Using modern and fossil pollen data for climate and human influence reconstructions in China

JIANYONG LI
Assessing the numerical relationships among pollen spectra, human influence and climate is a critical prerequisite for obtaining reliable and robust pollen-based quantitative reconstructions for past climatic variations and human-induced environmental changes. Here, modern pollen data from small river catchments in human-influenced mountainous ecosystems of northern China are collected to investigate the pollen/vegetation/land-use relationships. The effects of human influence index (HII) in comparison with key climate variables on influencing modern pollen distribution are evaluated using a continental-scale dataset. The importance of different climate variables accounting for modern pollen distribution is quantified for identifying the determinant variables that can be reconstructed in different bioclimatic regions of China. Furthermore, pollen-based calibration models for HII and climate parameters are developed and applied to a set of fossil pollen data for quantitative human influence and climate reconstructions. In addition, the bias caused by long-term and intensive human influence on vegetation in quantitative pollen-based palaeoclimatic reconstructions is examined, and pollen-based climate reconstructions are used to validate the reliability of climate model simulations.

The results reveal that pollen assemblages from surface sediment samples in the Tuoliang and Qipanshan catchments from northern China are dominated by Artemisia, Chenopodiaceae, Pinus and Selaginella sinensis. Pollen types indicative of human activities such as Cerealia-type are common but not abundant. Artemisia and Chenopodiaceae percentages decline with decreasing altitude, whereas Pinus and S. sinensis percentages increase, suggesting that saccate Pinus pollen and S. sinensis spores can be transported further than non-saccate pollen types and that pollen sorting is taking place over water transportation. There is a positive linear relationship between percentages of Cerealia-type pollen and proportions of farmland, whereas correlation between percentages of trees, shrubs and herbs pollen and proportions of woodland, scrubland and grassland respectively is poor. This can be attributed to the differences in pollen productivity, deposition and dispersal. These findings may provide the basis for a modeling approach to reconstruct past land-cover changes at small catchment scale in northern China.

The effects of HII on accounting for modern pollen distribution are smaller than annual precipitation (PANN) or annual average temperature (TANN) in most regions of China, with the exception of east-central China that has a significant pollen–HII relationship. The importance of six climate variables (PANN; TANN; actual/potential evapotranspiration ratio, Alpha; 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 geographical distribution of modern pollen data differs among regions in China. Hydrological variables are more important than temperature-related variables in north-western and -eastern China and the Tibetan Plateau,
whilst temperature-related variables are dominant in east-central and southern China. This suggests that the critical climate variables which can be optimally reconstructed from pollen data vary regionally in China. These regional differences are common to other large-scale surface pollen datasets and need to be taken into considerations in pollen-based climate reconstructions.

The calibration models for HII, PANN and TANN show good statistical performances and are applied to a fossil pollen sequence from Lake Tianchi in east-central China. The inferred HII increases abruptly at 1100 cal. yr BP and remains high until present-day. Correspondingly, the reconstructed PANN shows a sudden decrease since 1100 cal. yr BP, whereas TANN displays an increase. However, other independent palaeoclimatic records mostly contradict the reconstructed climatic trends for the last 1100 years, suggesting that the Tianchi climate reconstructions for this period are biased by the intensification of human influence. Such a bias probably occurs in other palaeoclimatic reconstructions based on pollen data from other regions where human impact on vegetation is also intensive and long-lasting.

The calibration models for the summer monsoon precipitation ($P_{jja}$) also exhibit good predictive performances and are applied to a set of fossil pollen data from monsoonal China. The Holocene $P_{jja}$ reconstructions from northern China are stacked to one record, showing that $P_{jja}$ increases since 9500 cal. yr BP, reaches the maximum at 7000–4000 cal. yr BP, and falls down to present-day. The reconstructions are compared with $P_{jja}$ simulations from the Kiel Climate Model (KCM). The model output and reconstructions are different for the early-Holocene for which the model suggests higher $P_{jja}$. The simulated $P_{jja}$ changes for the Holocene summer monsoon maximum (HSMM) are smaller than the reconstructed changes. The increasing or decreasing $P_{jja}$ trends before or after the HSMM are more pronounced in the reconstructions than in the model results. Other palaeoclimatic data show substantial Holocene monsoon precipitation changes. This suggests that the KCM underestimates the magnitude of the summer monsoon precipitation changes.

The results obtained in this work highlight the necessity for critically assessing the modern pollen/human-influence/climate relationships at various spatial scales and can be used to improve the correctness and robustness of pollen-based climate and human influence reconstructions in China and other regions of the world with large-scale surface pollen datasets. The pollen-based calibration models and their forthcoming applications can be employed to further quantitatively explore the past environmental changes and climatic dynamics and to validate the accuracy of climate model output.
Acknowledgements

I would like to express my sincere gratitude to all of those who have supported me over the last three years for completing my PhD study. First of all, I am deeply indebted to my supervisor, Prof. Heikki Seppä, for giving me the exquisite opportunity to study at the University of Helsinki. He has offered me enormous trust, guidance and encouragement during this work. I am also grateful to my former supervisor, Prof. Qinghai Xu from Hebei Normal University, China, for his constant support throughout the three years. Special thanks go to all co-authors of my papers for their fruitful contributions, suggestions and collaborations.

I sincerely thank my thesis reviewers, Prof. Rachid Cheddadi and Dr. Anne Bjune, for their encouraging comments on my thesis. I also wish to acknowledge all researchers and staff in the Department of Geosciences and Geography for creating a friendly and supportive working atmosphere. I sincerely thank Dr. Sakari Salonen for his useful discussions about numerical analyses, Niina Kuosmanen for her great assistance in the teaching, my officemates (Marttiina Rantala and Ferhat Kaya) for their joyful moments in the office, and all my Chinese friends (Cheng Zhu, Xia Yu, Yuan Shang, Yurui Zhang, Bin Wang and all the others) in Helsinki for their company and help in the last three years. My parents and brothers are also deeply thanked for their love and support.

Finally, I appreciate the financial support from the China Scholarship Council and several research projects from the Academy of Finland and the National Natural Science Foundation of China.
Contents

Abstract ...........................................................................................................................3
Acknowledgements ......................................................................................................5
List of original publications .........................................................................................7
Abbreviations ...............................................................................................................9
List of tables and figures ................................................................................................10

1 Introduction .............................................................................................................11
2 Materials and methods .............................................................................................14
   2.1 Study area .....................................................................................................14
   2.2 Data collection .............................................................................................15
   2.3 Numerical analyses ......................................................................................17
3 Summary of original publications .............................................................................18
   3.1 Paper I .......................................................................................................18
   3.2 Paper II ......................................................................................................19
   3.3 Paper III ......................................................................................................20
   3.4 Paper IV .......................................................................................................21
4 Discussion ..............................................................................................................22
   4.1 Potentials and limitations of quantitative pollen-based climate
       and human influence reconstructions in China ...............................................22
   4.2 Future prospects .........................................................................................25
5 Conclusions ............................................................................................................26

References ..................................................................................................................27

Publications I-IV
List of original publications

This thesis is based on the following publications:


The publications are referred to in the text by their roman numerals.
Author’s contribution

I The study was planned by J.Y. Li and Q.H. Xu. J.Y. Li conducted the field work with help from L.Y. Hun and M.Y. Li. Dr. J.S. Liu extracted and interpreted the MODIS data. J.Y. Li performed the laboratory work, the pollen analysis, and the numerical analyses. J.Y. Li prepared the manuscript with comments and contributions from Q.H. Xu, M.J. Gaillard, H. Seppä and Y.C. Li.

II The study was planned by J.Y. Li and H. Seppä. Q.H. Xu, Z. Zheng, H.Y. Lu, Y.L. Luo, Y.C. Li and C.H. Li contributed the modern pollen data. Y. Zhao contributed the fossil pollen data. J.Y. Li performed the statistical analyses and prepared the manuscript with comments and contributions from the other authors.

III The study was planned by J.Y. Li and H. Seppä. Q.H. Xu, Z. Zheng, H.Y. Lu, Y.L. Luo, Y.C. Li and C.H. Li contributed the modern pollen data. J.Y. Li performed the numerical analyses and prepared the manuscript with comments and contributions from the other authors.

IV The study was planned by J.Y. Li and H. Seppä. Q.H. Xu, Z. Zheng, H.Y. Lu, Y.L. Luo, Y.C. Li and C.H. Li contributed the modern pollen data. J. Ni contributed the fossil pollen data. L.Y. Jin and X.J. Zhang provided the KCM model output. J.Y. Li performed the WA-PLS reconstructions. L. Ilvonen and L. Holmström conducted the Bayesian reconstructions. J.Y. Li and L. Ilvonen prepared the manuscript with comments and contributions from the other authors.
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMS</td>
<td>accelerator mass spectrometry</td>
</tr>
<tr>
<td>BRT</td>
<td>boosted regression trees</td>
</tr>
<tr>
<td>CCA</td>
<td>canonical correspondence analysis</td>
</tr>
<tr>
<td>CRU</td>
<td>Climate Research Unit</td>
</tr>
<tr>
<td>DCA</td>
<td>detrended correspondence analysis</td>
</tr>
<tr>
<td>ECC</td>
<td>east-central China</td>
</tr>
<tr>
<td>GDD5</td>
<td>annual sum of the growing degree days above 5°C</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloric acid</td>
</tr>
<tr>
<td>HF</td>
<td>hydrofluoric acid</td>
</tr>
<tr>
<td>HII</td>
<td>human influence index</td>
</tr>
<tr>
<td>HSMM</td>
<td>Holocene summer monsoon maximum</td>
</tr>
<tr>
<td>KCM</td>
<td>Kiel Climate Model</td>
</tr>
<tr>
<td>MODIS</td>
<td>moderate resolution imaging spectroradiometer</td>
</tr>
<tr>
<td>MTCO</td>
<td>mean temperature of the coldest month</td>
</tr>
<tr>
<td>MTWA</td>
<td>mean temperature of the warmest month</td>
</tr>
<tr>
<td>NEC</td>
<td>northeastern China</td>
</tr>
<tr>
<td>NWC</td>
<td>northwestern China</td>
</tr>
<tr>
<td>PANN</td>
<td>annual average precipitation</td>
</tr>
<tr>
<td>$P_{\text{jja}}$</td>
<td>summer (June, July and August) precipitation</td>
</tr>
<tr>
<td>$r^2$</td>
<td>coefficient of determination</td>
</tr>
<tr>
<td>RDA</td>
<td>redundancy analysis</td>
</tr>
<tr>
<td>RMSEP</td>
<td>root-mean-square error of prediction</td>
</tr>
<tr>
<td>SC</td>
<td>southern China</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>TANN</td>
<td>annual average temperature</td>
</tr>
<tr>
<td>$T_{\text{djf}}$</td>
<td>winter temperature</td>
</tr>
<tr>
<td>$T_{\text{jja}}$</td>
<td>summer temperature</td>
</tr>
<tr>
<td>TP</td>
<td>Tibetan Plateau</td>
</tr>
<tr>
<td>VIF</td>
<td>variance inflation factors</td>
</tr>
</tbody>
</table>
List of tables and figures

Table 1  Details of the fossil pollen datasets used in this work, page 16
Fig. 1  Locations of modern and fossil pollen sets and two catchments, and distributions of PANN and TANN in China, page 14
Fig. 2  Distribution of human influence index (HII) in China, page 16
Fig. 3  Correlation between pollen percentages and land use proportions in the two catchments, page 18
Fig. 4  Comparison of HII, PANN and TANN reconstructions from the Tianchi pollen record with other related records, page 20
Fig. 5  Percentages of variation (%) in pollen data explained by each climate variable alone in RDA and CCA for WC and all subsets, page 21
Fig. 6  Comparison of pollen-based stacked $P_{jja}$ record for northern China with KCM-based $P_{jja}$ simulations and other related climate records, page 22
1. Introduction

Pollen data have been commonly utilized as a tool for quantitatively reconstructing past climate, vegetation and land-use/land-cover changes in different regions of the world (e.g., Huntley and Prentice, 1988; Prentice et al., 1996; Sugita et al., 1999; Broström et al., 2004; Nielsen, 2004; Jiang et al., 2006; Brewer et al., 2008; Poska et al., 2008; Seppä et al., 2009; Birks et al., 2010; Xu et al., 2010a; Nielsen et al., 2012; Cui et al., 2014). The feasibility and reliability of such reconstructions rely on our ability to understand the numerical relationships among modern pollen data, vegetation distribution, land-use/land-cover types and key climate parameters. Therefore it is essential to investigate these relationships along major environmental gradients of climate and land-use at local, regional and continental scales.

For the pollen-based vegetation/land-cover/land-use reconstructions, improved understandings of signal of human influence in pollen data are needed. Human influence is increasingly pervasive and has so far altered more than 75% of the ice-free landscapes on the earth (Ellis and Ramankutty, 2008), and large-scale human-induced landscape transformations have taken place not only in the present time but also over the last several millennia, particularly after the rapid development of Neolithic agriculture (e.g., Ruddiman, 2003; Kirch, 2005; Goudie, 2006; Dearing, 2006; Ellis et al., 2013). As a result, most ecosystems within the terrestrial biosphere have been affected by various human activities and can be regarded as ‘anthropogenic biomes’ (Alessa and Chapin III, 2008; Ellis and Ramankutty, 2008; Ellis et al., 2010). Therefore, understanding the relations between modern pollen data and land-use/land-cover distributions plays an important role in attaining reliable pollen-based reconstructions for past human-influenced landscape dynamics.

In Europe studies focusing on the relationships between modern pollen spectra and different vegetation types affected by human activities have been widely carried out. For example, in the late 1930s, Firbas (1937) was one of the earliest pioneers who utilized pollen data to investigate human influence on vegetation. Since the 1980s, several important monographs have been published, such as “The palynology of archaeological sites” by Dimbleby (1985), “Anthropogenic indicators in pollen diagrams” by Behre (1986), and “The cultural landscape–past, present and future” by Birks et al. (1988). Over the last decades, more attention has been paid to detect the signal of human influence on vegetation using modern pollen data from different cultural landscapes, such as farmland, pasture, rangeland and wasteland (e.g., Hicks, 1993; Gaillard et al., 1994; Carpelan and Hicks, 1995; Hjelle, 1997; Court-Picon, 2005; Mazier et al., 2008; Brun, 2011). By contrast, in China only few studies have so far been conducted for exploring the relationships between modern pollen spectra and human-affected vegetation. For example, Liu et al. (2006) and Zhang et al. (2010) analyzed surface pollen assemblages from soils under different vegetation types influenced by human activities in northern China, and revealed that Chenopodiaceae pollen percentages can reflect the degree of human-caused vegetation degradation. Li et al. (2008) recognized many pollen taxa that can be correlated with human activities such as Apiaceae, Amaryllidaceae, Brassicaceae, Cannabaceae, Convolvulaceae, Poaceae, Solanaceae, Fabaceae and Rosaceae. Schlütz and Lehmkühl (2009) identified several pollen types such as Anemone-type, Trollius and Cichoriodae, which are indicative of various grazing activities in the Tibetan Plateau. However, there has been a lack of studies focusing on the relationships among land-use/land-cover, vegetation and
pollen assemblages based on surface sediment samples from irrigation pools that are one of the best potential pollen archives recording modern vegetation and land-use in small river catchments from human-influenced mountainous ecosystems, which acts as an important step for the quantitative reconstructions of human-induced land cover changes at a local spatial scale.

For the pollen-based climate reconstructions, there are still some ongoing challenges that can affect the reliability and accuracy of the reconstructions. One important issue is that long-lasting human impact on vegetation exerts confounding effects on palaeoclimatic reconstructions based on pollen data, which can potentially lead to unreliable and biased results (e.g., Birks and Seppä, 2004; St. Jacques et al., 2008; Xu et al., 2010b). Large-scale surface pollen datasets are usually prepared for developing quantitative pollen-based calibration models, and mostly collected for representing natural vegetation and avoiding strong human disturbance (e.g., Whitmore et al., 2005; Bjune et al., 2010; Birks et al., 2010; Herzschuh et al., 2010). However, as mentioned above, human impact on ecosystems can be found in most regions of the world and has existed over the last several thousand years. Therefore it is difficult to eliminate the effect of human influence. For quantitatively reflecting the degree of human influence, a global-scale database of human influence index (HII) has been constructed by synthesizing the weight sum of several data networks such as human population distribution, land-cover/land-use, infrastructures and accessibility (Sanderson et al., 2002). It has been employed in ecological studies for modeling the spatial distribution of vegetation types (Greve et al., 2011) and animal species (e.g., Hu and Jiang, 2011; Fløjgaard et al., 2011). For palaeoecology, the availability of the HII database enables us to assess numerically the effects of HII on influencing surface pollen distribution and to establish inference models for HII which could be utilized to quantitatively reconstruct past human influence changes. China has the largest human population in the world and a long agriculture history that can be traced back to the early-Holocene (e.g., Ruddiman et al., 2008; Lu et al., 2009; Ellis et al., 2010). Human impact on vegetation in China has intensified since the mid-Holocene and has been especially intensive in North China where forest coverage has decreased by 50% after the mid-Holocene (e.g., Ren, 2000, 2007). Therefore it seems possible that HII is a significant factor affecting both modern and fossil pollen data in China and can thus cause bias in the pollen-based palaeoclimatic reconstructions.

Another important issue in using pollen data for quantitative climate reconstructions is to evaluate the importance of climate variables affecting the geographical distribution of surface pollen data to identify the dominant variables that can be reconstructed from fossil pollen data (e.g., Finsinger et al., 2007; Herzschuh et al., 2010; Birks et al., 2010; Salonen et al., 2012a, 2014). In China, many studies have been performed to assess the importance of climate variables that account for the distribution of surface pollen data at different regional scales (e.g., Li et al., 2007; Luo et al., 2010; Herzschuh et al., 2010; Lu et al., 2011; Zhao et al., 2012a). For example, Lu et al. (2011) identified annual precipitation, annual average temperature, July average temperature and relative humidity as the significant climate variables determining the pollen distribution in the Qinghai–Tibet Plateau. Luo et al. (2010) recognized annual precipitation as the most important climate variable in northwestern China. However, there has not been a comprehensive and methodologically consistent study for estimating quantitatively the importance of individual climate variables for modern pollen data at both regional and continental scales in China. Recently, Zheng et al. (2008) has com-
piled a continental-scale surface pollen dataset spanning all major vegetation and climatic zones of China, making it possible to explore these issues in the whole of China.

In addition, new quantitative climate reconstructions are in demand in China for exploring the long-term changes of the Asian monsoon climate and for validating the output of climate models. The summer monsoon precipitation plays a vital role in supporting agricultural production and natural ecosystem of China. Understanding the Holocene variability of the summer monsoon precipitation is important for predicting the future monsoon processes. In southern China the variations of the summer monsoon precipitation during the Holocene have been mostly reconstructed from speleothem δ18O records (e.g., Wang et al., 2001, 2005; Dykoski et al., 2005; Cai et al., 2010; Jiang et al., 2012), showing that the precipitation increased since the beginning of the Holocene, obtained the maximum level in the early-Holocene, and declined to present-day. This general pattern is different in northern China where palaeoclimatic records show that the summer monsoon precipitation attained its highest level in the middle-Holocene (e.g., Peng et al., 2005; Feng et al., 2006; Yang et al., 2011). However, these proxy studies are mostly qualitative. Pollen data have been employed for quantitatively reconstructing the Holocene climate changes in China (e.g., Jiang et al., 2006; Xu et al., 2010a), mostly focusing on the variables such as annual average temperature, annual average precipitation and July average temperature. For numerically investigating the magnitude of changes of the Holocene summer monsoon precipitation, it is necessary to obtain quantitative reconstructions from pollen data. Besides proxy data, the coupled climate models can produce numerical palaeoclimatic data and have been used for exploring the variability of the Holocene summer monsoon precipitation (e.g., Liu et al., 2003; Chen et al., 2010; Dallmeyer et al., 2013; Jin et al., 2014). Comparisons of proxy-based reconstructions and model simulations are a useful method for assessing the reliability of the model results and for identifying the main forcing mechanism of past climate changes (e.g., Masson et al., 1999; Braconnot et al., 2007; Renssen et al., 2009; Mauri et al., 2014).

In this work, modern pollen samples are compiled from different regions of China, relevant statistical approaches are applied to analyze the relationships among pollen, land-cover/land-use, human influence and climate on local, regional and continental scales, and quantitative reconstructions for past climate and human influence changes based on a set of fossil pollen data are performed. The objectives are (1) to explore the relationships among modern pollen data, vegetation and land-use based on surface sediment samples from irrigation pools in hydrologically-defined river catchments from the Taihang Mountains in northern China for understanding how surface pollen data from the pool sediments reflect the source vegetation and human land-use in the catchments (Paper I); (2) to evaluate the effects of HII on accounting for surface pollen distribution compared to key climate variables over continental China and to assess the bias caused by human impact on pollen-based quantitative palaeoclimatic reconstructions by comparing the climate reconstructions with a numerical HII reconstruction and other palaeoclimatic records (Paper II); (3) to quantify the importance of individual climate variables for the distribution of modern pollen data at both assemblage and taxonomic levels for recognizing the most important climate variables with low collinearity that can be optimally reconstructed for different bioclimatic regions of China (Paper III); and (4) to develop pollen-based transfer functions for the East Asian summer monsoon precipitation and apply them to fossil pollen datasets
for quantitative Holocene monsoon precipitation reconstructions compared with model simulations (Paper IV).

2. Materials and methods

2.1. Study area

The study area is located in China that extends from 3°52′N to 53°33′N and from 73°40′E to 135°2′30″E and includes eight major vegetation units: cold-temperate needleleaf forest and temperate mixed needleleaf and deciduous broadleaf forest in northeastern China (NEC); warm-temperate deciduous broadleaf forest in east-central China (ECC); subtropical broadleaf evergreen forest and tropical monsoonal rainforest in southern China (SC); temperate steppe and desert in northwestern China (NWC); and alpine vegetation on the Tibetan Plateau (TP) (Fig. 1; Hou, 2001). Most regions of China are dominated by the Asian monsoon circulation that is characterized by a pronounced seasonal variability in prevailing winds (Domrös and Peng, 1988). In wintertime the prevailing winds are northerly and north-westerly (winter monsoon) that bring dry, cold air masses from Mongolia and Siberia, whereas in summertime south-westerly and south-easterly winds (summer monsoon) carry moist, warm air masses from the Indian and Pacific oceans (Domrös and Peng, 1988). The extreme northwestern part of China is controlled by the middle-latitude westerlies. The climatic features differ largely between regions, with annual average temperature varying from >20 °C in SC to < -20 °C in TP and annual average precipitation ranging from >2400 mm in SC to <100 mm in NWC (Fig. 1).
2.2. Data collection

A set of 62 surface sediment samples in Paper I were collected from irrigation pools within two small river catchments, Tuoliang catchment with 26 samples from 13 pools and Qipanshan catchment with 36 samples from 10 pools, situated in the Taihang Mountains with a temperate deciduous broad-leaved forest in ECC (Fig. 1a). Each sample has an average depth of 1-2 cm measured from the sediment surface for representing the recent years of pollen sedimentation. To distinguish pollen of planted cereals from that of wild grasses, Poaceae pollen grains were categorized into two size groups (<36 and ≥36 μm), with the assumption that Poaceae pollen grains with the longest axis ≥ 36 μm may originate from cereals (Lan and Xu, 1996). All surface pollen samples were treated following the standard HCl–NaOH–HF procedure (Fægri et al., 1989).

A total of 1374 surface pollen samples spanning all major vegetation and climatic zones of China used in Papers II, III and IV were originally compiled by Zheng et al. (2008) (Fig. 1a). Most pollen samples in this dataset were obtained from surface soils, because lakes and peats are scarce or absent in many parts of China. A detailed description including the site information, deposition types, data contributors, laboratory processing, data quality control, taxonomic standardization, and pollen characteristics has been provided by Zheng et al. (2008, 2014). All pollen percentages (Papers I–IV) were computed based on the number of all terrestrial pollen grains. Five previously analyzed fossil pollen records from four lakes (Tianchi in ECC, Daihai in NWC, Qinghai in TP and Chaohu in SC) and one peatland (Qasq in NWC) were used in this work (Papers II and IV)(Fig. 1a). Pollen data for Qinghai, Qasq and Chaohu were acquired from the eastern Asian late-Quaternary pollen dataset compiled by Cao et al. (2013), for Tianchi from Zhao et al. (2010) and for Daihai from Xiao et al. (2004) and Xu et al. (2010a). The chronology of these fossil records was achieved by AMS radiocarbon dating. The total number of AMS 14C dates per record varies between 4 and 19 (Table 1). All ages were calibrated to calendar years according to the IntCal09 calibration dataset (Reimer et al., 2009). The age-depth model for Tianchi was established by utilizing a second-order polynomial function (Zhao et al., 2010). The age-depth models for other records were reconstructed by using the Clam package (Blaauw, 2010) in R (R Development Core Team, 2012).

Modern land-use and vegetation-cover data in Paper I were identified and extracted from the MODIS remote sensing images for the Taihang Mountains with a 30 m × 30 m spatial resolution. Modern climate data used in Papers II–IV were taken from the WorldClim database with a 30-arc-second spatial resolution (Hijmans et al., 2005) and the Climate Research Unit (CRU) database with a 10-arc-second spatial resolution (New et al., 2002). Seven climate variables were selected and calculated using GIS-based techniques, including annual average precipitation (PANN), actual and potential evapotrans-
piration ratio (Alpha), annual average 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), and summer (Jun, July and August) average precipitation ($P_{ja}$). The temperature data were corrected using a lapse-rate of 0.6 °C per 100 meters for accounting for the elevation discrepancies between the observed and the WorldClim-based data. Modern human influence index (HII) data for the 1374 samples in Paper II were obtained from the HII database with a 30-arc-second spatial resolution (Fig. 2; Sanderson et al., 2002; WCS/CIESIN, 2005). This database was established by combining many indices that quantitatively represent the degree of human impact, including human population-density, infrastructures such as land-use and built-up regions, and accessibility including roads, navigable rivers and coastlines. The value of HII ranges between 0 and 64 (WCS/CIESIN, 2005).

### Table 1. Details of the fossil pollen datasets used in this work.

<table>
<thead>
<tr>
<th>Site</th>
<th>La. (°)</th>
<th>Lo. (°)</th>
<th>Al. (m)</th>
<th>Archive</th>
<th>Dating method</th>
<th>Number of dates</th>
<th>Re. (years)</th>
<th>Reference</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tianchi</td>
<td>35.26</td>
<td>106.31</td>
<td>2430</td>
<td>Lake</td>
<td>$^{14}$C</td>
<td>19</td>
<td>85</td>
<td>Zhao et al. 2010</td>
<td>II</td>
</tr>
<tr>
<td>Daihai</td>
<td>40.58</td>
<td>112.67</td>
<td>1220</td>
<td>Lake</td>
<td>$^{14}$C</td>
<td>8</td>
<td>215</td>
<td>Xiao et al. (2004)</td>
<td>IV</td>
</tr>
<tr>
<td>Qasq</td>
<td>40.67</td>
<td>111.13</td>
<td>1000</td>
<td>Peat</td>
<td>$^{14}$C</td>
<td>4</td>
<td>90</td>
<td>Wang et al. (1997)</td>
<td>IV</td>
</tr>
<tr>
<td>Qinghai</td>
<td>36.67</td>
<td>100.52</td>
<td>3200</td>
<td>Lake</td>
<td>$^{14}$C</td>
<td>6</td>
<td>70</td>
<td>Liu et al. (2002)</td>
<td>IV</td>
</tr>
<tr>
<td>Chaohu</td>
<td>31.56</td>
<td>117.39</td>
<td>10</td>
<td>Lake</td>
<td>$^{14}$C</td>
<td>7</td>
<td>100</td>
<td>Wu et al. (2008)</td>
<td>IV</td>
</tr>
</tbody>
</table>

La., latitude; Lo., longitude; Al., altitude; Re. resolution

![Figure 2. Distribution of human influence index (HII) in China (WCS/CIESIN, 2005)](image-url)
2.3. Numerical analyses

Three ordination methods, detrended correspondence analysis (DCA), canonical correspondence analysis (CCA) and redundancy analysis (RDA), were employed for assessing the quantitative relationships among surface pollen assemblages and environmental or climatic variables (Papers II and III). DCA was utilized to detect the spatial patterns of pollen assemblages along main environmental or climatic gradients of the study area and to determine the use of unimodal-based CCA or linear-based RDA in further ordination analyses by calculating the gradient lengths of surface pollen assemblages that are expressed as standard deviation (SD) units of compositional turnover (ter Braak and Šmilauer, 2012). CCAs and RDAs using each variable as the only constraining variable at a time were run to estimate quantitatively how much of the variance in pollen assemblage data was accounted for by each variable alone. Variance inflation factors (VIF) value of each variable was measured to examine the degree of collinearity among the explanatory variables. Monte Carlo permutation tests involving 999 permutations were carried out for evaluating the statistical significance of each variable. Prior to ordination analyses, pollen percentage data were square-root transformed for optimizing the ‘signal’ to ‘noise’ ratio and stabilizing the variance (Prentice, 1980). All DCAs, CCAs and RDAs were conducted using the VEGAN package (Oksanen et al., 2013) in R.

Pearson’s correlation analysis was utilized to assess the relationships between pollen percentages and vegetation proportions in Paper I. Boosted regression trees (BRTs; Elith et al., 2008) were employed to model the response of individual pollen taxa to each climate variable or HII and to quantify the relative importance of individual climate variables or HII that account for the distribution of each pollen taxon in Papers II and III. BRTs incorporate two algorithms (regression trees and boosting function) and are insensitive to outliers. BRTs can deal with different types of environmental variables without prior data transformation and handle non-linear relationships (Elith et al., 2008). BRTs analyses for selected abundant and common pollen types were performed by utilizing the DISMO package (Hijmans et al., 2012) in R.

Quantitative climate and HII reconstructions in Papers II and IV were performed from the five fossil pollen sequences using modern pollen-based calibration models developed with weighted-averaging partial least squares (WA-PLS) regression (ter Braak and Juggins, 1993) or a Bayesian regression method. Leave-one-out cross-validation (Birks et al., 1990) was employed for evaluating the performance of the WA-PLS models. Performance statistics including coefficient of determination ($r^2$) between measured and predicted values, root-mean-square error of prediction (RMSEP), and maximum bias were estimated for each WA-PLS model. The suitable WA-PLS components used for the reconstructions were selected according to the $t$-test (van der Voet, 1994). All WA-PLS models and reconstructions were conducted with the RIOJA package (Juggins, 2012) in R. The Bayesian method is Bummer originally developed in Vasko et al. (2000). The Bayesian approach has been utilized for pollen-based palaeoclimatic reconstructions in Europe (e.g., Korhola et al., 2002; Haslett et al., 2006; Salonen et al., 2012b; Holmström et al., 2014) but not in China before this work. As the computational time makes the leave-one-out cross-validation impracticable for the Bayesian model, a simplified procedure with 10-fold cross-validation was used for assessing the Bayesian model performance. Performance statistics including $r^2$, RMSEP and maximum bias were calculated for the Bayesian model. All computations in terms of the Bayesian model and reconstruc-
3. Summary of original publications

3.1. Paper I

In Paper I, the objective was to investigate the relationships among modern pollen, vegetation cover, and human land-use in hydrologically-defined river catchments from mountainous ecosystems influenced by human activities and to provide a basis for reconstructing the pollen/vegetation/land-use relationships at small catchment scale. For this purpose, palynological analysis for 62 surface sediment samples from small irrigation pools in the Tuoliang and Qipanshan catchments from the Taihang Mountains in northern China was carried out to understand how pollen assemblages from the pool sediments correlate with the vegetation cover and land-use data.

The results show that pollen and spores such as *Artemisia*, Chenopodiaceae, *Pinus* and *S. sinensis* are the dominant components and pollen taxa that could be from cultivated plants such as cereals, Brassicaceae, Fabaceae, Solanaceae, Apiaceae and Cucurbitaceae occur frequently with low content in most pollen assemblages from the two catchments. With the decrease of altitude, pollen percentages of *Artemisia* and Chenopodiaceae gradually decline, whereas *Pinus* and *S. sinensis* percentages rise, implying that saccate *Pinus* pollen and large *S. sinensis* spores can be transported downstream over longer distances and that pollen-sorting is happening in the process of water transportation. Both constrained and unconstrained cluster analysis show that pollen assemblages from the sampled irrigation pools can be divided into three groups, corresponding to the upstream, midstream and downstream parts of the river catchments. There are differences in pollen assemblages among the three groups, but differences among the pools in each group are small, suggesting that these three groups may correspond relatively well to the potential pollen source areas.

![Figure 3](image-url)  
*Correlation of pollen percentages of tree, shrub, herb, crop and cereal with land-use proportions of woodland, scrubland, grassland and farmland in the Tuoliang and Qipanshan catchments.*
of the pools of each group. In addition, the largest and deepest pools have the highest pollen concentrations, probably because these pools reduce water velocity, contain deep layers of deposits with abundant silt and organic matter, and thus reduce the amount of pollen grains that can be transported to the downstream pools.

It is found that percentages of pollen types from cultivated crops such as vegetable and cereals increase with the decline of altitude and the increase of human influence and farmland areas from the upstream to downstream parts of the catchments (Fig. 3). Vegetation-cover proportions of woodland and scrubland are higher than percentages of tree and shrub pollen, while grassland proportions are lower than herb pollen percentages, because of the differences in pollen productivity, representation, dispersal and deposition in terms of different pollen types. Moreover, correlation analysis between vegetation/land-use proportions and their corresponding pollen percentages reveals that vegetation proportions of woodland, scrubland and grassland are mostly poorly correlated with pollen percentages of trees, shrubs and herbs, whilst there is a positive linear correlation between pollen percentages of cereals and proportions of farmland in the catchments (Fig. 3).

### 3.2. Paper II

In Paper II, we tested a timely and important issue regarding the potential source of bias caused by long-term and intensive human influence on palaeoclimatic reconstructions using pollen data. Modern datasets over continental China for pollen, human influence index (HII) and two climatic parameters (PANN and TANN) were used for constructing quantitative pollen-based inference models for these variables with WA-PLS regression. The relative influence of HII on affecting surface pollen distribution was evaluated by comparison to PANN or TANN at the assemblage level by employing constrained ordination methods (CCA or RDA) and at the taxonomic level by using BRTs. The inference models were applied to a pollen sequence covering the past 6200 years from Tianchi Lake in east-central China (ECC) for quantitative human influence and climate reconstructions. The accuracy of the climate reconstructions was assessed by comparing them with the HII reconstruction and other palaeoclimatic records.

It is found from the ordination results that the relative influence of HII is smaller than PANN or TANN on affecting the spatial distribution of modern pollen assemblages, with the exception of ECC where the relationship between HII and pollen assemblages is both ecologically and statistically significant. BRTs results show that the relative importance of HII affecting the spatial distribution of individual pollen taxa differs among regions and is mostly smaller than PANN or TANN. Correspondingly, the pollen-based inference models for HII in most regions of China have poor statistical performances compared with PANN or TANN, except for in ECC where the pollen–HII model displays a good performance. The reconstructed HII based on the Tianchi pollen record is low from 6200 to 2900 cal. yr BP, begins to increase steadily between 2900 and 1100 cal. yr BP, and rises suddenly 1100 years ago (Fig. 4). The HII variations are roughly comparable with Holocene regional population and charcoal changes. By contrast, the climate reconstructions show a sudden PANN decrease and a TANN increase since 1100 cal. yr BP. However, other palaeoclimatic records are mostly inconsistent with the reconstructed variations of PANN or TANN over the past 1100 years, indicating that the pollen-based climatic reconstructions for the past 11 centuries are biased because of the rise of human impact on vegetation distribution in the surrounding areas of Lake Tianchi. It seems
likely that this problem also occurs in the Holocene climatic reconstructions based on pollen data from other regions affected by the intensive and long-term human impact on ecosystems.

3.3. Paper III

In Paper III, the focus was to assess quantitatively the importance of individual climate parameters affecting the geographical distribution of modern pollen data and to identify the most important climate parameters that can be reconstructed from fossil pollen data in different climatic regions of China. To this end, the Chinese continental-scale modern pollen dataset was employed and subsequently divided into five climatically and ecologically different subsets corresponding to different regional vegetation zones. The importance of six bioclimatic parameters including PANN, Alpha, TANN, MTWA, MTCO and GDD5 for the spatial variation of modern pollen data was quantified using BRTs at the taxonomic level and constrained ordination techniques (CCA or RDA) at the assemblage level.

In BRTs analyses, taxon-specific responses to the six climate parameters were found to be individualistic and exhibit a regional diversity. In ordination analyses, the minimal set of the most important climate variables with low levels of collinearity (variance inflation factors value < 4) that mostly accounts for the highest amounts of the spatial variations in modern pollen assemblages were regionally different (Fig. 5). In general, in north-western and north-eastern China and the Tibetan Plateau the most important

Figure 4. Comparison of (a) HII, (d) PANN and (e) TANN reconstructions from the Tianchi pollen record with other relevant records: (b) the Tianchi charcoal record (Zhao et al., 2010); (c) the human population record for the past 4000 years of Gansu Province (Fang and Zhang, 2007); (f) the reconstructed precipitation index (PI) over the past 1800 years of north-central China (Tan et al., 2011); (g) the synthesized temperature anomaly (TA) for the past 2000 years in China (Ge et al., 2013); and the δ¹⁸O records from (h) Heshang (Hu et al., 2008) and (i) Dongge (Wang et al., 2005) Caves.
climate variable is PANN and the importance of hydrological variables (PANN or Alpha) is greater than that of thermal variables. However, in east-central and southern China MTWA and MTCO are the dominant variables, and the thermal variables such as MTWA, MTCO, TANN and GDD5 are mostly more important than hydrological variables. These results suggest that the selection of important climate variables used in pollen-based climate reconstructions is essential and that the climate variables with low collinearity that can be optimally reconstructed in China differ across regions. These regional differences in pollen–climate relationships are typical to other large-scale modern pollen datasets that exist in other regions of the world and need to be taken into consideration in the pollen-based climatic reconstructions.

Figure 5. Percentages of variation (%) in pollen data that are explained by each variable in RDA and CCA, with this variable used as the only constraining variable for WC and all subsets.

3.4. Paper IV

In paper IV, new pollen-based transfer functions for the East Asian summer monsoon precipitation ($P_{jja}$) were constructed with WA-PLS and a Bayesian regression technique. These transfer functions were applied to a set of fossil pollen data from monsoonal China to reconstruct the $P_{jja}$ changes over the last 9500 years. The results were compared with $P_{jja}$ simulations using a coupled ocean–atmosphere–sea ice general circulation model (the Kiel Climate Model, KCM). The results show that both the WA-PLS and Bayesian transfer functions display good statistical performances and robust predictive power. The WA-PLS reconstructions from northern China are statistically significant, being mostly consistent with the Bayesian reconstructions. All WA-PLS and Bayesian reconstructions from northern China were stacked to one summary record,
suggesting that $P_{jja}$ increases rapidly from 9500 cal. yr BP onwards, reaches its highest level at 7000–4000 cal. yr BP, and decreases to present (Fig. 6). This general pattern is in accordance with previous synthesized records for the Holocene moisture changes in northern China. The comparison with the KCM output indicates that the range of the absolute $P_{jja}$ values is roughly comparable between our reconstructions and the simulations. However, the magnitude of the $P_{jja}$ changes and its overall trends are different. For the early-Holocene (9500–8000 cal. yr BP), the model suggests higher $P_{jja}$ than the reconstructions. During the Holocene summer monsoon maximum (HSMM), $P_{jja}$ was about 70–110 mm above present in the stacked record, whereas it was only about 20–60 mm higher than present in the model results. The rising or declining $P_{jja}$ patterns before or after the HSMM in the reconstructions are more pronounced and fluctuating than in the simulations. Other palaeoclimatic data from monsoonal China indicate substantial changes of the monsoon precipitation throughout the Holocene. This suggests that the KCM underestimates the amplitude of the Holocene monsoon precipitation changes in comparison with the reconstructions.

![Figure 6](image.png)

**Figure 6.** Comparison of (a) the pollen-based stacked $P_{jja}$ record for northern China with the KCM-based $P_{jja}$ simulations (Jin et al., 2014) for (b) northern and (c) southern China, and the speleothem $\delta^{18}O$ records from (d) Henshang (Hu et al., 2008) and (e) Dongge (Dykoski et al., 2005) Caves in southern China.

4. Discussion

4.1. Potentials and limitations of quantitative pollen-based climate and human influence reconstructions in China
On the basis of analyzing modern numerical pollen–climate and pollen–human influence relationships as well as performing many pollen-based quantitative reconstructions, this thesis provides some perspectives on how to employ pollen data for past climate and human influence reconstructions in China.

From the viewpoint of climate reconstructions, the continental-scale Chinese surface pollen dataset used in this work (Papers II, III and IV) has some strengths and weaknesses. Most of the pollen–climate calibration models based on this dataset generally display good statistical performances in cross-validation tests (Fig. 2 in Papers II and IV), mostly because that they capture major climatic and vegetation gradients of China from tropical region in the south to boreal region in the north and from forests in the east to deserts in the west. For this reason, it is also possible to employ this dataset for quantitatively reconstructing climatic changes for the last glacial period and the early-Holocene. However, this dataset has also several disadvantages which limit its robustness. For example, because of accessibility difficulty and human disturbance, some modern pollen samples were selectively and unevenly collected from different geographical areas of China (e.g., Li et al., 2007; Xu et al., 2010b; Zheng et al., 2014), leading to spatial gaps in some areas such as western Tibet, whereas other areas such as north-central China are over-represented. This regionally aggregated distribution pattern likely increases the degree of spatial autocorrelation for pollen–climate calibration models (Telford and Birks, 2009). Moreover, due to the scarcity of lakes in most regions of China, most of the samples in our dataset were obtained from surface soils and a lower number of samples from peats, moss polsters, dust traps, etc. (e.g., Li et al., 2007; Luo et al., 2010; Lu et al., 2011, Zheng et al., 2014). This probably increases reconstruction errors when the calibration models are applied to fossil pollen sequences from lake sediment cores (Birks et al., 2010), because pollen spectra from lakes, soils, peats and moss polsters have different relevant pollen source areas and hence reflect differently the surrounding vegetation (e.g., Li et al., 2005; Zhao et al., 2009; Zhao and Herzschuh, 2009; Birks et al., 2010). However, some studies show that model errors do not rise or even decrease when surface samples from multiple sedimentary environments are assembled for constructing pollen-based calibration models for some climate variables (e.g., Goring et al., 2010).

Paper II highlights that the pollen-based climatic reconstructions for the past 1100 years from the Tianchi pollen record in east-central China are biased due to the rise of human impact (Fig. 4). This situation also occurs in another reconstruction of the Holocene climate using a fossil pollen record from the Anyang region in the east part of east-central China (Xu et al., 2010b), because this record is affected by human deforestation, showing an increase of herb pollen and a drop of tree pollen after the mid-Holocene. In addition to east-central China, many Holocene pollen sequences from other areas of eastern China are characterized by an increase of herb pollen such as Poaceae, Artemisia, Plantago, Chenopodiaceae and Urticaceae in connection with a reduction of tree pollen such as Quercus, Betula, Ulmus, Pinus and Cyclobalanopsis in the middle- and late-Holocene because of human influence (e.g., Shen et al., 2006; Chen et al., 2009; Cao et al., 2010; Yue et al., 2015). It is therefore clear that the quantitative reconstructions of the Holocene climate based on these pollen sequences would be biased by human impact and must be treated with caution. Paper III emphasizes that the climate variables with low collinearity which can be optimally reconstructed from pollen data are regionally different in China (Fig. 5). As Birks et al. (2010) note, identifying
the critical climatic determinants is an essential step in quantitative pollen-based climate reconstructions. In addition, a high degree of collinearity among explanatory variables is a problem in quantitative climatic reconstructions based on biological proxy data and can lead to spurious results (Birks et al., 2010; Juggins and Birks, 2011; Juggins, 2013). For instance, Salonen et al. (2014) showed that in northern Europe pollen-based winter temperature ($T_{d_j}$) reconstructions are nearly interchangeable with summer temperature ($T_{j_s}$) reconstructions and may therefore be spurious, due to the past changes of a correlated, ecologically significant determinant ($T_{j_s}$) accounting for the spatial variability in surface pollen distribution. However, the confounding effect of the correlated climate variables has often been ignored in China. For example, Jiang et al. (2006) used the same surface and fossil pollen datasets from northern China for simultaneously reconstructing MTCO, MTWA, TANN, PANN, GDD5 and Alpha. Wen et al. (2010) applied WA-PLS to the same pollen datasets in northern China to concurrently reconstruct MTCO, MTWA, TANN and PANN. The results of Paper III demonstrate that some of these climate variables are strongly correlated and not identified as the most ecologically significant determinants. These problems should be considered in the future quantitative palaeoclimatic reconstructions based on pollen data in China.

As regards to human influence reconstructions, Paper I sheds lights on how surface pollen data from irrigation pool sediments reflect human-influenced land-use and vegetation distribution at the small catchment scale in northern China. Our results indicate that there exist no linear relationships between pollen percentages of trees, shrubs and herbs and percentage covers of woodland, scrubland and grassland, while Cerealae-type pollen percentages are linearly correlated to farmland proportions (Fig. 3), suggesting complex pollen–land-use or vegetation relationships. Furthermore, some pollen taxa that may stem from crop plants such as cereal, Brassicaceae, Fabaceae, Solanaceae, Apiaceae and Cucurbitaceae can be used as indicators for reconstructing human-caused land-use changes in other areas of China. In terms of the small hydrologically-defined water-bodies studied in Paper I, it has been demonstrated in previous studies that the pollen is mainly transported by river water (Pennington, 1979; Bonny and Allen, 1984; Fall, 1987; Kong et al., 2000; Xu et al., 2004, 2005), so such sites are unsuitable to use the LOcal Vegetation Estimation (LOVE) model or other related models developed by Sugita (2007a, b) which assume that the pollen is transported by wind only. Therefore more detailed studies on the pollen/land-use/vegetation relationships at small spatial scale are needed in China for developing a quantitative model for human-influenced landscape reconstructions.

Paper II provides new insights into how to quantitatively extract the underlying signal of human influence from modern and fossil pollen data by using the high-resolution human influence index (HII) dataset. Our results manifest that the pollen-based transfer function for HII in east-central China (ECC) yields reasonable cross-validation performance (Fig. 2 in Paper II) and that the Tianchi HII reconstruction is mostly consistent with historical population and charcoal records (Fig. 4). However, the pollen-based HII reconstruction has many limitations which need to be considered in its future applications. The HII gradient is highly complex compared to the simple and smooth southeast-northwest precipitation gradient or south-north temperature gradient (Figs. 1 and 2), because the global HII database was constructed by combining fine-grained data networks with strong local variability including human population distributions, built-up regions, roads, railroads, and vari-
ous land-use/land-cover (WCS/CIESIN, 2005). This feature makes it difficult to develop a robust large-scale pollen-based inference model for HII. Moreover, HII has smaller effects on accounting for modern pollen distribution than climate variables except for ECC. This makes it impossible to quantitatively reconstruct HII in most regions of China (except ECC) based on our surface pollen dataset. The Tianchi HII reconstruction indicates that the rise of HII is predominantly caused by the increase of herb pollen (e.g., *Artemisia*, Poaceae and Chenopodiaceae) and the reduction of tree pollen (e.g., *Betula*, *Pinus* and *Picea*) (Figs 3 and 4 in Paper II). These herb pollen types are associated with high HII values and can be regarded as anthropogenic indicators in the deforested areas of eastern China (e.g., Li et al., 2008; Li et al., 2015), whilst they represent natural vegetation in the steppe or desert areas of arid and semi-arid China where HII reconstructions from pollen data are unlikely to be realistic and feasible. The last-glacial period and the early-Holocene are characterized by cold and dry climate and high percentage of herb pollen types. The reconstructed HII values for these periods would be very high and therefore problematic, particularly when there is no strong evidence of human activity. Additionally, herb pollen types can be present throughout most of the fossil pollen records in China, and hence the pollen-inferred HII values would not go down to zero and could not be interpreted to reveal human influence on ecosystems.

In terms of quantitative reconstruction methods, Paper IV uses a Bayesian multinomial regression model which has not been tested before this work for pollen-based palaeoclimatic reconstructions in China according to our knowledge. In contrast, Bayesian-based approaches have been utilized for quantitatively reconstructing past climate changes in Europe (e.g., Vasko et al., 2000; Toivonen et al., 2001; Korhola et al., 2002; Haslett et al., 2006; Salonen et al., 2012b). Our results show that the Bayesian model for the summer monsoon precipitation ($P_{jja}$) performs slightly better than the WA-PLS model in terms of performance statistics (Fig. 2 in Paper IV). The reason can be that Bayesian modeling offers many noticeable advantages in comparison with the so-called frequentist reconstruction methods such as WA-PLS (e.g., Vasko et al., 2000; Birks et al., 2010; Salonen et al., 2012b). For example, ecological information and processes over time and space could be explicitly embedded to the Bayesian model and expressed as a posterior density function. The structure of the Bayesian calibration procedure is transparent because of the utilization of probability distributions instead of point estimates and can thus provide more information with respect to the taxon-climate responses, model behavior and uncertainty. The Bayesian framework has great potential and can be further tested and applied for pollen-based quantitative climate reconstructions in China. However, the challenge is that the Bayesian reconstructions require special computer software and higher computational burden especially when large-scale surface and fossil datasets are involved (e.g., Birks et al., 2010). In addition, the results of Paper IV show that the Bayesian reconstructions for $P_{jja}$ are mostly consistent with the WA-PLS reconstructions. The use of multi-model reconstructions suggests that testing different reconstruction models is essential for assessing the reconstructions and extracting the underlying consistent features (e.g., Salonen et al., 2014; Holmström et al., 2014).

### 4.2. Future prospects

For climate reconstructions based on the continental-scale Chinese modern pollen dataset, more samples are needed for reducing the spatial gaps. In recent years, a large number of avail-
able pollen samples from surface soils or lake sediments have been published in, for example, the Qinghai-Tibet Plateau (Zhao et al., 2009; Herzschuh et al., 2010; Lu et al., 2011), the Loess Plateau (Zhao et al., 2012b), north-western China (Zhao et al., 2012a; Wang et al., 2014), north-eastern China (Wen et al., 2013) and southern China (Zhu et al., 2008; Fang et al., 2015). These samples can be potentially incorporated into our dataset for minimizing the current geographical gaps and covering more detailed vegetation types and climatic spaces, thereby possibly improving the robustness of our dataset. In addition, a continental-scale fossil pollen dataset over eastern Asia has been recently constructed (Cao et al., 2013) and can be used for large-scale quantitative climate reconstructions coupled with the Chinese surface pollen dataset. Moreover, there are still some important methodological issues that can be examined for further improving the reliability and accuracy of pollen-based climate reconstructions in China. For example, the use of our surface pollen dataset for reconstructing climate changes over longer timescales such as the late glacial or interglacial period that are less analogous to modern climate needs to be tested in China (e.g., Herzschuh et al., 2010; Mu et al., 2015). The effects of calibration-set or taxon selection on the model performances and reconstructions can be evaluated (e.g., Bjune et al., 2010; Xu et al., 2010; Juggins et al., 2015). Critical validation of the quantitative reconstructions using statistical tests (Telford and Birks, 2011) and multi-proxy comparisons is increasingly important (e.g., Birks et al., 2010) and should be relied on.

For human influence reconstructions, many surface pollen samples are needed from various human-disturbed areas such as forest clearing, cultivated land and wasteland (e.g., Yang et al., 2010; Li et al., 2015) in eastern China where modern HII values are high. These samples together with the samples from natural vegetation would make it possible to cover long HII gradient for developing robust pollen-based calibration models for HII (Li et al., 2015). To apply HII inference models for reconstructing the regional-scale temporal and spatial patterns of past human influence on vegetation, a large dataset consisting mainly of high-resolution fossil pollen records which can reveal in detail the historical processes of human impact on landscape can be constructed for eastern China. For effectively validating the HII reconstructions, it is necessary to obtain other proxy data indicative of human influence such as charcoal (macro-charcoal and micro-charcoal) and crop plant macrofossil (e.g., leaves, fruits and seeds) records and historical document materials (e.g., Li et al., 2008). Additionally, a dataset that summaries and describes all anthropogenic pollen indicators for different cultural landscapes such as farmland, pasture and rangeland needs to be compiled for China. To better understand the history of human influence on environmental changes, a comprehensive cooperation among scientists from palaeoclimatology, palaeoecology and archaeology is required.

5. Conclusions

The major results and implications of this work can be summarized as follows:

- Modern pollen/land-use/vegetation relationships in the two small river catchments of northern China are found to differ noticeably in terms of different pollen and land-use types. In general, percentages of tree and shrub pollen are lower than vegetation proportions of woodland and scrubland, whereas percentages of herb and crops pollen are higher than proportions of grassland and farmland. There are no linear relationships between pollen percentages and vegetation proportions for trees, shrubs
and herbs in relation to woodland, scrubland and grassland, respectively. In contrast, pollen percentages of cereals are significantly and linearly correlated to proportions of farmland. In addition, the pollen/land-use/vegetation relationships vary strongly from upstream to downstream areas along the human-disturbed land-use gradients in the catchments. These local-scale complex pollen/land-use relationships can be ascribed to the significant differences in pollen dispersal, deposition and productivity and land-use distribution characteristics within the catchments affected by human influence.

- The quantitative effect of human influence index (HII) accounting for the distribution of modern pollen data at both assemblage and taxonomic levels is mostly smaller than climate variables in different geographical regions of China, with the exception of east-central China (ECC). In addition, the quantifiable importance of key climate variables explaining modern pollen distribution patterns varies regionally in China on regional-to continental-scales. Hydrological variables are found to be more important than temperature-related variables in north-western and north-eastern China and the Tibetan Plateau, while temperature-related variables are more important in east-central and southern China. Taxon-specific responses to the climate variables are found to be highly individualistic. These results show that HII can be potentially reconstructed from pollen data in ECC and that the climate variables which could be optimally reconstructed from pollen data differ significantly in terms of different bioclimatic regions of China.

- Pollen-based calibration models for HII in ECC and climate parameters such as PANN, TANN and \( P_{\text{jja}} \) over continental China generally have good statistical performances and prediction powers. These models are applied to a set of fossil pollen sequences for quantitative HII and climate reconstructions. HII reconstruction based on the Tianchi pollen record shows that, HII increases since the middle-Holocene, rises suddenly at 1100 cal. yr BP, and remains high until present. PANN and TANN reconstructions over the last 1100 years from the Tianchi pollen data are mostly inconsistent with other independent palaeoclimatic records, suggesting that the Tianchi climate reconstructions are strongly biased by the abrupt rise of human impact. This indicates that climate reconstructions from human-influenced pollen records should be treated with caution. \( P_{\text{jja}} \) reconstructions with two different calibration models from three fossil records in northern China are consistent and stacked to a summary curve which shows that \( P_{\text{jja}} \) increases substantially since 9500 cal. yr BP, reaches the highest extent at 7000–4000 cal. yr BP, and decreases to present-day. This highlights the importance of ensemble climate reconstructions with different calibration models and fossil sites for identifying the underlying robust palaeoclimatic characteristics. In addition, the \( P_{\text{jja}} \) variations in the reconstructions are more pronounced and fluctuating than in the KCM simulations. Other palaeoclimatic records indicate strong Holocene monsoon precipitation changes. This indicates that the KCM underestimates the monsoon precipitation changes and the pollen-based reconstructions are more reliable.

References


Behre, K.E., 1986. Anthropogenic Indicators in Pollen
Diagram. Balkema, Rotterdam.


Jin, L.Y., Schneider, B., Park, W., Latif, M., Khon, M,


Juggins, S., 2013. Quantitative reconstructions in palaeolimnology: new paradigm or sick science? Quaternary Science Reviews 64, 20–32.


Sugita, S., 2007b. Theory of quantitative reconstruction
of vegetation II: all you need is LOVE. Holocene 17, 243–25.


