Quaternary Research 83 (2015) 287–297

Contents lists available at ScienceDirect

Quaternary Research

journal homepage: www.elsevier.com/locate/yqres

Assessing the importance of climate variables for the spatial distribution of modern pollen data in China

Jianyong Li [a], Qinghai Xu [b,c,d,e,*], Zhuo Zheng [f], Houyuan Lu [g], Yunli Luo [h], Yuecong Li [c,d], Chunhai Li [i], Heikki Seppä [a]

[a] Department of Geosciences and Geography, University of Helsinki, Gustaf Hällströmin katu 2a, P.O. Box 64, Helsinki 00014, Finland
[b] Institute of Nihewan Archaeology Research, Hebei Normal University, East Road of Southern 2nd Ring, Shijiazhuang 050024, China
[c] College of Resources and Environment Science, Hebei Normal University, East Road of Southern 2nd Ring, Shijiazhuang 050024, China
[d] Hebei Key Laboratory of Environmental Change and Ecological Construction, Hebei Normal University, East Road of Southern 2nd Ring, Shijiazhuang 050024, China
[e] MOE Key Laboratory of Western China's Environmental System, Lanzhou University, Southern Tianshui Road, Lanzhou 730000, China
[f] Department of Earth Sciences, Sun Yat-sen University, Xingang Xi Road, Guangzhou 510275, China
[g] Key Laboratory of Cenozoic Geology and Environment, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beitucheng Western Road, Beijing 100029, China
[h] Institute of Botany, Chinese Academy of Sciences, Xiangshan Nanxincun, Beijing 100093, China
[i] State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, East Beijing Road, Nanjing 210008, China

ARTICLE INFO

Article history:
Received 22 August 2014
Available online 7 February 2015

Keywords:
Constrained ordination techniques
Boosted regression trees
Spatial distribution
Climate variables
Climate reconstructions
Modern pollen data
China

ABSTRACT

To assess the importance of climate variables for the distribution of modern pollen data in China, we present a continental-scale dataset consisting of 1374 samples. Boosted regression trees and constrained ordination techniques are employed to quantify the importance of six climate variables (annual precipitation, PANN; actual/potential evapotranspiration ratio, Alpha; mean annual temperature, TANN; mean temperature of the warmest month, MTWA; mean temperature of the coldest month, MTCO; annual sum of the growing degree days above 5°C, GDD5) for the distribution of individual pollen taxa and modern pollen assemblages. The results show that taxon-specific responses to the climate variables display a wide regional diversity and that the climate variables with low collinearity that best account for the spatial variability of modern pollen assemblages differ regionally. PANN is the most important variable in northwestern and northeastern China and the Tibetan Plateau, while MTWA and MTCO are the dominant variables in east-central and southern China. This suggests that the climate variables that can be optimally reconstructed from fossil pollen data vary in different bioclimatic regions of China. This feature is typical to many continental-scale modern pollen datasets and needs to be considered in pollen-based climate reconstructions.

Introduction

Assessing the quantitative relationships between modern pollen spectra and climate data is an essential prerequisite for pollen-based climate inferences and reconstructions (e.g., Seppä et al., 2004; Whitmore et al., 2005; Shen et al., 2006; Finsinger et al., 2007; Minckley et al., 2008; St. Jacques et al., 2008; Williams and Shuman, 2008; Salonen et al., 2012; Zhao et al., 2012; Schäbitz et al., 2013; Wen et al., 2013; Tian et al., 2014). In China, many studies have been conducted to analyze the numerical relationships among surface pollen samples and present-day climate parameters at different regional scales (e.g., Shen et al., 2006; Li et al., 2007; Luo et al., 2010; Xu et al., 2010; Lu et al., 2011; Xiao et al., 2011; Zhao et al., 2012), but there has not been a comprehensive and methodologically consistent study at both regional and continental scales.

Recently, Zheng et al. (2008) compiled a spatially extensive modern pollen dataset that spans the whole of China from the tropical and subtropical forests in the south to the temperate steppes and deserts and boreal coniferous forest in the north. This dataset consists of 1374 modern pollen spectra mostly collected from surface soil samples, because lakes and peatlands are scarce or absent in many parts of China, particularly in the steppe and desert areas. This dataset enables us to investigate the numerical pollen–climate relationships in different bioclimatic regions of China.

One essential step in the use of pollen data for quantitative climate reconstructions is to estimate the importance of climate variables affecting the spatial distribution of modern pollen data to select the most important climate variables that can optimally be reconstructed from fossil pollen data (e.g., Klemm et al., 2013; Wen et al., 2013; Salonen et al., 2014; Tian et al., 2014). In China, previous studies on this issue have been mainly restricted to regional surface pollen datasets from the

* Corresponding author at: East Road of Southern 2nd Ring, Shijiazhuang 050024, Hebei Province, China.
E-mail address: xuqinghai@mail.hebtu.edu.cn (Q. Xu).

http://dx.doi.org/10.1016/j.yqres.2014.12.002
0033-5894/

Tibetan Plateau and northwestern China. For example, Lu et al. (2011) found that annual precipitation, annual mean temperature, July mean temperature, and relative humidity are significant climatic predictors of the spatial distribution of surface pollen data in the Qinghai–Tibet region. Luo et al. (2010) investigated the pollen–climate relationships using the integrated surface pollen dataset in northwestern China and concluded that the most important climate variable is annual precipitation. However, these regional datasets are not sufficient to reflect the full-scale spatial variability of modern climate and pollen data from other regions of China, because regional climatic gradients are shorter than continental climatic gradients.

To estimate quantitatively the importance of climate variables that account for the distribution of modern ecological data at the assemblage level, the commonly used statistical techniques are constrained ordination methods that combine a multivariate ordination of species composition data with a constrained regression, such as redundancy analysis (RDA) that assumes a linear response, and canonical correspondence analysis (CCA) that assumes a unimodal response (ter Braak and Šmilauer, 2002). These techniques have been widely used to estimate quantitatively how much of the variation in modern pollen assemblages can be accounted for by individual climate variables and to test the statistical significance of climate variables (e.g., Birks, 1995, 1998; Li et al., 2007; Fréchette et al., 2008; Legendre and Birks, 2012; Schäbitz et al., 2013; Salonen et al., 2014; Tian et al., 2014).

To evaluate quantitatively the importance of climate variables that account for the distribution of modern ecological data at the species or taxonomic level, boosted regression trees (BRTs) have been recently demonstrated to be an effective approach for quantifying the influence of environmental variables for the distribution of various animal and plant species (e.g., Elith et al., 2006, 2008; De'ath, 2007; Aertsen et al., 2011; Chakona and Swartz, 2012; Kint et al., 2012; Palm et al., 2013). In comparison with other approaches, BRTs have the advantage that they are able to fit any combination of linear, sigmoidal, unimodal, bimodal, etc. response models, deal with different types of explanatory variables, model interactive effects among explanatory variables, incorporate missing values, and do not require eliminating outliers and prior transformations of data (Elith et al., 2008; Salonen et al., 2014). However, few studies have so far been conducted to utilize BRTs for assessing the importance of individual climate variables for the distribution of modern pollen data at each taxonomic level. For instance, Salonen et al. (2012) applied BRTs to the northern European surface pollen dataset and identified summer temperature as the most important climate variable for most taxa. Zhao et al. (2012) employed BRTs to analyze the relationships among the *Artemisia* and Chenopodiaceae (A/C) pollen ratio, annual mean temperature (TANN) and precipitation (PANN) in northwestern China and found that PANN is more important than TANN in explaining the A/C ratio as a moisture indicator. Li et al. (2014) applied BRTs to estimate the importance of human influence index (HII) and climate parameters for the modern pollen distribution in China and showed that climate variables are more important than HII for the distribution of most taxa.

In this study, we employ constrained ordination techniques and boosted regression trees to the continental-scale Chinese modern pollen dataset (Zheng et al., 2008), in combination with high-resolution modern climate data, to (1) estimate quantitatively the importance of individual climate variables for the distribution of modern pollen assemblages and individual pollen taxa on regional to continental scales, and to (2) identify the most important climate variables with low collinearity that can be optimally reconstructed from fossil pollen data in different bioclimatic regions of China.

Study area

The territory of China (latitude 3°52′N–53°33′N, longitude 73°40′E–135°2′30″E) embraces a large geographical area with complex topography and various vegetation zones (Fig. 1). The climate is typically governed by the well-pronounced monsoon system in East Asia, because China is strongly influenced by the two prominent seasonal high-pressure systems in the Northern Hemisphere, corresponding to the Siberian–Mongolian high pressure in winter and the Pacific high pressure in summer (Domrös and Peng, 1988). The prevailing winds in winter are northwesterly and northerly (winter monsoon) with cold and dry air masses from the inland continental areas such as Siberia and Mongolia, while in summer southeasterly winds (summer monsoon) from the Pacific Ocean bring warm and moist air masses (Domrös and Peng, 1988). The climatic characteristics differ significantly across regions. The annual mean temperature ranges from >20°C in southern China to < −20°C in the Tibetan Plateau and the annual precipitation ranges from >2400 mm in southeastern China to <100 mm in northwestern China.

Eight main vegetation zones are defined in China (e.g., Chang, 1983; Hou, 1983, 2001; Fang et al., 2002) (Fig. 1). Cold-temperate needleleaf forest (I) is located in the extreme north of China and consists chiefly of boreal conifers such as *Picea obovata*, *Abies sibirica*, *Larix gmelinii*, and *Pinus koraiensis*. Temperate mixed needleleaf and deciduous broadleaf forest (II) exists in northeastern China, where the common plant species are *Pinus koraiensis*, *Abies nephrolepis*, *Picea jezoensis*, *Quercus mongolica*, *Tilia amurensis*, and *Betula platyphylla*. Warm-temperate deciduous broadleaf forest (III) lies in east-central China with common species such as *Quercus mongolica*, *Quercus liaotungensis*, *Quercus aliena*, *Salix matsuduna*, and *Populus simorill*. Subtropical broadleaf evergreen forest (IV) occurs in southern China with *Quercus variabilis*, *Castanopsis sclerophylla*, *Castanea sequinii*, *Platycarya strobilacea*, *Lithocarpus cleistocarpus*, and *Liquidambar formosana* as the common species. Tropical monsoon forest and rainforest region (V) is located in the extreme south of China and is dominated by *Terminalia hainanensis*, *Ficus altissima*, *Vatica astrotricha*, *Altingia chinenesis*, *Dacrydium pierrei*, and *Lithocarpus fenzelianus*. Temperate steppe (VI) lies in the dry and semi-dry areas in northwestern China and is composed primarily of herbaceous species such as *Artemisia frigid*, *Artemisia desertorum*, *Stipa breviflora*, *Stipa glareosa*, *Leymus chinensis*, and *Achnatherum splendens*. *Pinus*, *Quercus*, *Betula*, *Ulmus*, *Salix*, *Corylus* and *Populus* are the major arboreal plants on hills. Temperate desert (VII) is distributed in inland continental China with sparse vegetation that is characterized by herbaceous species such as *Salsola abrotanoides*, *Sympegma regelii*, *Haloxylon ammodendron*, *Ceratoides latens*, *Ajania fruticulosa*, and *Nitraria tangutorum*. High-cold vegetation (VIII) covers the Tibetan Plateau, where the common species are *Carex moorcroftii*, *Kobresia royleana*, *Artemisia wellbyi*, *Artemisia gmelinii*, *Stipa purpurea*, and *Ceratoides compacta*.

Materials and methods

Modern pollen and climate data

The Chinese modern pollen dataset is a subset of the East Asian pollen database (Zheng et al., 2008, 2014) and comprises 1374 surface samples, of which 92% were collected from surface soils, 1% from peat surfaces, 2% from moss polsters, and 5% by using the dust flux approach (Cour, 1974; Cour et al., 1999) (Fig. 1). It spans all major vegetation and climatic zones of continental China. A comprehensive description of the site information, deposition types, data contributors, laboratory processing and pollen characteristics has been presented by Zheng et al. (2008, 2014). We calculated pollen percentage values based on the number of terrestrial pollen grains. Our dataset contains pollen types with extremely low occurrences and many zero values. To reduce the noise in the dataset, 75 common pollen taxa that occur at least in 5% of the total pollen samples were chosen for all further analyses. The minimum number of occurrences of these pollen taxa is 69 samples.

The modern climate variables that we selected are annual precipitation (PANN), actual/potential evapotranspiration ratio (Alpha, α), annual mean temperature (TANN), mean temperature of the warmest month (MTWA), mean temperature of the coldest month (MTCO), and annual