Climatic Change (2010) 98:113–131
DOI 10.1007/s10584-009-9622-2

# Use of the Köppen–Trewartha climate classification to evaluate climatic refugia in statistically derived ecoregions for the People's Republic of China

**Barry Baker · Henry Diaz · William Hargrove · Forrest Hoffman**

Received: 13 February 2008 / Accepted: 27 May 2009 / Published online: 29 July 2009


**Abstract** Changes in climate as projected by state-of-the-art climate models are likely to result in novel combinations of climate and topo-edaphic factors that will have substantial impacts on the distribution and persistence of natural vegetation and animal species. We have used multivariate techniques to quantify some of these changes; the method employed was the Multivariate Spatio-Temporal Clustering (MSTC) algorithm. We used the MSTC to quantitatively define ecoregions for the People's Republic of China for historical and projected future climates. Using the Köppen–Trewartha classification system we were able to quantify some of the temperature and precipitation relationships of the ecoregions. We then tested the hypothesis that impacts to environments will be lower for ecoregions that retain their approximate geographic locations. Our results showed that climate in 2050, as

B. Baker
Natural Resource Ecology Lab, B256, Colorado State University, Fort Collins, CO 80523, USA

H. Diaz
NOAA/ESRL/CIRES, 325 Broadway, Boulder, CO 80305, USA
e-mail: Henry.F.Diaz@noaa.gov

W. Hargrove
Eastern Forest Threat Assessment Center, USDA Forest Service,
Southern Research Station, Asheville, NC 28804-3454, USA
e-mail: hnw@geobabble.org

F. Hoffman
Computer Science & Mathematics Division, Oak Ridge National Laboratory,
P.O. Box 2008, Oak Ridge, TN 37831-6016, USA
e-mail: forrest@climatemodeling.org

*Present Address:*
B. Baker (✉)
The Nature Conservancy, Canyonlands Research Center, 820 Kane Creek Blvd.,
P.O. Box 1329, Moab, UT 84532, USA
e-mail: Barry.Baker@Colostate.edu

projected from anthropogenic forcings using the Hadley Centre HadCM3 general circulation model, were sufficient to create novel environmental conditions even where ecoregions remained spatially stable; cluster number was found to be of paramount importance in detecting novelty. Continental-scale analyses are generally able to locate potentially static ecoregions but they may be insufficient to define the position of those reserves at a grid cell-by-grid cell basis.

## 1 Introduction

In the last two decades of the 20th century the rate of warming (0.3°C/10 years) over China has been considerably greater than the global value of 0.19°C/10 years (Wang and Gong 2000). Current fully-coupled general circulation models project that this trend will continue into the future with mean annual temperatures increasing 1.8–3.5°C by 2050 and 2–7°C by 2100 (Cubash et al. 2001). Changes such as these will likely result in novel combinations of climate and topo-edaphic factors that will have substantial impacts on the distribution and persistence of natural vegetation and animal species (Hobbs et al. 2006; Ohlemüller et al. 2006; Ricciardi 2007; Williams and Jackson 2007; Williams et al. 2007).

Recently multivariate statistical techniques have been utilized to quantify these changes (Coulston and Riitters 2005; Hargrove and Hoffman 1999; Metzger et al. 2005; Wang and Price 2007; Williams et al. 2007). One such method, the Multivariate Spatio-Temporal Clustering (MSTC) algorithm, developed by Hargrove and Hoffman (2004), was used to demonstrate how abiotic environmental domains in the continental United States, Alaska and portions of western Canada move through space and time under various climate scenarios (Hargrove and Hoffman 2004; Hoffman et al. 2005; Saxon et al. 2005). These abiotic domains have similar characteristics with respect to bioclimatic, edaphic, and topographic properties and have been referred to as "climate-dynamic domains" (Saxon et al. 2005) or quantitatively defined ecoregions (*sensu* Hargrove and Hoffman 2004). We will refer to these domains as ecoregions throughout the rest of this paper.

Once calculated, the ecoregions are mapped back into geographic space to demonstrate shifts in their spatial locations. Additionally a single metric, the Euclidean distance between ecoregion centroids, has been used to demonstrate the magnitude of change between any two ecoregions (Saxon et al. 2005). Although this metric is one of the few that can be used as a quantitative index of environmental change, it lacks the ability to communicate the nature of the change without additional detailed analyses.

Climate classification schemes provide an efficient method for capturing climatic variables and seasonality into a single metric. One such accepted method, the Köppen–Geiger classification system (Köppen 1931, 1936), has been widely used to describe the potential distribution of natural vegetation based on climatic thresholds thought to drive critical physiological processes (Bailey 1996; Kottek et al. 2006). One of the main advantages of this type of classification scheme is that it is easy to use with a variety of data sets and model outputs. The classification has been used to detect recent changes in global and regional climate regimes (Diaz and Eischeid 2007; Fraedrich et al. 2001; Wang and Overland 2005) and to categorize simulation results from general circulation models for paleo (Guetter and Kutzbach 1990) and future climates (Gnanadesikan and Stouffer 2006; Kalvová et al. 2003).

This paper builds on previous work to develop methods that enable conservation practitioners to anticipate climatic change (Saxon et al. 2005). For this study we used a modified Köppen classification (Trewartha and Horn 1980) method for quantifying the climatic portion of the ecoregions and to test the assumption made by Saxon et al. (2005) that impacts will be lower from climatic change for ecoregions, which retain their locations and will provide refugia for species.

## 2 Methods

### 2.1 Multivariate spatio-temporal clustering

We used the MSTC technique (Hargrove and Hoffman 2004), which is based on a *k*-means clustering algorithm (Hartigan 1975), to quantitatively define ecoregions for the People's Republic of China under historical climatic conditions and eight future predicted climate scenarios. In general, the algorithm starts with *k*, 500 for this study, randomly selected unique map cells (pixels) used as initial "seed" centroids. An iterative process tests the Euclidean distance from each map cell to every centroid, classifying it to the closest existing centroid. After all the map cells are classified, a new centroid position is calculated as the mean of all coordinates of each map cell classified to that centroid. Thus, the centroids move through the data space while the map cells remain fixed. The classification converges and the iterations stop when fewer than a fixed number of map cells change their ecoregion assignment from the last iteration (we used <0.05%). Map cells are then re-united with their geographic coordinates and assembled back into the map along with their final ecoregion classification assignments.

We included 17 environmental variables (Table 1) from three distinct abiotic categories (edaphic, topographic, and bioclimatic) that exert a strong influence on the geographic distribution of flora and fauna and net primary production (Box 1981; Lugo et al. 1999; Neilson 1995; Prentice et al. 1992; Prentice 1990; Woodward 1987). All data sets used in this analysis were spatially distributed on a 4 $km^2$ (1.25 arcmin) data grid.

Map-based gridded climatic data used to calculate the bioclimatic variables consisted of monthly time-series for maximum and minimum monthly temperature (MMxT and MMnT) and precipitation (PPT). These geographic layers were developed from weather station data (~2,500 stations) which were interpolated to a 4 $km^2$ grid using the PRISM model (Daly et al. 2002, 1994) and a subset of the Tyndall Centre TYN SC 2.0 climatic data set (Mitchell et al. 2004), which were consistent with the data sets used by Saxon et al. (2005).

The TYN SC 2.0 data included monthly output (MMxT, MMnT, and PPT) for the period 2001–2100. These data were generated by two general circulation models (GCMs) from the Intergovernmental Panel on Climate Change Third Assessment Report (Cubash et al. 2001)—the Hadley Centre's general circulation model HadCM3 (Gordon et al. 2000; Pope et al. 2000) and the Parallel Climate Model (PCM) developed by the National Center for Atmospheric Research (Washington et al. 2000). Both models were run into the future using two different forcing scenarios, A1FI and B1, from the IPCC Special Report on Emissions Scenarios (Nakićenović and Swart 2000).

**Table 1** Results from the principal component analysis and factor loadings after the varimax rotation of the standardized input variables for all 500 ecoregion centroid values

| | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| Eigenvalue | 6.43 | 3.51 | 2.78 |
| Cumulative percent | 37.8 | 58.5 | 74.8 |
| Rotated factor pattern | | | |
| Edaphic variables | | | |
| Available water capacity | 0.164 | −0.137 | 0.803 |
| Bulk density ($g/cm^3$) | 0.303 | −0.158 | 0.760 |
| Total carbon ($g/m^2$) | −0.171 | 0.086 | 0.844 |
| Total nitrogen ($g/m^2$) | −0.169 | 0.092 | 0.844 |
| Topographic variables | | | |
| Elevation (m) | −0.890 | −0.058 | −0.159 |
| Compound topographic index | 0.485 | 0.501 | 0.017 |
| Annual potential solar insolation ($kW/m^2$) | 0.210 | 0.190 | 0.083 |
| Bioclimatic variables | | | |
| Annual maximum temperature (°C) | 0.962 | 0.044 | 0.032 |
| Annual minimum temperature (°C) | 0.616 | −0.641 | −0.173 |
| Annual precipitation (mm) | 0.185 | −0.939 | 0.066 |
| Potential evapotranspiration (mm) | 0.891 | −0.303 | −0.071 |
| Precipitation of coldest quarter (mm) | 0.194 | −0.704 | −0.014 |
| Precipitation/potential evapotranspiration | −0.241 | −0.896 | 0.118 |
| Precipitation of warmest quarter (mm) | 0.085 | −0.844 | 0.109 |
| Temperature of coldest quarter (°C) | 0.639 | −0.624 | −0.175 |
| Temperature of warmest quarter (°C) | 0.968 | 0.019 | 0.025 |
| Kira's warmth index | 0.901 | −0.399 | −0.056 |

Signs and magnitude of the loadings indicate warming on factor 1, drying on factor 2, and edaphic parameters load on factor 3 as soil fertility

Over the course of the next 100 years, the A1FI and B1 SRES scenarios correspond with higher and lower emissions, respectively. The A1FI represents a world with fossil fuel-intensive economic growth such that atmospheric concentrations of carbon dioxide ($CO_2$) should rise from 380 ppm to 580 ppm in 2050 and 940 ppm in 2100. The B1 scenario represents higher economic growth based on a shift away from fossil fuels such that atmospheric $CO_2$ should reach 482 ppm by 2050 and 550 ppm by 2100.

We used the 30-year (1961–1990) monthly climatology for MMxT, MMnT, and PPT from the PRISM dataset to represent historical climate. For the future climate, we calculated monthly averages for MMxT, MMnT, and PPT over the 30-year period 1961–1990 for each of the four coarse-scale GCM scenarios at each grid cell over the entire spatial domain. We calculated anomalies (differences for the temperature variables and as ratios for precipitation) between the 30-year average and the corresponding month for each of the four future GCM time-series. The anomaly for each variable interpolated to the high-resolution grid using a bilinear interpolation procedure. The interpolated anomalies were applied back to the observed PRISM climatology (1961–1990 monthly averages) to create a high-resolution future monthly time-series (2001–2100). We calculated monthly averages for two future time periods 2050 (representing the 30-year average for 2041–2070) and 2080 (representing the 30-year average for 2071–2100) for each of the four downscaled scenarios, thus giving us eight future scenarios.

The non-climate variables were derived from the 1:1,000,000 Chinese gridded (4 $km^2$) soil database (Shi et al. 2004) and the GTOPO30 digital elevation model (U. S. Geological Survey 1996), which we resampled to a 4 $km^2$ resolution grid. For this study we assumed the spatial distribution of the edaphic and topographic variables will be identical to historical conditions under all future scenarios. Clearly, changes in climate and atmospheric concentrations of $CO_2$ have direct and indirect effects on soil nutrients and influence ecosystem dynamics at various temporal and spatial scales (Morgan et al. 2007; Shaver et al. 2000; Weltzin et al. 2003). Saxon et al. (2005), however, maintained that soil carbon and nitrogen were slowly changing variables; therefore, it was necessary that we make the same assumption in order to test whether areas with lower magnitude of change would provide refugia.

We standardized the values for all seventeen variables (Table 1) across the nine climate scenarios such that each variable had a mean of zero and a standard deviation of one. We conducted a principal components analysis (PCA) to remove the effects of cross-correlations. We performed the MSTC on the resulting data to simultaneously divide the $\sim 4.25 \times 10^8$ data points into 500 ecoregions. We chose this level of cluster division, since it was able to resolve several large-scale regions such as the Taklamakan Desert and Chengdu Basin while retaining the spatial detail of Tian Shan and Himalayan mountain ranges. We mapped the data back to their original geographic locations and each ecoregion was assigned a unique color.

We calculated the distance in environmental space between historical and future ecoregions by first performing a principal components analysis (PCA) with a varimax rotation on the centroid values for each of the ecoregions (Table 1). We then calculated the Euclidean distance in three-dimensional environmental space (weighted sum of the first three PCA rotated factors) between the centroids to create the "magnitude of change" from historical conditions for each of the future scenarios (Saxon et al. 2005). The greater the value for this metric, the larger the multivariate changes in the complex topographic, edaphic, and bioclimatic space.

### 2.2 Köppen–Trewartha

We used a modified Köppen–Trewartha climate classification (Trewartha and Horn 1980) to determine the distribution of climatic types and to describe the average climate of each ecoregion and to evaluate the magnitude of change between historical and future ecoregions. This modified classification identifies six main climate groups (Table 2), five of which are based on the great thermic zones (A, C, D, E, and F) and a dry group (B) that cuts across the first four thermic zones. The A-, C-, and D-zones are further subdivided based on the seasonality of temperature and precipitation.

The Köppen–Trewartha climatic types were obtained by applying the classification rules to each grid cell of the 30-year PRISM climatology as well as to all future scenarios. For example, a grid cell where the average monthly temperature for 4 to 7 months is greater than 10°C, the average temperature of the coldest month is <0°C, precipitation in the wettest summer month is higher than ten times that of the driest winter month, and the warmest month is <22°C would be classified as a temperate, continental, cool, dry winter climatic type (DCwa). For the clustered data, we calculated the climatic types by applying the Köppen–Trewartha classification rules to the mean monthly temperature and precipitation values for each ecoregion.

**Table 2** Classification rules and descriptions of the Köppen climatic types as modified by Trewartha

| Class | Type | Subtype | Description | Rules |
|---|---|---|---|---|
| A | | | Tropical | Coolest month >18°C |
| | Ar | | Rainy (tropical broadleaf evergreen rain forest) | 10 to 12 months wet |
| | Am | | Tropical monsoonal forest | $6 > P_{min} > (250 - P_{year}) / 25$ |
| | Aw | | Tropical deciduous forests/woodland | Winter dry > 2 dry months |
| B | | | Dry climates | Evaporation > precipitation |
| | BS | | Semi-arid | Annual rainfall < twice desert limit of precipitation |
| | | BSh | Tropical-subtropical shrubland | 8 or more months >10°C |
| | | BSk | Temperate-boreal steppe | Fewer than 8 months >10°C |
| | BW | | Arid or desert | Annual rainfall < desert limit of precipitation |
| | | BWh | Tropical-subtropical desert | 8 or more months >10°C |
| | | BWk | Temperate-cold desert | Fewer than 8 months >10°C |
| C | | | Subtropical climates | 8 to 12 months >10°C |
| | Cw | | Subtropical winter dry season | Winter dry season |
| | | Cwa | Mixed broadleaf deciduous and needleleaf forest | Warmest month >22°C |
| | | Cwb | Needleleaf evergreen and broadleaf evergreen forests | Warmest month <22°C |
| | Cf | | Subtropical humid | Driest month >30 mm |
| | | Cfa | Long and short needleleaf evergreen and broadleaf deciduous forests and evergreen broadleaf shrub understory | Warmest month >22°C; no distinct dry season |
| | | Cfb | Needleleaf evergreen and deciduous forest | Warmest month <22°C; no distinct dry season |
| D | | | Temperate climates | 4 to 7 months >10°C |
| | DC | | Temperate continental climate | Coldest month <0°C |
| | DCs | | Summer dry season | Summer dry season |
| | | DCsa | Mixed evergreen and deciduous forests | Warmest month >22°C |
| | | DCsb | Mixed evergreen and deciduous forests | Warmest month <22°C |
| | DCw | | Winter dry season | Winter dry season |
| | | DCwa | Mixed deciduous and needleleaf evergreen forests | Warmest month >22°C |
| | | DCwb | Needleleaf evergreen forests | Warmest month <22°C |
| | DCf | | Humid continental | Driest month >30 mm |
| | | DCfa | Mid-latitude grassland, broadleaf deciduous forests and woodlands, mixed evergreen and broadleaf forests | Warmest month >22°C; no distinct dry season |
| | | DCfb | Needleleaf evergreen and mixed needleleaf-deciduous forest | Warmest month <22°C; no distinct dry season |
| E | | | Needleleaf deciduous forest and tundra woodland | 1 to 3 months >10°C |
| F | | | Polar | |
| | Ft | | Tundra, high altitude steppe | 0°C ≤ warmest month < 10°C |
| | Fi | | Perpetual frost | All months <0°C |

## 3 Results

Results from the MSTC yielded 500 unique clusters or quantitatively defined ecoregions. Because the MSTC analyzed all nine scenarios simultaneously it was theoretically possible for one or more of the nine scenarios to contain all 500 ecoregions. However, none of the nine scenarios individually contained more than 496 ecoregions. For purposes of brevity and ease of discussion we present results from two of the nine possible climate scenarios. The historical scenario is represented

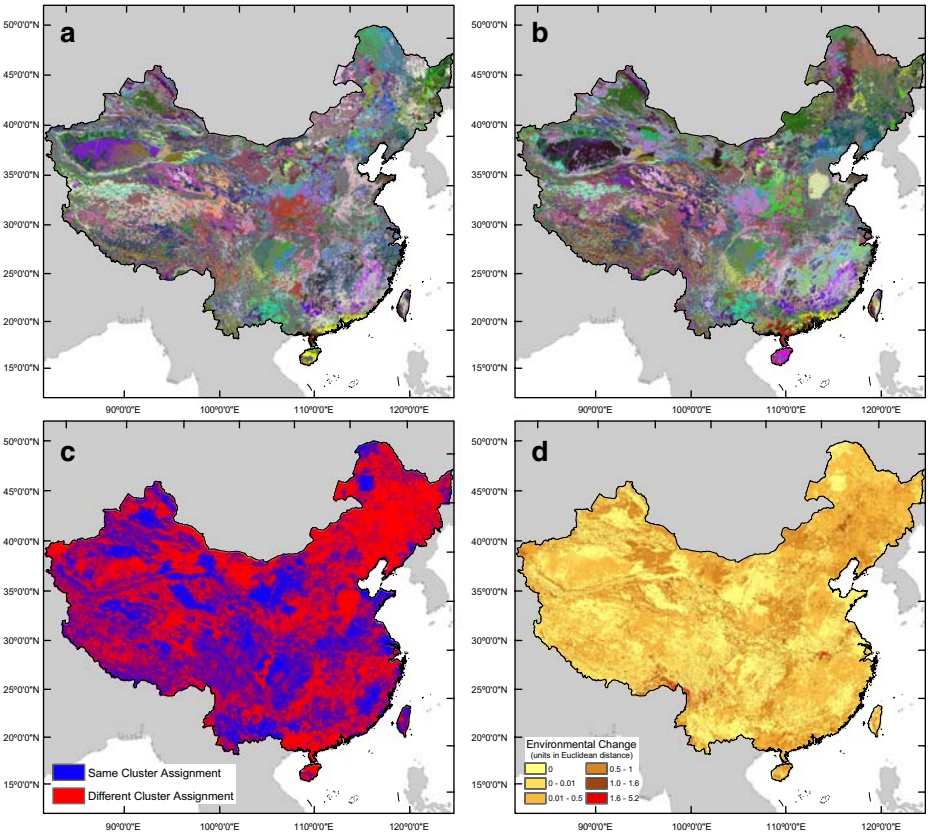


**Fig. 1** Quantitatively derived ecoregions from the Multivariate Spatio-temporal clustering of edaphic, topographic and climatic variables for the People's Republic of China and the magnitude of environmental change under two climate scenarios. **a** Ecoregions under historical climatic conditions (30-year average 1961–1990), **b** ecoregions for the 2050s time period (30-year average 2041–2070) under the HadCM3 A1FI scenario. *Colors* are consistent between ecoregions, **c** spatial representation of the change in pixel cluster assignment (*blue*—no change; *red*—new cluster number) between historical conditions and the 2050s time period (30-year average 2041–2070) under the HadCM3 A1FI scenario, **d** the magnitude of environmental change, based on Euclidean distance, from historical conditions to conditions for the 2050s time period (30-year average 2041–2070) under the HadCM3 A1FI scenario

by the PRISM climatology while the 2050 HadCM3 A1FI scenario represents one potential future climate.

We found that changes in precipitation and temperature simulated by the 2050 HadCM3 A1FI scenario were sufficient to cause shifts in the spatial distribution of the majority of ecoregions (Fig. 1a, b). Under the historical climate scenario, the MSTC algorithm grouped 458 ecoregions. However, in the future time period, shifts in bioclimatic variables resulted in the creation of 39 novel environmental conditions and the loss of one historical ecoregion resulting in the formation of 496 unique ecoregions.

Further analyses revealed that 60% of all pixels changed cluster or ecoregion assignment with the simulated climatic change (Fig. 1c). We found that the values for the magnitude of change ranged from 0 to 5.2 when we compared the historical climate to the future under the 2050 HadCM3 A1FI scenario and were spatially heterogeneous (Fig. 1d).

Results from the PCA of the ecoregion centroids showed that 74.8% of the variation could be accounted for in the first three principal components (Table 1). The loading pattern revealed that the first principal component was highly related to temperature, while decreasing precipitation and soil fertility corresponded to the second and third components, respectively.

Applying the Köppen–Trewartha classification to each pixel of the PRISM climatology resulted in 18 climatic types (Fig. 2a). The Köppen–Trewartha classification captures the broad scale climate patterns previously described Kottek et al. (2006) and Trewartha and Horn (1980). However, due to the relatively high-resolution climatic data used in this study, we were able to resolve climatic types for geographical settings such as the hot-dry valleys (BSh—arid shrubland and steppe climatic type) of the Yarlong Tsangpo (Brahmaputra), Lancang (Mekong) and Jinsha (upper Yangtze) Rivers (Chang 1981, 1983; Jin and Ou 2000) in the Tibetan Autonomous Region and northwestern Yunnan Province (Fig. 2a).

Comparison of the MSTC clustered climatic types (Fig. 2b) with the PRISM climatology (Fig. 2a) generally shows good general agreement in broad patterns of climatic types. However, not all of the Köppen–Trewartha climatic types were represented when we applied the classification to the average climate of the clustered ecoregions (Fig. 2b and d). The tropical rainforest (Ar) was not present under the historical climate scenario and neither the subtropical shrubland (BSh) nor subtropical desert (BWh) types were represented under either scenario.

As with the comparisons above, the future scenarios (un-clustered and clustered) showed broad-scale agreement in the spatial distribution of climatic types (Fig. 2c and d). All climatic types represented in the historical PRISM climatology remained in the un-clustered future HadCM3 A1FI scenario.

We chose a single ecoregion (cluster number 59) located in northeastern China to demonstrate the effects of climatic change on the spatial distribution of ecoregions and to test the assumptions that geographic stationarity of an ecoregion and small values of the magnitude of change metric corresponds to refugia. Under the future scenario, the spatial area of ecoregion 59 increased two-fold from $\sim 21.3 \times 10^3$ km$^2$ (5,324 pixels) to $\sim 48.7 \times 10^3$ km$^2$ (12,170 pixels) resulting mostly from a northern shift in extent (red shading, Fig. 3). Additionally we found that pixels in the future domain of the ecoregion 59 tended to be warmer and wetter as well as being distributed over a greater elevation range and less fertile, as measured by the

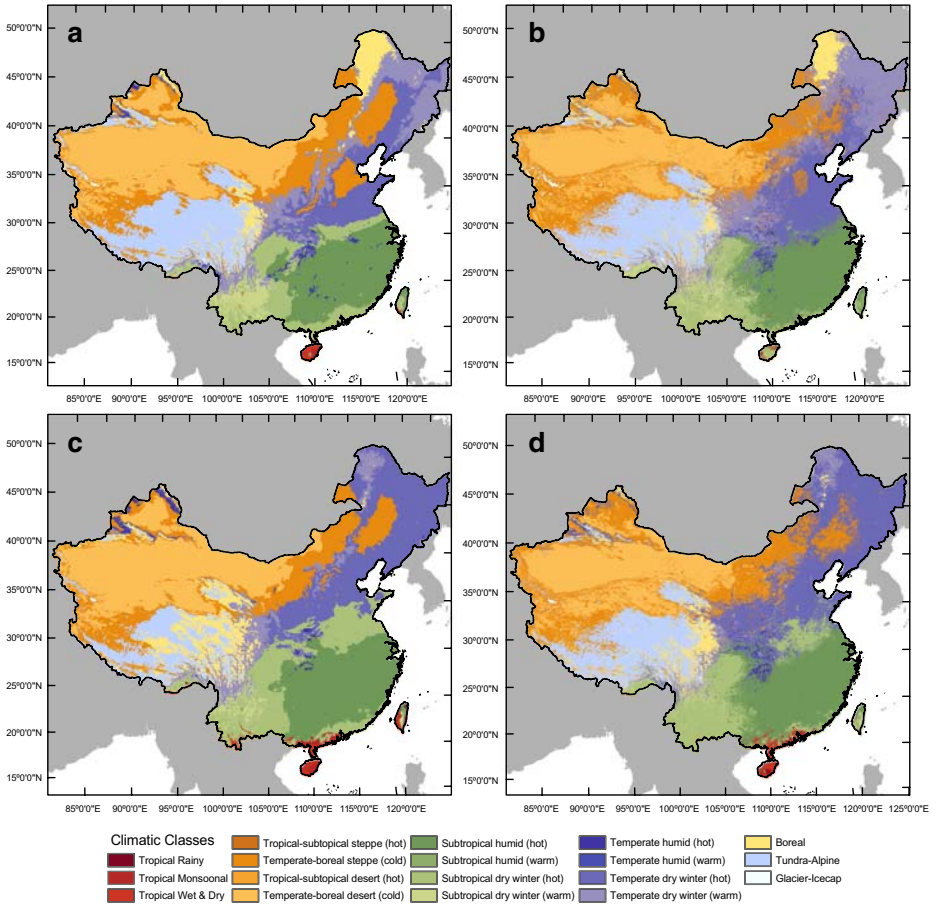


**Fig. 2** Spatial representation of Köppen–Trewartha climate classification for the People's Republic of China. **a** Climate classification of the historical climate for the individual cells of the PRISM climatology (30-year average 1961–1990). **b** Climate classification of the clustered ecoregions under historical climatic conditions (30-year average 1961–1990). **c** Climate classification of the individual cells of the climate data for the 2050s time period (30-year average 2041–2070) under the HadCM3 A1FI scenario. **d** Climate classification of the 500 ecoregions for the 2050s time period (30-year average 2041–2070) under the HadCM3 A1FI scenario. Differences between maps in the *left* and *right columns* are due to the addition of edaphic and topographic factors as well as regional generalization by Multivariate Spatio-Temporal Clustering (MSTC)

carbon to nitrogen ratio, when compared to the historical distribution of pixels (Fig. 4). The majority of the core area (4,815 pixels) of this ecoregion, however, remained geographically intact (blue shading, Fig. 3). Consequently, the amount of environmental change, as represented by the magnitude of change metric, was equal to 0.0 for these cells.

For the historical scenario, ecoregion 59 was classified climatically as type E, a boreal climatic type (Fig. 2b), which is consistent with the current vegetation types found in this region (Chen et al. 2003; Fang and Yoda 1989; Liu et al. 2003, 2006;

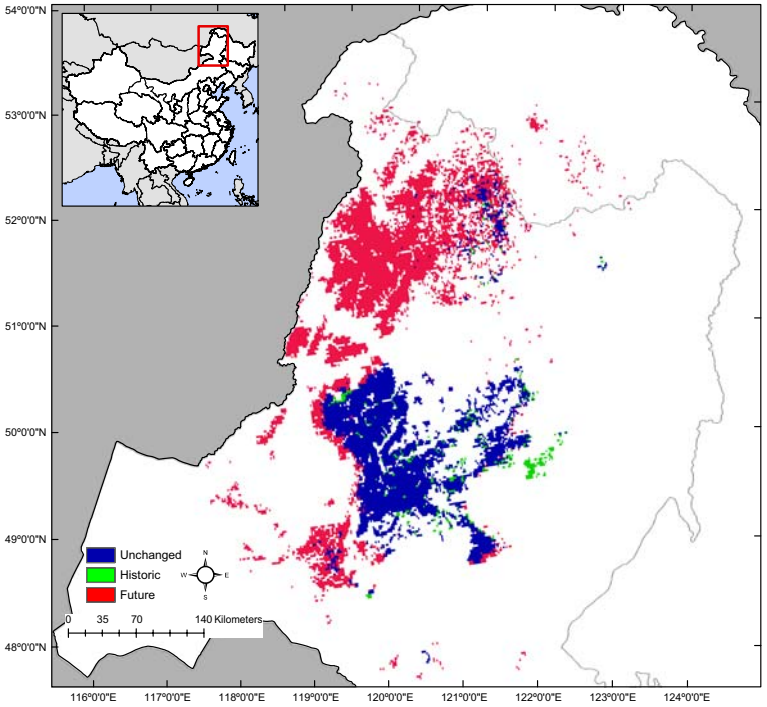


**Fig. 3** Change in spatial extent of ecoregion 59 under the historical climate (30-year average 1961–1990) and the 2050s (30-year average 2041–2070) HadCM3 A1FI scenarios. *Green* represents areas that occur only under the historical climate. *Red* represents areas that occur only under the future scenario. Locations in *blue* indicate persistence of ecoregion and magnitude of change equal to 0

Yong and Feoli 1991). We found that the climatic variability within the entire spatial domain of ecoregion 59 was represented by three climatic types (BSk—cold semi-arid steppe, DCwb—cold temperate forest, and E—boreal forest). The majority of the pixels (~99%) were classified as the boreal climatic type (E) and had a median elevational distribution of 803 m. The distributions of the other two types (DCwb and BSk) represented fewer pixels and were usually found at lower elevations (Table 3).

Under the 2050 A1FI scenario the average climate over the entire future spatial extent of the ecoregion became warmer and wetter (Fig. 4a and b). Consequently the climatic type for the ecoregion switched from the E—boreal forest type to a cold temperate forest type (Fig. 2d) with ~51% of the pixels being classified as DCwb (Table 3).

Analysis of the climatic variability within the ecoregion showed a shift in the composition of the original climatic types (BSk, DCwb, and E) and the addition of a warmer temperate forest type climate (DCwa) as well as DCfb, which represents a cooler temperate forest with no distinct dry season. The new climatic type, DCwa, represented approximately 47% of the northern latitudinal pixels in the future distribution of the ecoregion (Table 3). The other three climatic types, BSk, DCfb, and E, represented ~2.5% of the total area. Although the number of pixels classified

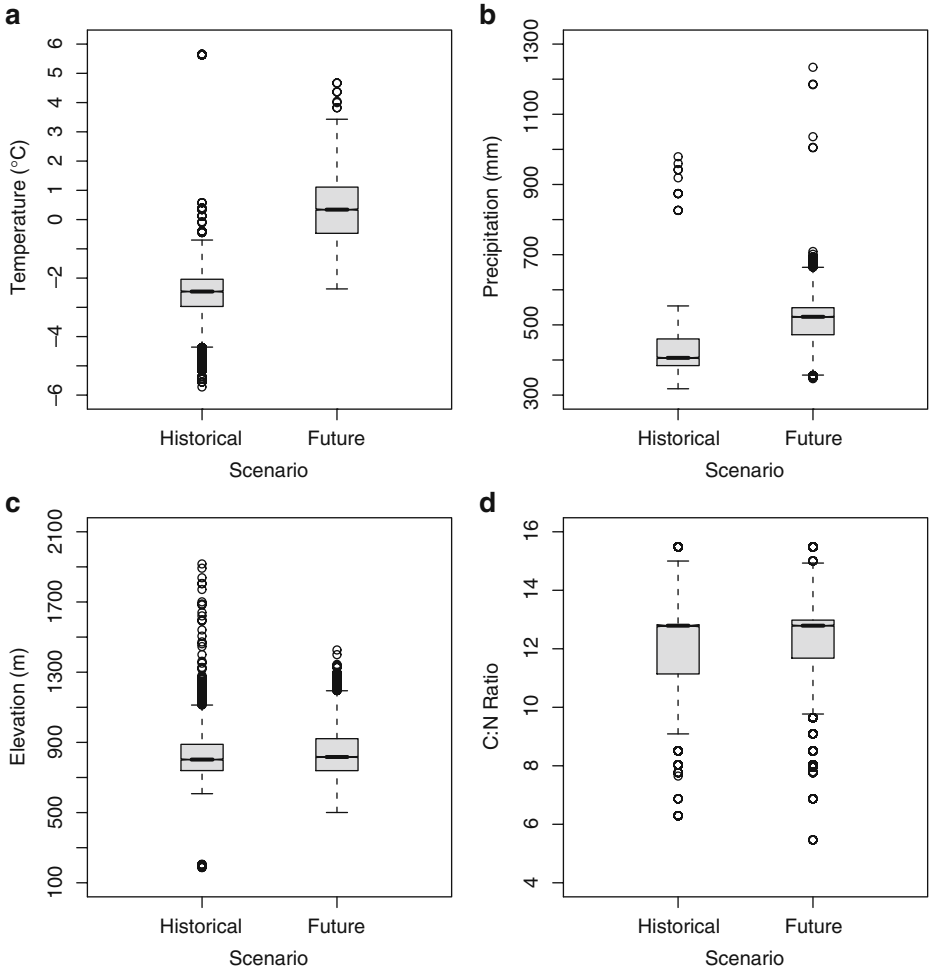


**Fig. 4** Boxplots showing median, quartiles (25–75%), minima and maxima of mean annual temperature (**a**), annual precipitation (**b**), elevation (**c**), and carbon to nitrogen ratio (**d**) for all pixels in ecoregion 59 under historical and future (2050s HadCM3 A1FI) climate scenarios

**Table 3** The number of pixels, elevational and latitudinal distribution of climatic types for the entire spatial extents of ecogregion 59 under historical and future climate scenarios

| Type | Historical | | | | | Future | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Number of pixels | Elevation (m) | | Latitude (°N) | | Number of pixels | Elevation (m) | | Latitude (°N) | |
| | | Median | Range | Median | Range | | Median | Range | Median | Range |
| BSk | 29 | 258 | 187–794 | 44 | 44–48 | 222 | 672 | 541–794 | 48 | 48–49 |
| DCfb | 0 | | | | | 6 | 968 | 906–1,426 | 42 | 42 |
| DCwa | 0 | | | | | 5,703 | 755 | 500–1,140 | 49 | 47–52 |
| DCwb | 47 | 624 | 614–701 | 49 | 49 | 6,176 | 906 | 577–1,426 | 51 | 46–53 |
| E | 5,248 | 803 | 607–1,917 | 49 | 42–51 | 66 | 1,125 | 983–1,334 | 51 | 47–51 |

**Table 4** Comparison of the elevational distribution and number of cells (in parentheses) for each climatic type contained in ecoregion number 59 under historical and future conditions where pixel cluster assignment did not change

| Type | Future | | | |
|---|---|---|---|---|
| Historical | DCwa | DCwb | DCfb | E |
| BSk | 785 m (3) | | | |
| DCwb | 634 m (46) | | | |
| E | 780 m (3,601) | 947 m (1,138) | 1,039 m (4) | 1,131 m (23) |

as BSk—cold semi-arid steppe increased, the climatic type represented a small fraction (1.8%) of the total area in the northern portions of the ecoregion.

When we examined the 4,815 pixels of ecoregion 59 that did not change cluster assignment (blue shading, Fig. 3) we discovered only 23 pixels of the higher elevational boreal climatic type (E) were retained under the future climate forcing scenario (Table 4). All other pixels were classified as a subcategory of a continental temperate forest type (DC), which resulted from a significant shift (paired Wilcoxon signed rank test, $p < 0.001$), from colder, drier to warmer, wetter climatic conditions (Fig. 5).

## 4 Discussion and conclusions

### 4.1 Discussion

Unlike other approaches to ecoregional delineation (Bailey 1996; Omernik 1995) by using the MSTC algorithm we were able to objectively define unique clusters

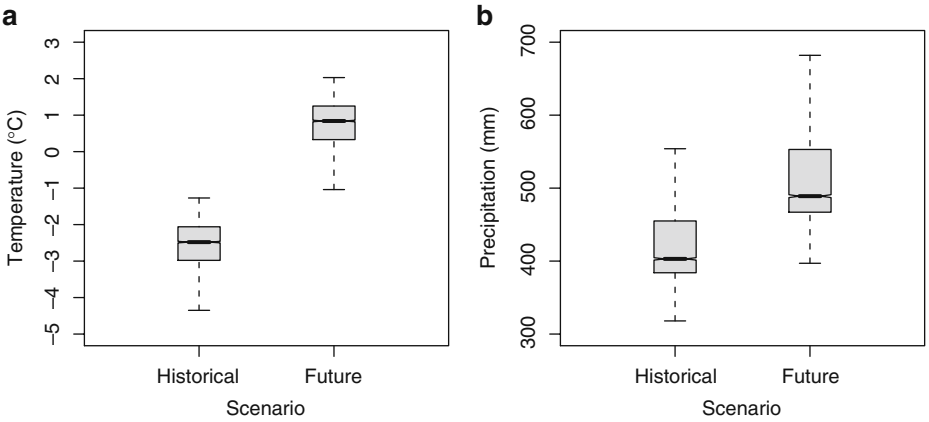


**Fig. 5** Boxplots showing median, quartiles (25–75%), minima and maxima of average temperature (**a**) and annual precipitation (**b**) for all pixels in ecoregion 59 which persisted in the same spatial location under historical and future (2050s HadCM3 A1FI) climate scenarios. The differences in both average temperature (**a**) and annual precipitation were significant at $p < 0.001$ 95% confidence level using the paired Wilcoxon signed rank test

or ecoregions to demonstrate how abiotic variables change through space and time under a projected future climate forcing. Changes in climate by 2050, from projected anthropogenic forcings, resulted in the creation of novel combinations of climate and topo-edaphic factors. Moreover, as in other recent studies that demonstrated the evolution of novel climates (Ohlemüller et al. 2006; Williams and Jackson 2007; Williams et al. 2007) under future scenarios, the changes presented in this study were sufficient to create environmental conditions that are not currently present anywhere in China today. By using the MSTC approach to ecoregional delineation we are able to provide some insight into the potential disruption of current habitat structure and species distributions under future climatic change scenarios. To completely analyze the impact of climatic shifts on distributions of plants and animals, stationary topographic and edaphic factors must be considered in addition to strictly climatic variables.

The Köppen–Trewartha classification provided an additional method for visualizing and interpreting how climate space and consequently vegetation space shifted under a future scenario. Using this method we were able to document that the spatial patterns of climatic change resulted in a northern migration of warmer climatic types and an increase in the elevational limits of forests (Table 3) as well as a slight expansion in the high latitude desert and arid shrubland regions in northwestern China (Fig. 2c and d). These shifts were consistent with observations of impacts on vegetation from recent climatic change (Baker and Moseley 2007; Diaz et al. 2003; Gou et al. 2007; Gu et al. 2007) and future predictions from more process based vegetation models (He et al. 2005; Leng et al. 2008; Song et al. 2004, 2005).

The *k*-means algorithm homogenizes and partitions the environmental variance equally across each of the statistically derived ecoregions, such that each had similar amounts of heterogeneity (Estivill-Castro and Yang 2004); thus we were able to compare any two ecoregions for any given time period, as demonstrated by Hargrove and Hoffman (2004). However, one of the main challenges with the *k*-means algorithm lies in determining, *a priori*, a suitable number of clusters (Estivill-Castro and Yang 2004). For example, choosing a relatively small number of clusters ($k$) can result in the failure to discriminate small unique areas as regions separate from their larger-scale "parent" ecoregions.

The influence of the size of $k$ became obvious when we applied the Köppen–Trewartha classification to the 500-cluster data set. Even though we were able to reproduce the broad spatial pattern of climates existing in the un-clustered data in both the historical and future scenarios (Fig. 2), three climatic types (Ar, BSh, and BWh) were not represented in the clustered analysis. The reason for the omission was that these three types (Ar, BSh, and BWh) represented a small percentage (0.02%, 0.07%, and 0.005% respectively) of the total area and were not identified as distinct ecoregions when the "within cluster" climate was averaged over the spatial domain.

Clearly, the correspondence between the clustered and un-clustered climatic types would have been better if more clusters (larger $k$ value) had been chosen. Requesting fewer clusters resulted in more broadly defined groups and allowed for greater within-ecoregion environmental variation. A larger value of k-clusters would have resulted in more rigorously defined homogenous ecoregions that would have exhibited less within-region variability. We should note that regardless of the $k$ value, the statistical classification process was uniformly applied such that the

environmental heterogeneity was equal across all ecoregions formed anywhere in the maps.

The number of clusters also had a strong influence on the amount of intra-cluster environmental variability in data space as demonstrated by the range of values for mean annual temperature (−5.77°C to 5.66°C) and annual precipitation (318 mm to 1234 mm) within ecoregion 59 (Fig. 4). By using the Köppen–Trewartha classification scheme we were able to meaningfully bound that variability thus showing change in the potential distribution of habitats.

In general, ecoregion 59 was classified as a northern (∼42°N–53°N) forested ecoregion that had a wide elevational range. The amount of variation in climatic data space for the ecoregion was large enough to range from a cold high latitude steppe/shrubland (BSk) to boreal forest (E) climatic types (Table 3). Furthermore, when we examined how the variability was partitioned between the historical and future scenarios, we found that the climatic characteristics of the ecoregion were quite different. In fact the ecoregion became more environmentally diverse under the future scenario (Table 3). More importantly, even though the amount of intra-cluster environmental variance for the pixels of ecoregion 59 did not cross a statistical threshold which would push them into another cluster, the amount of internal variation was sufficient for nearly all the pixels to cross a climatic threshold according to Köppen–Trewartha, demonstrated by the shift of three climatic types to five climatic types (Table 3), thus possibly increasing the number of potential habitats within the ecoregion.

The three historical climatic types (BSk, DCwb, and E) that did persist into the future were spatially distributed in the higher elevational and latitudinal regions of ecoregion 59 (Table 3). Additionally, only 23 of the original 5,248 pixels that were classified as a boreal forest type were able to persist in the same spatial location (Table 4). This clearly indicates that the majority of species currently found in this ecoregion would be displaced and forced to migrate to new locations, forced to adapt to the new environmental conditions, or become locally extinct.

Saxon et al. (2005) maintained that the areas of greatest risk from climatic change are locations where domains disappear and different domains take their place. However, looking simply at domain stability, either "*in situ*" or by the magnitude of change metric may not be sufficient to guarantee the persistence of species or habitats as we have shown for ecoregion 59 (Table 4). Even though the historical extent of the ecoregion remained largely intact, changes in temperature and precipitation regimes, as defined by Köppen–Trewartha, suggested projected life-form level changes. If the ecological amplitude of native species is presumed stationary, then the assumption that species in this portion of the ecoregion will be least at risk or provide potential refugia from climatic change does not hold as maintained by Saxon et al. (2005) given the number of ecoregions ($k = 500$) used in the study.

As we stated above, the number of clusters or division influenced the amount of environmental variability that was distributed within each ecoregion. Choosing too few clusters (small values of $k$) or "under dividing" would homogenize the total amount of environmental variation across fewer ecoregions, as may have been the case with Saxon et al. (2005), would result in the conclusion that climatic refugia would exist for a particular species or assemblage when in fact they did not thus resulting in a Type I error. Conversely, "over division" (larger values of $k$) would

result in the conclusion that there were no refugia when in fact there were, a Type II error. In this case, a species' range might have been adequately represented by multiple quantitative ecoregions, any one of which could have provided adequate habitat. In climatic shift analyses such as the one performed here or in Saxon et al. (2005), the Type I error has more severe consequences since the suggestion of refugia may cause unjustified assurance of a species' persistence. Therefore, selecting larger values of $k$ would provide a more conservative strategy for identifying potential refugia and the conservation of species.

Additionally we suggest that using the Köppen–Trewartha classification can be used as an incomplete, partial "yardstick" to ensure that enough regions have been created to adequately discriminate the level of climatic variation relevant to plant and animal species. Because most of the environmental variability, which loaded on the first two rotated factors (Table 1), was in terms of climatic variability (temperature and precipitation respectively) one can assume that, once these are satisfied by the level of division, so will the other non-climatic variables included in the MSTC. The regionalization of present and future China produced here appears to be slightly coarser than the divisions utilized in Köppen–Trewartha (Fig. 2a and b), and so may be slightly insufficient. We believe that the regionalization of China would compare favorably with Köppen–Trewartha if it were repeated with 800 to 1,000 ecoregions. There is little cost (other than computation time) to over dividing beyond this critical level of discrimination; the additional division would merely result in a one-to-many relationship between the range of a particular animal or plant and the resulting quantitative ecoregions, as opposed to the minimum resolution of a one-to-one relationship.

### 4.2 Conclusions

The use of clustering methods, like ecoregions themselves, assumes that applying a certain optimum level of generalization provides some measure of value. The value may stem from increased understanding of ecological conditions that is made possible by dividing a multivariate set of conditions into a discrete set of kinds or types of environments such as those in the Köppen–Trewartha classification. It may be easier to recognize, understand, and remember that certain sets of plants and animals can be associated with each of these environmental types. Recognition of such simplified environmental sets should aid in the conservation of these associated plants and animals.

As with all generalizations, however, scrutiny at sufficiently fine resolution may reveal fine scale details that have been homogenized together. Such a tradeoff between simplicity and generalization must be optimized. Thus the results from our study suggest that a two-step process may be necessary for determining the ecological impacts of predicted climatic change. The Köppen–Trewartha classification alone is insufficient for determining ecological regions, since it does not consider non-climatic factors in defining potentially suitable habitat. For example, it produces fewer environmental groups than the 500 quantitative ecoregions produced by the MSTC. However, the Köppen–Trewartha classification has been refined so that it produces a level of generalization that is congruent with the level of discrimination used by ecologists when they examine plant and animal associations.

We suggest that when the MSTC is used, at a particular level of *k*-divisions, to produce an initial environmental stratification, a subsequent application of the Köppen–Trewartha classification should then be used to ensure that all major and minor ecologically relevant climate groupings have been successfully discriminated. If not the MSTC procedure should be repeated, requesting a greater number of *k* groups. In this way, the ecological impact analysis can be ensured to have sufficient resolution to discern all climatic combinations presently recognized as ecologically relevant. Only then should the analysis proceed to estimate the ecological impacts of the forecasted climatic shift.

Our new findings suggest that the ecological impacts of climatic shifts may be more severe than originally suggested by Saxon et al. (2005). Because their continental scale analyses was based only on 500 clusters, their analysis may have included generalizations that lumped together ecological differences that may have been relevant for the survival of species following the climatic shifts. This level of generalization may have caused Saxon et al. (2005) to overestimate the number and extent of refugia for plants and animals in the United States; therefore, underestimating the severity of the impacts from projected climatic change. Repeating their analysis with additional levels of division, followed by testing with Köppen–Trewartha to assure that an ecologically sufficient level of climatic discrimination remains, could produce a more accurate estimate of potential ecological impacts to guide future conservation efforts.

Understanding the appropriate level of ecoregional resolution is of paramount importance when interpreting the results of broad climatic change impact analyses like Saxon et al. (2005). A continental scale analysis, while able to generally locate potentially static refuges or reserves, may still be insufficient to define the position of those reserves at a cell-by-cell scale. The level of ecoregional generalization, as set in this study, may be such that few specific individual locations remain unaffected by the climatic shift. As the number of resolved ecoregions increases, however, the value of the ecoregional generalization decreases and the computation load increases. The prognostic ability of climatic change science at its current state of maturity may be insufficient to accurately provide such high-resolution forecasts.

In summary, the MSTC algorithm provides conservation planners with a robust alternative for delineating dynamic ecoregional boundaries. When these statistically defined boundaries are related to ecological threshold responses, such as the Köppen–Trewartha schema used in this paper, then a meaningful classification of the physical characteristics of the ecoregions can be developed, thus providing insights into species persistence and aiding in the development of conservation networks in an uncertain future.

**Acknowledgements** We wish to thank Chris Daly at Oregon State University, Shi Xuezheng from the Chinese Academy of Science's Institute of Soil Science, Tim Mitchell and the Tyndall Centre, East Anglia for providing access to data and scenarios; Bruce Godfrey for sharing some of the GIS code used in the analyses; Oak Ridge National Laboratory for developing the MSTC, guidance on interpretation and use of the computing facilities; The Nature Conservancy's China Program for partial financial and logistical support of this project; Mike Heiner and Dominique Bachelet for their thoughtful comments on earlier drafts of the manuscript; and finally, the two anonymous reviewers for their comments which increased the clarity of this manuscript.

## References


Bailey RG (1996) Ecosystem geography. Springer, New York

Baker BB, Moseley RK (2007) Advancing treeline and retreating glaciers: implications for conservation in Yunnan, P.R. China. Arct Antarct Alp Res 39:200–209

Box EO (1981) Macroclimate and plant forms: an introduction to predictive modeling in phytogeography. Dr. W. Junk Publishers, The Hague

Chang DHS (1981) The vegetation zonation of the Tibetan Plateau. Mt Res Dev 1:29–48

Chang DHS (1983) The Tibetan Plateau in relation to the vegetation of China. Ann Mo Bot Gard 70:564–570

Chen X, Zhang X-S, Li B-L (2003) The possible response of life zones in China under global climate change. Glob Planet Change 38:327–337

Coulston JW, Riitters KH (2005) Preserving biodiversity under current and future climates: a case study. Glob Ecol Biogeogr 14:31–38

Cubash U, Meehl GA, Boer GJ, Stouffer RJ, Dix M, Noda CA, Senior CA, Raper S, Yap KS (2001) Projections of future change. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Climate Change 2001: the scientific basis. Contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge

Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J Appl Meteorol 33:140–158

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

Diaz HF, Eischeid JK (2007) Disappearing "alpine tundra" Köppen climatic type in the western United States. Geophys Res Lett 34:L18707. doi:10.1029/2007GL031253

Diaz HF, Eischeid JK, Duncan C, Bradley RS (2003) Variability of freezing levels, melting season indicators, and snow cover for selected high-evelation and continental regions in the last 50 years. Clim Change 59:33–52

Estivill-Castro V, Yang J (2004) Fast and robust general purpose clustering algorithms. Data Min Knowl Disc 8:127–150

Fang J-Y, Yoda K (1989) Climate and vegetation in China II. Distribution of main vegetation types and thermal climate. Ecol Res V4:71–83

Fraedrich K, Gerstengarbe FW, Werner PC (2001) Climate shifts during the last century. Clim Change 50:405–417

Gnanadesikan A, Stouffer RJ (2006) Diagnosing atmosphere–ocean general circulation model errors relevant to the terrestrial biosphere using the Köppen climate classification. Geophys Res Lett 33:L22701. doi:10.1029/2006GL028098

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16:147–168

Gou X, Chen F, Jacoby G, Cook E, Yang M, Peng J, Zhang Y (2007) Rapid tree growth with respect to the last 400 years in response to climate warming, northeastern Tibetan Plateau. Int J Climatol 27:1497–1503

Gu Z, Chen J, Shi P, Xu M (2007) Correlation analysis of Normalized Different Vegetation Index (NDVI) difference series and climate variables in the Xilingole steppe, China from 1983 to 1999. Front Biol China 2:218–228

Guetter PJ, Kutzbach JE (1990) A modified Köppen classification applied to model simulations of glacial and interglacial climates. Clim Change 16:193–215

Hargrove WW, Hoffman FM (1999) Using multivariate clustering to characterize ecoregion borders. Comput Sci Eng 1:18–25

Hargrove WW, Hoffman FM (2004) Potential of multivariate quantitative methods for delineation and visualization of ecoregions. Environ Manage 34:S39–S60

Hartigan JA (1975) Clustering algorithms. Wiley, New York

He HS, Hao Z, Mladenoff DJ, Shao G, Hu Y, Chang Y (2005) Simulating forest ecosystem response to climate warming incorporating spatial effects in north-eastern China. J Biogeogr 32: 2043–2056

Hobbs RJ, Arico S, Aronson J, Baron JS, Bridgewater P, Cramer VA, Epstein PR, Ewel JJ, Klink CA, Lugo AE, Norton D, Ojima D, Richardson DM, Sanderson EW, Valladares F, Vila M,

Zamora R, Zobel M (2006) Novel ecosystems: theoretical and management aspects of the new ecological world order. Glob Ecol Biogeogr 15:1–7
Hoffman FM, Hargrove Jr WW, Erickson DJ III, Oglesby RJ (2005) Using clustered climate regime to analyze and compare predictions from fully coupled general circulation models. Earth Interact 9:1–27
Jin ZZ, Ou XK (2000) Jinshajiang vegetation of dry-hot valleys, Yunnan and Sichuan. In: Jin ZZ, Ou XK (eds) Yuanjiang, Nujiang, Jinshajiang, Lancangjiang vegetation of dry-hot valley. Yunnan University Press and Yunnan Science & Technology Press, Kunming
Kalvová J, Halenka T, Bezpalcová K, Nemešová I (2003) Köppen climate types in observed and simulated climates. Stud Geophys Geod 47:185–202
Köppen W (1931) Grundriss der Kilmakunde. Walter de Gruyter, Berlin
Köppen W (ed) (1936) Das Geographische System der Klimate. Gerbrüder Bonträger, Berlin
Kottek M, Grieser J, Beck C, Rudolf B, Rubel F (2006) World map of the Köppen–Geiger climate classification updated. Meteorol Z 15:259–263
Leng W, He HS, Bu R, Dai L, Hu Y, Wang X (2008) Predicting the distributions of suitable habitat for three larch species under climate warming in Northeastern China. For Ecol Manage 254: 420–428
Liu JY, Zhuang DF, Luo D, Xiao X (2003) Land-cover classification of China: integrated analysis of AVHRR imagery and geophysical data. Int J Remote Sens 24:2485–2500
Liu R, Liang S, Liu J, Zhuang D (2006) Continuous tree distribution in China: a comparison of two estimates from moderate-resolution imaging spectroradiometer and Landsat data. J Geophys Res 111. doi:10.1029/2005JD006039
Lugo AE, Brown SL, Dodson R, Smith TS, Shugart HH (1999) The Holdridge life zones of the conterminous United States in relation to ecosystem mapping. J Biogeogr 26:1025–1038
Metzger MJ, Bunce RGH, Jongman RHG, Mucher CA, Watkins JW (2005) A climatic stratification of the environment of Europe. Glob Ecol Biogeogr 14:549–563
Mitchell TD, Carter TR, Jones PD, Hulme M, New M (2004) A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100) Working Paper 55. Tyndall Centre for Climate, Norwich
Morgan JA, Milchunas DG, LeCain DR, West M, Mosier AR (2007) Carbon dioxide enrichment alters plant community structure and accelerates shrub growth in the shortgrass steppe. Proc Natl Acad Sci 104:14724–14729
Nakićenović N, Swart R (eds) (2000) Emissions scenarios special report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
Neilson RP (1995) A model for predicting continental-scale vegetation distribution and water balance. Ecol Appl 5:362–385
Ohlemüller R, Gritti ES, Sykes MT, Thomas CD (2006) Towards European climate risk surfaces: the extent and distribution of analogous and non-analogous climates 1931–2100. Glob Ecol Biogeogr 15:395–405
Omernik JM (1995) Ecoregions: a spatial framework for environmental management. In: Davis W, Simon TP (eds) Biological assessment and criteria: tools for water resource planning and decision making. Publishing, Boca Raton
Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parameterizations in the Hadley Centre climate model—HadCM3. Clim Dyn 16:123–146
Prentice KC (1990) Bioclimatic distribution of vegetation for general circulation model studies. J Geophys Res 95:11811–11830
Prentice IC, Cramer W, Harrison SP, Leemans R, Monserud RA, Solomon AM (1992) A global biome model based on plant physiology and dominance, soil properties and climate. J Biogeogr 19:117–134
Ricciardi A (2007) Are modern biological invasions an unprecedented form of global change? Conserv Biol 21:329–336
Saxon E, Baker B, Hargrove W, Hoffman F, Zganjar C (2005) Mapping environments at risk under different global climate change scenarios. Ecol Lett 8:53–60
Shaver GR, Canadell J, Chapin FS, Gurevitch J, Harte J, Henry G, Ineson P, Jonasson S, Melillo J, Pitelka L, Rustad L (2000) Global warming and terrestrial ecosystems: a conceptual framework for analysis. Bioscience 50:871–882
Shi XZ, Yu DS, Warner ED, Pan XZ, Peterson GW, Gong ZG, Weindorf DC (2004) Soil database of 1:1,000,000 digital soil survey and reference system of the Chinese genetic soil classification system. Soil Surv Horiz 45:129–136

Song M, Zhou C, Hua O (2004) Distribution of dominant tree species on the Tibetan Plateau under current and future climate scenarios. Mt Res Dev 24:166–173
Song M, Zhou C, Hua O (2005) Simulated distribution of vegetation types in response to climate change on the Tibetan Plateau. J Veg Sci 16:341–350
Trewartha GT, Horn LH (1980) An introduction to climate. McGraw-Hill, New York
U. S. Geological Survey (1996) GTOPO3. http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html. Cited 15 July 2005
Wang S, Gong D (2000) Enhancement of the warming trend in China. Geophys Res Lett 27: 2581–2584
Wang M, Overland JE (2005) Detecting arctic climate change using Köppen climate classification. Clim Change 67:43–62
Wang A, Price DT (2007) Estimating global distribution of boreal, temperate, and tropical tree plant functional types using clustering techniques. J Geophys Res 112:G01024. doi:10.1029/2006JG000252
Washington WM, Weatherly JW, Meehl GA, Semtner AJ Jr, Bettge TW, Craig AP, Strand WG Jr, Arblaster JM, Wayland VB, James R, Zhang Y (2000) Parallel climate model (PCM) control and transient simulations. Clim Dyn 16:755–774
Weltzin JF, Loik ME, Schwinning S, Williams DG, Fay PA, Haddad BM, Harte J, Huxman TE, Knapp AK, Lin G, Pockman WT, Shaw MR, Small EE, Smith MD, Smith SD, Tissue DT, Zak JC (2003) Assessing the response of terrestrial ecosystems to potential changes in precipitation. Bioscience 53:941–952
Williams JW, Jackson ST (2007) Novel climates, no-analog communities, and ecological surprises. Front Ecol Environ 5:475–482
Williams JW, Jackson ST, Kutzbach JE (2007) Projected distributions of novel and disappearing climates by 2100 AD. PNAS 104:5738–5742
Woodward FI (1987) Climate and plant distribution. Cambridge University Press, Cambridge
Yong SC, Feoli E (1991) A numerical phytoclimatic classification of China. Int J Biometeorol 35:76–87