ClimateWNA—High-Resolution Spatial Climate Data for Western North America

TONGLI WANG
Centre for Forest Conservation Genetics, Department of Forest Sciences, The University of British Columbia, Vancouver, British Columbia, Canada

ANDREAS HAMANN
Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada

DAVID L. SPITTLEHOUSE
Competitiveness and Innovation Branch, Ministry of Forests, Lands and Natural Resource Operations, Victoria, British Columbia, Canada

TREVOR Q. MURDOCK
Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia, Canada

(Manuscript received 25 February 2011, in final form 22 June 2011)

ABSTRACT

This study addresses the need to provide comprehensive historical climate data and climate change projections at a scale suitable for, and readily accessible to, researchers and resource managers. This database for western North America (WNA) includes over 20,000 surfaces of monthly, seasonal, and annual climate variables from 1901 to 2009; several climate normal periods; and multimodel climate projections for the 2020s, 2050s, and 2080s. A software package, ClimateWNA, allows users to access the database and query point locations, obtain time series, or generate custom climate surfaces at any resolution. The software uses partial derivative functions of temperature change along elevation gradients to improve medium-resolution baseline climate estimates and calculates biologically relevant climate variables such as growing degree-days, number of frost-free days, extreme temperatures, and dryness indices. Historical and projected future climates are obtained by using monthly temperature and precipitation anomalies to adjust the interpolated baseline data for the location of interest. All algorithms used in the software package are described and evaluated against observations from weather stations across WNA. The downscaling algorithms substantially improve the accuracy of temperature variables over the medium-resolution baseline climate surfaces. Climate variables that are usually calculated from daily data are estimated from monthly climate variables with high statistical accuracy.

1. Introduction

In the context of global climate change, spatial databases of historical climate and predictions of climate change have become increasingly important for resource managers, scientists, and policy makers (Fowler et al. 2007). To support a broader community of nonspecialist users, several efforts have been made to provide gridded surfaces of climate data, Internet-based tools that visualize projected climate change, and software packages to query large climate databases. Examples include the WorldClim program (Hijmans et al. 2005), the ClimateBC software package (Wang et al. 2006a), the Climex model (Sutherst et al. 2007), the Canadian Climate Change Scenarios Network (McKenney et al. 2006), the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) Climate Group (Daly et al. 2008), and the Climate Wizard Internet tools (Girvetz et al. 2009).

An important problem that all these efforts face is that the underlying climate databases come in various spatial and temporal resolutions that need to be converted to...
scales meaningful for the end user. High temporal resolution can be important for agricultural and engineering applications that often require daily data to calculate important variables—such as growing degree-days or the length of the frost-free period—or to assess the probability of extreme climate events that structures have to withstand or that may endanger a crop (Schlenker et al. 2007). High spatial resolution climate data are often required for ecological research in montane ecosystems where climate conditions and species compositions can change substantially at scales of a few hundreds of meters in elevation (Hamann and Wang 2006; DeLong et al. 2010).

Climate change projections from general circulation models are typically provided at a coarse resolution with spatial resolutions of 200 km or more and monthly time steps. To provide climate data with higher spatial and temporal resolutions for research and management applications, sophisticated downscaling methods have been developed that include weather generators, statistical approaches, or dynamic downscaling (Fowler et al. 2007; Bürger et al. 2009; Maurer et al. 2009). However, downscaling to high spatial and temporal resolution produces databases that are often inaccessible to nonspecialists because of the file size associated with the thousands of surfaces for multiple climate variables, historical periods of interest, and multimodel projections for the future.

It is important to distinguish between downscaling from general circulation models to finer spatial or temporal scales (Fowler et al. 2007), and downscaling from interpolated weather station data to improve estimates of the climate in topographically complex environments (Hamann and Wang 2005). The mountain ranges of western North America (WNA) are perhaps one of the climatically most heterogeneous regions of the world, with environments that include temperate rain forests, deserts, and dry parkland and subarctic, alpine, boreal, and Mediterranean chaparral ecosystems. While acknowledging the need for improved downscaling algorithms for future projections from general circulation models, our principal objective in this paper is to provide accessible, high-resolution climate data that properly reflect the often large climatic gradients that drive ecosystem differentiation at small spatial scales in western North America.

We also attempt to strike a different balance with respect to accuracy versus accessibility of climate data than do other efforts. To take full advantage of the high-resolution and high-quality baseline data generated through our downscaling algorithms and to keep the overall size of the database small, we rely on the “delta approach,” where historical data and future projections are expressed as a difference (also referred to as “delta” or “anomaly”) from a common reference period (often the 1961–90 normals). The delta surfaces are generated at relatively coarse resolution and are then overlaid onto a high-resolution climate baseline data corresponding to the common reference period. This approach is not without issues, but has been successfully applied particularly for regions or historical time periods with sparse weather station coverage (Mitchell and Jones 2005). For climate change projections, the delta approach is considered a simple statistical downscaling method and it appears to perform as well as sophisticated downscaling methods in producing mean characteristics (Fowler et al. 2007), which are what we deal with in this study. We use the monthly climate data to estimate biologically relevant climate variables that are usually calculated from daily data. Finally, we use partial derivative functions of temperature change along elevation gradients to provide high-resolution climate data for mountainous terrain.

This work builds on previous research and climate databases that we have developed for western Canada (Hamann and Wang 2005; Wang et al. 2006a; Mbogga et al. 2009). These databases and associated software packages have been widely used to address problems in natural resource management, urban development, engineering, ecology, and genetics with over 80 scientific papers published to date. In the current paper we present a spatially expanded version of our software and database that covers North America west of 100° longitude, including the continent’s mountainous terrain where our downscaling solutions provide the greatest benefits. Conscious of the fact that not all aspects of the climate database are meaningful at high spatial resolutions, we contribute a discussion with examples of how the database is used in various applications. We also provide an assessment of the statistical accuracy of derived climate variables, downscaling algorithms, and climate surfaces against observations from 3353 weather stations.

2. Methods

a. Baseline climate data

We primarily rely on monthly climate data for the 1961–90 normal period generated by PRISM (Daly et al. 2002) for our reference climate grid (2.5 arc min). The PRISM data were developed using an approach that incorporates weather station data, a digital elevation model, and expert knowledge of climate patterns such as rain shadows, coastal effects, orographic lift, and temperature inversions over topographically delineated “facets” (Daly et al. 2002). PRISM data have clear advantages over other products in reflecting these effects. However, we found that the pattern of temperature inversion in some areas in the northern part of the prairies in winter was not well predicted by PRISM. Also, PRISM
data are not available for the Northwest Territories and Nunavut. Therefore, the Australian National University Spline (ANUSPLIN; Hutchinson 1989) was used to inter-
poilate weather station records to create 2.5 arc min gridded 1961–90 monthly temperature and precipitation nor-
mals for these regions. To have seamless integration with the PRISM data, PRISM data at the border regions were also included as training data for the ANUSPLIN surfaces. The complete baseline dataset comprised monthly maximum and minimum temperatures and monthly to-
surfaces. The complete baseline dataset comprised monthly maximum and minimum temperatures and monthly total precipitation for a total of 36 basic monthly climate vari-
ables, plus the mean elevation of each grid cell.

b. Historical data and future projections

Monthly temperature and precipitation data for 1901–
2002 used in ClimateWNA are based on Mitchell and
Jones's (2005) interpolated historical data at 0.5° × 0.5° resolution [Climatic Research Unit Time Series, version
2.1 (CRU TS 2.1)]. We subtracted the 1961–90 average
from the gridded surfaces of individual years and months
to recover the original anomaly surfaces (deviations from
the 1961–90 normals). The same methodology was used
to develop the anomaly surfaces covering western North
America for 2003–06 (Mbogga et al. 2009) and up to 2009
(this paper). Temperature was expressed as a difference
in degrees Celsius, and the delta surfaces for precipita-
tion were calculated as percentage difference from the
1961–90 normal values; for example, −50% represents
half of and 200% represents 2 times the normal pre-
cipitation value for a particular month.

We use climate change projections of the phase 3 of
the Coupled Model Intercomparison Project (CMIP3)
multimodel dataset that represents Atmosphere–Ocean General Circulation Models (AOGCMs) from the Inter-
governmental Panel on Climate Change (IPCC) Fourth
Assessment (Solomon et al. 2007). Three emission sce-
narios (A1B, A2, and B1) were included for most of the
12 AOGCMs used. As with the historical data, tem-
perature and precipitation were calculated as anomalies
from the 1961–90 reference period. We used anomalies
for 30-yr normal periods 2011–40, 2041–70, and 2071–
2100 (hereinafter referred to as 2020s, 2050s, and 2080s),
which were obtained from the Pacific Climate Impacts
Consortium (online at http://www.pacificclimate.org).
Because different climate models use different grids and
different spatial resolutions, we interpolated the gridded
anomaly data to a standardized 1° latitude by 1° longi-
tude grid using ANUSPLIN (Hutchinson 1989).

ClimateWNA uses the same algorithms as our pre-
vious software packages (Wang et al. 2006a; Mbogga
et al. 2009) to integrate both historical and future cli-
data with the delta method. Anomaly grids are
interpolated using bilinear interpolation in run time to
avoid step artifacts at grid boundaries, and the differ-
cence is added to the baseline climate normal data to
arrive at the final climate surface.

c. Downscaling of baseline data

ClimateWNA uses a combination of bilinear inter-
polation and elevation adjustment to downscale the
baseline climate data to specific points of interest with
known elevations (Wang et al. 2006a). For each point of
interest, the program first extracts the monthly tem-
perature, precipitation, and elevation for the four closest
grid cells and then calculates bilinear-interpolated cli-
mate and elevation values for that point. Subsequently,
a lapse-rate-based elevation adjustment is carried out
based on the elevation difference between the location
of interest [e.g., a recorded elevation or a high-resolution
digital elevation model (DEM) value], and the elevation
value interpolated from the four neighboring tiles of the
baseline dataset.

The appropriate lapse rates for this elevation adjust-
ment vary by spatial location, elevation, and climate
variable of interest. Following Wang et al. (2006a), poly-
nomial functions for lapse rates were developed for each
monthly temperature variable using the gridded baseline
dataset of 923 176 points (not weather station data)
based on latitude, longitude, and elevation and their com-
binations and transformations. The polynomial functions
were developed separately for three latitudinal bands:
north of 60°N, between 47° and 60°N, and south of 47°N.
To avoid steps in predicted climate data at these bound-
aries, data used for developing the elevation adjustment
functions were extended to neighboring sections by 2° in
latitude. We took the partial derivative of each of these
functions with respect to elevation to obtain equations
for elevation adjustments, which are the rate of change
in a climate variable in response to a change in elevation
for any given latitude, longitude, and elevation within
these three latitudinal bands.

In agreement with our previous study (Wang et al.
2006a), precipitation is only adjusted by the bilinear in-
terpolation of the gridded data. This does not indicate
a lack of elevation precipitation gradients but rather that
they are at a larger scale than temperature and are al-
ready adequately captured within the PRISM data. This
is likely because in mountainous terrain the precipi-
tation is often dominated by the upper-level topography
(Daly et al. 2002).

d. Derived climate variables

In addition to the 36 basic monthly climate variables
described above, ClimateWNA produces 12 monthly, 16
seasonal, and 21 annual variables, derived from the 36 basic
variables. The monthly and seasonal variables include
minimum, maximum, and average temperatures, and precipitation. Of the 21 annual climate variables, 8 are directly calculated from the monthly data while the remainders are derived by equations. The directly calculated variables are mean annual temperature (MAT), mean coolest-month temperature (MCMT), mean warmest temperature (MWMT), continentality (TD, which is the difference between MWMT and MCMT), mean annual precipitation (MAP), mean May-to-September precipitation (MSP), annual heat-to-moisture index (AHM), and summer heat-to-moisture index (SHM). Climate variables representing the winter period of a particular year are calculated based on the monthly variables for January and February in year of interest and for December in the previous year.

Climate variables such as degree-days and frost-free periods are usually calculated from daily observations. However, we previously obtained excellent agreement between derived variables estimated from monthly data and those calculated from daily data, at least for climate normal periods and multiyear averages (Wang et al. 2006a). Similar success in calculating degree-days was obtained by Schlenker et al. (2007). Because some of these formulas were somewhat region specific, we develop new formulas in this paper to estimate degree-days below 0°C (DD < 0), degree-days above 5°C (DD > 5), degree-days below 18°C (DD < 8), degree-days above 18°C (DD > 18), number of frost-free days (NFFD), frost-free period (FFP), beginning of FFP (bFFP), ending date of FFP (eFFP), estimated extreme minimum temperature over a 30-yr normal period (EMT), and proportion of precipitation as snow (PAS). Daily climate data to develop equations for these climate variables were obtained from the Daily Global Historical Climatology Network (http://www.ncdc.noaa.gov). We used 1650 stations without missing years for the 1961–90 normal period covering the study area for the development of equations. An additional 1196 stations that contained up to five missing years were used for validation. Equations for the derived climate variables are provided in the appendix.

This paper also includes two derived climate variables, reference atmospheric evaporative demand ($E_{\text{ref}}$) and a climatic moisture deficit (CMD), that were not part of our previous publications. CMD is the sum of the monthly difference between $E_{\text{ref}}$ and precipitation. CMD is a useful measure of the moisture needed for vegetation growth that must be met from other sources than rain (e.g., soil moisture, irrigation) to avoid the impact of drought. If $E_{\text{ref}}$ is less than precipitation then the monthly CMD is zero (in this case the precipitation minus $E_{\text{ref}}$ is a climatic moisture surplus). The reference evaporation is for a grass surface with no soil moisture restriction (Allen et al. 1998) and is usually calculated with the Penman–Monteith equation (Allen et al. 1998; Shuttleworth 1993). However, this equation requires solar radiation, wind speed, and humidity data, which are not available in our database. We evaluated five temperature-based approaches (Thornthwaite, Hargreaves, Linacre, Hamon, and a vapor deficit method) for calculating monthly $E_{\text{ref}}$. Testing was done with the 1961–90 normals for 56 weather stations distributed across western North America west of 100°W, chosen because they have monthly normals of sunshine hours, or solar radiation, as well as air temperature and precipitation data. A constant wind speed of 2 m s$^{-1}$ was assumed for all stations and if the average monthly air temperature was less than 0°C, then $E_{\text{ref}}$ was assumed to be zero.

To assess the temperature-based methods for climate conditions not currently observed, they were also evaluated against the Penman–Monteith equation under a climate change projection of extreme warming, specifically the third climate configuration of the Met Office Hadley Centre Unified Model (HadCM3) A2 run1 for 2080s. The projection has a 4°C–6°C increase in mean annual temperature depending on the location of the test sites. The change in summer precipitation ranged from −25% to +25%.

e. Statistical evaluation

The accuracy of climate variables estimated with the described methods and algorithms was assessed against weather station observations. Observed monthly normals of the primary climate variables for the reference period (1961–90) were obtained from 3353 weather stations across the entire study area (Fig. 1). Climate normals for 1921 stations in the United States were obtained from the National Oceanic and Atmospheric Administration (NOAA) and for 1432 stations in western Canada from Environment Canada archives. We calculated the amount of variance $R^2$ as well as mean absolute errors (MAE) for the interpolated baseline climate data (2.5 arc min–resolution PRISM and ANUSPLIN grids), as well as the enhanced climate estimates using the described downscaling algorithms and elevation adjustments of the ClimateWNA software package.

The quality of historical climate data generated with the delta approach was evaluated through comparisons between observed and predicted values for three temperature and two precipitation variables for one year from every 10 years starting from 1901 to cover the period between 1901 and 2009. In total, 12 years were evaluated. The climate variables include minimum temperature in January (Tmin01), maximum temperature in July (Tmax07), MAT, MAP, and MSP. The amount of variance in observed values from weather stations explained by ClimateWNA-generated estimates was used to represent the agreement between the observed and predicted values.
3. Results

a. Lapse-rate-based downscaling

The effects of ClimateWNA’s downscaling algorithms are illustrated for MAT for a mountain area in south Washington State (Fig. 2a). Downscaling of MAT baseline data (Fig. 2b) through bilinear interpolation (Fig. 2c) increases the spatial resolution of the data, but does not reflect climate gradients associated with the topography. Elevation-adjusted MAT (Fig. 2d) reflects the effects of topography but step artifacts at the boundaries between adjacent tiles are clearly visible. This is because the differences in climate data among the neighboring tiles are not entirely driven by elevation (Daly et al. 2002). Therefore, elevation adjustment alone cannot generate seamless surfaces across tiles, and a combination of bilinear interpolation and elevation adjustment (Fig. 2e) provides the best results. The downscaled temperature estimates reflect temperature gradients associated with topography shown by the DEM and satellite image (Figs. 2a,f). This run-time adjustment improves the accuracy of moderate-resolution climate data in complex terrain without a large increase in the size of the database.

These visual improvements are also reflected in statistical accuracy of climate estimates (Fig. 3). Improvements in the variance explained in weather station data relative to the 2.5 arc min–resolution baseline data are most pronounced for average and maximum temperature estimates. Improvements were greater during the summer months than for the winter months. ClimateWNA eliminated the majority (between 57% and 65%) of the variance unexplained in maximum temperatures for April through October. While variance explained is a sensitive measure to provide a relative comparison of the improvement in accuracy, mean absolute errors describe the accuracy improvements in the units of the climate variables. In absolute terms, improvements in accuracy can be seen in all variables. Mean absolute errors improved by approximately 0.1, 0.3, and 0.2 for monthly minimum, maximum, and average temperatures, respectively. This represents an error reduction of up to 40% for monthly maximum temperatures and 30% for monthly average temperatures. Bilinear interpolation had also a small positive effect on precipitation estimates.

b. Derived climate variables

The statistical accuracy for climate variables that are usually calculated from daily data, but estimated from monthly temperature and precipitation data in this study, are shown in Table 1. This represents an independent validation of estimates against weather stations that were not included in model development. We found that the formulas and algorithms for the expanded data coverage presented in this paper required only minor adjustments in parameters for DD < 0, DD > 5, DD < 18, NFFD, and PAS, while new estimation approaches had to be developed for DD > 18, bFFF, eFFF, FFP, and EMT (see appendix). The R² values for the linear relationships between derived and calculated variables were generally very high and MAE values were very low (Table 1). However, it should be kept in mind that this evaluation represents estimates for a long-term 1961–90 period and not for individual years.

For DD > 18 and EMT, the nonlinear functions with a single independent variable used in previous work (Wang et al. 2006a) were not able to represent their relationships with any of the primary climate variables. However, using mean temperatures in June, July, August and their transformations as predictors, we were able to develop an accurate polynomial function to estimate DD > 18. Similarly, a polynomial function was developed for EMT using minimum temperatures in both December and January as predictors (see appendix). The variables bFFF, eFFF, and FFP were the most difficult variables to derive using monthly data. Strong relationships were obtained for bFFF with a polynomial function based on...
NFFD and minimum monthly temperatures in April, May, and June, and for eFFP with a polynomial function based on NFFD and minimum monthly temperatures in June, July, September, and October. FFP is calculated by subtracting bFFP from eFFP.

For estimation of evaporation, the Hargreaves equation (Hargreaves and Sammi 1982; Shuttleworth 1993) had the best agreement of the five temperature-based methods tested (data not shown), which is in agreement with Shuttleworth’s (1993) assessment. A latitude correction of the reference evaporation $E_{\text{ref}} = E_{\text{Harg}}[(1.18 - 0.0067) \times \text{latitude}]$ improved estimates further. Evaporation was underestimated by up to 20% at two high-elevation sites (near mountain tops). Here the surface layer climate is strongly affected by the free atmosphere (Daly et al. 2008; McCutchan and Fox 1986), which reduces the daily temperature range, a parameter in the Hargreaves equation. On a monthly basis there was a tendency for

**Fig. 2.** Effects of the downscaling processes in ClimateWNA shown in a mountainous area of south Washington State (centered on 45°35′N, 121°31′W). (a) DEM at 90 m; (b) MAT generated by PRISM at 4 km; (c) interpolated MAT using bilinear interpolation; (d) elevation-adjusted MAT using nearest data values from PRISM data and elevation adjustment function developed in this study; (e) ClimateWNA downscaled MAT using a combination of bilinear interpolation and elevation adjustment; (f) ClimateWNA downscaled MAT overlaid to a satellite image to show the trend of MAT along topography in the upper-left part of this area.
$E_{\text{ref}}$ to be overestimated in winter and spring and underestimated in late summer and autumn because the annual temperature does not fully follow annual trends in solar radiation and humidity (cf. Donohue et al. 2010). For the 1961–90 normals, the standard errors for $E_{\text{Har}}$ and CMD were 31 and 26 mm, respectively, when compared with calculations from the Penman–Monteith equation. Under the large climate warming scenario the annual $E_{\text{ref}}$ had a standard error of 50 mm with a 12-mm underestimate bias, while values for the CMD were 46 and 23 mm, respectively.

c. Historical and future periods

The delta approach we use to add low- or medium-resolution anomalies to high-resolution climate baseline data is illustrated in Fig. 4. For historical data, we use 0.5° gridded anomalies, while for future projections the original resolution of the GCMs output varies. For example, the Canadian model GCM3 has an approximately 300-km resolution (Fig. 4a). ClimateWNA uses a standardized 1° to import GCM anomalies as comma-separated text files, and further downscales the data to the target resolution with simple bilinear interpolation to avoid step artifacts (Fig. 4b). Then, the anomalies are added onto the baseline data at the same resolution (Fig. 4c) to generate the final surface (Fig. 4d). With this approach, the original baseline data (absolute values for 1961–90 normal period) of the historical data and future projections are replaced by scale-free climate data generated by ClimateWNA.

To assess the accuracy of the delta approach for historical data, we have to rely on a nonindependent test because the same weather station data that were used for developing the climate baseline data as well as the historical anomalies are used for the evaluation. This

![Fig. 3. Reduction in (top) variance unexplained and (bottom) mean absolute errors for (left to right) monthly maximum (Tmax), minimum (Tmin), and average (Tave) air temperatures, and monthly total precipitation (Prec). The gray plus black bars represent the accuracy of the 2.5 arc min-resolution baseline data (PRISM). The black bars represent estimates from ClimateWNA, and gray indicates the improvement.](image)

<table>
<thead>
<tr>
<th>Climate variable</th>
<th>$R^2$</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree-day $&lt; 0^\circ$C ($^\circ$C $\times$ days)</td>
<td>0.998</td>
<td>41</td>
</tr>
<tr>
<td>Degree-day $&gt; 5^\circ$C ($^\circ$C $\times$ days)</td>
<td>0.999</td>
<td>46</td>
</tr>
<tr>
<td>Degree-day $&lt; 18^\circ$C ($^\circ$C $\times$ days)</td>
<td>0.999</td>
<td>39</td>
</tr>
<tr>
<td>Degree-day $&gt; 18^\circ$C ($^\circ$C $\times$ days)</td>
<td>0.993</td>
<td>33</td>
</tr>
<tr>
<td>No. of frost-free days (days)</td>
<td>0.992</td>
<td>6</td>
</tr>
<tr>
<td>Beginning date of frost-free period (yearday)</td>
<td>0.947</td>
<td>8</td>
</tr>
<tr>
<td>Ending date of frost-free period (yearday)</td>
<td>0.971</td>
<td>5</td>
</tr>
<tr>
<td>Frost-free period (days)</td>
<td>0.969</td>
<td>12</td>
</tr>
<tr>
<td>Extreme min temperature ($^\circ$C)</td>
<td>0.958</td>
<td>2.5</td>
</tr>
<tr>
<td>Precipitation as snow (mm)</td>
<td>0.920</td>
<td>75</td>
</tr>
</tbody>
</table>
test still allows for a relative comparison of whether the delta method results in comparable accuracy to the 1961–90 normal period. Historical years between 1901 and 2009 were in good agreement with observations from weather stations (Fig. 5). For the three temperature variables examined (Tmin01, Tmax07, and MAT); $R^2$ values were above 0.9 for most cases, particularly for MAT; $R^2$ values were all above 0.95, suggesting a high quality of predicted temperature variables. The $R^2$ values varied between 0.74 and 0.91 for MSP and 0.79 and 0.97 for MAP, which were generally lower than that for temperatures. The quality of predicted historical data varied with time, with the best quality between 1961–90, which was associated with the number of weather stations available for generating the interpolated historical anomalies.

4. Discussion

a. Lapse-rate-based downscaling

Downscaling of temperature data to high-resolution based on environmental lapse rates had consistently positive effects on the accuracy of estimated climate data. Environmental lapse rates vary spatially and temporally and also depend on the type of climate variable measured.
(Sinha 1995). The improvement in monthly minimum temperatures is relatively small in terms of variance explained and mean absolute error compared to maximum temperatures (Fig. 3). Improvements were also generally more pronounced over the summer months. This implies weaker environmental lapse rates for minimum temperatures and winter temperatures. Presumably, winter temperature inversions and cold air drainage may help moderate the minimum temperatures at higher elevations relative to those at lower elevations, resulting in weaker environmental lapse rates for those variables.

We derive regional lapse rates for individual temperature variables directly from surfaces of interpolated temperature data by calculating partial derivatives of climate values as they change along elevation gradients. No weather station records are involved in the derivation of local lapse rates, and therefore our evaluation against station records represents an independent validation of improvements in statistical accuracy. We recommend that lapse-rate-based elevation adjustments should always be used when temperature-based climate values are estimated at point locations of interest with a well-documented elevation value. Examples may be latitude, longitude, and elevation records for species census data in ecology. Even if latitude and longitude are not recorded with high accuracy (e.g., only to the nearest minute), the lapse-rate-based elevation adjustment can still yield more accurate climate estimates for such sample points than could be accurately characterized by gridded data of any resolution.

b. The delta approach for historical data

With the delta approach, we overlay low- to medium-resolution anomalies onto high-resolution 1961–90 baseline climate data to generate estimates for individual years or periods. This approach allows us to take the full advantage of the interpolated and lapse-rate-adjusted baseline data to produce thousands of high-resolution derivative surfaces for historical time periods (as well as for future periods) without massively increasing the size of the database. Because their original baseline data (at a resolution of 0.5°) are replaced with the spatially more accurate interpolated and lapse-rate-adjusted baseline data, the amount of error associated with their baseline data is expected to be reduced, which is relatively large and more difficult to model than with anomalies (Mbobga et al. 2009; Mitchell and Jones 2005). Although we did not carry out an explicit comparison in this paper, Mbobga et al. (2009) have previously shown that this improves the statistical accuracy of historical data. Assessment of annual variables for individual years in this paper as well as the previous study (Mbobga et al. 2009) showed that prediction accuracy is high for temperatures, except for the first third of the twentieth century. However, precipitation variables showed considerable variation in statistical precision among individual years. Predictions for longer periods (5, 10, and 30 years) are expected to be more reliable than those for individual years where stochasticity in weather conditions plays a much greater role.

To implement the delta approach, we chose the 1961–90 normals as a common reference period because this time interval corresponds to the period with the most extensive weather station network (Fig. 5). The 1961–90 normals also represent a reference period prior to the recent pronounced anthropogenic warming signal. More recent climate normal grids are therefore not necessarily superior to serve as reference period for the delta method, just because they are more recent. Note that alternate normal periods can easily be generated at high accuracy for all climate variables with this approach by overlaying a medium-resolution anomaly surfaces representing other normal periods.

c. Downscaling of future projections

The realism of climate change projection for the future, including various downscaling approaches, cannot be directly validated. It is possible that climate change may fundamentally alter local weather patterns at small scales (Fyfe and Flato 1999; Fowler et al. 2007), and such changes are obviously not accounted for with the delta approach or other statistical downscaling methods. Our main objectives are to reduce the errors associated with GCM baseline data for the reference period 1961–90 and to avoid large step artifacts at boundaries of large-scale GCM grid cells. Although our approach does not “improve” the GCM-projected changes at the local scale, our delta approach can substantially reduce the amount of error associated with the GCM baseline data. For example, the average in MAT of 20 GCM predictions for Vancouver is 6.0°C cooler than the observation from weather stations for the reference period 1961–90 because of the low resolution. Consequently, the average projection of these GCMs for 2020s is still 4.1°C cooler than the observation for the reference period. ClimateWNA can eliminate the majority of this error and generate reasonable projections for future periods. Many studies use the GCM-predicted changes (anomalies) to avoid this error; however, projections with absolute values are important for many climate-change-related studies including projections of bioclimatic envelopes for ecosystems and species ranges in future periods.

In addition, the ClimateWNA software package provides a convenient platform to integrate downscaled GCM data from other downscaling methods including dynamical [regional climate model (RCM)] or statistical
approaches (e.g., Maurer et al. 2007, 2009; Bürger et al. 2009). Users can add other gridded data as text files in data directories for anomaly products of both 0.5° and 1° resolution; this is explained in more detail in the next section.

d. Software and data access

The software package ClimateWNA that we provide can be used interactively to query climate data for a single location. Second, multiple locations can be processed based on a comma-separated values (CSV) text file that contains the geographic coordinates and elevation for multiple locations. Third, the program can generate time series data for single or multiple locations of interest. The ClimateWNA software and the associated climate database are available at no charge and can be downloaded anonymously (http://www.genetics.forestry.ubc.ca/cfcg/climate-models.html). No installation is required and the program has been tested on all versions of Microsoft Windows. With relative limited use of disk space (approximately 500 MB for the uncompressed database), ClimateWNA provides easy access to over 20 000 scale-free surfaces of monthly, seasonal, and annual climate variables from 1901 to 2009; multiple climate normal periods; and future climate projections from 12 GCMs and three emission scenarios for the 2020s, 2050s, and 2080s. An Internet-based application, restricted to single location processing, is also available (http://www.genetics.forestry.ubc.ca/cfcg/ClimateWNA_web).

The multilocation processing function can be used to generate custom climate surfaces at any resolution and any projection by advanced users (e.g., Fig. 6). This requires some basic GIS work that involves the conversion of a DEM to a text file consisting of latitude–longitude–elevation triplets for processing by ClimateWNA. The resulting tabular climate data can be imported to GIS applications.

Another feature of ClimateWNA for advanced users is that all underlying climate databases are provided in plain text files in CSV format that can be opened in Excel or any other software package. For example, custom historical periods can be generated by averaging the anomaly values of individual years; Gray et al. (2011) used a custom 1997–2006 period relevant to their study. Similarly, AOGCM anomalies files can be averaged to arrive at mean climate projection. ClimateWNA accepts custom anomaly files at 0.5° and 1° resolution that can be placed in separate folders and appear automatically as menu choices based on the file name, For example, Roberts and Hamann (2011) created anomalies for periods between 6000 and 21 000 years before present for a paleoecology study.

e. Applications

Like its predecessors, ClimateWNA has numerous applications in ecology, hydrology, agriculture and urban studies. For example, estimating the effect of climate change on heating and cooling demand, which is

Fig 6. Maps of three major climate variables in WNA for 1971–2000 normals created from data produced by ClimateWNA. Presented are (left to right) MAT, MAP, and mean AHM.
proportional to heating degree-days (DD < 18°C) and cooling degree-days (DD > 18°C) (Christenson et al. 2006), is readily accomplished with the software. Similarly, estimates of growing degree-days, chilling degree-days, and the frost-free period that are important in agricultural applications are readily available. We have used our methodology to provide visualization of projected future climates to aid in community planning and for informing the general public about climate change (http://pacificclimate.org/tools-and-data/plan2adapt).

Climate-informed resource management applications that have used our software include the delineation of forest seed planning zones (Wang et al. 2006a; Hamann et al. 2011), mapping forest site index (Coops et al. 2011), and assessing the probability of pest outbreaks (Campbell et al. 2007).

Other ecological research applications include bioclimate envelope modeling to investigate how species habitat may shift geographically under projected climate change (e.g., Hamann and Wang 2006; Mbogga et al. 2010), ecosystem classification (DeLong et al. 2010), and habitat conservation (Rose and Burton 2011). Genetic studies that investigate how organisms are adapted to the environments in which they occur rely on accurate climatic characterization of the source environment of sampled material (Wang et al. 2006b, 2010; O’Neill et al. 2008). The time series functionality of our software packages also provides a convenient tool for historical biological research, which relies on past climate fluctuations and biological records to establish correlative relationships between climate and species response, for example, interpretation of tree ring chronologies or pollen assemblages to detect and predict plant–climate relationships (Goring et al. 2009; Miyamoto et al. 2010; Messaoud and Chen 2011; McLane et al. 2011a,b).

f. Limitations

While the software package provides convenient access to climate data at any resolution, it is important to discuss limitation of the database for end users. Mean absolute errors of climate estimates for all monthly, seasonal, annual, and derived climate variables are small for 30-yr normal periods, but the error increases for shorter time periods as previously discussed. A second note of caution applies to areas with sparse coverage of weather stations (Fig. 1). It is impossible to assess the statistical accuracy of climate surfaces for areas that lack station coverage, such as high montane and Arctic environments. It has been pointed out that different interpolation techniques applied to northern Canada and montane environments produce rather different climate surfaces that nevertheless converge at weather station locations to produce identical statistical accuracies (Mbogga et al. 2010). Finally, all interpolated climate surfaces used in this study are ultimately based on standard weather stations (shaded temperature sensors located at 1.5-m height in open areas). Consequently, microclimate driven by small-scale topography such as aspect, slope, and frost pockets, and geographic features such as rivers and lakes are not captured. Vegetation cover can also influence the local climate; for example, forest canopies are usually cooler during the day and warmer during the night than open areas. The reference evaporation is calculated for a grass surface with no soil moisture restrictions. Values for other surfaces such as forest or lakes and for locations with limited soil moisture storage capacity or topographic shading will likely be different. Therefore, when ClimateWNA is used to generate climate data at fine scales and for specific locations, it is necessary to be aware that while temperature changes along elevation gradients may be accurately captured, other effects of local topography and vegetation are not.

Acknowledgments. We thank Peter Gould at the USDA Forest Service, Pacific Northwest Research Station, for providing and processing daily climate data for this study. We also acknowledge David Bronaugh for preparing the additional GCM anomaly files for all future projections. This study was funded by Forest Science Program under the Forest Investment Account, the Future Forest Ecosystems Scientific Council, and the Ministry of Forests, Lands and Natural Resource Operations, British Columbia, Canada. Additional funding was provided by the NSERC/Discovery Grant RGPIN-330527-07 and the NSERC/CRD Grant CRDPJ 349100-06.

APPENDIX

Equations for Estimation of Derived Variables

The equation for DD < 0°C (DD0) is

\[
DD0 = \sum_{m=1}^{12} (0 - T_m) \times N_m \text{ for } T_m < 0°C
\]

\[
DD0 = DD0 \times 0.9661 + 179.37.
\]

If \( T_1 \) or \( T_{12} \gtrless 0 \), then

\[
DD0 = 184.79 - 64.13 \times T_1 + 7.7103 \times T_1^2 - 3.165 \times T_1^3,
\]

where \( T_m \) is the average temperature for month \( m \) (°C), and \( N_m \) is the number of days in month \( m \). These also apply to the following equations.
For DD > 5°C (DD5),
\[
DD5 = \sum_{m=1}^{12} (T_m - 5) \times N_m \quad \text{for} \quad T_m \geq 5°C
\]
\[
DD5 = DD5 \times 0.982 + 136.05.
\]

For DD < 18°C (DD_18),
\[
DD_18 = \sum_{m=1}^{12} (18 - T_m) \times N_m \quad \text{for} \quad T_m \leq 18°C
\]
\[
DD_18 = DD_18 \times 0.9875 + 105.92.
\]

If DD_18 < 0, then DD_18 = 0.

The equation for DD > 18°C (DD18) is
\[
DD18 = 152.19 - 15.28176 \times T_6 + 0.05559 \times T_6^3
\]
\[
- 2.10012 \times T_5^2 + 0.12639 \times T_5^3
\]
\[
- 2.812 \times 10^{-5} \times T_7^5.
\]

If DD18 < 0 or \(T_7 < 14\), then DD18 = 0.

For EMT,
\[
EMT = -18.09995 + T_{\min(1)}^2 \times 2.14095
\]
\[
+ T_{\min(1)}^2 \times 0.06836 + T_{\min(12)}^2 \times (-0.04771)
\]
\[
+ \frac{TD^2}{100} \times 0.00306,
\]

where \(T_{\min(1)}\) and \(T_{\min(12)}\) are the average minimum temperatures (°C) for January and December, respectively. TD is the temperature difference (°C) between mean warmest and coldest monthly temperature.

For NFFD,
\[
NFFD = \sum_{m=1}^{12} \left(1 + 1.15 \times e^{[-0.4 \times T_{\min(m)}]}\right) N_m
\]
\[
NFFD = 3.3144 + 1.0114 \times NFFD
\]

If NFFD < 0 then NFFD = 0,

where \(T_{\min(m)}\) is the average minimum temperature (°C) for month \(m\).

The equations for FFP, bFFP, and eFFP are
\[
bFFP = 124.9495 + T_{\min(4)}^2 \times (-1.7581) + T_{\min(5)}^2 \times (-11.87934) + T_{\min(6)}^2 \times 2.09433
\]
\[
+ T_{\min(4)}^2 \times (-0.3746) + T_{\min(4)}^3 \times 0.01482 + T_{\min(5)}^3 \times 0.06751 + T_{\min(4)}^4 \times 0.00123
\]
\[
+ T_{\min(5)}^4 \times (-0.00266) + T_{r(4)} \times 5.21934 + T_{r(4)}^2 \times (-0.16101)
\]
\[
+ NFFD^3 \times (-7.19) \times 10^{-6} + NFFD^4 \times 5.976953 \times 10^{-8}
\]
\[
+ NFFD^5 \times (-1.2266) \times 10^{-10}.
\]

If bFFP < 0, then bFFP = 0. For eFFP,
\[
eFFP = 231.6577 + T_{\min(9)} \times 8.87656 + T_{\min(6)}^2 \times (-0.05996) + T_{\min(7)}^2 \times (-0.0751)
\]
\[
+ T_{\min(10)}^2 \times 0.20123 + T_{\min(9)}^3 \times (-0.026) + T_{\min(7)}^3 \times 9.435 \times 10^{-5}
\]
\[
+ T_{\min(9)}^4 \times 6.7816 \times 10^{-4} + T_{\min(10)}^4 \times (-2.9319) \times 10^{-12} + NFFD^5 \times 7.94 \times 10^{-12}.
\]

If eFFP > 365, then eFFP = 365. For FFP,
\[
FFP = eFFP - bFFP.
\]

If FFP < 0, then FFP = 0. If FFP > 365, then FFP = 365,
where \(T_{r(4)} = T_{\max(4)} - T_{\min(4)}\).

For PAS,
\[
PAS(1) = 1/[1 + e^{(-T_4 + 2.9901)/(-2.50353)}] \times P(1)
\]
\[
PAS(2) = 1/[1 + e^{(-T_4 + 1.3948)/(-2.0004)}] \times P(2)
\]
\[
PAS(3) = 1/[1 + e^{(-T_4 - 0.5473)/(-1.5719)}] \times P(3)
\]
where \( \text{PAS}(m) \), \( T_r(m) \), and \( P(m) \) are the PAS (mm), average air temperature \( ^\circ \text{C} \) and the total precipitation (mm) for the \( m \)th month.

The reference monthly evaporation from the Hargreaves equation \( E_{\text{Harg}}(m) \) mm is (Hargreaves and Sammi 1982; Shuttleworth 1993) given by

\[
E_{\text{Harg}}(m) = \begin{cases} 
0.0023 \times d \times S_0 \times [T_m + 17.8] \times T_r^{0.5} & T_m \geq 0 \\
0 & T_m < 0 
\end{cases}
\]

where \( d \) is the number of days in the month; \( S_0 \) is the water equivalent of the radiation above the atmosphere (mm day\(^{-1}\)) at the latitude of the site for the day of the year in the middle of the month; \( T_m \) is the monthly mean daily temperature \( ^\circ \text{C} \); and \( T_r(m) \) is the mean daily temperature range \( ^\circ \text{C} \), that is, the difference between the monthly mean maximum and minimum temperatures. [Note that the square root on \( T_r(m) \) appears to have been omitted in Shuttleworth (1993).]

Correction of \( E_{\text{Harg}}(m) \) for latitude to give monthly reference evaporation \( E_{\text{ref}}(m) \) is

\[
E_{\text{ref}}(m) = E_{\text{Harg}}(m) \times (1.18 - 0.0065 \times \text{latitude}),
\]

where the latitude is in degrees. The quantity \( E_{\text{ref}}(m) \) is summed over the months of the year to give \( E_{\text{ref}} \) output by ClimateWNA.

The annual climatic moisture deficit is the sum on the monthly values [\( \text{CMD}(m) \), mm] calculated from the monthly reference evaporation [\( E_{\text{ref}}(m) \), mm] and precipitation [\( P(m) \), mm] as

\[
\text{CMD}(m) = \begin{cases} 
0 & E_{\text{ref}}(m) \leq P(m) \\
E_{\text{ref}}(m) - P(m) & E_{\text{ref}}(m) > P(m)
\end{cases}
\]

REFERENCES


