

RESEARCH ARTICLE

ClimateAP: an application for dynamic local downscaling of historical and future climate data in Asia Pacific

Tongli WANG (✉)^{1,2}, Guangyu WANG², John L. INNES², Brad SEELY², Baozhang CHEN³

¹ Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

² Department of Forest Resource Management, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

³ State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

Abstract While low-to-moderate resolution gridded climate data are suitable for climate-impact modeling at global and ecosystems levels, spatial analyses conducted at local scales require climate data with increased spatial accuracy. This is particularly true for research focused on the evaluation of adaptive forest management strategies. In this study, we developed an application, ClimateAP, to generate scale-free (i.e., specific to point locations) climate data for historical (1901–2015) and future (2011–2100) years and periods. ClimateAP uses the best available interpolated climate data for the reference period 1961–1990 as baseline data. It downscales the baseline data from a moderate spatial resolution to scale-free point data through dynamic local elevation adjustments. It also integrates and downscales the historical and future climate data using a delta approach. In the case of future climate data, two greenhouse gas representative concentration pathways (RCP 4.5 and 8.5) and 15 general circulation models are included to allow for the assessment of alternative climate scenarios. In addition, ClimateAP generates a large number of biologically relevant climate variables derived from primary monthly variables. The effectiveness of the local downscaling was determined based on the strength of the local linear regression for the estimate of lapse rate. The accuracy of the ClimateAP output was evaluated through comparisons of ClimateAP output against observations from 1805 weather stations in the Asia Pacific region. The local linear regression explained 70%–80% and 0%–50% of the total variation in monthly temperatures and precipitation, respectively, in most cases. ClimateAP reduced prediction error by up to 27% and 60% for monthly temperature and precipitation, respectively, relative to the original baselines data. The improvements for baseline portions of historical and future

were more substantial. Applications and limitations of the software are discussed.

Keywords biologically relevant climate variables, downscaling, dynamic local regression, future climate, historical climate

1 Introduction

High-quality climate data are essential for conducting climate change impact analyses. Accordingly, a large volume of climate data has become available through various sources^[1–4] in recent years. However, these data sets are mostly in grid format with low-to-moderate spatial resolutions. While such data are suitable for climate impact modeling and analyses conducted at global and ecosystems scales, modeling studies are increasingly shifting toward local scales often with the goal of developing practical and effective adaptive management strategies. Considering the high levels of heterogeneity in climate in mountainous areas, such gridded climate data are often not adequate to meet the requirements for such studies. Thus, the demand for high-resolution and high-quality spatial climate data are rapidly increasing^[5].

Historical climate data, often from a specific reference period (e.g., 1961–1990), are widely utilized to support a wide range of research focused on understanding the relationships between climate variables and plant health and productivity. They are also essential for building climate niche models for examining the distribution of specific ecosystems and their components^[6,7] and for testing mechanistic models designed to predict plant growth response to climate^[8,9]. Further, the veracity of analyses of plant-climate relationships is dependent on the degree to which the climate conditions reflected in historical data match the actual climate conditions where the plants or ecosystems reside^[10,11]. For example, when

Received April 14, 2017; accepted September 20, 2017

Correspondence: tongli.wang@ubc.ca

conducting research with a climate component, it is desirable that climate data are derived from a nearby weather station(s) as such data would be the most accurate and reliable. However, as the number of weather stations is limited, the locations for research projects are often far away and have considerably different climate conditions from weather stations^[2,12]. Spatial interpolation techniques are often used to predict climate conditions for these locations or for developing spatial climate data sets to cover large areas. Statistical methodologies applied to interpolate climate data, including kriging, bilinear and spline interpolations, are primarily based on distances from nearby weather stations. Yet, considering the degree to which topography variations and other factors can influence local climates over relatively short distances, the accuracy of interpolated climate data obtained through these methods is often not adequate.

A number of algorithms and software packages have been developed to facilitate the spatial interpolation of climate data. The ANUSPLIN software developed at the Australian National University employs thin plate smoothing splines to improve on distance-based interpolation, and has been widely used. For example, WorldClim hosts gridded climate data for the entire globe using this approach^[2]. The algorithm used in ANUSPLIN is a purely statistical approach and its accuracy often limited for the areas with sharp changes in elevation^[1]. Another widely used interpolation method is the Parameter-elevation Regressions on Independent Slopes Model (PRISM)^[13] developed at the Oregon State University. PRISM uses a combination of statistical techniques and knowledge-based methods to make adjustments to account for the effect of factors including rain shadows, coastal proximity, and temperature inversions on climate interpolations. PRISM climate data are generally regarded as the highest quality spatial climate data sets currently available. Interpolated climate data are available for the United States and some other regions, including China, at moderate resolutions.

Spatialized historical climate data sets are essential for supporting research geared toward understanding the response of plants, communities and ecosystems to past and ongoing climate change. The most widely used historical climate data are from the Climate Research Unit in the UK, which provides global time series for monthly climate data at the spatial resolution of 30 arcmin (about 50 km × 50 km) available from 1901 to 2015^[3].

Although the availability of climate data has improved considerably during the last decade, some challenges remain for non-meteorological users. For example, the process required to obtain climate data for locations of interest can be onerous. Specific software and knowledge of geographic information systems are typically required to process the data. This is particularly true when climate data for historical and future periods are derived from different sources and with different formats, variables and resolutions. In addition, climate data generated by meteorologists

often lack climate variables that are relevant for biological applications.

Several climate databases and tools have been developed to address these needs including those created by McKenney et al.^[7] and Rehfeldt et al.^[14] using ANUSPLIN for North America. With these tools, users must submit coordinates files for specific locations to request climate data. The web-based Climate Wizard Tool^[15] covers the entire world and allows users to obtain climate data for a location. However, there is no downscaling involved and it can be time consuming to assemble climate data for multiple locations. ClimateBC represents another tool that has been widely used for its effectiveness in downscaling, inclusion of both historical and future data, and its user-friendly interface^[16]. ClimateBC was initially developed for western Canada and was subsequently expanded to cover the western North America (ClimateWNA)^[17] and finally encompassed the whole of North America (ClimateNA)^[18]. Along with expansion in coverage, its functionality has also been enhanced, particularly by shifting the downscaling algorithm from a partial-derivative-function based approach^[16] to a dynamic-local-regression approach^[18].

Using the dynamic-local-regression approach designed for ClimateNA, we developed a climate model to represent the eastern Asia Pacific region called ClimateAP. While the fundamental approach is similar to that employed in ClimateNA, ClimateAP differs considerably in terms of the baseline data, the fact that it covers distinct terrains and topography, and the fact that it crosses an equatorial region. Considering these differences, the effectiveness of the downscaling algorithms and the accuracy of the output of ClimateAP need to be evaluated to provide confidence for its use in supporting climate change research in the Asia Pacific region. Accordingly, the primary objective of the work presented here is to describe the key algorithms employed in the model and to evaluate the accuracy of its output through comparisons against historical climate data. The evaluation includes detailed assessments of derived biologically relevant climate variables, the performance of regionally specific functions, and the efficacy of methods to incorporate daily weather station data in estimating these climate variables.

2 Data and methods

2.1 Baseline data

Baseline gridded climate data for the reference period 1961–1990 are required to drive the ClimateAP. PRISM data represent the highest quality data but were only available for China and Mongolia. Consequently, it was necessary to use monthly climate data from WorldClim^[2] for other countries within the AP region (Fig. 1). Since the WorldClim data set had a complete coverage for the region,

it was used as the base data set, with the PRISM data substituted for China and Mongolia. To avoid steps at the contact points between the data sets, inverse-distance-weighted averages were applied to a 20 km overlapping zone of the two data sets. The baseline climate data were provided at the resolution of 0.25×0.25 arcmin (about $4 \text{ km} \times 4 \text{ km}$). The climate variables obtained from these data sources included three primary climate variables: monthly minimum temperatures ($T_{\min 01}$ – $T_{\min 12}$), monthly maximum temperatures ($T_{\max 01}$ – $T_{\max 12}$) and monthly precipitation ($\text{Pre}01$ – $\text{Pre}12$). The mean elevation of each grid cell was also included.

2.2 Historical climate data

Monthly temperature and precipitation data for 1901–2015 (version CRU TS 3.24) generated by the Climatic Research Unit (CRU) at the University of East Anglia^[3,19] were used to represent the historical climate data. According to Mitchell and Jones^[19], the original data

were developed based on anomalies relative to the reference period 1961–1990, although absolute values were delivered for each individual year at the resolution of $0.5^\circ \times 0.5^\circ$. To facilitate the delta downscaling approach (described below) in ClimateAP, the data were converted back to anomalies for each year by subtracting 1961–1990 normals.

2.3 Future climate data

The climate data for future periods and years were derived from a variety of general circulation models (GCM) included in the CMIP5 project in the IPCC Fifth Assessment Report^[20] and downloaded from <http://cmip-pcmdi.llnl.gov/cmip5>. GCM data are available at various spatial resolutions, ranging from $0.75^\circ \times 0.75^\circ$ to $2.85^\circ \times 2.85^\circ$. The GCM data were converted to a resolution of $1^\circ \times 1^\circ$ using bilinear interpolation to simplify the process of integration into ClimateAP. Monthly climate data from two emission scenarios (RCP 4.5 and 8.5) and 15 GCMs^[21]

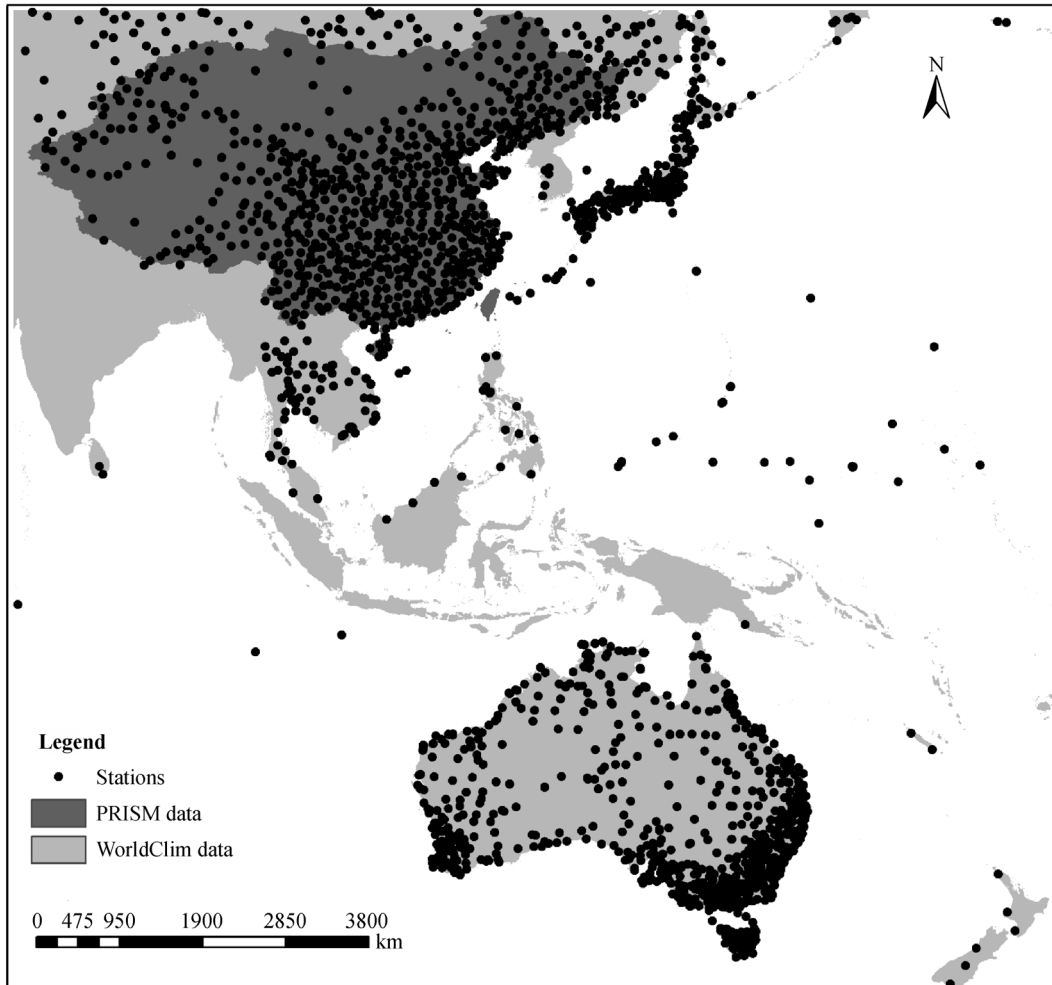


Fig. 1 The distribution of the baseline climate data (from PRISM and WorldClim) and weather stations that were used to derive biologically relevant climate variables and to evaluate the output of ClimateAP.

were included in ClimateAP for three normal periods 2011–2040, 2041–2070 and 2071–2100. An ensemble data set based on all the 15 GCMs was generated for each of the two RCPs and the three periods. In addition, future monthly data were included for individual years 2011–2100 from three GCMs for the two RCPs. As for the historical data, future monthly values were converted into anomalies relative to the 1961–1990 normals for the purpose of downscaling using the delta approach described below.

2.4 Downscaling of the baseline climate data

The downscaling algorithm in ClimateAP was adopted from ClimateNA^[18]. It uses a combination of bilinear interpolation and local regression approaches to downscale the baseline monthly grid data (4 km × 4 km) to scale-free point data. The bilinear interpolation functions to interpolate values for locations between midpoints of the four neighbor grids, while the local linear regression is used to estimate lapse rates for the location of interest to account for elevation effects. In the case of the local regression, ClimateAP retrieves monthly climate data and elevation values from nine closest neighbors of the location of interest and calculates differences in climate variables and elevation between all 36 possible unique pairs of the nine data points. A simple linear regression of the differences in a climate variable on the difference in elevation is performed and the slope of the regression is used as the lapse rate for the climate variable at that specific location. A lapse rate is estimated for each of the 36 monthly primary climate variables at each location of enquiry. Given that the local linear regressions are dynamically performed for every location of interest, the process is called dynamic local regression or dynamic local downscaling. To avoid over-adjustment due to a weak local regression, the adjusted amount is weighted by the R^2 value of the regression.

2.5 Integration and downscaling of historical and future climate data

Both historical and future monthly climate data were integrated and downscaled using a delta method following Wang et al.^[18]. With this approach, the historical data (and future projections) are split into two sections: one section for the reference period (i.e., 1961–1990 normal period) and another section for the anomaly period (e.g., the period between the reference period and 2015). For the anomaly section, historical and future monthly climate anomaly grids are interpolated to continuous surfaces using bilinear interpolation at runtime. For the reference-period section, the scale-free baseline monthly climate normal data generated by ClimateAP are used to replace the original historical and future climate data. After adding the interpolated anomaly section to the scale-free reference

section, ClimateAP produces scale-free climate data for historical and future years or periods. Since the baseline data downscaled by ClimateAP are at much higher accuracy than that of the historical and future projections, this process is designed to improve the prediction accuracies for both historical and future climate data over their original values.

2.6 Calculated and derived climate variables

The baseline data contain 36 primary monthly climate variables including monthly maximum ($T_{\max 01}$ – $T_{\max 12}$) and minimum ($T_{\min 01}$ – $T_{\min 12}$) temperatures and precipitation ($PPT01$ – $PPT12$). ClimateAP generates many additional biologically relevant climate variables based on these primary climate variables (Table 1). Some of these variables, such as mean annual temperature (MAT) and mean annual precipitation, can be directly calculated using the primary monthly climate variables. Others, such as degree-days (DD) above 5°C ($DD > 5$) and number of frost-free days, require daily climate data. Since daily data are not available in ClimateAP, daily climate data from weather stations in the region were used to calculate these climate variables and then related back to monthly climate data. The functions reflecting such relationships were used in ClimateAP to estimate those climate variables. Specifically, these relationships were modeled using daily climate data from 1805 weather stations located throughout the Asia Pacific region from the Daily Global Historical Climatology Network. The distribution of the weather stations is shown in Fig. 1.

Due to the wide range of variation in climate variables in the Asia Pacific region, no single linear, polynomial or nonlinear function could adequately reflect the relationships between DD and monthly climate variables. Thus, piecewise functions constructed using combinations of linear and nonlinear functions were applied to model these relationships^[18]. The DD variables included $DD < 0$, > 5 , < 18 and > 18 . The general form of the DD piecewise functions (DD_m) is:

$$DD_m = \begin{cases} \text{if } T_m > k, \frac{a}{1 + e^{-\left(\frac{T_m - T_0}{b}\right)}} \\ \text{if } T_m \leq k, c + \beta T_m \end{cases} \quad (1)$$

where T_m is the monthly mean temperature for the m month; and k , a , b , T_0 , c and β are the six parameters to be optimized. The annual values of the DD variables were calculated by summing up the monthly values. The same approach was applied to other derived monthly climate variables described below.

For the number of frost-free days ($NFFD$) and precipitation as snow (PAS), a sigmoid function was used to model the relationship between these monthly variables and monthly temperatures:

Table 1 Climate variables directly calculated (Calculated) using the primary climate variables included in the baseline data and functionally derived (Derived) based on the relationships between the climate variables calculated using daily climate data and monthly climate variables from the 1805 weather stations shown in Fig. 1

Category	Variable short name	Variable long name	Availability at monthly and seasonal levels
Calculated	MAT	Mean annual temperature	Yes
	MWMT	Mean warmest month temperature	Yes
	MCMT	Mean coldest month temperature	Yes
	TD	Continentially, temperature difference between MWMT and MCMT	No
	MAP	Mean annual precipitation	Yes
	AHM	Annual heat-moisture index (MAT + 10)/(MAP/1000))	No
Derived	DD < 0 (°C)	Degree-days below 0°C	Yes
	DD > 5 (°C)	Degree-days above 5°C	Yes
	DD < 18 (°C)	Degree-days below 18°C	Yes
	DD > 18 (°C)	Degree-days above 18°C	Yes
	NFFD (day)	Number of frost-free days	Yes
	EMT (°C)	Extreme minimum temperature over a 30-year period	Yes
	EXT (°C)	Extreme maximum temperature over a 30-year period	Yes
	PAS (mm)	Precipitation as snow	Yes
	Eref	Hargreaves reference evaporation	Yes
	CMD	Climatic moisture deficit	Yes

$$NFFD_m(\text{or } PAS) = \frac{a}{1 + e^{-\left(\frac{T_m - T_0}{b}\right)}} \quad (2)$$

where T_m is the monthly minimum temperature for the m month; and a , b and T_0 are the three parameters to be optimized.

For extreme minimum (EMT) and maximum (EXT) temperature over a 30-year period, polynomial functions were used as follows:

$$EMT = a + bT_{\min 01} + cT_{\min 01}^2 + dT_{\min 12}^2 + eTD^2 \quad (3)$$

$$EXT = a + bT_{\max 07} + cT_{\max 07}^2 + dT_{\max 08} + eT_{\max 08}^2 + fTD \quad (4)$$

where a , b , c , d , e and f are the parameters to be optimized; $T_{\min 01}$ and $T_{\min 12}$ are monthly minimum temperature for January and December; $T_{\max 07}$ and $T_{\max 08}$ are monthly maximum temperature for July and August, respectively; and TD is continentality.

The other two derived climate variables, Hargreaves reference evaporation and a climatic moisture deficit were directly adopted from Wang et al.^[18] due to a lack of observations to evaluate these variables and the fact that the methods used to estimate them are rather generic^[22].

2.7 Statistical evaluations of ClimateAP outputs

The accuracies of the climate variables generated by

ClimateAP were evaluated against observations from weather stations. Observed monthly normals of the primary climate variables for the reference period 1961–1990 were calculated based on the daily climate data from the 1805 weather stations across the entire study area shown in Fig. 1. The amount of variance explained by the ClimateAP outputs and prediction errors, defined as the mean absolute error, were used to evaluate the accuracy of the climate variables generated by ClimateAP for the baseline data.

For the evaluations of the historical data downscaled with the delta approach, we compared the predicted values using CRU, GCMs and ClimateAP against observations for the three primary monthly variables for the baseline normal period 1961–1990. As in the delta approach, the baseline part of the historical and future climate data was replaced by the scale-free baseline data generated by ClimateAP. The amount of improvement in the scale-free baseline data relative to the baseline part of the historical and future climate data provides a conservative evaluation of the predictions of historical and future climate data.

3 Results

3.1 Effectiveness of the local downscaling

The dynamic local linear regressions explained 70%–80% of the total variation in monthly maximum and minimum temperatures, and 0%–50% in monthly precipitation

variables among the nine neighboring pixels in most cases (i.e., medians) (Fig. 2). These results suggest the local downscaling algorithm is effective for temperature and moderate for precipitation variables. An assessment of seasonal results shows that the downscaling was more effective for warmer months than for cooler months. Spatially, the downscaling was more effective in mountainous areas ($R^2 > 0.85$ in most cases), where elevation adjustment is critical, than in flat areas. The seasonal and spatial patterns were similar for areas both north and south of the equator (Fig. 2).

The local regression showed a negative linear relationship between the changes in elevation and the changes in monthly minimum and maximum temperatures in most cases. However, positive relationships were also found in winter months associated with temperature inversions that often occur in mountainous areas during the winter. The relationship between precipitation and elevation was mostly positive. The benefit of downscaling for refining gridded climate data in mountainous regions is illustrated in Fig. 3. When considered for the whole AP region, the effect of downscaling applied in ClimateAP was relatively small in terms of the amount of total variance explained for temperatures (less than 1%) and precipitation (0%–12%) due to the large amount of total variation in the region. However, the improvement was substantial on prediction

errors (Fig. 4). The downscaling reduced prediction errors by 16%–27% for monthly maximum temperatures and up to 60% for monthly precipitation. The improvement on monthly minimum temperatures was relatively small (1%–6%).

The amount of variance in observed climate variables explained by ClimateAP predictions were high, between 97.3% and 99.3% for monthly minimum temperatures, between 97.7% and 99.6% for monthly maximum temperatures, and between 81.6% and 93.9% for monthly precipitation (Table 2). The prediction errors were between 1.0 and 1.4°C for monthly minimum, between 0.7 and 1.0°C for maximum temperatures, and between 22 and 30 mm for monthly precipitation.

3.2 Accuracies of derived climate variables

The fits of the piecewise and nonlinear functions were strong for all derived climate variables. Comparisons between the derived climate variables that were estimated from monthly climate variables and those that were calculated from observed daily climate data indicated that the derived climate variables were highly accurate based on the amount of variance explained (mostly over 99%) and the prediction errors listed in Table 3.

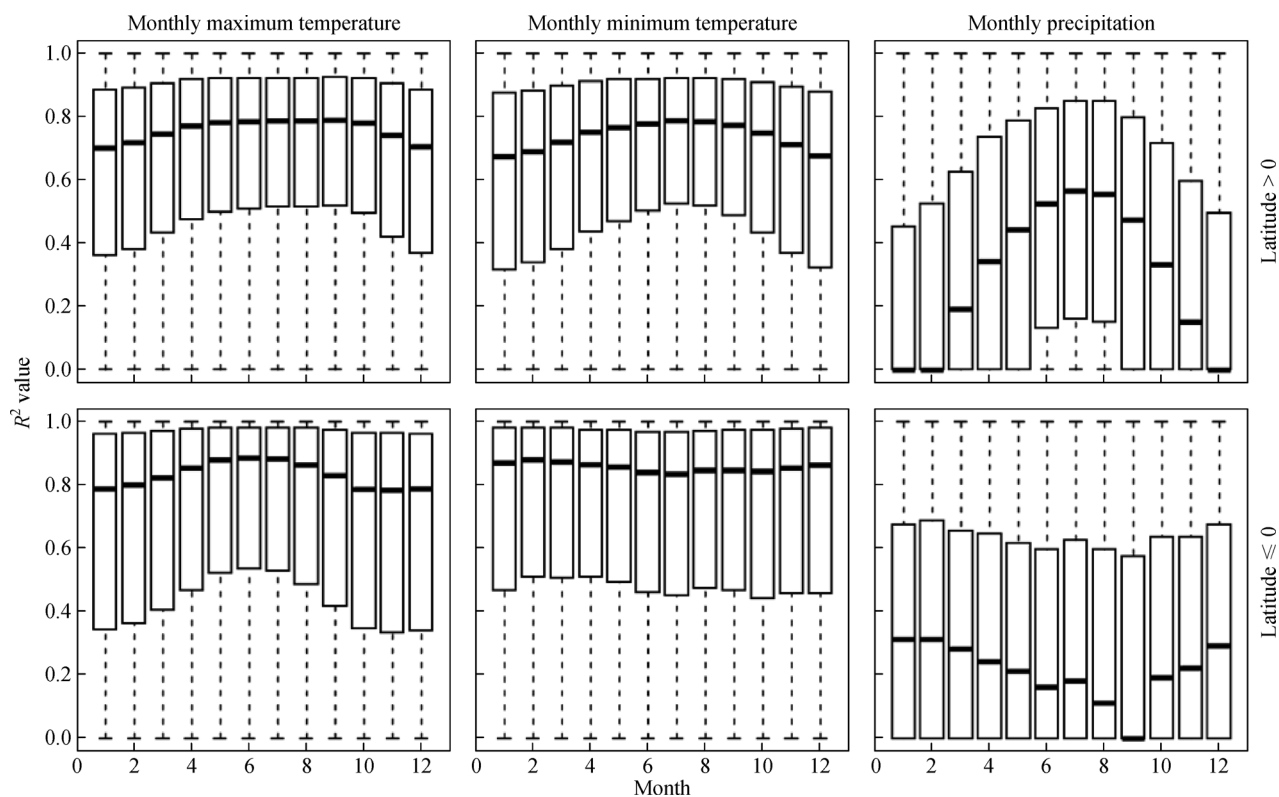


Fig. 2 Proportions of variance explained by local linear regressions in total variation among the nine neighboring pixels for the three primary monthly climate variables separated by the equator. The extent of the box indicates the 25th and 75th percentiles. The horizontal solid lines inside the boxes indicate the medians.

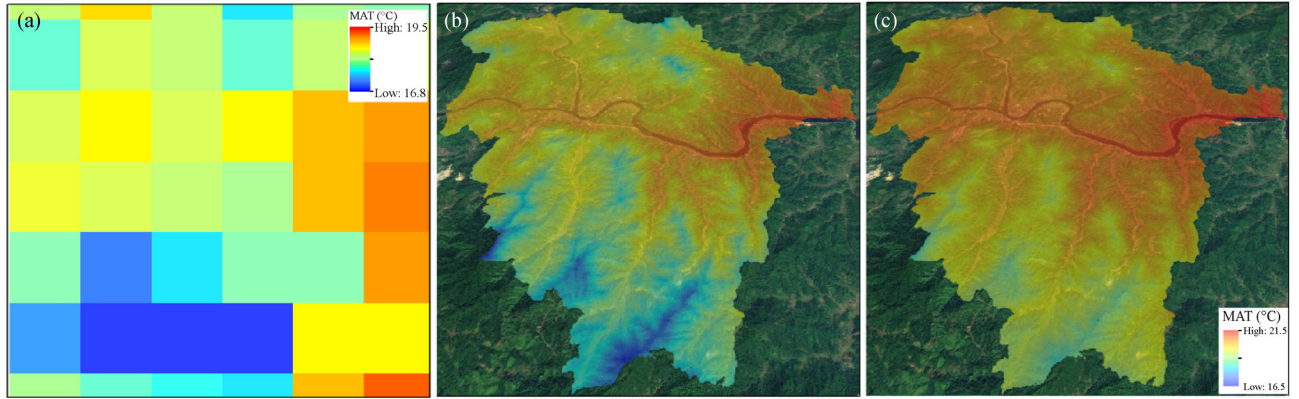


Fig. 3 Illustration of the effectiveness of the local downscaling applied in ClimateAP shown in a mountainous area in Fujian Province (117.50° E and 26.52° N). (a) Mean annual temperature (MAT) of the baseline data at 4 km × 4 km; (b) downscaled MAT (90 m × 90 m) by ClimateAP and overlaid on a satellite geographic image for the reference period; (c) downscaled future MAT for the period of 2041–2070 based on the 15 GCMs' ensemble for the RCP 4.5.

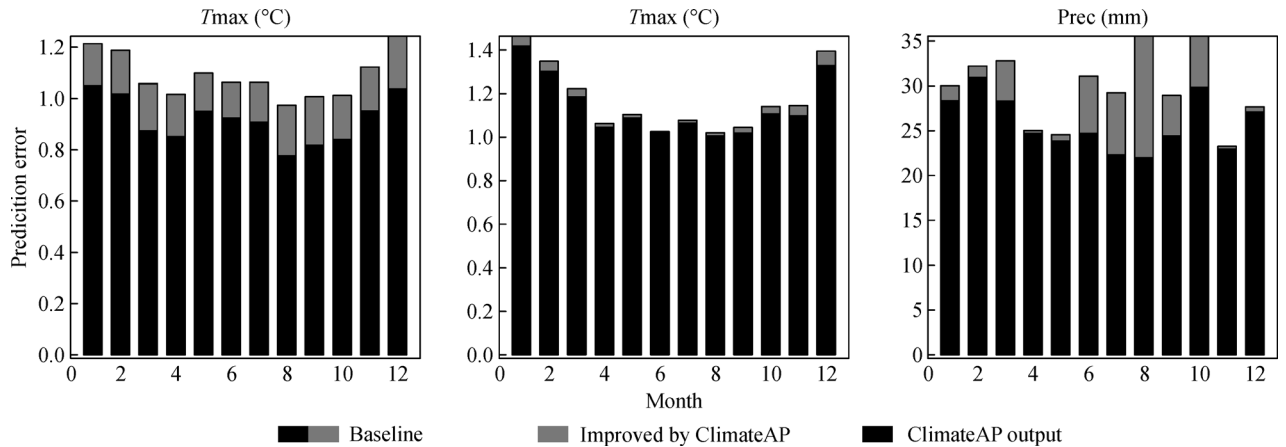


Fig. 4 Comparisons in prediction errors between baseline data (i.e., directly from PRISM and WorldClim) and ClimateAP output for monthly maximum temperature ($T_{\max 01}$ – $T_{\max 12}$), monthly minimum temperature ($T_{\min 01}$ – $T_{\min 12}$) and monthly precipitation ($P_{\text{prec}01}$ – $P_{\text{prec}12}$).

3.3 Improvements of historical and future climate data

ClimateAP output showed substantial improvements for historical and future climate data relative to the original CRU and GCM data for the baseline period (1961–1990 normal) (Fig. 5). On average, the prediction errors for CRU data were reduced by 0.5°C (31%) for monthly minimum temperatures, 0.8°C (47%) for monthly maximum temperatures, 21 mm (44%) for monthly precipitation. The relative improvement for GCM output was even greater as prediction errors were reduced by 2°C (67%) for monthly temperatures and 35 mm (56%) for monthly precipitation on average.

4 Discussion

High-quality and high-resolution climate data are essential

for facilitating climate impact studies and applications. ClimateAP allows users to extract and downscale gridded climate data to scale-free climate data for specific point locations through a dynamic local downscaling approach. It also generates a diverse series of biologically relevant derived climate variables at monthly, seasonal and annual time steps. Evaluations of the downscaling process and comparisons of output against observations suggest that the dynamic local downscaling approach implemented in ClimateAP is effective for reducing interpolation error in baseline data. The level of improvement in prediction accuracy is comparable to that determined for ClimateNA^[18]. Historical monthly climate data from CRU and future monthly climate data from GCMs for individual years or normal periods can also be effectively downscaled using the delta method by taking advantage of the scale-free climate data for the reference period generated by ClimateAP.

Table 2 The amount of variance in observed climate variables explained by ClimateAP predictions and their prediction errors

Variable	Variance explained/%*	Prediction error/°C	Variable	Variance explained/%	Prediction error/°C	Variable	Variance explained/%	Prediction error/mm
Tmin01	99.2	1.4	Tmax01	99.6	1.0	Pre01	88.2	28
Tmin02	99.3	1.3	Tmax02	99.5	1.0	Pre02	85.2	30
Tmin03	99.0	1.2	Tmax03	99.4	0.9	Pre03	88.4	28
Tmin04	98.3	1.0	Tmax04	98.8	0.8	Pre04	84.8	24
Tmin05	97.3	1.1	Tmax05	97.7	0.9	Pre05	89.1	23
Tmin06	97.9	1.0	Tmax06	98.2	0.9	Pre06	92.4	24
Tmin07	98.2	1.1	Tmax07	98.6	0.9	Pre07	93.7	22
Tmin08	98.3	1.0	Tmax08	98.9	0.7	Pre08	93.9	22
Tmin09	97.8	1.0	Tmax09	98.4	0.8	Pre09	91.9	24
Tmin10	98.0	1.1	Tmax10	98.8	0.8	Pre10	86.5	29
Tmin11	99.0	1.1	Tmax11	99.3	0.9	Pre11	85.6	23
Tmin12	98.6	1.3	Tmax12	99.3	1.0	Pre12	81.6	27

Note: * ClimateAP predictions all exceeded the significance level of $P < 0.0001$.

Table 3 The amount of variance in observed climate variables explained by monthly climate variables and their prediction standard errors

Variable	Variance explained/%		Prediction standard error	
	Monthly	Annual	Monthly	Annual
DD < 0 (°C)	94.4 – 99.9	99.8	0.5 – 8.2	21.7
DD > 5 (°C)	98.7 – 100.0	99.8	3.2 – 27.2	97.9
DD < 18 (°C)	99.3 – 100.0	99.8	5.2 – 13.9	73.4
DD > 18 (°C)	99.2 – 99.8	99.9	5.0 – 10.8	41.5
NFFD (day)	90.5 – 99.8	98.5	0.5 – 1.6	7.1
EMT (°C)		99.3		2.1
EXT (°C)		86.5		1.5

Note: All derived climate variables exceeded the significance level of $P < 0.0001$.

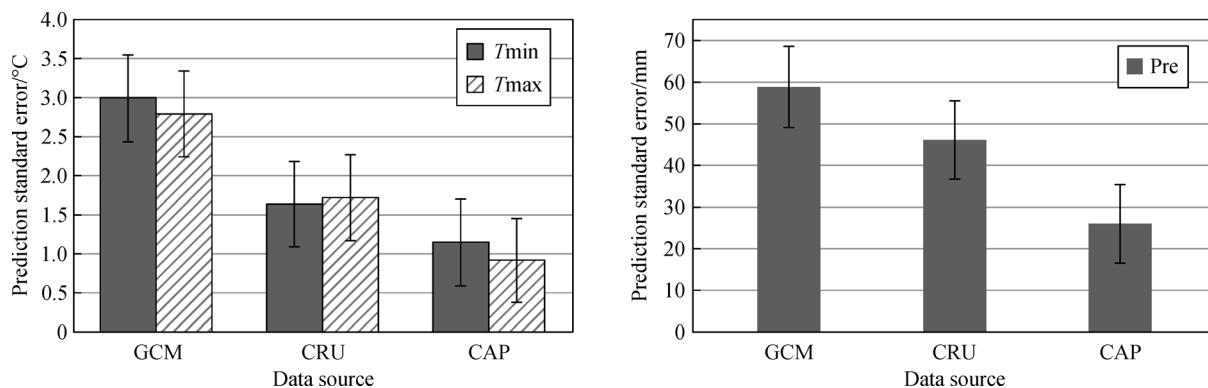


Fig. 5 Comparisons in prediction errors among three data sources for the reference normal period 1961–1990: IPCC GCM predictions (GCM), Climate Research Unit (CRU) and ClimateAP (CAP) output. The error bars are standard errors indicating the variation among the 12 months.

4.1 Downscaling and prediction accuracy for the reference period

The baseline data used in ClimateAP from PRISM^[23] and

WorldClim^[2] are at relatively high spatial resolution (4 km × 4 km) in comparison to historical climate data from CRU^[3] (50 km × 50 km) and projected future climate data from GCMs^[20] (up to 250 km × 250 km). Elevation has

been identified as the dominant factor driving the local variation in climate within each pixel of the baseline climate data^[1]. Results of this study have confirmed this. Thus, the effectiveness of the downscaling depends largely on the accuracy of the estimated lapse rate, which varies with latitude at a large scale and local site factors at a fine scale. Although partial derivative functions have been used successfully for elevation-related temperatures adjustment, they are not feasible for precipitation^[24]. In this study, a dynamic local linear regression approach developed for ClimateNA^[18] was applied to ClimateAP. Results of the analyses show that the lapse rates estimated through local linear regressions were able to capture such variation in monthly temperature and precipitation variables among the neighboring pixels, including temperature inversion in winter months, which is considered difficult to model^[7,25].

The accuracy of ClimateAP predictions for temperatures were high based on the amount of variance explained (97%–99%) and prediction errors (0.7–1.4°C) (Table 2). The prediction accuracy was considerably lower for precipitation than for temperatures in terms of both the amount of variance explained (82%–94%) and prediction errors (22–30 mm). The accuracy of ClimateAP predictions was slightly lower than those observed for ClimateNA^[18], despite the fact that the relative improvement over the baseline data by elevation adjustment was comparable or even greater in ClimateAP. This suggests that the quality of the baseline data in ClimateNA is better, as would be expected. PRISM data have been updated for the USA, but not for China. In addition, the number of weather stations used for developing these baseline data was much greater in North America than in the Asia Pacific region.

4.2 Calculated and derived climate variables

The baseline climate data used in ClimateAP have 36 primary monthly climate variables (T_{min01} – T_{min12} , T_{max01} – T_{max12} and $Pre01$ – $Pre12$). ClimateAP calculates and derives many additional biologically relevant climate variables at monthly, seasonal and annual time steps (Table 1). These additional climate variables have been widely used in modeling forest ecosystem and species climate niches^[6,10,26], population response to climate^[8,27,28] and evolutionary genomics^[29]. The most frequently used climate variables include MAT, annual heat-moisture index, continentality or temperature difference, cooling and growing DD (< 0 and > 5 , respectively), and precipitation as snow. The ability to generate these biologically relevant climate variables significantly expands the utility of ClimateAP.

4.3 Integration and downscaling of historical and future climate variables

ClimateAP integrates both historical and future climate

data through a delta method. While the delta method is relatively simple, it appears to perform as well as sophisticated downscaling methods in producing mean characteristics^[30]. With this method, the baseline portion of the historical and future climate data are replaced by the scale-free data generated by ClimateAP. As the accuracy of the scale-free baseline data are much higher than that in the historical or future climate data, this replacement not only increase the spatial resolution, but also considerably improves the accuracy of the historical (31%–47% for temperatures and 44% for precipitation) and future climate data (67% for temperatures and 56% for precipitation).

For the anomaly section of the historical and future climate data, ClimateAP bilinearly interpolates the gridded data into point data at runtime. Although this process is not likely to improve the accuracy of the predictions, it smoothens steps across the original grid cells. Moreover, errors associated with the anomalies are typically small as they are less sensitive to topography and their spatial variations are relatively easy to model in comparison to absolute values for historical climate data^[19]. The same applies for projections of future climate data. Through a combination of the replacement of the baseline section and the bilinear interpolation to the anomaly section, ClimateAP is able to generate seamless climate surfaces (Fig. 3c). This makes it possible to develop adaptive forest management strategies at the local scale for a changing climate.

4.4 Applications and limitations

Modeling climate niches for ecosystems and species requires climate data that are consistent with the actual locations of the vegetation in question^[31,32]. The scale-free climate data generated by ClimateAP allows it to meet this requirement. In addition, the availability of a large number of calculated and derived biologically relevant climate variables will considerably increase the flexibility of models to address such objectives. ClimateAP has been used to model the climate niche for four major forest tree species in Asia Pacific^[33].

Due to the difficulty in accessing climate data in the past, scientists have been using geographic variables, such as latitude, longitudes and elevation as substitutes for climate variables in experimental design and data analysis. However, results obtained from such studies are limited to specific locations. In contrast, a relationship between plant performance and specific climate variables should theoretically remain consistent at different locations. ClimateAP significantly improves access to spatially distributed, climate variables for use in a wide range of applications including climate based experimental design, modeling climate niches for ecosystems and species, and population responses to climate. In particular, the time-series function in ClimateAP allows users to generate monthly climate time series for multiple locations and

multiple years for historical and future years between 1901 and 2100. In the absence of such a tool, it may take several months to obtain and process the same amount of data. ClimateAP can be used to generate climate surfaces. ClimateAP produces output in a comma-separated-values format that can be imported easily to ArcGIS to generate maps. The resolution of the map depends on the resolution of the input file. It is particularly useful to generate high-resolution climate maps for specific areas to develop management strategies accounting for locally variable climate conditions.

ClimateAP predictions can account for more than 97% of the total variation in temperatures and more than 82% in precipitation for the 1961–1990 normals. However, the standard prediction errors in term of absolute values are still considerably large, that is 0.7–1.4°C for temperatures and 22–30 mm for monthly precipitation. This can be much greater when it comes to specific locations and specific years (instead of the 30-year normal periods). The prediction errors are expected to be smaller in the regions with more weather stations than the regions where weather stations are lacking. The dynamic local elevation adjustment is effective to downscaling the baseline data. However, it is not able to correct all the errors associated with the baseline data or the anomalies of the historical and future climate data. In addition, although ClimateAP can generate spatial climate surface at very fine resolution or for specific locations, the microclimate affected by aspect, slope and vegetation types at micro-scale are not reflected in the predictions. Therefore, caution is required when interpreting the climate data at very fine spatial resolution.

5 Conclusions

ClimateAP uses the best available climate data as baseline climate data and downscales it to a scale-free format with improved prediction accuracy using a dynamic local downscaling approach. It incorporates historical and future climate data into a single package and improves prediction accuracy using a delta approach by taking advantage of the high-quality scale-free baseline data generated by the application. Additions of calculated and derived biologically relevant climate variables make the application more useful for a variety of applications in forest modeling. In addition, ClimateAP allows users to generate climate data at any spatial resolution, which is particularly important for development of adaptive forest management strategies at local scales. However, it should be noted that errors inherent in the original baseline and anomaly climate data cannot be totally corrected by the application. The desktop version (intensive users) and a Google Map based web version (occasional users) of ClimateAP can be accessed at UBC server (<http://asiapacific.forestry.ubc.ca/research-approaches/climate-modeling>).

Acknowledgements This study was funded by a research grant “Adaptation of Asia-Pacific Forests to Climate Change” (APFNet/2010/PPF/001) funded by the Asia-Pacific Network for Sustainable Forest Management and Rehabilitation.

Compliance with ethics guidelines Tongli Wang, Guangyu Wang, John L. Innes, Brad Seely, and Baozhao Chen declare that they have no conflict of interest or financial conflicts to disclose.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Daly C, Halbleib M, Smith J I, Gibson W P, Doggett M K, Taylor G H, Curtis J, Pasteris P P. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, 2008, **28** (15): 2031–2064
- Hijmans R J, Cameron S E, Parra J L, Jones P G, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 2005, **25**(15): 1965–1978
- Harris I, Jones P D, Osborn T J, Lister D H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 Dataset. *International Journal of Climatology*, 2014, **34**(3): 623–642
- Thornton P E, Thornton M M, Mayer B W, Wei Y, Devarakonda R, Vose R S, Cook R B. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. *ORNL Distributed Active Archive Center*; 2016.
- Hamann A, Wang T L, Spittlehouse D L, Murdock T Q. A comprehensive, high-resolution database of historical and projected climate surfaces for western North America. *Bulletin of the American Meteorological Society*, 2013, **94**(9): 1307–1309
- Wang T L, Campbell E M, O'Neill G A, Aitken S N. Projecting future distributions of ecosystem climate niches: uncertainties and management applications. *Forest Ecology and Management*, 2012, **279**: 128–140
- McKenney D W, Hutchinson M F, Papadopol P, Lawrence K, Pedlar J, Campbell K, Milewska E, Hopkinson R F, Price D, Owen T. Customized spatial climate models for North America. *Bulletin of the American Meteorological Society*, 2011, **92**(12): 1611–1622
- Wang T, O'Neill G A, Aitken S N. Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecological Applications*, 2010, **20**(1): 153–163
- Rehfeldt G E, Ying C C, Spittlehouse D L, Hamilton D A Jr. Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecological Monographs*, 1999, **69**(3): 375–407
- Hamann A, Wang T. Potential effects of climate change on ecosystem and tree species distribution in British Columbia. *Ecology*, 2006, **87**(11): 2773–2786
- Rehfeldt G E, Crookston N L, Sáenz-Romero C, Campbell E M. North American vegetation model for land-use planning in a changing climate: a solution to large classification problems. *Ecological Applications*, 2012, **22**(1): 119–141
- Daly C, Smith J W, Smith J I, McKane R B. High-resolution spatial

- modeling of daily weather elements for a catchment in the Oregon Cascade Mountains, United States. *Journal of Applied Meteorology and Climatology*, 2007, **46**(10): 1565–1586
13. Daly C, Gibson W P, Taylor G H, Johnson G L, Pasteris P. A knowledge-based approach to the statistical mapping of climate. *Climate Research*, 2002, **22**: 99–113
 14. Rehfeldt G E, Crookston N L, Sáenz-Romero C, Campbell E M. North American vegetation model for land-use planning in a changing climate: a solution to large classification problems. *Ecological Applications*, 2012, **22**(1): 119–141
 15. Girvetz E H, Zganjar C, Raber G T, Maurer E P, Kareiva P, Lawler J J. Applied climate-change analysis: the climate wizard tool. *PLoS One*, 2009, **4**(12): e8320
 16. Wang T, Hamann A, Spittlehouse D L, Aitken S N. Development of scale-free climate data for western Canada for use in resource management. *International Journal of Climatology*, 2006, **26**(3): 383–397
 17. Wang T, Hamann A, Spittlehouse D L, Murdock T. ClimateWNA—high-resolution spatial climate data for Western North America. *Journal of Applied Meteorology and Climatology*, 2012, **51**(1): 16–29
 18. Wang T, Hamann A, Spittlehouse D, Carroll C. Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One*, 2016, **11**(6): e0156720
 19. Mitchell T D, Jones P D. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, 2005, **25**(6): 693–712
 20. IPCC. Climate Change 2014: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2014
 21. Taylor K E, Stouffer R J, Meehl G A. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 2012, **93**(4): 485–498
 22. Allen R G, Pereira L S, Raes D, Smith M. Crop evapotranspiration—guidelines for computing crop water requirements. *FAO Irrigation and Drainage Paper FAO56, U.N. Food and Agriculture Organization, Rome*, 1998
 23. Daly C, Gibson W P, Taylor G H, Johnson G L, Pasteris P. A knowledge-based approach to the statistical mapping of climate. *Climate Research*, 2002, **22**(2): 99–113
 24. Wang T L, Hamann A, Spittlehouse D L, Murdock T Q. ClimateWNA-high-resolution spatial climate data for Western North America. *Journal of Applied Meteorology and Climatology*, 2012, **51**(1): 16–29
 25. Daly C, Halbleib M, Smith J I, Gibson W P, Doggett M K, Taylor G H, Curtis J, Pasteris P A. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology*, 2008, **28**(15): 2031–2064
 26. Rehfeldt G E, Jaquish B C, Lopez-Upton J, Saenz-Romero C, St Clair J B, Leites L P, Joyce D G. Comparative genetic responses to climate for the varieties of *Pinus ponderosa* and *Pseudotsuga menziesii*: realized climate niches. *Forest Ecology and Management*, 2014, **324**: 126–137
 27. Wang T, Hamann A, Yanchuk A, O'Neill G A, Aitken S N. Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biology*, 2006, **12**(12): 2404–2416
 28. Chakraborty D, Wang T, Andre K, Konner M, Lexer M J, Matulla C, Weißenbacher L, Schueler S. Adapting Douglas-fir forestry in Central Europe: evaluation, application, and uncertainty analysis of a genetically based model. *European Journal of Forest Research*, 2016, **135**(5): 919–936
 29. Yeaman S, Hodgins K A, Lotterhos K E, Suren H, Nadeau S, Degner J C, Nurkowski K A, Smets P, Wang T, Gray L K, Liepe K J, Hamann A, Holliday J A, Whitlock M C, Rieseberg L H, Aitken S N. Convergent local adaptation to climate in distantly related conifers. *Science*, 2016, **353**(6306): 1431–1433
 30. Fowler H J, Blenkinsop S, Tebaldi C. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 2007, **27**(12): 1547–1578
 31. Wang T, Campbell E M, O'Neill G A, Aitken S N. Projecting future distributions of ecosystem climate niches: uncertainties and management applications. *Forest Ecology and Management*, 2012, **279**: 128–140
 32. Roberts D R, Hamann A. Predicting potential climate change impacts with bioclimate envelope models: a palaeoecological perspective. *Global Ecology and Biogeography*, 2012, **21**(2): 121–133
 33. Wang T, Wang G, Innes J, Nitschke C, Kang H. Climatic niche models and their consensus projections for future climates for four major forest tree species in the Asia-Pacific region. *Forest Ecology and Management*, 2016, **360**: 357–366