Quaternary Science Reviews 178 (2017) 37–53

Contents lists available at ScienceDirect

Quaternary Science Reviews

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

Impacts of the spatial extent of pollen-climate calibration-set on the absolute values, range and trends of reconstructed Holocene precipitation

Xianyong Cao [a], Fang Tian [a, b, *], Richard J. Telford [c], Jian Ni [d], Qinghai Xu [e], Fahu Chen [f], Xingqi Liu [g], Martina Stebich [h], Yan Zhao [i, j], Ulrike Herzschuh [a, k, l, **]

[a] *Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Research Unit Potsdam, Telegrafenberg A43, Potsdam, 14473, Germany*
[b] *Max Planck Institute for Meteorology, KlimaCampus, Bundesstrasse 53, Hamburg 20146, Germany*
[c] *EECRG, Department of Biology, University of Bergen, Bjerknes Centre for Climate Research, PO Box 7803, N-5020 Bergen, Norway*
[d] *College of Chemistry and Life Sciences, Zhejiang Normal University, Yingbin Avenue 688, 321004 Jinhua, China*
[e] *College of Resources and Environment Sciences, Hebei Normal University, Shijiazhuang 050024, China*
[f] *Institute of Tibetan Plateau Research, Chinese Academy of Science, Beijing 100101, China*
[g] *Beijing Key Laboratory of Resource Environment and GIS, College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China*
[h] *Senckenberg Research Station for Quaternary Palaeontology, Am Jakobskirchhof 4, D-99423 Weimar, Germany*
[i] *Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Datun Road 11, Beijing 100101, China*
[j] *University of Chinese Academy of Sciences, Yuquan Road 19, Beijing 100049, China*
[k] *Institute of Earth and Environmental Sciences, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany*
[l] *Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany*

ARTICLE INFO

Article history:
Received 7 July 2017
Received in revised form
20 October 2017
Accepted 21 October 2017

Keywords:
Analogue quality
Statistical significance
Cross-validation
Holocene
Climate reconstruction
WA-PLS
MAT

ABSTRACT

Pollen-based quantitative reconstructions of past climate variables is a standard palaeoclimatic approach. Despite knowing that the spatial extent of the calibration-set affects the reconstruction result, guidance is lacking as to how to determine a suitable spatial extent of the pollen-climate calibration-set. In this study, past mean annual precipitation (P_{ann}) during the Holocene (since 11.5 cal ka BP) is reconstructed repeatedly for pollen records from Qinghai Lake (36.7°N, 100.5°E; north-east Tibetan Plateau), Gonghai Lake (38.9°N, 112.2°E; north China) and Sihailongwan Lake (42.3°N, 126.6°E; north-east China) using calibration-sets of varying spatial extents extracted from the modern pollen dataset of China and Mongolia (2559 sampling sites and 168 pollen taxa in total). Results indicate that the spatial extent of the calibration-set has a strong impact on model performance, analogue quality and reconstruction diagnostics (absolute value, range, trend, optimum). Generally, these effects are stronger with the modern analogue technique (MAT) than with weighted averaging partial least squares (WA-PLS). With respect to fossil spectra from northern China, the spatial extent of calibration-sets should be restricted to radii between ca. 1000 and 1500 km because small-scale calibration-sets (<800 km radius) will likely fail to include enough spatial variation in the modern pollen assemblages to reflect the temporal range shifts during the Holocene, while too broad a scale calibration-set (>1500 km radius) will include taxa with very different pollen-climate relationships.

1. Introduction

Quantitative climate reconstructions using modern pollen-climate relationships have been widely applied to investigate past climate change. This approach is based on various assumptions. Among others, it is assumed that pollen taxa recorded in the calibration-set have similar ecological requirements as those in the fossil spectra (Birks et al., 1990, 2010; Juggins and Birks, 2012). As

---

* Corresponding author. Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Research Unit Potsdam, Telegrafenberg A43, Potsdam, 14473, Germany.

** Corresponding author. Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Research Unit Potsdam, Telegrafenberg A43, Potsdam, 14473, Germany.

E-mail addresses: fang.tian@awi.de (F. Tian), Ulrike.Herzschuh@awi.de (U. Herzschuh).

https://doi.org/10.1016/j.quascirev.2017.10.030

pollen are commonly only identified to genus level (most tree and shrub taxa) or family level (most of the herbaceous taxa), the selection of the spatial extent of the calibration-set defines which plant species (and sub-species) determines the taxon-environment relationship of a certain pollen taxon (Williams and Shuman, 2008). For example, it has been concluded that the obvious regional differences in pollen taxa-climate relationships between the eastern Tibetan Plateau and central-northern China of Cyperaceae, *Picea* and *Pinus* originate from the specific plant species ensemble covered by a pollen taxon (i.e. *Kobresia* vs. *Carex*; *Picea asperata* vs. *Picea meyeri*; *Pinus densata* vs. *Pinus tabuliformis*, respectively) (Tian et al., 2017). Likewise, *Quercus* is restricted to the wet end of the moisture gradient in eastern North America, while it prefers arid/semi-arid climate conditions in the western part (Williams and Shuman, 2008). Furthermore, the pollen-climate relationship may also be dependent on the length of the climatic gradient. For example, Poaceae has a significant relationship with P_{ann} and the mean temperature of the warmest month (increasing and decreasing, respectively) in a regional-scale modern pollen data-set from Mongolia revealed by Huisman-Olff-Fresco modelling (Tian et al., 2014), while it shows no response to these two climatic variables at a continental scale in eastern Asia (including Mongolia; Cao et al., 2014). Accordingly, the spatial extent of a calibration-set needs to be thoroughly considered. On the one hand, a calibration-set with too small a spatial extent risks that the temporal changes in the fossil pollen spectra cannot be covered by the spatial variation in the modern pollen assemblages (without high quality modern analogues for fossil pollen samples), but on the other hand, too large a spatial extent of the calibration-set will weaken or even distort taxon-climate relationships because various plant species that differ in their climate envelopes are represented by a single pollen taxon.

Weighted averaging partial least squares (WA-PLS) and the modern analogue technique (MAT) are the most common methods used for quantitative climate reconstruction based on pollen data (Juggins and Birks, 2012). The prediction strength of these methods is usually assessed by the coefficient of determination between observed and predicted environmental values (r^2) and the root mean square error of prediction (RMSEP) as derived by cross-validation. The spatial extent of a calibration-set may affect these parameters, particularly r^2 (Birks, 1998), and a large spatial extent of the calibration-set will raise RMSEP since the pollen assemblage heterogeneity increases (Kucera et al., 2005; Cao et al., 2014). For weighted-averaging based methods, the inevitable overestimate and underestimate at environmental gradient ends ("edge effects") have lower impact for the reconstruction when the target environmental gradient is long and the spatial extent of the calibration-set is large (Birks, 1998; Birks and Seppä, 2004). The basic requirement of analogue-based methods is that similar biological assemblages are deposited under similar environmental conditions (Juggins and Birks, 2012). However, the risk for multiple analogues (i.e. similar modern biological assemblages were deposited under different environmental conditions) increases with the spatial extent of the calibration-set, raising the prediction error (Birks et al., 2010). Hence, it is recommended to tailor the calibration-set with respect to spatial extent or environmental gradient for the specific reconstruction (e.g. Birks and Seppä, 2004; Birks et al., 2010; Juggins and Birks, 2012; Chen et al., 2015).

Besides the cross-validation indices (r^2 and RMSEP), the tailored calibration-set should also be assessed by the analogue quality and pass the statistical significance test for the target reconstruction. It can be assumed that quantitative climate reconstructions for fossil pollen assemblages that have good analogues in the calibration-set are more reliable than those that have no close analogues (Birks et al., 1990). It also needs to be verified whether the reconstructed climate, inferred by the application of a modern pollen-climate model of a certain spatial extent, best explains a sufficient amount of the variance in the fossil data by assessment with a statistical significance test (Telford and Birks, 2011).

Eastern continental Asia is a suitable region to investigate methodological aspects of pollen-based quantitative climate reconstruction. First, it is a hotspot for the establishment and application of pollen-climate relationships aiming to reconstruct climate as a contribution to the highly debated topic of the spatial-temporal variability of the East Asian Summer Monsoon (e.g. Herzschuh, 2006; Wang et al., 2010; Liu et al., 2015). In the past decade, numerous pollen-environment calibration-sets at both a regional scale (e.g. Shen et al., 2006; Herzschuh et al., 2010; Xu et al., 2010; Tian et al., 2014) and a sub-continental scale have been applied (e.g. Stebich et al., 2015; Li et al., 2016, 2017a, 2017b). Second, eastern continental Asia is characterised by complex climate, topography and vegetation (Hou, 1983; Tao and Chen, 1987), which harbour long climate gradients in biogeographically different regions, such as north-central China and the eastern Tibetan Plateau. Hence, pollen-based reconstructions from eastern continental Asia should be particularly sensitive to the selection of the spatial extent of the calibration-set.

A sub-continental scale modern pollen-climate dataset from eastern continental Asia (including China and Mongolia) has been established, and P_{ann} has been confirmed as the most important climatic determinant of pollen distribution (Cao et al., 2014), which has also been revealed by regional pollen-climate calibration-sets (Shen et al., 2006; Herzschuh et al., 2010; Xu et al., 2010; Tian et al., 2014). In this study, we employ the modern pollen-climate dataset from eastern continental Asia (Cao et al., 2014) to reconstruct P_{ann} for three high-resolution Holocene pollen records from northern China (Qinghai Lake, Liu et al., 2002; Gonghai Lake, Chen et al., 2015; and Sihailongwan Lake, Stebich et al., 2015) repeatedly, using calibration-sets of varying spatial extent. In particular, we investigate the impact of the spatial extent of the calibration-set on the pollen-climate model performance, analogue quality and reconstruction diagnostics (including absolute values, range, trend, and optimum).

2. Data and methods

2.1. Pollen and climate datasets

The modern pollen dataset from China and Mongolia comprises 2559 sampling sites (including soil-surface, lake sediment-surface, dust and moss samples; Fig. 1; Cao et al., 2014) and 168 major pollen taxa (those present in at least 3 samples and with $\geq$3% in at least one sample; Cao et al., 2014). Pollen names have been taxonomically harmonised. Canonical correspondence analysis suggests that P_{ann} has the strongest influence on pollen composition across the entire dataset (Cao et al., 2014). Around 80% of P_{ann} in monsoonal areas comprises summer rainfall (as revealed by the climate data of the 2559 pollen sites), and thus, given the reservoir function of soil, P_{ann} should be comprehensive in representing the moisture situation in northern China. Hence, we selected P_{ann} as the target variable for our quantitative reconstructions, which means the reconstructions are directly comparable with previous pollen-climate studies.

To explore the impact of the spatial extent of the calibration-set on reconstruction, three high-quality target fossil pollen records were selected that have previously been used for climate reconstructions. They originate from different regions in northern China within the area marginally affected by the East Asian