shown in Table V, the IDW scheme in Scenario 4 performed very poorly, greatly under-predicting all of the high-elevation stations.

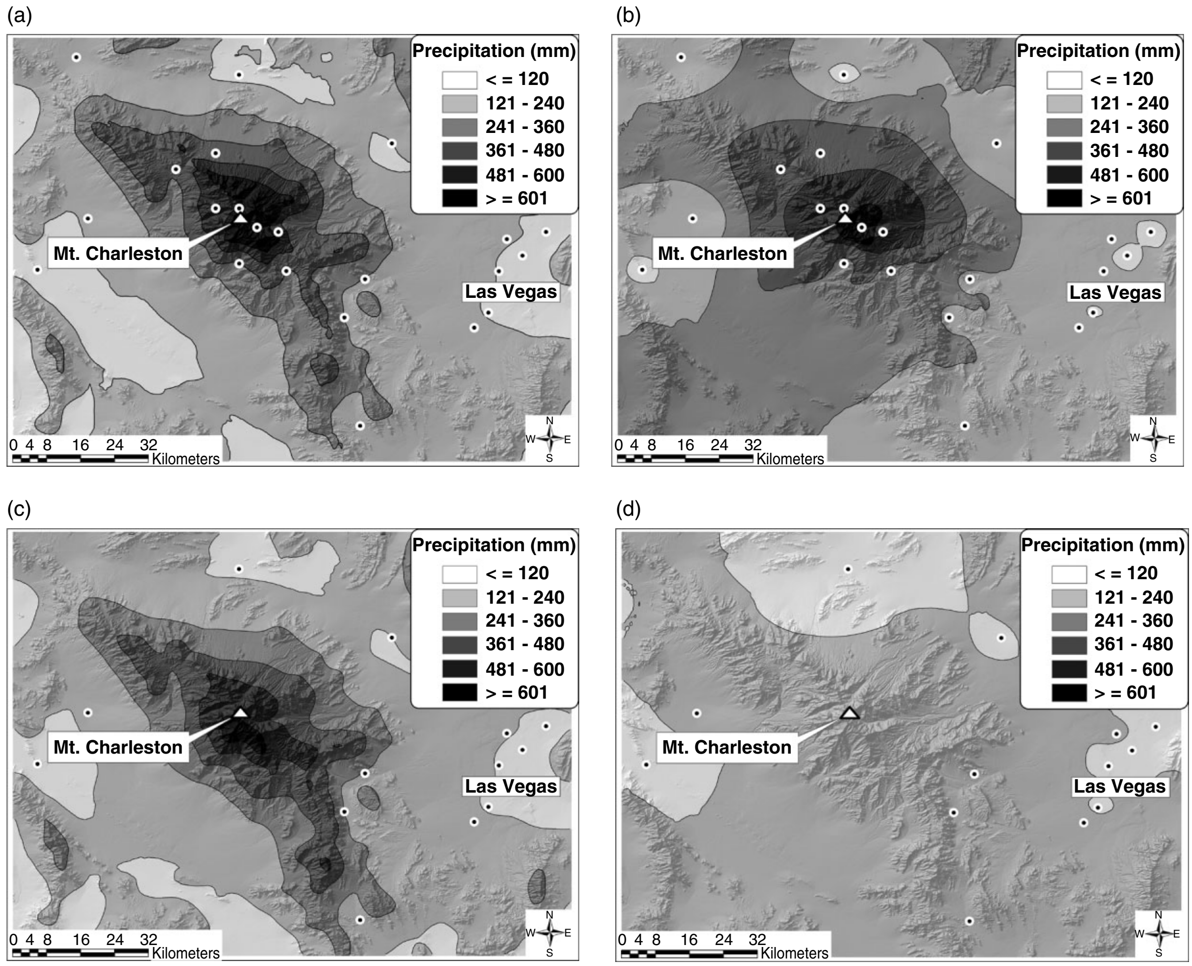

Figure 1. Maps of 1971–2000 mean annual precipitation for Spring Mountains/Las Vegas and its vicinity, Nevada, interpolated under various configurations of PRISM and station data: (a) Scenario 1: full PRISM configuration using all stations; (b) Scenario 2: PRISM configured to simulate IDW interpolation, using all stations; (c) Scenario 3: full PRISM configuration using stations below 1500 m only; (d) Scenario 3: PRISM configured to simulate IDW interpolation, using stations below 1500 m only. Dots indicate locations of stations used in the interpolation

In contrast, the full PRISM model in Scenario 3 predicted the high-elevation stations reasonably well, with only a slight negative bias.

Ironically, the above example illustrates a general tendency for interpolation analyses that use comprehensive station data to report greater cross-validation errors than studies that use relatively little data. However, the comprehensive error estimates are clearly more realistic. Comprehensive data sets are more likely to include stations in remote, or otherwise 'difficult' locations that challenge the interpolation system when omitted, but help produce the best spatial climate data set when present. In addition, no one measure or estimation method gives a complete picture of interpolation error. The best methods are those that provide maximum independence from the model and data used to produce the spatial climate data set (station data suffer from significant errors, as well, but cannot be discussed in any detail here). Climate data sets may be independently evaluated by assessing their consistency with other spatial elements, such as stream flow, vegetation patterns, and related climate elements (e.g. snow pack), or observations not available at the time of interpolation (Daly *et al*., 2002; Milewska *et al*., 2005; Simpson *et al*., 2005).

Because the goal is (or should be) to produce climate maps that best represent the state of human knowledge, the most useful source of verification is probably that of experts who can integrate information from disparate sources to give definite, justifiable feedback on how well the model results reproduce their best knowledge, from a variety of perspectives. This feedback can be in the form of evaluation of the spatial patterns and magnitudes of the mapped climate values, as well as insight into station data quality issues (Daly and Johnson, 1999). If the intent of the developer is to invite wide usage of their data set, conducting an expert review should be an integral part of the development process. Unfortunately, this is rarely done. Reasons range from the need to minimize costs and increase expediency, to a failure to recognize the importance of the process. Expert review is a nontrivial exercise, involving significant time and resources, and contact with reviewers working in diverse disciplines. Depending on the geographic extent of the data set and other considerations, the number of reviewers will often be many times greater than the number of referees for a submitted journal manuscript. The results of such reviews may require the developers to make significant changes in their methodology, repeat the development process, and possibly repeat the review process, if necessary.

5. CONCLUSIONS AND USER GUIDELINES

On the basis of the above information, some general conclusions and guidelines for the user of spatial climate data sets are summarized below.

Terrain and water bodies, where they occur, often have major effects on spatial climate patterns. The importance of terrain and coastal effects is lowest at scales of 100 km and greater, and become greatest at less than 10 km. Except in densely populated regions of developed countries, the typical station spacing of more than 100 km is likely to be insufficient to directly represent climate patterns caused by the major climate forcing factors. Interpolation techniques that develop relationships between existing station data and explanatory physiographic variables such as elevation, coastal proximity, etc. for which much higher-density information is available, have the potential to better represent the actual climate patterns.

Regions without major terrain features and are at least 100 km from climatically important coastlines usually have the simplest spatial climate patterns. As such, they can probably be handled by all the methods discussed, including IDW and ordinary kriging, if sufficient station data exist to represent the major circulation patterns. Situations characterized by significant terrain features, but with no climatically important coastlines or rain shadows, and a well-mixed atmosphere (little cold air drainage) can be reasonably handled by methods that explicitly account for elevation effects; these include ANUSPLIN, Daymet, PRISM, and regional regression models. Examples of these conditions are precipitation and summer daily maximum temperature in an inland region. Regions having significant terrain features, and also significant coastal effects, rain shadows, or cold air drainage and inversions are difficult to map accurately. They are best handled by a system such as PRISM, configured and evaluated by experienced climatologists. Examples include minimum temperature in complex terrain, and maximum temperature and precipitation in mountainous and/or maritime-influenced areas.

Several additional spatial climate-forcing factors are most important at scales of less than 1 km, but may also have effects at larger scales. These factors, which include small-scale slope and aspect, riparian zones, and land use/landcover, are typically not accounted for in spatial climate interpolation. Further improvements in spatial climate interpolation should involve accounting for these factors. However, this will require not only basic, quantitative knowledge of the effects of these factors on climate, but high-quality, fine-grid data sets describing these factors that can be used as input.

Spatial climate data sets represent a significant source of error in any analysis that uses them as input. There is no one satisfactory method for quantitatively estimating error in spatial climate data sets, because the field that is being estimated is unknown between data points. Error estimates based on model assumptions are useful in a relative sense only, and cannot be compared to those of other models. Cross-validation errors are limited to locations for which stations exist. They can be compared among models, but only when all of the parameters of the interpolation are identical. Perhaps the best overall way to assess error is to use a combination of approaches, involve data that are as independent from those used in the analysis as possible, and use common sense in the interpretation of results.

Developers of spatial climate data sets have a responsibility to frankly communicate the strengths and limitations of their data sets; 'overselling' the data increases the danger that the information will be used in inappropriate ways, possibly causing misleading results and conclusions. Developers are encouraged to conduct expert reviews of their draft data sets, which is probably the single most effective way to improve data set quality. As resources allow, developers should consider developing tools to aid users in determining the usefulness of the data sets for their applications. These would include serving the data in a usable map form on the Internet, and allowing users to see what data were used in the interpolation. In the end, the development of high-quality spatial climate data sets is a nontrivial activity, requiring a significant, long-term commitment of expertise, time, and resources. Given the extremely large number and diversity of scientific studies that depend on these data sets, such commitments should be easily justified from a cost-benefit perspective.

Although not discussed here, station data suffer from significant errors. The care and effectiveness of procedures used in their QC can spell the difference between a clean, relatively accurate data set, and one that is fraught with errors and inconsistencies. Good QC takes much time and effort, usually more than the interpolation itself. Interestingly, effective QC systems often use spatial interpolation methods to obtain an expected value to which the station value is compared. In many ways, the QC and interpolation of station data involve the same processes and should be developed and performed together.

Users are encouraged to think critically when evaluating a spatial climate data set for their needs. None are perfect, but many are useful for a variety of regions and applications, if their limitations and assumptions are understood and respected. Take great care when reading descriptions of interpolation methods that state or infer that they 'handle sparse data well,' as this says nothing about the realism and accuracy of the spatial climate fields. Determine whether an expert review has been performed on the data set, and who reviewed it. Use the guidelines in this paper and pursue the literature cited to become educated on how the major spatial climate forcing factors were accounted for during the construction of a spatial climate data set. When reporting on an analysis, state how deficiencies in the spatial climate data set may affect the results.

ACKNOWLEDGEMENTS

The ideas and opinions in this paper, while expressed as my own, have been influenced greatly by many stimulating discussions over the years with knowledgeable people who recognize the importance of spatial climate data sets, and therefore share a commitment to excellence in their creation. Foremost of these are Greg Johnson, Ron Neilson, Phil Pasteris, Kelly Redmond, and George Taylor, and their contributions are gratefully recognized. Producing spatial climate data sets is a group effort, and in my case is made possible only through the unique talents and dedicated service of Matthew Doggett, Wayne Gibson, Mike Halbleib, Eileen Kaspar, and Joseph Smith at The PRISM Group. Finally, I would like to thank Andrew Comrie, who first suggested I write this paper and provided crucial feedback during the manuscript preparation process, and Cort Willmott and two anonymous referees, who provided very thoughtful reviews that improved the manuscript appreciably.

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.

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

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

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.

Brown DP, Comrie AC. 2002. Spatial modeling of winter temperature and precipitation in Arizona and New Mexico, USA. *Climate Research* **22**: 115–128.

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.

Cressie N. 2003. *Statistics for Spatial Data*, Revised edn. John Wiley and Sons: New York.

Dai A, Fung IY, Del Genio AD. 1997. Surface observed global land precipitation variations during 1900–88. *Journal of Climate* **10**: 2943–2962.

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

Daly C, Johnson GL. 1999. PRISM spatial climate layers: their development and use. short course on topics in applied climatology. *79th Annual Meeting of the American Meteorological Society*, American Meteorological Society: Dallas, TX, 10–15 January, available at http://www.ocs.orst.edu/prism/prisguid.pdf.

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, 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, Gibson WP, Doggett M, Smith J, Taylor G. 2004. Up-to-date monthly climate maps for the conterminous United States. *Proceedings of 14th AMS Conference on Applied Climatology, 84th AMS Annual Meeting Combined Preprints*. American Meteorological Society: Seattle, WA, January 13–16, 2004, Paper P5.1, http://ams.confex.com/ams/pdfpapers/71444.pdf.

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**: 4, 497–504.

Dingman SL, Seely-Reynolds DM, Reynolds RC III. 1988. Application of kriging to estimate mean annual precipitation in a region of orographic influence. *Water Resources Bulletin* **24**: 29–339.

Dodson R, Marks D. 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate Research* **8**: 1–20.

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.

Funk C, Michaelsen J, Verdin J, Artan G, Husak G, Senay G, Gadain H, Magadazire T. 2002. The collaborative historical African rainfall model: description and evaluation. *International Journal of Climatology* **1**: 47–66.

Garen DC, Johnson GL, Hanson CL. 1994. Mean areal precipitation for daily hydrologic modeling in mountainous terrain. *Water Resources Bulletin* **30**: 481–491.

Geiger R. 1965. *The Climate Near the Ground*. Harvard University Press: Cambridge, MA.

Goodale C, Aber J, Ollinger S. 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. *Climate Research* **10**: 35–49.

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.

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

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

Hevesi JA, Istok JD, Flint AL. 1992. Precipitation estimation in mountainous terrain using multivariate geostatistics. 1. Structural analysis. *Journal of Applied Meteorology* **31**: 661–676.

Hijmans RJ, Cameron S, Parra JL, Jones PJ, Jarvis A. 2006. 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 Systems* **9**: 385–403.

Isaaks EH, Srivastava RM. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press: New York.

Johansson B, Chen D. 2005. Estimation of areal precipitation for runoff modelling using wind data: a case study in Sweden. *Climate Research* **29**: 53–61.

Lawton RO, Nair US, Pielke RA Sr, Welch RM. 2001. Climatic impact of tropical lowland deforestation on nearby montane cloud forests. *Science* **294**: 584–587.

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.

Legates DR, Willmott CJ. 1990a. Mean seasonal and spatial variability of global surface air temperature. *Theoretical and Applied Climatology* **41**: 11–21.

Legates DR, Willmott CJ. 1990b. Mean seasonal and spatial variability in gauge-corrected, global precipitation. *International Journal of Climatology* **10**: 111–127.

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

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.

Matheron G, Cahiers du Centre de Morphologie Mathematique. 1971. *The Theory of Regionalized Variables and its Applications*. Ecole des Mines: Fountainebleau, France.

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.

New M, Hulme M, Jones PD. 2000. Representing twentieth century space-time climate variability. Part II: development of 1901–1996 monthly grids of terrestrial surface climate. *Journal of Climate* **13**: 2217–2238.

New M, Lister D, Hulme M, Makin I. 2002. A high-resolution data set of surface climate over global land areas. *Climate Research* **21**: 1–15.

Oke TR. 1978. *Boundary Layer Climates*. Methuen: London; 372.

Ollinger SV, Aber JD, Federer CA, Lovett GM, Ellis J. 1995. Modeling physical and chemical climatic variables across the Northeastern U.S. for a geographic information system. U.S. Forest Service General Technical Report, U.S. Department of Agriculture: NE-191, Radnor, PA.

Pandey GR, Cayan D, Dettinger MD, Georgakakos KP. 1999. A hybrid orographic plus statistical model for downscaling daily precipitation in northern California. *Journal of Hydrometeorology* **1**: 491–506.

Phillips DL, Dolph J, Marks D. 1992. A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agricultural and Forest Meteorology* **58**: 119–141.

Pielke RA Sr. 2001. Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall. *Reviews of Geophysics* **39**: 151–177.

Pitman AJ, Narisma GT, Pielke RA Sr, Holbrook NJ. 2004. The impact of land cover change on the climate of southwest western Australia. *Journal of Geophysical Research* **109**: D18109, doi:10.1029/2003JD004347.

Plantico MS, Goss LA, Daly C, Taylor G. 2000. A new U.S. climate atlas. *Proceedings of 12th AMS Conference on Applied Climatology*. American Meteorological Society: Asheville, NC; 247–248, May 8–11, http://www.ncdc.noaa.gov/oa/about/cdrom/climatls2/info/atlasad.html.

Renka RJ. 1984. Interpolation of data on the surface of a sphere. *ACM Transactions on Mathematical Software* **10**: 417–436.

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.

Smith RB, Barstad I, Bonneau L. 2005. Orographic precipitation and Oregon's climate transition. *Journal of the Atmospheric Sciences* **62**: 177–191.

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.

US Geological Survey. 2005. USGS Digital Elevation Model (DEM) guide. http://edc.usgs.gov/guides/dem.html. Active as of 18 October 2005.

Wahba G, Wendelberger J. 1980. Some new mathematical methods for variational objective analysis using splines and cross validation. *Monthly Weather Review* **108**: 1122–1143.

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

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

Willmott CJ, Johnson ML. 2005. Resolution errors associated with gridded precipitation fields. *International Journal of Climatology* **25**: 1957–1963; (Published Online: 7 Oct 2005, DOI: 10.1002/joc.1235).

Willmott CJ, Rowe CM, Philpot WD. 1985a. Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *American Cartographer* **18**: 2249–2251.

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

Willmott CJ, Robeson SM, Feddema JJ. 1994. Estimating continental and terrestrial precipitation averages from rain-gauge networks. *International Journal of Climatology* **14**(4): 403–414.

