Furong Li

Pollen productivity estimates and pollen-based reconstructions of Holocene vegetation cover in northern and temperate China for climate modelling
Pollen productivity estimates and pollen-based reconstructions of Holocene vegetation cover in northern and temperate China for climate modelling
POLLEN PRODUCTIVITY ESTIMATES
AND POLLEN-BASED
RECONSTRUCTIONS OF HOLOCENE
VEGETATION COVER IN NORTHERN
AND TEMPERATE CHINA FOR CLIMATE
MODELLING

FURONG LI

LINNAEUS UNIVERSITY PRESS
Abstract


Model projections of future climate change require that coupled climate-vegetation models are developed and validated, i.e. these models should be able to reproduce past climate and vegetation change. Records of pollen deposited in lake bottoms and peat bogs can provide the information needed to validate these models. The aim of this thesis was i) to explore the modern relationships between pollen and vegetation in northern and temperate China and estimate pollen productivity of major plant taxa, and ii) to use the results of i) to produce the first reconstruction of plant cover in China over the last 10 000 years for the purpose of climate modelling. A study of the modern pollen-vegetation-climate relationships was performed in northwestern China (Paper I). Pollen productivity for 18 major plants of cultural landscapes in central-eastern China was estimated (Paper II). Based on a synthesis and evaluation of all existing estimates of pollen productivity in the study region, a standard dataset of pollen productivity for 31 plant taxa is proposed (Paper III). This dataset was used to achieve pollen-based REVEALS reconstructions of plant cover over the last 10 000 years in 35 regions of northern and temperate China (Paper IV).

The major findings can be summarized as follows. Paper I: Annual precipitation (P ann) is the major climatic factor influencing pollen assemblages, followed by July precipitation (P Jul). The shared effect of combinations of two climatic factors explains a larger portion of the variation in pollen data than individual variables. Paper II: Of the 16 reliable pollen productivities estimated, the estimates for 8 taxa are new, Castanea, Cupressaceae, Robinia/Sophora, Anthemis type/Aster type, Cannabis/Humulus, Caryophyllaceae, Cruciferae, and Galium type. Trees have in general larger pollen productivity than herbs. Paper III: Of the total 31 taxa for which estimates of pollen productivity are available in China, 13 taxa have more than 1 value. All or most of these values are similar for Artemisia, Cyperaceae, Larix, Quercus and Pinus. Eight taxa have very variable estimates. Paper IV: The REVEALS plant percentage-cover strongly differs from the pollen percentages, and they provide new important insights on past changes in plant composition and vegetation dynamics.

Key words: climate, pollen-vegetation relationships, relative pollen productivity, REVEALS model, anthropogenic land-cover change, Holocene, China

Pollen productivity estimates and pollen-based reconstructions of Holocene vegetation cover in northern and temperate China for climate modelling
Doctoral dissertation, Department of Environmental Science, Linnaeus University, Kalmar, 2016

ISBN: 978-91-88357-52-6
Published by: Linnaeus University Press, 351 95 Växjö
Printed by: DanagårdLiTHO AB, 2016
Abstract


Model projections of future climate change require that coupled climate-vegetation models are developed and validated, i.e. these models should be able to reproduce past climate and vegetation change. Records of pollen deposited in lake bottoms and peat bogs can provide the information needed to validate these models. The aim of this thesis was i) to explore the modern relationships between pollen and vegetation in northern and temperate China and estimate pollen productivity of major plant taxa, and ii) to use the results of i) to produce the first reconstruction of plant cover in China over the last 10 000 years for the purpose of climate modelling. A study of the modern pollen-vegetation-climate relationships was performed in northwestern China (Paper I). Pollen productivity for 18 major plants of cultural landscapes in central-eastern China was estimated (Paper II). Based on a synthesis and evaluation of all existing estimates of pollen productivity in the study region, a standard dataset of pollen productivity for 31 plant taxa is proposed (Paper III). This dataset was used to achieve pollen-based REVEALS reconstructions of plant cover over the last 10 000 years in 35 regions of northern and temperate China (Paper IV). The major findings can be summarized as follows. Paper I: Annual precipitation (P ann) is the major climatic factor influencing pollen assemblages, followed by July precipitation (P Jul). The shared effect of combinations of two climatic factors explains a larger portion of the variation in pollen data than individual variables. Paper II: Of the 16 reliable pollen productivities estimated, the estimates for 8 taxa are new, Castanea, Cupressaceae, Robinia/Sophora, Anthemis type/Aster type, Cannabis/Humulus, Caryophyllaceae, Cruciferae, and Galium type. Trees have in general larger pollen productivity than herbs. Paper III: Of the total 31 taxa for which estimates of pollen productivity are available in China, 13 taxa have more than 1 value. All or most of these values are similar for Artemisia, Cyperaceae, Larix, Quercus and Pinus. Eight taxa have very variable estimates. Paper IV: The REVEALS plant percentage-cover strongly differs from the pollen percentages, and they provide new important insights on past changes in plant composition and vegetation dynamics.

Key words: climate, pollen-vegetation relationships, relative pollen productivity, REVEALS model, anthropogenic land-cover change, Holocene, China
致我的家人

To My Parents with Love
致我的家人

To My Parents with Love
ABSTRACT ......................................................................................................... 5
SAMMANFATTNING ........................................................................................ 7
LIST OF PAPERS .............................................................................................. 11
1. INTRODUCTION .......................................................................................... 13
   1.1 Background ............................................................................................. 13
   1.2 Pollen vegetation relationship studies ..................................................... 15
       1.2.1 Empirical and semi-quantitative comparative approaches to study
       the relationship between pollen and vegetation ....................................... 16
       1.2.2 Models of the relationship between pollen and vegetation ............. 17
   1.3 Climate studies in China and debated issues ........................................... 20
   1.4 Holocene vegetation history in China ..................................................... 21
   1.5 Archaeological studies on the origin of Neothilic agriculture and
   related human-induced land use .................................................................... 22
2. OBJECTIVES ................................................................................................ 25
3. STUDY REGION ........................................................................................... 27
   3.1 The Jungar Basin and the western Altay Mountains ............................... 28
   3.2 The cultural landscape of Shandong ....................................................... 28
   3.3 Synthesis of the relative pollen productivity (RPP) estimates available
   in China ......................................................................................................... 28
   3.4 The REVEALS application ..................................................................... 29
4. MATERIAL AND METHODS ..................................................................... 31
   4.1 Site selection, sampling and vegetation survey (Papers I and II) ............ 31
   4.2 Theory of the ERV- and REVEALS models .......................................... 32
       4.2.1 The ERV model and its three submodels ........................................ 32
       4.2.2 Pollen source area ........................................................................... 33
       4.2.3 The REVEALS model (Paper IV) .................................................. 34
   4.3 Data analysis in papers I, II and III ......................................................... 34
       4.3.1 Pollen assemblages and numerical methods (Paper I) .................... 34
       4.3.2 ERV-model runs (Paper II) ............................................................. 35
       4.3.3 Synthesis of available RPPs in China (Paper III) ........................... 35
   4.4 Data collection, chronology and REVEALS runs (Paper IV) ................. 35
       4.4.1 Study sites, data sources, and chronology ...................................... 35
       4.4.2 Grouping of sites and time windows .............................................. 36
       4.4.3 The REVEALS model runs and presentation of results ................. 37
5. RESULTS-SUMMARY OF PAPERS ........................................................... 39
   5.1 PAPER I (Li et al., in press): Modern pollen-climate relationships in
   north Xinjiang, northwestern China: implications for pollen-based
   reconstruction of Holocene climate .............................................................. 39
   5.2 PAPER II (Li et al., in revision): Relative Pollen Productivity (RPP)
   estimates for major plant taxa in cultural landscapes of central eastern
   China ............................................................................................................. 40
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>5</td>
</tr>
<tr>
<td>SAMMANFATTNING</td>
<td>7</td>
</tr>
<tr>
<td>LIST OF PAPERS</td>
<td>11</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>13</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>13</td>
</tr>
<tr>
<td>1.2 Pollen vegetation relationship studies</td>
<td>15</td>
</tr>
<tr>
<td>1.2.1 Empirical and semi-quantitative comparative approaches to study the relationship between pollen and vegetation</td>
<td>16</td>
</tr>
<tr>
<td>1.2.2 Models of the relationship between pollen and vegetation</td>
<td>17</td>
</tr>
<tr>
<td>1.3 Climate studies in China and debated issues</td>
<td>20</td>
</tr>
<tr>
<td>1.4 Holocene vegetation history in China</td>
<td>21</td>
</tr>
<tr>
<td>1.5 Archaeological studies on the origin of Neolithic agriculture and related human-induced land use</td>
<td>22</td>
</tr>
<tr>
<td>2. OBJECTIVES</td>
<td>25</td>
</tr>
<tr>
<td>3. STUDY REGION</td>
<td>27</td>
</tr>
<tr>
<td>3.1 The Jungar Basin and the western Altay Mountains</td>
<td>28</td>
</tr>
<tr>
<td>3.2 The cultural landscape of Shandong</td>
<td>28</td>
</tr>
<tr>
<td>3.3 Synthesis of the relative pollen productivity (RPP) estimates available in China</td>
<td>28</td>
</tr>
<tr>
<td>3.4 The REVEALS application</td>
<td>29</td>
</tr>
<tr>
<td>4. MATERIAL AND METHODS</td>
<td>31</td>
</tr>
<tr>
<td>4.1 Site selection, sampling and vegetation survey (Papers I and II)</td>
<td>31</td>
</tr>
<tr>
<td>4.2 Theory of the ERV- and REVEALS models</td>
<td>32</td>
</tr>
<tr>
<td>4.2.1 The ERV model and its three submodels</td>
<td>32</td>
</tr>
<tr>
<td>4.2.2 Pollen source area</td>
<td>33</td>
</tr>
<tr>
<td>4.2.3 The REVEALS model (Paper IV)</td>
<td>34</td>
</tr>
<tr>
<td>4.3 Data analysis in papers I, II and III</td>
<td>34</td>
</tr>
<tr>
<td>4.3.1 Pollen assemblages and numerical methods (Paper I)</td>
<td>34</td>
</tr>
<tr>
<td>4.3.2 ERV-model runs (Paper II)</td>
<td>35</td>
</tr>
<tr>
<td>4.3.3 Synthesis of available RPPs in China (Paper III)</td>
<td>35</td>
</tr>
<tr>
<td>4.4 Data collection, chronology and REVEALS runs (Paper IV)</td>
<td>35</td>
</tr>
<tr>
<td>4.4.1 Study sites, data sources, and chronology</td>
<td>35</td>
</tr>
<tr>
<td>4.4.2 Grouping of sites and time windows</td>
<td>36</td>
</tr>
<tr>
<td>4.4.3 The REVEALS model runs and presentation of results</td>
<td>37</td>
</tr>
<tr>
<td>5. RESULTS-SUMMARY OF PAPERS</td>
<td>39</td>
</tr>
<tr>
<td>5.1 PAPER I (Li et al., in press): Modern pollen-climate relationships in north Xinjiang, northwestern China: implications for pollen-based reconstruction of Holocene climate</td>
<td>39</td>
</tr>
<tr>
<td>5.2 PAPER II (Li et al., in revision): Relative Pollen Productivity (RPP) estimates for major plant taxa in cultural landscapes of central eastern China</td>
<td>40</td>
</tr>
</tbody>
</table>
5.3 PAPER III (Li et al., submitted): A review of relative pollen productivity (RPP) estimates for major plant taxa of northern and temperate China and implications for long-term quantitative reconstruction of past plant cover ................................................................. 41

5.4 PAPER IV (Li et al., manuscript): Pollen-based REVEALS quantitative reconstructions of Holocene plant abundance in northern and temperate China .................................................................................... 42

6. DISCUSSION and CONCLUSIONS ............................................................... 45

6.1 The comparative approach: the ability of modern pollen assemblages in distinguishing vegetation types (PAPER I) .................................................. 45

6.2 The mechanistic-modeling approach: Relative Pollen Productivity estimates (PAPERS II and III) ...................................................................... 46

6.2.1 Relevant source area of pollen (RSAP) .......................................... 46

6.2.2 Relative Pollen Productivity estimates (RPPs) ............................... 47

6.2.3 The possible factors behind between-study differences in the estimates of relevant source area of pollen (RSAP) and relative pollen productivity (RPP) ................................................................. 48

6.2.4 Selection of RPP values for the standardized RPP datasets ........... 49

6.3 Holocene vegetation abundance (PAPER IV) ........................................ 50

6.3.1 Differences between pollen percentages and pollen-based REVEALS estimates of plant cover ................................................................. 50

6.3.2 Spatio-temporal Holocene forest dynamics and landscape openness: climate or human induced? ............................................................ 51

ACKNOWLEDGEMENTS ............................................................................... 53

LITERATURE CITED ...................................................................................... 57
ABSTRACT

Reconstruction of past changes in climate and vegetation cover is needed to understand the present climate and vegetation, and project their future changes under assumed future conditions such as climate warming and anthropogenic land-cover change. Projections require that climate and vegetation models are developed and validated, i.e. these models should be able to reproduce past climate and vegetation change to be useful for projections of future changes. Records of pollen deposited in lake bottoms or peat bogs are widely used to reconstruct climate and vegetation, and can provide the information needed to test and validate the models. The aim of this thesis is i) to explore the relationships between modern pollen, vegetation and climate in northern and temperate China based on own studies and syntheses of results from the literature and discuss their use in reconstructions of past climate and plant-cover, and ii) to produce the first reconstruction of plant cover in China for the purpose of climate modelling. A study of the modern pollen-vegetation-climate relationships was performed in northwestern China (Jungar desert and Altay Mountains) using principal components analysis, redundancy analysis, Monte Carlo permutation tests, and variation partitioning. The results were evaluated based on a synthesis of earlier studies in northern and temperate China (Paper I). Relative pollen productivity (RPP) for 18 major plants of cultural landscapes in central-eastern China (Shandong province) was estimated from modern pollen and related vegetation cover using the ERV-model (Paper II). Based on a synthesis and evaluation of all existing RPP estimates in northern and temperate China (10 study areas), standard datasets of RPPs for 31 plant taxa is proposed (Paper III). The standardized RPP dataset from Paper III was used to achieve a pollen-based REVEALS reconstruction of plant cover in northern and temperate China (Paper IV). The major findings of this thesis can be summarized as follows. Paper I: Annual precipitation ($P_{ann}$) is the major climatic factor influencing pollen assemblages, followed by July precipitation ($P_{Jul}$). A variation partitioning analysis suggests that the shared effect of a) January precipitation ($P_{Jan}$) and $P_{Jul}$, b) $P_{Jan}$ and $P_{ann}$,
c) $P_{\text{jul}}$ and annual temperature ($T_{\text{ann}}$), and d) $T_{\text{ann}}$, January temperature ($T_{\text{Jan}}$), and July temperature ($T_{\text{jul}}$) explains a larger portion of the variation in pollen data than the individual effect of each variable. The *Artemisia*/*Chenopodiaceae* relationship is a strong index of aridity and *Artemisia*/*Gramineae* might be a useful index of $P_{\text{ann}}$ and $P_{\text{jul}}$. Paper II: Of the 16 RPPs estimated in the Shandong province, the estimates for 8 taxa are new, i.e. *Castanea*, Cupressaceae, Robinia/Sophora, Anthemis type/Aster type, *Cannabis/Humulus*, Caryophyllaceae, Cruciferae, and *Galium* type are new for China. Trees have in general larger estimates than herbs, except *Robinia/Sophora* (tree with low value: 0.78) and *Artemisia* (herb with high value: 24.7). Paper III: Of the total 31 taxa for which RPP estimates are available in China, 13 taxa have more than 1 value (2 values for 7 taxa, and >2-5 values for 6 taxa). Within these 13 taxa, there are only 5 taxa for which all or most of the available values are similar, i.e. *Artemisia*, Cyperaceae, *Larix*, *Quercus* and *Pinus*. The other 8 taxa have either very variable estimates, i.e. Amaranthaceae/Chenopodiaceae and Ranunculaceae, or as many low as high values, i.e. *Juglans*, *Betula*, *Ulmus*, Lamiaceae, Fabaceae, Asteraceae. To set up a standard dataset of RPPs for the 31 taxa, a number of rules were applied and expert knowledge helped to select the estimates to be included. Paper IV: REVEALS-based reconstructions of past plant cover were performed for 35 regions of northern and temperate China. The REVEALS-based values of plant cover strongly differ from the pollen percentages and provide new insights on past changes in plant composition and vegetation dynamics. As in Europe, pollen percentages generally underestimate the cover of herbs in the vegetation, except for *Artemisia* that is overrepresented by pollen. As expected, human-induced deforestation is highest in eastern China with 3 major phases at ca. 5500, 3000 and 1000 calibrated years before present. It is also in eastern China that the cover of trees was highest during mid Holocene from ca. 8000 in some regions, mainly between ca. 6000-5000 BP and until ca. 3000 calibrated years BP. However, disentangling human-induced from climate-induced vegetation openness remains a challenge and will require thorough comparison of the REVEALS reconstructions of land cover with historical and archaeological information on the impact of settlements and other human activities.

KEY WORDS: climate, pollen-vegetation relationships, relative pollen productivity, REVEALS model, anthropogenic land-cover change, Holocene, China.
SAMMANFATTNING

För att kunna förstå både den nutida kopplingen mellan klimat och vegetation, samt göra bedömningar av framtida scenarier för global uppvärmning och av mänskans påverkad vegetation och täckningsgrad, är det helt nödvändigt att rekonstruera och studera historiska klimat- och vegetationsförändringar. Strategier för att möta framtida miljöproblem kräver att klimat- och vegetationsmodeller utvecklas och valideras, d.v.s. modellerna ska kunna reproducerera historiska klimat- och vegetationsförändringar för att vara användbara för framtida prognoser. Pollenarkiv från sjösediment eller torvskikt i mossor används huvudsakligen för att rekonstruera forntida klimat- och vegetationsförhållanden. De kan ge de nödvändig data som behövs för att testa och validera modellerna. Syftet med detta avhandlingsarbete var att: 1) utforska förhållandet mellan pollen, vegetation och klimatet i norra och tempererade delen av Kina, baserat på egna studier och en syntes av tidigare publicerade resultat med avsikten att utnyttja sambanden mellan klimat och vegetationstäckning. 2) att generera den första rekonstruktionen av vegetationstäckning i Kina för att sedan utnyttja dessa i klimatmodelleringar. En undersökning av nutida pollen-/vegetation/klimatsamband utfördes i Jungaröknen och Altajbergen, nordvästra Kina, med hjälp av numeriska analyser (PCA, principal component analysis; RDA, redundancy analysis, Monte Carlo permutation tests; variation partitioning). Resultaten utvärderades genom jämförelse med syntesen av tidigare publicerade data från norra och tempererade delen av Kina (artikel I). Pollenproduktivitet för 18 växter, vilka dominerar kulturlandskapet i östra mellersta Kina (Shandong provinsen), beräknades med hjälp av nutida pollenprov (pollen som samlas i mossor) och vegetationsinventeringar kring varje pollenprov. Beräkningen gjordes med ”ERV-modellen” som uttrycker relationen mellan växternas täckningsgrad och pollen som sprids från växterna och deponeras på marken (artikel II). Baserat på en syntes och utvärdering av all tillgängliga beräkningar av pollenproduktivitet i norra och tempererade delen av Kina (10 undersökningsområden) skapades ett standarddataset av pollenproduktivitet.
arkeologiska fynd är nödvändig för att vidare bekräfta tolkningarna som presenteras och diskuteras i denna studie.

Översättning från engelska av Geoffrey Lemdahl.
LIST OF PAPERS

This thesis is based on the following Papers, referred to in the text by their roman numerals:


Paper I and II are printed with the kind permissions from the Holocene and Vegetation History and Archaeobotany.
LIST OF PAPERS

This thesis is based on the following Papers, referred to in the text by their roman numerals:


Paper I and II are printed with the kind permissions from the *Holocene* and *Vegetation History and Archaeobotany*. 
Additional work outside the scope of this thesis


1. INTRODUCTION

1.1 Background

Global and regional climate models (AOGCMs and RCMs) are used for producing scenarios of future climate changes, and they provide a basis for assessing the impacts and efficiency of different mitigation and adaptation strategies. Vegetation, ecosystems, and human land-use are inherent parts of the climate system, interacting with the atmosphere and influencing climate development at multiple scales by various processes, for example biogeophysical and biogeochemical processes. Biogeophysical processes act on a more regional to local scale and influence climate by modulating exchanges of energy and water vapor between land and atmosphere through changes in albedo, surface roughness, and evapotranspiration of plants. Biogeochemical processes, on the other hand, act on a global scale and affect e.g. sources and sinks of greenhouse gases (CO2, CH4), aerosols, pollutants and other gases (figure 1). These processes/feedbacks between land-surface and atmosphere may be either positive, amplifying changes in climate, or negative, attenuating climate change, and they represent a major source of uncertainty in projections of climate change under rising greenhouse gas concentrations in the atmosphere (Meehl et al. 2007). Hence, the incorporation of land surface components to account for feedbacks and refine projections of global climate change is a current priority in the climate modelling community, but depends on the availability of data to parameterize and evaluate the models.

Palaeoecological and palaeoclimatic data, provide unique records of how vegetation, land use and climate have co-varied in the past, on time scales relevant to vegetation processes and global change projections. In addition, vegetation models, such as the dynamic vegetation model LPJ-GUESS (Smith et al., 2011) also provide important information about the interaction between vegetation and climate.
LPJ-GUESS (Smith et al. 2001) is a dynamic, process-based vegetation model optimized for application across a regional grid that simulates vegetation dynamics based on climate data input. The modelling results suggested that climate variability (magnitude and temporal structure of variability) is important to vegetation dynamics and that correctly simulated vegetation changes in response to climate variability requires a realistic “baseline” simulation of plant community composition. However, the dynamic vegetation models simulate potential natural vegetation (i.e. climate-driven vegetation). Hence, these models do not take into account the human-induced land cover. In contrast, the anthropogenic land-cover change (ALCC) scenarios HYDE (Klein Goldewijk et al., 2010) and KK10 (Kaplan et al., 2009), which are based on historical population estimates and land-use modeling, have been produced to assess the effects of ALCC on past climate. However, the different ALCC scenarios show large discrepancies in land-cover estimates. Therefore, these existing ALCC scenarios need to be evaluated, and the best way to do this is by using empirical data such as palaeoecological data (Gaillard et al., 2010; Boyle et al., 2011).
Producing quantitative land-cover reconstructions based on pollen data has long been a challenge due to the non-linear relationship between pollen percentages and vegetation abundances (Fagerlind, 1952). Quantitative reconstructions of past vegetation are not only important for evaluating vegetation-climate models and the ALCC scenarios, but also for answering questions related to e.g. human-environment interactions, biodiversity, and nature conservation. Therefore, it is important to develop tools for more robust quantitative palaeoenvironmental reconstructions (Seddon et al., 2014). One such tool is the mechanistic model REVEALS developed by Sugita (2007a). It has the advantage to include both pollen productivity estimates and models of pollen dispersal and deposition in order to account for differences in pollen dispersal and deposition between plant species, the spatial structure of vegetation/land cover, and size and type (bog or lake) of sedimentary basins (Sugita, 2007a).

In this thesis, we used pollen data and the REVEALS model to quantitatively reconstruct past vegetation in northern and temperate China. This thesis will also demonstrate the use of the models’ integration within climate research and archaeological research. The application of integrated land-surfaces from REVEALS (Sugita, 2007a), LPJ-GUESS (Smith et al., 2011) and ALCCs (Kaplan et al. 2010) in regional climate modelling will lead to new assessments of the possible effect of land-use and vegetation changes on climate, e.g. effects of deforestation and afforestation, vegetation shifts, and changes in forest structure. Refined climate models and empirical land-cover reconstructions will shed new light on controversial hypotheses of past climate change and human impacts (e.g. Ruddiman, 2003; Gaillard et al. 2010; Kaplan et al. 2010).

1.2 Pollen vegetation relationship studies

Pollen analysis has been used to infer the long-term dynamics of vegetation since von Post (1916) presented and showed its potential in representing the surrounding vegetation. Pollen analysis is regarded as a remote sensing instrument and used to define the composition of vegetation, today and in the past, that is “the present is the key to the past” (Webb et al., 1978). Like any other sensing instruments, the reliability of the proxy based reconstruction depends on how well the proxy data are interpreted and translated. So the basis of pollen-based vegetation and climate reconstruction is the present-day relationship between pollen rain, vegetation and climate.

In an attempt to produce reliable vegetation reconstructions, empirical and theoretical studies were conducted in the last decades to better understand the relationship between pollen and vegetation. One method involves comparisons of fossil pollen records with modern pollen samples (surface sediment) from different vegetation types in an attempt to find a modern analogue for the past
vegetation. Other methods involve pollen-vegetation relationship models, correction factors estimated from modern pollen assemblages and surrounding vegetation of the modern sampling sites (Davis 1963; Andersen, 1970; Parsons and Prentice, 1981).

1.2.1 Empirical and semi-quantitative comparative approaches to study the relationship between pollen and vegetation

The interpretation of fossil pollen records has been, and still is, largely qualitative or semi-quantitative. The pollen-vegetation relationship studies are mainly based on the comparison of the similarity of modern pollen assemblages with the modern vegetation based on the presence, absence and abundances of indicator species. The analogue techniques are based on the comparison of fossil pollen records with samples taken from analogous environments with known vegetation (e.g. Overpeck et al., 1985). In recent decades, the ordination methods, for instance discriminant analysis (DA), detrended correspondence analysis (DCA), principal components analysis (PCA), redundancy analysis (RDA) and other multivariate numerical techniques, have been widely used in describing the major characteristics of vegetation types and distinguish pollen assemblages from different environment (e.g. Gaillard et al., 1992, 1994; Court Picon et al., 1999; Shen et al., 2006; Luo et al., 2009; Li et al., 2011; Lu et al., 2011; Zhang et al., 2012; Zhao et al., 2012a).

The strategy of the analogue method is to collect environmental information at the site where modern pollen is collected, and then compare the fossil pollen assemblages with modern ones. For each fossil pollen assemblage, one may have several analogue modern pollen assemblages, and the environment variables at the analogue sites are weighted to reconstruct the environment represented by the fossil pollen assemblage. This method was tested and evaluated in China for climatic variables using 1127 surface pollen samples including 61 pollen taxa (Zheng et al., 2009), and it greatly improved the application of the method for palaeoclimatic studies in China (Shen et al., 2006; Jiang et al., 2006; 2010).

These traditional qualitative and semi-quantitative studies provide palynologists important means for the interpretation of fossil pollen records. However, the relationship between pollen and vegetation is non-linear when pollen percentage data is used (Fagerlind, 1952) and the relationships are complex due to a lot of factors related with representation of vegetation by pollen: inter-taxonomic differences in pollen production, dispersal and deposition, variation in spatial structure of vegetation, and differences in size and type of the sediment basin (Sugita, 2007a). These factors are not taken into account in the traditional methods of pollen analysis. In addition, before interpreting a pollen record, one needs to have some idea of the pollen source area of the studied site. Therefore, model-based methods play an essential role in correcting for the factors that influence the pollen representation of
vegetation and in linearizing the pollen-vegetation relationship (Davis, 1963; Andersen, 1970; Parsons and Prentice, 1981).

1.2.2 Models of the relationship between pollen and vegetation

The R-value model
The first specific correction method was proposed by Iversen in 1954 (Iversen, 1949; Faegri and Iversen, 1950). In his study, Iversen divided European tree species into three groups and defined their ratios according to pollen production found in modern pollen. Later, a model called the R-value model was developed by Davis (1963), and in her first model she described that the major factors affecting the representation of pollen taxa in a pollen assemblage were the pollen production of individual species and the distance between the plant and the sampling point of pollen deposition. She calculated a correction factor, the R-value. The R-value allows the transformation of pollen percentages into abundances of a taxon in the source vegetation within a specified area. So the R-value is taxon- and site specific, and it is described by the following function:

\[ r_{i,k} = \frac{P_{i,k}}{V_{i,k}} \] (equation 1)

where

- \( P_{i,k} \) is the pollen proportion of taxon \( i \) at site \( k \) (in the pollen assemblage),
- \( V_{i,k} \) is the vegetation proportion of taxon \( i \) at site \( k \) (in the surrounding vegetation within a specified area),
- \( r_{i,k} \) is the R-value for taxon \( i \) at site \( k \).

In spite of the ingenious idea behind Davis’ R-value model, it has an essential practical limitation, which is the lack of a specific spatial scale of the calculated plant abundance: from how large an area do the pollen derive from? The R-values will vary depending on the size of the vegetation survey, and the vegetation characteristics of the geographical region. Therefore, the application of the R-value model has been limited.

The Andersen’s model
As explained above, the major limitation of Davis’ R-value model is the lack of a clearly defined source area of pollen. This limitation is solved in later models by the inclusion of a background pollen component (Andersen, 1970; Webb et al., 1981; Bradshaw and Webb, 1985; Prentice and Webb, 1986). The model proposed by Andersen (1970) can be expressed as below:

\[ y_{i,k} = \alpha_i x_{i,k} + \omega_i \] (equation 2)

where
$y_{i,k}$ is the pollen deposition rates (PDR) of taxon $i$ at site $k$, 
$x_{i,k}$ is the absolute abundance of taxon $i$ within a fixed distance $k$, 
$\alpha_i$ is a representation term (the slope in figure 2), 
$\omega_i$ is a background component term (the intercept in figure 2) 
$\alpha_i$ and $\omega_i$ are taxon-specific constants;

Fig. 2. The pollen-vegetation relationship for taxon $i$ according to the model of Andersen (1970). The slope of the relationship is the pollen representation factor (or pollen production of taxon $i$) and the $y$-intercept represents the background pollen for taxon $i$.

The rationale of Andersen’s model is that different pollen types have different dispersal ranges, and therefore different source areas. The model was designed for a study where samples were collected from a closed forest, and the pollen data were expressed as semi-absolute values while vegetation data were expressed as absolute values. However, this model is works with absolute data, which is not always easy to get. Usually, we are dealing with percentages.

The Extended R-value (ERV) model
In Andersen’s equation, $\alpha_i$ is a constant representation factor for taxon $i$ (Andersen, 1970), but when vegetation data is expressed in percentage, estimation of $\alpha_i$ is not possible due to the Fagerlind effect (Fagerlind, 1952). Pollen percentages of any given taxon $i$ also depend on the relative abundances of other taxa. Variation in the site dependent factors among sites such as different vegetation is the cause of the Fagerlind effect (Prentice and Webb, 1986). The Fagerlind effect is related to the phenomenon that an increase in vegetation abundance of taxon $i$ will not necessarily correspond to
yi,k is the pollen deposition rates (PDR) of taxon i at site k, xi,k is the absolute abundance of taxon i within a fixed distance k, αi is a representation term (the slope in figure 2), ωi is a background component term (the intercept in figure 2). αi and ωi are taxon-specific constants.

Fig. 2. The pollen-vegetation relationship for taxon i according to the model of Andersen (1970). The slope of the relationship is the pollen representation factor (or pollen production of taxon i) and the y-intercept represents the background pollen for taxon i. The rationale of Andersen’s model is that different pollen types have different dispersal ranges, and therefore different source areas. The model was designed for a study where samples were collected from a closed forest, and the pollen data were expressed as semi-absolute values while vegetation data were expressed as absolute values. However, this model is works with absolute data, which is not always easy to get. Usually, we are dealing with percentages.

The Extended R-value (ERV) model

In Andersen’s equation, αi is a constant representation factor for taxon i (Andersen, 1970), but when vegetation data is expressed in percentage, estimation of αi is not possible due to the Fagerlind effect (Fagerlind, 1952). Pollen percentages of any given taxon i also depend on the relative abundances of other taxa. Variation in the site dependent factors among sites such as different vegetation is the cause of the Fagerlind effect (Prentice and Webb, 1986). The Fagerlind effect is related to the phenomenon that an increase in vegetation abundance of taxon i will not necessarily correspond to an increase in pollen percentage due to the effect of other taxa and the closed system of percentages, the effect becomes progressively stronger with increasing percentages. The “site dependant factor” of Prentice and Webb (1986) overcomes the Fagerlind effect. αi cannot be estimated, but the R-value of taxa i at site k can be calculated. As Davis (1963) pointed out, the R-value of species i at site k is not expected to be the same at different sites, but the ratios of R-values, for example the ratio of R-value of taxon i and R-value of taxon j is expected to be constant from site to site:

\[ \frac{r_{i,k}}{r_{j,k}} = \frac{(p_{i,k}/v_{i,k})}{(p_{j,k}/v_{j,k})} = \frac{(y_{i,k}/v_{i,k})}{(y_{j,k}/v_{j,k})} = \frac{\alpha_i}{\alpha_j} \]


\[ \frac{r_{i,k}}{r_{j,k}} = \frac{(p_{i,k}/v_{i,k})}{(p_{j,k}/v_{j,k})} = \frac{(y_{i,k}/v_{i,k})}{(y_{j,k}/v_{j,k})} = \frac{\alpha_i}{\alpha_j} \]

The ERV model includes two submodels with different assumptions for the background components (Parsons and Prentice, 1981; Prentice and Parsons, 1983). In addition, distance weighted vegetation abundances were used in the ERV model by considering pollen deposition in a sediment basin as the “pollen samples’ view the landscape”. Therefore, the plants growing closer to
the sampling site contribute with more pollen than the plants growing further away (Webb et al., 1981). Empirical and modelling studies found that there are several factors that influence the pollen representation of vegetation (Figure 3). Nevertheless, when the vegetation data are distance-weighted, spatial pattern of source plants are taken into account and, therefore, the relative pollen productivities of taxa (relative to a reference taxon) can be estimated.

**The Prentice-Sugita model**

Several definitions of the source area of pollen have been proposed (e.g. Davis, 1963; Andersen, 1970; Jacobson and Bradshaw, 1981; Prentice 1985, 1988). Most suggested models estimate pollen deposition at a point in the centre of a sedimentary basin (Jacobson and Bradshaw, 1981; Prentice, 1985, 1988). However, sediment cores are often taken from lakes, where mixing of pollen occurs within the lake before it is deposited on the bottom of the lake. Sugita (1993) developed a model in which pollen deposition is integrated over the entire lake surface. Moreover, this model allows estimating the source area of pollen for various lake sizes.

### 1.3 Climate studies in China and debated issues

Oxygen Isotope records from Dongge Cave (Wang et al., 2005) and Sanbao Cave (Shao et al., 2006; Wang et al., 2008) show a strong Asian monsoon in the early Holocene, followed by a progressive, long-term decrease since 7 ka. Moreover, studies found that different regions of China exhibit various climate patterns during the Holocene. An et al. (2000) found that the Eastern Asian Monsoon is asynchronous in different regions of China and, therefore, precipitation reached a maximum 10 ka-8 ka in north-eastern China, 10 ka-7 ka in north-central and northern east-central China, ca. 7 ka-5 ka in the middle and lower reaches of the Yangtze River, ca. 3 ka in southern China, and ca. 11 ka in south-western China. Moreover, a synthesis based on pollen and diatom assemblages, lake levels and geochemistry data from eleven lakes from arid and semi-arid regions showed that climate was driven by westerlies and indicated a dry early Holocene, a wetter early to mid-Holocene, and a moderately wet late Holocene, a development that is pit of phase compared to the climate characteristic of regions driven by monsoon processes (Chen et al., 2008). Recent climate studies on monsoons demonstrated that the maximum wet period occurred in the region dominated by Indian Monsoon during early Holocene (8 ka) (e.g. Tibetan Plateau and southwest China) (Herzschuh, 2006; Wang et al., 2010). In contrast, areas which are dominated by the South-East Asian monsoon (SE Monsoon) and the westerlies (in north-western and north-central China) show maximum wet conditions in mid-Holocene (Herzschuh, 2006; Wang et al., 2010). Moreover, the monsoon patterns and climate dynamics inferred from oxygen isotope-ratio cycles are under debate in
several studies (e.g. Liu et al., 2015). Pollen-inferred climate studies are regarded as one of the most reliable ones within the proxies that were used so far (e.g. Herzschuh, 2006; Zhao et al., 2009a, 2009b; Zhao and Yu 2012.) There is a large number of empirical and modelling studies on Holocene climate change and climate-vegetation feedbacks in China (e.g. Fu et al., 2008). For instance, using the regional climate model RegCM2, Jiang and Zhang (2006) show that vegetation and ocean feedbacks can amplify the response of the East Asian climate to the Earth’s orbital parameters and atmospheric CO₂ concentration in mid-Holocene. Ni et al. (2006) carried out a series of nine simulations using the Lund–Potsdam–Jena Dynamic Global Vegetation Model (LPJ-DGVM) aiming to explore the impacts of climate variability and Holocene changes in variability (as simulated by the Fast Ocean-Atmosphere Model, FOAM) on vegetation in three forest-dominated regions of China. This study suggests that the distribution of major trees in northeast and east China changed greatly in terms of northward shifts of vegetation boundaries by 200-500 km. The forest–grassland boundary shifted 200-300 km westwards from today’s boundary, while the tree line in the Tibetan Plateau region was 300-500 m higher than its present elevation.

1.4 Holocene vegetation history in China

In the last decades, many studies have been conducted in China to reconstruct past vegetation at the broad regional scale using multiple fossil pollen records (e.g. Yu et al., 2000; Ren and Beug, 2002; Ren, 2007; Zhao et al., 2009a, 2009b; Ni et al., 2010, 2014; Zhao and Yu, 2012; Cao et al., 2015; Tian et al., 2016). The first attempt at mapping past vegetation based on pollen records was performed by Ren and Beug (2002) who produced vegetation maps for five trees (Picea/Abies, Pinus, Betula, Quercus, and Ulmus) and two herb taxa (Artemisia and Chenopodiaceae) based on 142 pollen records. The maps show significant changes of land cover in the study area through the Holocene: in the south-eastern region, arboreal pollen (AP) generally expanded in the early Holocene, reaching their maximum at 6 or 4 ka BP, and then decreased during the late Holocene. The evident drop in AP percentages, and therefore the decline of forests, occurred after 6 ka BP in the south-easternmost regions, especially in the middle and lower reaches of the Yellow River, and may have been caused by the expansion of farming since the Yang Shao Culture period. Ren (2007) mapped Holocene forest cover in China, and found an increase of AP percentages in the pollen diagrams during early Holocene (10 ka BP to 6 ka BP), and a maximum peak level at 6 ka BP in all regions, except in north-eastern China where the peak is observed at 8 ka BP. A marked decline of AP percentages was found after 6 ka BP in all regions except in northeast China. In contrast, AP percentages in northeast China increased significantly from 6 ka BP and reached a peak level at 2 ka BP. Ren (2007) concluded that the AP
percentage changes before 6 ka BP were due to climate change, and the marked drop of AP after 6 ka BP to human activities. The biomisation work of Yu et al. (1998) based on 8 taxa from a selection of 112 pollen records with an assumed weak human impact is the first quantitative vegetation reconstruction in China. The authors found that the forest zones in northeast China were systematically shifted northwards at 6 ka BP, while the area of desert and steppe vegetation was reduced compared to present. On the Tibetan Plateau, forest vegetation extended to higher elevations and the tundra areas became smaller than today. After this study, several subsequent attempts were conducted to improve the biome-based vegetation reconstructions (Yu et al., 2000; Ni et al., 2010). A recent study on biome distribution based on 19 natural biomes and one anthropogenic biome found that the natural modern biome patterns show a good agreement with the actual, modern vegetation but failed to reconstruct correctly the anthropogenic biome due to non-species level of pollen identification and lack of classification of anthropogenic biomes and land-cover (Ni et al., 2014).

Among the syntheses of pollen records China those by Zhao et al. (2009a) for the east monsoon region, Zhao and Yu (2012) for the monsoon margin region and Zhao et al. (2009b) for the arid, semi-arid region of China are very useful. These syntheses reveal that abrupt changes occurred in pollen assemblages at 11–10 ka, 6–5 ka and 2–1 ka BP at most study sites. Cao et al. (2015) inferred the spatial and temporal distributions of 14 key arboreal taxa for East Asia based on the frequencies (presence, dominance) of those taxa in 251 pollen records and found that the abundances of most trees decreased since the late mid-Holocene, whereas no obvious changes were observed from pollen percentages in their spatial extent. Using the same pollen dataset as Cao et al. (2015) and the Modern Analogue Technique (MAT), Tian et al. (2016) demonstrated very strong temporal changes of forest cover on the eastern margin of the Tibetan Plateau, less strong changes in the forest-steppe transition zones in north-central China and the west part of the Tibetan Plateau since 9 ka BP. In contrast, no notable changes in northeast China were observed from 9 ka BP to present except an increase of tree cover north of 47° at 6 ka BP.

1.5 Archaeological studies on the origin of Neolithic agriculture and related human-induced land use

The human impact on land cover started as early as mid-Holocene in Europe (e.g. Trondman et al., 2015). In China, cultivation of broomcorn millet and foxtail millet goes back to around 10 ka in northern China (Chang, 1986; Liu et al., 2007; Barton et al., 2009; Lu et al., 2009), and rice cultivation began at 7.5 ka in the coastal wetlands of eastern China (Zong et al., 2007) to then spread in the middle and lower reach of the Yangtze River as well as
throughout southern China during the late Holocene (Gong et al., 2007). Archaeological evidence demonstrates that the existence of Neolithic agriculture at the northeast margins of the Tibetan Plateau goes back to 5.2 ka BP (Barton, 2016), and the sustained settlement of humans on the Tibetan Plateau at altitudes > 2500 m up to over 4000 m a.s.l. started from ca. 3.6 ka (Chen et al. 2015). However, the role of human activity on long-term land-cover change is not yet fully understood. There are several methods available to infer the history of land-cover after 1700 AD (Gaillard et al., 2010), while reconstruction is much more difficult for older times. Recently scenarios of anthropogenic land-cover and land-use changes were inferred from human-population density data and assumptions on cleared land per person. For instance, Liu and Tian (2010) reconstructed land use and land cover change (LULCC) in China for the last 300 years based on high resolution satellite-image data and historical archives. Their study find that the forest cover decreased by 22% while cropland increased by 42% and urban areas (including urban and rural settlements, factories, quarries, mining and other urban built-up land) strongly expanded during the last 300 years.
This thesis is a contribution to the Swedish strategic research area "ModElling the Regional and Global Earth systems - MERGE" (hosted by the Faculty of Science at Lund University) and adopts the same research strategy as the Swedish research project "LAND cover – CLIMate interactions in NW Europe during the Holocene - LANDCLIM" (Gaillard et al., 2010). The overall objective of the LANDCLIM project was to reconstruct Holocene land cover using pollen data in order to better understand Holocene land cover - climate interactions and to improve vegetation and climate models for better projections of future changes. The major steps of LANDCLIM are:

1. To use Sugita's model REVEALS (Sugita, 2007a) to estimate the Holocene cover of major taxa using well dated pollen records.
2. Compare the pollen- inferred REVEALS land-cover reconstructions with simulated vegetation cover from LPJ-GUESS (Smith et al., 2001) and with existing anthropogenic land-cover change (ALCC) scenarios (e.g. Kaplan et al., 2009; Klein Goldewijk et al., 2010).
3. Simulate climate for Europe at selected, contrasting Holocene time periods in terms of human-induced landscape openness – using a regional climate model (RCA3; Samuelsson et al., 2011) forced by lateral boundary conditions and sea-surface temperatures from different global climate models together with regional land surface boundary conditions from LPJ-GUESS, existing ALCC scenarios, REVEALS land-cover reconstructions, and relevant integrations between them.
4. Evaluate the possible effects of two different historical processes (compared with a baseline of present-day land cover) on the climate development, e.g. via the influence of forested vs. non-forested land cover on energy and water fluxes between the atmosphere and land surface: climate-driven changes in vegetation, and human-induced changes in land cover.

The major objective of this thesis is to achieve quantitative pollen-based reconstructions of Holocene vegetation cover for northern and temperate China using the REVEALS model in order to provide empirical descriptions...
2. OBJECTIVES

This thesis is a contribution to the Swedish strategic research area “ModElling the Regional and Global Earth systems - MERGE” (hosted by the Faculty of Science at Lund University) and adopts the same research strategy as the Swedish research project “LAND cover – CLIMate interactions in NW Europe during the Holocene - LANDCLIM” (Gaillard et al., 2010). The overall objective of the LANDCLIM project was to reconstruct Holocene land cover using pollen data in order to better understand Holocene land cover-climate interactions and to improve vegetation and climate models for better projections of future changes. The major steps of LANDCLIM are:

1. To use Sugita’s model REVEALS (Sugita, 2007a) to estimate the Holocene cover of major taxa using well dated pollen records.
2. Compare the pollen-inferred REVEALS land-cover reconstructions with simulated vegetation cover from LPJ-GUESS (Smith et al., 2001) and with existing anthropogenic land-cover change (ALCC) scenarios (e.g. Kaplan et al., 2009; Klein Goldewijk et al., 2010).
3. Simulate climate for Europe at selected, contrasting Holocene time periods in terms of human-induced landscape openness – using a regional climate model (RCA3; Samuelsson et al., 2011) forced by lateral boundary conditions and sea-surface temperatures from different global climate models together with regional land surface boundary conditions from LPJ-GUESS, existing ALCC scenarios, REVEALS land-cover reconstructions, and relevant integrations between them.
4. Evaluate the possible effects of two different historical processes (compared with a baseline of present-day land cover) on the climate development, e.g. via the influence of forested vs. non-forested land cover on energy and water fluxes between the atmosphere and land surface: climate-driven changes in vegetation, and human-induced changes in land cover.

The major objective of this thesis is to achieve quantitative pollen-based reconstructions of Holocene vegetation cover for northern and temperate China using the REVEALS model in order to provide empirical descriptions
of vegetation cover that can be used in climate modelling in a similar way as it was performed in Europe (Trondman et al., 2015). Estimates of pollen productivity for the plant taxa to be reconstructed are required for the application of REVEALS. Therefore, this thesis focuses on studying pollen-vegetation relationships and estimating relative pollen productivity (RPP) for major taxa in cultural landscapes of eastern China in order to produce quantitative reconstructions of past land cover in temperate China using pollen records, values of RPP from earlier studies and this thesis, and the REVEALS model. All existing RPP datasets for China, including the RPP produced in this thesis, are reviewed in order to set up a reliable RPP dataset to be used for the application of REVEALS. More specifically, the aims of this thesis are to:

(1) Study the relationship between pollen and vegetation using traditional methods, i.e. ordination techniques and other numerical methods to study the relationships between modern pollen assemblages (expressed as pollen percentages) and related vegetation and climate in order to evaluate whether modern pollen assemblages reflect present-day vegetation types correctly, and whether climate variables can be inferred from pollen with confidence (PAPER I).

(2) Estimate the relative pollen productivity (RPP) and relevant source area of pollen (RSAP) for major plant taxa in cultural landscapes of eastern China using the ERV model, in order to complement earlier RPP studies in China that focused on desert, steppe, meadow and forest vegetation (PAPER II).

(3) Review the methodology used for calculating RPP in previous studies that are available for China and evaluate all the existing RPP (i.e. the RPP from previous studies and the RPP from this study). Moreover, we study the factors that influence the RPP values and the size of the RSAP in China. Taking into account all possible variables that may influence RSAP, a dataset of best RPP and fall speeds of pollen (FSP, another important parameter for the application of REVEALS) was established for REVEALS-based land-cover reconstructions in China (PAPER III).

(4) Use the REVEALS model (Sugita, 2007a) to estimate the Holocene regional cover of major taxa using well-dated Holocene pollen records from northern and temperate China, and discuss the obtained land-cover reconstructions in terms of the impact of climate change and human impact on vegetation dynamics, highlighting the insights that quantitative REVEALS reconstructions add to the interpretation of pollen percentage data (PAPER IV).
3. STUDY REGION

The study regions related for each Paper are compiled in Figure 4. There are two small study regions for case studies (Paper I and Paper II), synthesis available relative pollen productivities in northern China (Paper III) and pollen-based reconstruction of Holocene vegetation abundances in northern and temperate China (Paper IV).

Fig 4. Study region and study areas for Papers I to IV. Paper I: Eastern edge of the Jungar Basin and western Altay Mountains; Paper II: cultural landscapes of the Shandong province; Paper III: from west to east, Tibetan Plateau, Xinglong Mountains, Alashan Plateau (western Inner Mongolia), Taiyue Mountains, Central Inner Mongolia, Eastern Inner Mongolia, Shandong Province, Northeastern Inner Mongolia, and two areas in the Changbai Mountains. Paper IV: temperate and temperate-subtropical vegetation zones of China, north of latitude 30°.
3.1 The Jungar Basin and the western Altay Mountains

The study of the relationship between pollen assemblages, modern vegetation, and climate was conducted along a transect from the eastern edge of the Jungar Basin to the Altay Mountains in the west (44-49° N, 85-90° E) (Paper I, Figure 1). The 66 modern pollen sample sites are located at altitudes ranging from 500 to 2500 m a.s.l. and characterized by continuous temperature and precipitation gradients. The mean annual precipitation and temperature range from 120 mm to 400 mm and from 1°C to 5°C, respectively, with the maximum mean monthly precipitations and temperatures of the year occurring during the period June-September and the lowest mean monthly temperatures of the year occurring in January (Li, 1991). According to the Editorial Board of Xinjiang Vegetation (1978) and the Vegetation Atlas of China (Hou, 2001), the study region covers four vegetation types, i.e. forest, alpine meadow (henceforth referred as meadow), steppe, and desert. Moreover, marsh areas occur within the forested areas (Paper I, Figure 1). Detailed information on the sites and their vegetation is described in Paper I Table I.

3.2 The cultural landscape of Shandong

In order to calculate relative pollen productivities for common taxa in cultural landscapes, the low mountains region of Shandong (Paper II figure 1) was selected because of the traditional agriculture practices and landscape structures still existing today. The Shandong province is one of the most important agricultural provinces located in the lower reach of the Yellow River drainage basin in central-eastern China. The overall landscape is the result of long-term human modification of the low mountains into cultivated terrasses (fruits trees and crops) still partly managed in a traditional way, and the flat low land characterized by modern agriculture with large cultivated fields and planted trees (Figure 5). The study region is ca. 100km (longitudinal distance) x 200km (latitudinal distance) large, and located between 35°00’ - 36°30’ N and 117°00’ - 118°30’ E. The elevations are < 500m a.s.l except for some mountains > 1000m a.s.l. high.

3.3 Synthesis of the relative pollen productivity (RPP) estimates available in China

Paper III is a synthesis of all available RPP estimates in China from ten studies. The landscape varies among the studies and all the studies are conducted mainly in natural vegetation region, Except the RPP estimates from Shandong, in which the samples were collected in cultural landscapes.
3.1 The Jungar Basin and the western Altay Mountains

The study of the relationship between pollen assemblages, modern vegetation, and climate was conducted along a transect from the eastern edge of the Jungar Basin to the Altay Mountains in the west (44-49º N, 85-90º E) (Paper I, Figure 1). The 66 modern pollen sample sites are located at altitudes ranging from 500 to 2500 m a.s.l. and characterized by continuous temperature and precipitation gradients. The mean annual precipitation and temperature range from 120 mm to 400 mm and from 1 ºC to 5 ºC, respectively, with the maximum mean monthly precipitations and temperatures of the year occurring during the period June-September and the lowest mean monthly temperatures of the year occurring in January (Li, 1991). According to the Editorial Board of Xinjiang Vegetation (1978) and the Vegetation Atlas of China (Hou, 2001), the study region covers four vegetation types, i.e. forest, alpine meadow (henceforth referred as meadow), steppe, and desert. Moreover, marsh areas occur within the forested areas (Paper I, Figure 1). Detailed information on the sites and their vegetation is described in Paper I Table I.

3.2 The cultural landscape of Shandong

In order to calculate relative pollen productivities for common taxa in cultural landscapes, the low mountains region of Shandong (Paper II figure 1) was selected because of the traditional agriculture practices and landscape structures still existing today. The Shandong province is one of the most important agricultural provinces located in the lower reach of the Yellow River drainage basin in central-eastern China. The overall landscape is the result of long-term human modification of the low mountains into cultivated terrasses (fruits trees and crops) still partly managed in a traditional way, and the flat low land characterized by modern agriculture with large cultivated fields and planted trees (Figure 5). The study region is ca. 100km (longitudinal distance) x 200km (latitudinal distance) large, and located between 35°00' - 36°30' N and 117°00' - 118°30' E. The elevations are < 500m a.s.l except for some mountains > 1000m a.s.l high.

3.3 Synthesis of the relative pollen productivity (RPP) estimates available in China

Paper III is a synthesis of all available RPP estimates in China from ten studies. The landscape varies among the studies and all the studies are conducted mainly in natural vegetation region, except the RPP estimates from Shandong, in which the samples were collected in cultural landscapes.

3.4 The REVEALS application

RPP estimates are needed for the application of the REVEALS model. All the available RPP (PAPER III) in China are from northern and temperate China, therefore the region was selected for the first attempt at REVEALS-based reconstructions of Holocene vegetation abundance in China. In addition, the REVEALS model requires pollen counts. Unfortunately, no pollen counts data were available from the boreal forest in northeastern China (vegetation Zone I). The REVEALS model was applied fossil pollen records from 95 lake and
bog sites located in sixth vegetation zones in northern and temperate China, (Paper IV, figure 1).
The vegetation zones in this study follow the classification by Wu (1980). They are II. Coniferous-broadleaved deciduous mixed forest in northeastern China; III. Temperate deciduous forest in central eastern China; IV. Subtropical broadleaved evergreen and deciduous forest in central-southern China; VI. Temperate steppe in central-northern China; VII. Temperate desert in northwestern China; and VIII. Highland vegetation on the Tibetan Plateau.
4. MATERIAL AND METHODS

4.1 Site selection, sampling and vegetation survey (Papers I and II)

Pollen-vegetation-climate relationships (Paper I)
In order to get continuous precipitation, temperate and elevation gradients, surface soil samples and moss polsters were collected along a transect from the eastern edge of the Jungar basin in the south to the western Altay Mountains in the north. For each site, ten sub-samples collected within an area of 1m² were mixed into one single sample. The vertical projection of stems and leaves for herbs and shrub and tree crowns on the ground of dominant plants were estimated within an area of 50m x 50m in forest and 20m x 20 m in desert, steppe and meadow landscapes.

Relative pollen productivity in cultural landscapes (Paper II)
The sites were selected according to three criteria:
1. Sites should include taxa characteristic of cultural landscapes;
2. Vegetation should be relatively homogenous in terms of taxa composition and spatial distribution, because the ERV model assumes constant “background pollen” for all sites included in the analysis;
3. Sites should be randomly distributed in the landscape because random distribution of sites in the study area is necessary for good results of the maximum likelihood method (Broström et al., 2005).
Before vegetation survey, moss polsters were collected within and area of 1 m² for each site and soil particles were removed. Detailed vegetation survey was conducted within 0-100 m of the pollen sample following Bunting et al. (2013). For extraction of vegetation data beyond 100 m satellite images within an area of 1500 m around each pollen site were downloaded from Google Earth professional, which were then georeferenced using the open source geographic information system and ArcGIS Desktop (ArcView 10.0). Landscape/vegetation features (18 in total) were assigned by using maximum likelihood classification with ArcView (Arc View 10.0) based on the satellite
image. Plant composition in each landscape unit was inferred from the field surveys.

4.2 Theory of the ERV- and REVEALS models

4.2.1 The ERV model and its three submodels

There are three ERV sub-models. The difference between them is the assumption chosen to describe the “background” pollen component. In sub-model 1, the background component represents the proportion of pollen coming from beyond the source area of pollen (Parsons and Prentice 1981). In sub-model 2, the background component represents the ratio of pollen proportion coming from beyond the source area of pollen to the total sum of plant abundance within the area of the vegetation-data survey (Parsons and Prentice 1981). In sub-model 3, the background component represents the amount of pollen loading coming from beyond the relevant source area of pollen (Sugita 1993; Sugita 1994). In sub-models 1 and 2, the relative cover (in percentage or proportion) of each plant taxon (harmonized with a pollen morphological type) is used. In sub-model 3, the absolute cover of each plant taxon (in m$^2$/m$^2$) is used.

The Sugita-Prentice model (ERV sub-model 3)

The basic equation of the Sugita-Prentice model is as below:

$$y_{i,k} = \alpha_i \cdot \psi_{i,k} \quad \text{(Equation 1)}$$

Where:
- $y_{i,k}$ is the pollen loading of taxon $i$ at site $k$
- $\alpha_i$ is the relative pollen productivity of taxon $i$
- $\psi_{i,k}$ is the distance-weighted plant abundance of taxon $i$ at site $k$, i.e.:

$$\psi_{i,k} = \int_{R}^{Z} x_i(z) \cdot g_i(z) \cdot dz \quad \text{(Equation 2)}$$

Where:
- $x_i(z)$ is the cover of taxon $i$ at distance $z$
- $g_i(z)$ is the species-specific pollen-dispersal function at distance $z$
- $R$ is the radius of the sedimentary basin
- $Z$ is the distance from the sampling site to the pollen producing vegetation.

$\psi_{i,k}$ can be divided into two parts, the first part representing the vegetation within the RSAP and the second part representing the vegetation beyond the RSAP. Equation (1) becomes:
There are three ERV sub-models. The difference between them is the assumption chosen to describe the “background” pollen component. In sub-model 1, the background component represents the proportion of pollen coming from beyond the source area of pollen (Parsons and Prentice 1981). In sub-model 2, the background component represents the ratio of pollen proportion coming from beyond the source area of pollen to the total sum of plant abundance within the area of the vegetation-data survey (Parsons and Prentice 1981). In sub-model 3, the background component represents the amount of pollen loading coming from beyond the relevant source area of pollen (Sugita 1993; Sugita 1994). In sub-models 1 and 2, the relative cover (in percentage or proportion) of each plant taxon (harmonized with a pollen morphological type) is used. In sub-model 3, the absolute cover of each plant taxon (in m²/m²) is used.

The Sugita-Prentice model (ERV sub-model 3)

The basic equation of the Sugita-Prentice model is as below:

\[ y_{i,k} = P_i \cdot \int_{0}^{Z_{RSAP}} x_i(z) \cdot g_i(z) \cdot dz + P_i \cdot \int_{Z_{RSAP}}^{Z_{\text{max}}} x_i(z) g_i(z) dz \]  

(Equation 3)

Where

- \( P_i \) is the relative pollen productivity of taxon \( i \)
- \( Z_{RSAP} \) is the distance from the pollen sampling point to the relevant source area of pollen (RSAP). See below for the definition of RSAP:
- \( Z_{\text{max}} \) is the maximum distance from the pollen sampling point with available vegetation data (if estimating \( P_i \) from field pollen and vegetation data) or maximum distance from the pollen sampling point from which pollen grains are coming.

4.2.2 Pollen source area

There are several ways to define pollen source area. When homogeneous vegetation is assumed, a “characteristic radius” can be estimated for each taxon and for a given sized site/basin by using taxon-specific fall speeds of pollen (Prentice, 1985, 1988). Hellman et al. (2008b) provide an example of the “characteristic radius” for Quercus (oak), i.e. when the basin radius is 100 m, the “50% characteristic radius” for Quercus is 2 km, which means that the radius of the area from which 50% of the total Quercus pollen comes from is 2 km, and the “70% characteristic radius” for Quercus is 13.6 km. The “characteristic radius” of a taxon depends on the fall speed of the pollen grains, hence, light pollen types that are transported over longer distances by wind than heavy pollen types will have a larger characteristic radius than the heavy ones, and therefore they will be better represented in larger basins (Prentice, 1985). Many empirical and modelling studies found that the source area increases as the basin size increases (Bradshaw and Webb, 1985; Prentice et al., 1987; Sugita, 1994). When estimating the “characteristic radius” of a taxon, the vegetation is assumed to be homogeneous (not patchy). However, in real landscapes, the vegetation is patchy. Sugita (1994) suggested another definition of pollen source area, the relevant source area of pollen (RSAP), which is especially applicable in mosaic, patchy vegetation. The RSAP was defined by Sugita (1994) as the distance from a given basin beyond which the correlation between pollen and vegetation do not improve and for which the background pollen (proportion of pollen coming from long distances, i.e. from beyond the RSAP) is constant between sites (Sugita, 1993). The RSAP can be estimated for a group of pollen sites by comparing pollen proportions with distance-weighted plant abundances along increasing distance from the sampling sites. The goodness-of-fit of the correlation between pollen loading and distance-weighted vegetation with the ERV model will increase with distance until the RSAP radius is reached. To estimate the RSAP, the
maximum likelihood method is applied. (More detailed information regarding the RSAP is described in Paper II Appendix 1.)

The moving-window linear-regression technique (Gaillard et al., 2008) is often used to identify the RSAP in an objective way (e.g. Hellman et al., 2009b). It is a statistical method that defines the RSAP distance (or radius) as the central point of a moving-window regression line where the slope of the line is statistically not different from zero. This method overcomes the subjectivity that may occur if the RSAP is identified by visual inspection of the log likelihood-scores curve.

4.2.3 The REVEALS model (Paper IV)

The “Regional Estimates of VEgetation Abundance from Large Sites” (REVEALS) model was developed by Sugita (Sugita, 2007a). It is an inverse form of the Extended R-value model (Parsons and Prentice, 1981; Prentice and Parsons, 1983) with the assumptions that RPP and background pollen components are constant over time. It accounts for inter-taxonomic differences in pollen productivity, pollen dispersal and deposition properties as well as the size and type of the sedimentary basin from which the pollen data is derived. Given that pollen counts, relative pollen productivity (RPP) estimates, and fall speed of pollen are available, vegetation abundances (proportion/percentages) can be estimated by applying the REVEALS model. Additional parameters needed for the application of REVEALS are the size and type of the sedimentary basin, wind speed, and atmospheric conditions.

For a more detailed description of the REVEALS model see e.g. Sugita (2007a) and Trondman et al. (2015).

4.3 Data analysis in papers I, II and III

4.3.1 Pollen assemblages and numerical methods (Paper I)

Pollen percentage and concentration diagrams were plotted (Paper I, Fig. 2). An unconstrained incremental sum of squares (ISS) cluster analysis was conducted in order to show which vegetation types that are best characterized by their pollen assemblages. Multivariate analyses were implemented in the program CANOCO (ter Braak and Šmilauer, 2002) to explore the relationships between modern pollen assemblages, vegetation, and environmental variables. The major purposes of the multivariate analyses were to: (i) assess how well pollen assemblages characterize the different vegetation types (Principal Component Analysis, PCA), (ii) analyse the relationships between pollen assemblages and environmental variables, climatic characteristics in particular (redundancy analysis, RDA, and Monte Carlo permutation tests), and (iii) quantify the relative percentage of variation in the pollen percentage data explained by individual or groups of environmental variables (variation partitioning, Borcard et al., 1992).
4.3.2 ERV-model runs (Paper II)

Data preparation for running the ERV model
For each set of vegetation data (0-100, 100-1500m), the mean absolute cover (m²/m²) of the plant taxa were calculated for each 1 meter increment. Fall speed of pollen for a first selection of 26 target taxa were estimated according to the Stoke’s law (Gregory, 1973) of particle settling velocity based on pollen morphological grain size and density of each type.

ERV model runs
A series of ERV-model runs were performed. Two different methods of vegetation distance-weighting were used; i) the taxon-specific weighting of Prentice’s model (Prentice, 1985, 1988), and ii) the inverse distance model. ERV sub-model 1 and 2 (Parsons and Prentice, 1981; Prentice and Parsons, 1983) and ERV sub-model 3 (Sugita, 1993) were run to compare the results between different combinations of sub-models and vegetation weighting methods. Two steps of the ERV-model were ran, the first run was based on 24 taxa, and the second, final run was based on 18 taxa with best gradients/linear relationships in the scatter plot of absolute vegetation related to pollen.

4.3.3 Synthesis of available RPPs in China (Paper III)
For each available RPP study in China, a description of the landscape structure was provided, and the methods for site distribution, sampling collection, vegetation data collection, and ERV-model runs were reviewed. The influencing factors that may affect the RPP and RSAP were discussed. Based on an evaluation of all RPP values available, a standardized RPP dataset was proposed for use in applications of the REVEALS model in temperate China. The RPP dataset can serve for other simulation models that have RPPs as input data (e.g. the Multiple Scenario Approach, MSA, Bunting et al., 2009). In order to set up a standard RPP dataset for the study region, we followed similar criteria as those of Mazier et al (2012) for standards 2 and 3 in Europe. Standard 1 was not useful in the case of this study.

4.4 Data collection, chronology and REVEALS runs (Paper IV)

4.4.1 Study sites, data sources, and chronology
The selection of study regions for the application of REVEALS is restricted by the availability of relative pollen productivity estimates (RPP) and fall speed of pollen (FSP). In China, RPP and FSP for 27 pollen taxa are now available from the western, northern, northeastern and central eastern parts of the country (Paper III). Therefore, the REVEALS-based vegetation reconstructions presented in this thesis (Paper IV) focus on the northern and
temperate parts of China. In total, 29 pollen types are included in the REVEALS runs presented in this thesis (RPPs for 27 plant taxa presented in Paper III, and two additional plant taxa, *Castanopsis* and *Cyclobalanopsis*, for which the RPPs of *Castanea* respectively *Quercus* are used. The chronology of the pollen records is essential and should be as good as possible to ensure that the pollen records that are used for the REVEALS runs are asynchronous at a regional scale. However, the restriction of selection criteria will decrease the available records significantly. There are a large number of pollen records available in Chinese pollen databases (Cao et al., 2013), however, many have too low resolution. The pollen records used in this thesis (Paper IV) are selected based on three criteria: the records should have (1) a reliable chronology with a minimum of 3 dating control points over the Holocene; (2) a high sampling resolution with a minimum of 500 years per sample; (3) raw pollen counts. Except for the records from Cao et al. (2013), some other records from individual data contributors, that are not included in the Cao et al. database, were also collected. A total of 95 lake and bog pollen records met the criteria and were selected for the REVEALS-based land-cover reconstructions. All chronologies used in this study were carefully checked to ensure their reliability, and for some sites new age-depth models were performed using the BACON program (Blaauw and Christen, 2011). The chronologies are given as kilo calendar years BP (abbreviated ka BP, or simply ka).

4.4.2 Grouping of sites and time windows

Sugita (2007a) demonstrated by simulations that the higher the pollen counts (i.e. more pollen grains per sample), the smaller the REVEALS standard errors. Studies in southern Sweden have also showed that the standard errors of the REVEALS estimates decrease with increasing size of pollen counts and increasing number of sites used for the REVEALS reconstruction (Hellman et al., 2008a and 2008b; Trondman et al., 2016). The REVEALS model was developed for regional vegetation reconstructions using pollen data from large lakes, however, simulations have shown that REVEALS can provide reliable estimates of the regional vegetation abundance even when multiple small sites (< ca. 50 ha) are used, although the standard errors are smaller when large lakes are used than when several smaller sites are used for the reconstruction (Sugita, 2007a; Trondman et al., 2016). Trondman et al. (2016) also showed that the distance between the small sites and vegetation boundaries is a determinant factor for the accuracy of the REVEALS-based vegetation reconstructions. Therefore, in order to obtain reliable REVEALS reconstructions, pollen sites located in the same biogeographical regions are grouped and REVEALS was run for each site group. In total, this resulted in 35 site groups. By grouping many sites together, i.e. using the “multiple site approach”, we maximize the number of pollen samples for each REVEALS run.
temperate parts of China. In total, 29 pollen types are included in the REVEALS runs presented in this thesis (RPPs for 27 plant taxa presented in Paper III, and two additional plant taxa, *Castanopsis* and *Cyclobalanopsis*, for which the RPPs of *Castanea* respectively *Quercus* are used. The chronology of the pollen records is essential and should be as good as possible to ensure that the pollen records that are used for the REVEALS runs are asynchronous at a regional scale. However, the restriction of selection criteria will decrease the available records significantly. There are a large number of pollen records available in Chinese pollen databases (Cao et al., 2013), however, many have too low resolution. The pollen records used in this thesis (Paper IV) are selected based on three criteria: the records should have (1) a reliable chronology with a minimum of 3 dating control points over the Holocene; (2) a high sampling resolution with a minimum of 500 years per sample; (3) raw pollen counts. Except for the records from Cao et al. (2013), some other records from individual data contributors, that are not included in the Cao et al. database, were also collected. A total of 95 lake and bog pollen records met the criteria and were selected for the REVEALS-based land-cover reconstructions. All chronologies used in this study were carefully checked to ensure their reliability, and for some sites new age-depth models were performed using the BACON program (Blaauw and Christen, 2011). The chronologies are given as kilo calendar years BP (abbreviated ka BP, or simply ka).

4.4.2 Grouping of sites and time windows

Sugita (2007a) demonstrated by simulations that the higher the pollen counts (i.e. more pollen grains per sample), the smaller the REVEALS standard errors. Studies in southern Sweden have also showed that the standard errors of the REVEALS estimates decrease with increasing size of pollen counts and increasing number of sites used for the REVEALS reconstruction (Hellman et al., 2008a and 2008b; Trondman et al., 2016). The REVEALS model was developed for regional vegetation reconstructions using pollen data from large lakes, however, simulations have shown that REVEALS can provide reliable estimates of the regional vegetation abundance even when multiple small sites (< ca. 50 ha) are used, although the standard errors are smaller when large lakes are used than when several smaller sites are used for the reconstruction (Sugita, 2007a; Trondman et al., 2016). Trondman et al. (2016) also showed that the distance between the small sites and vegetation boundaries is a determinant factor for the accuracy of the REVEALS-based vegetation reconstructions. Therefore, in order to obtain reliable REVEALS reconstructions, pollen sites located in the same biogeographical regions are grouped and REVEALS was run for each site group. In total, this resulted in 35 site groups. By grouping many sites together, i.e. using the "multiple site approach", we maximize the number of pollen samples for each REVEALS run.

We use the same time windows as in the LANDCLIM project, i.e. 0.1 ka-years’, 0.25 ka-years’, and 0.35 ka-years’ time windows around 0.05 ka BP, 0.2 ka BP, and 0.5 ka BP, respectively, and 0.5 ka-years’ time windows between 0.7 and 11.7 ka BP. When presenting the results we chose to assign each time window to ca. its mid-point, e.g. the time window 2.7-3.2 ka BP will be presented as 3 ka BP, the time window 5.7-6.2 ka BP will be presented as 6 ka BP etc.

4.4.3 The REVEALS model runs and presentation of results

For each site group the REVEALS model was run separately for lakes and bogs in each of the 35 site groups using Sugita’s pollen dispersal-deposition model for lakes (Sugita, 1993) and Prentice’s model for bogs (Prentice, 1985, 1988). When running the REVEALS model \(Z_{\text{max}}\) (the maximum spatial extent of the regional vegetation) was set to 100 km, and wind speed was set to 3 m/s. The RPP dataset used for the vegetation reconstructions is based on standard 3 in Li et al., in prep. A REVEALS mean estimate for lakes and bogs for each site group was produced using a “bog-lake fusion program” (Sugita, unpublished). The 29 reconstructed plant taxa were grouped into nine broader taxa groups, i.e. nine plant-functional types (PFTs)/ PFT groups that represent 4 biome groups according to the definition of biomes by Ni et al. (2010, 2014). The nine PFTs/PFT groups consist of three taxa groups and six individual taxa, and the REVEALS reconstructions of these PFTs are showed in synthesis diagrams for each site group (Paper IV, Figs. 2-7). REVEALS estimates with standard errors for the common taxa are presented for selected site groups (Paper IV, Figs. S2-S16).
5. RESULTS-SUMMARY OF PAPERS

5.1 PAPER I (Li et al., in press): Modern pollen-climate relationships in north Xinjiang, northwestern China: implications for pollen-based reconstruction of Holocene climate

The aim of this Paper is to examine the relationships between modern pollen, vegetation, and climate variables based on 66 surface pollen samples from forest, meadow, steppe and desert vegetation along a transect from the Jungar desert to the Altay Mountains in north Xinjiang, northwest China. The major objective is to provide guidance for the application of fossil pollen assemblages to vegetation and climate reconstructions in arid and semi-arid regions of the northern hemisphere. Surface soil and moss samples were collected for extracting pollen. The vegetation types, vegetation cover, and the major plant component around each sampling site were recorded in the field. The final vegetation zones were classified according to the field survey and major communities inferred from the vegetation atlas of China (Hou, 2001). Six related climate variables for each sampling site were interpolated from the nearest climate stations and employed for numerical analysis. Principal component analysis (PCA), Redundancy Analysis (RDA) and Variation Partitioning (VP) were used to explore the correlations between pollen taxa, pollen assemblages, and vegetation types. The major results of each analysis are listed below:

1) The pollen assemblages from different vegetation zones are characterized by different composition of pollen taxa: pollen assemblages from the forest zone are generally characterized by high proportions of tree pollen, the major taxa including *Pinus*, *Betula* and *Picea*. While the pollen assemblages from meadow are dominated by Gramineae Caryophyllaceae and Cyperaceae. *Artemisia* and Chenopodiaceae are the major components of pollen assemblages from steppe and desert, and the percentages of *Artemisia* in
5. RESULTS-SUMMARY OF PAPERS

5.1 PAPER I (Li et al., in press): Modern pollen-climate relationships in north Xinjiang, northwestern China: implications for pollen-based reconstruction of Holocene climate

The aim of this Paper is to examine the relationships between modern pollen rain, vegetation, and climate variables based on 66 surface pollen samples from forest, meadow, steppe and desert vegetation along a transect from the Jungar desert to the Altay Mountains in north Xinjiang, northwestern China. The major objective is to provide guidance for the application of fossil pollen assemblages to vegetation and climate reconstructions in arid and semi-arid regions of the northern hemisphere. Surface soil and moss samples were collected for extracting pollen. The vegetation types, vegetation cover, and the major plant component around each sampling site were recorded in the field. The final vegetation zones were classified according to the field survey and major communities inferred from the vegetation atlas of China (Hou, 2001). Six related climate variables for each sampling site were interpolated from the nearest climate stations and employed for numerical analysis. Principal component analysis (PCA), Redundancy Analysis (RDA) and Variation Partitioning (VP) were used to explore the correlations between pollen taxa, pollen assemblages, and vegetation types. The major results of each analysis are listed below:

1) The pollen assemblages from different vegetation zones are characterized by different composition of pollen taxa: pollen assemblages from the forest zone are generally characterized by high proportions of tree pollen, the major taxa including *Pinus*, *Betula* and *Picea*. While the pollen assemblages from meadow are dominated by Gramineae Caryophyllaceae and Cyperaceae. *Artemisia* and Chenopodiaceae are the major components of pollen assemblages from steppe and desert, and the percentages of *Artemisia* in...
steppe is much larger than that of desert, while Chenopodiaceae in desert has larger percentages than in steppe.

2) The unconstrained ISS analysis of pollen percentages shows that the first hierarchical classification separates the samples from desert from those from forest and steppe (although a few assemblages from desert were classified with assemblages from forest and steppe). The second hierarchical classification separates the pollen assemblages from forest from those from steppe and meadow.

3) The application of the semi-quantitative method of pollen ratios showed that the arboreal pollen (AP) percentages or the ratio of arboreal to non-arboreal pollen (AP/NAP) is an important tool for reconstruction of changes in landscape openness. The results show that the mean AP/NAP ratio is 1.39 in forest and 0.62 in marsh located in the forest, while it is lower than 0.03 in open vegetation. Therefore, the AP/NAP ratio can be used to distinguish forest and non-forested vegetation. Except for the AP/NAP ratio, the Artemisia/Chenopodiaceae ratio is an excellent aridity index with value <0.5 in desert and >1 in steppe. The Artemisia/Grainneae ratio is a valuable index of summer moisture.

4) The Principal Components Analysis (PCA) results show that the first two axes explain 48.4% of the variation in pollen data. The first PCA axis separates samples from forest from those from desert while axis 2 contrasts pollen assemblages from deserts and forest with those from meadow.

5) The results of the Redundancy Analysis (RDA) together with Monte Carlo permutation tests suggest that the vegetation types account for most of the variation in pollen assemblages from all variables, and $P_{\text{ann}}$ and $P_{\text{july}}$ explain most of the variation in pollen assemblages when all climate variables are included.

6) The Variation Partitioning (VP) results suggest that the vegetation and climate together account for 32.9% of the variation, while vegetation alone accounts for 13% and climate alone accounts for 2.8% of the variation. We conclude that pollen assemblages together with numerical semi-quantitative methods, such as pollen ratios, ordination methods (PCA, DCA and VP), is a promising tool to infer vegetation and major climate variables.

5.2 PAPER II (Li et al., in revision): Relative Pollen Productivity (RPP) estimates for major plant taxa in cultural landscapes of central eastern China

The aim of this study was to obtain pollen productivity estimates for major tree and herb taxa in cultural landscapes of central eastern China. For this purpose, the landscapes of the central and southern mountain region of the Shandong province were selected, because traditional agriculture practices and structures (cultivation on terraces) are still found in those areas. 36 moss
polster samples were collected for pollen analysis. Vegetation was extracted in concentric rings out to 1500 m radius surrounding each sampling point based on vegetation field surveys within 100 m from the pollen samples, and from satellite images beyond 1500 m. The Extended R-value (ERV) model was used to linearize the relationship between pollen and vegetation data. Pollen counts were obtained from the moss samples, vegetation data were distance-weighted, and fall speed of pollen were estimated using the Stoke’s law (Gregory, 1973) of particle settling velocity from measurements of the long and short axes of pollen grains of each pollen type. The relevant source area of pollen (RSAP) varies depending on the submodel used (92-173 m). The RSAP is estimated to ca. 150 m by submodel 3 that is considered to be the most reliable submodel when absolute vegetation abundance is used. The fall speeds of 23 pollen types (used in the first, exploratory ERV-model run) and the RPP for 18 taxa characteristic of the agricultural landscape were calculated. Among the 18 RPP estimates, the RPP estimates for Castanea, Cupressaceae, Robinia/Sophora, Anthemis/Aster type, Cannabis/Humulus, Caryophyllaceae, Cruciferae, and Galium type are the first and only values available for China so far. RPP values for Quercus, Pinus, and Artemisia are reasonably consistent in northern and temperate China and - for Quercus and Pinus - the values from this study are comparable with the mean RPP in northern and central Europe. However, there are estimates with standard errors larger than the RPP mean value for two taxa, and therefore they are not considered as reliable and were excluded from the RPP dataset. We are confident that 16 RPP values are reliable for all taxa except Juglans regia. Generally, trees (except Robinia/Sophora, 0.78) have higher RPP estimates than herbs (except Artemisia, 24.7). These RPP values are a great contribution to the RPP dataset algorithm in China and can be used for first quantitative reconstructions of Holocene vegetation abundance using the Land reconstruction Algorism of Sugita (2007a, 2007b) and other models requiring RPPs as input.

**5.3 PAPER III (Li et al., submitted): A review of relative pollen productivity (RPP) estimates for major plant taxa of northern and temperate China and implications for long-term quantitative reconstruction of past plant cover.**

The objective of this Paper was to synthesize the RPP studies (ten studies) available in northern and temperate China and set up a standardized RPP
dataset that can be used in first tentative reconstructions of the land cover in China using the REVEALS model. 

In this Paper we review ten studies that estimated the relative pollen productivity (RPP) of major plant taxa in northern and temperate China based on field data, i.e. modern pollen collected from moss polsters, surface soils, surface lake sediments or pollen traps and related vegetation around the pollen samples. The differences between studies in the estimates of relevant source area of pollen (RSAP), FSP, and RPP estimates for 31 plant taxa were assessed and causes behind the between-site differences were discussed. The major factors that account for the difference in RPP and RSAP can be summed up into two parts: the methodological and the environment differences. The differences in methods include site selection, pollen and vegetation data collection, choice of reference taxon, method used to estimate the RSAP, and differences in climate, landscape and vegetation spatial structure, and involved plant taxa between the study regions. A standard RPP dataset for 31 taxa (17 taxa with single values and 14 taxa with a mean RPP estimate based on 1-5 values) is proposed as a mean to attempt first model-based reconstructions of past plant cover using fossil pollen records and the Landscape Reconstruction Algorithm. LRA: REVEALS and LOVE models, Sugita 2007a, 2007b) and other model and simulation studies that require values of RPP.

5.4 PAPER IV (Li et al., manuscript): Pollen-based REVEALS quantitative reconstructions of Holocene plant abundance in northern and temperate China

This study is the first pollen-based REVEALS reconstruction of Holocene land-cover change at the regional scale in China. The major aim of this study is to quantify changes in anthropogenic land-cover changes, primarily deforestation, through time and space. The REVEALS model corrects biases caused by inter-taxonomic differences in pollen productivity and characteristics of dispersal and deposition, basin size and wind speed. We applied the REVEALS model using well dated pollen counts from 95 lake and bog records located in sixth of the eight vegetation zones in temperate China. 29 taxa are reconstructed and presented in synthesis figures (Paper IV, Figure 2-6) for each vegetation zone and nine PFTs/ PFT groups (three PFTs as individual taxa and six PFTs as taxa groups including together 26 taxa). The REVEALS-based values of plant cover strongly differ from the pollen percentages and provide new insights on past changes in plant composition and vegetation dynamics in temperate China. Pollen percentages generally underestimate the cover of herbs in the vegetation, except for *Artemisia* that is
overrepresented by pollen. Pollen percentages of herbs underestimate the cover of herbs by ca. 15-40% in the temperate steppes, temperate deciduous forest, subtropical broadleaved evergreen and deciduous forest, and temperate desert zones, and by 5-10% in some areas of the Tibetan Plateau. The REVEALS reconstructions exhibit vegetation changes that can be interpreted as human-induced from 7.5 or 7 ka in the temperate deciduous forest, temperate steppe, and on the Tibetan Plateau, from 6.5, 6, 5.5 or 5 ka in the subtropical region, the temperate steppe, and on the Tibetan Plateau, from 4.5 ka or later (4, 3.5 ka) in the temperate desert region and on the Tibetan Plateau, and from 2.5 ka in all vegetation zones except the northeastern conifer-broadleaved deciduous forest. Most of these changes correspond to known periods of human activities as documented by archaeologists. These results suggest that the published scenarios of Holocene anthropogenic land-cover change used in climate modelling often underestimate deforestation for several periods of the Holocene and may have to be adjusted. However, disentangling human-induced from climate-induced vegetation openness remains a challenge and will require thorough comparison of the REVEALS reconstructions of land cover with historical and archaeological information on the impact of settlements and other human activities.
6. DISCUSSION and CONCLUSIONS

6.1 The comparative approach: the ability of modern pollen assemblages in distinguishing vegetation types (PAPER I)

In Paper I, we show that appropriate numerical analyses including tests of the significance of the numerical axes and the environmental variables with the aids of application of pollen ratios ensure that correlations between pollen, vegetation and climate can be established with confidence, which is essential for the use of pollen assemblages for climate reconstructions.

By Using PCA and unconstrained ISS on pollen percentages data, we show that, in our study area, pollen assemblages can separate forest and meadow vegetation from desert and steppe (axis 1), and forest and desert from steppe and meadow (Figure 4A). However, discrimination between steppe and desert or meadow vegetation might be difficult. PCA is widely used to investigate the pollen-vegetation relationships to distinguish vegetation types from pollen assemblages (Yili Basin, Zhao and Li, 2013; Loess Plateau, Zhao et al. 2012a). Pollen ratios have been widely used to characterize vegetation types and infer climate characteristics in arid and semi-arid regions of the northern hemisphere (e.g. Herzschuh, 2007). Common used ones are as follows:

- aboreal pollen to non-aboreal pollen (AP/NAP) ratio to distinguish forest and non-forest;
- \textit{Artemisia} to Chenopodiaceae (A/C) ratio to infer vegetation types and moisture changes (El-Moslimany, 1990);
- the \textit{Ephedra}/Chenopodiaceae (Ep/C) ratio to separate steppes from deserts (Luo et al., 2009);
- the ratio of the sum of \textit{Artemisia} and Chenopodiaceae to Gramineae (A+C)/G to distinguish steppe from meadow steppe and forest steppe vegetation (Fowell et al., 2003);
- the ratio of \textit{Artemisia} to Gramineae (A/G) was applied as moisture change index (Tang et al., 2009).

Zhao et al. (2012b) evaluated the potential and limitations of the A/C ratio in arid and semi-arid China. Our study tested the use of pollen ratios and found that AP/NAP is $< 0.1$ in steppe and desert, and $> 0.2$ in forested vegetation; A/C is an excellent index of aridity and is $<0.5$ in deserts and $>1$ in steppes, and A/G is a valuable index of summer moisture;
6. DISCUSSION and CONCLUSIONS

6.1 The comparative approach: the ability of modern pollen assemblages in distinguishing vegetation types (PAPER I)

In Paper I, we show that appropriate numerical analyses including tests of the significance of the numerical axes and the environmental variables with the aids of application of pollen ratios ensure that correlations between pollen, vegetation and climate can be established with confidence, which is essential for the use of pollen assemblages for climate reconstructions.

By using PCA and unconstrained ISS on pollen percentages data, we show that, in our study area, pollen assemblages can separate forest and meadow vegetation from desert and steppe (axis 1), and forest and desert from steppe and meadow (Figure 4A). However, discrimination between steppe and desert or meadow vegetation might be difficult. PCA is widely used to investigate the pollen-vegetation relationships to distinguish vegetation types from pollen assemblages (Yili Basin, Zhao and Li, 2013; Loess Plateau, Zhao et al. 2012a)

Pollen ratios have been widely used to characterize vegetation types and infer climate characteristics in arid and semi-arid regions of the northern hemisphere (e.g. Herzschuh, 2007). Common used ones are as follows:

- aboreal pollen to non-aboreal pollen (AP/NAP) ratio to distinguish forest and non-forest;
- *Artemisia* to *Chenopodiaceae* (A/C) ratio to infer vegetation types and moisture changes (El-Moslimany, 1990);
- *Ephedra*/*Chenopodiaceae* (Ep/C) ratio to separate steppes from deserts (Luo et al., 2009);
- the ratio of the sum of *Artemisia* and *Chenopodiaceae* to *Gramineae* (A+C)/G to distinguish steppe from meadow steppe and forest steppe vegetation (Fowell et al., 2003);
- the ratio of *Artemisia* to *Gramineae* (A/G) was applied as moisture change index (Tang et al., 2009).

Zhao et al. (2012b) evaluated the potential and limitations of the A/C ratio in arid and semi-arid China. Our study tested the use of pollen ratios and found that AP/NAP is < 0.1 in steppe and desert, and > 0.2 in forested vegetation; A/C is an excellent index of aridity and is <0.5 in deserts and >1 in steppes, and A/G is a valuable index of summer moisture;
estimates of pollen productivity and fall speed are necessary. For such techniques quantification, other methods are required such as the Landscape percentage cover of the vegetation type or the individual plant taxa. For vegetation in the past, i.e. they do not provide estimates of the abundance or that are occurring in a region. But these techniques do not allow quantifying reconstructions of vegetation, i.e. to identify from pollen the vegetation types that are occurring in the past. This result is of interest as it quantifies the strong correlation between vegetation and climate in these ecosystems and, in turn, the validity of using pollen assemblages to reconstruct climate in the past.

This study shows that numerical analyses are very valuable for qualitative reconstructions of vegetation, i.e. to identify from pollen the vegetation types that are occurring in a region. But these techniques do not allow quantifying vegetation in the past, i.e. they do not provide estimates of the abundance or percentage cover of the vegetation type or the individual plant taxa. For quantification, other methods are required such as the Landscape Reconstruction Algorithm (LRA, Sugita, 2007a and b). For such techniques estimates of pollen productivity and fall speed are necessary.

6.2 The mechanistic-modeling approach: Relative Pollen Productivity estimates (PAPERS II and III)

6.2.1 Relevant source area of pollen (RSAP)
RSAP varies depending on the submodel used (92-173m) in the cultural landscape of Shandong. A RSAP of 145 m is found using the submodel 3, which is the most reliable submodel to use in our case. RSAPs of 90 to 173 m radius are very small compared to RSAP values obtained in earlier studies in China. For instance, Ge et al. (2015) estimated a RSAP radius of ca. 2100 m for soil samples in the meadow and steppe landscapes of central Inner Mongolia and Li et al. (2015) a RSAP of 2000-2500 m for moss polsters in the wood landscapes of the Changbai Mountains in north eastern China. In the study of Wang and Herzschuh (2011) in the alpine meadow vegetation of the Tibetan Plateau using pollen data from lake surface samples, the curve of likelihood function scores never reaches an asymptote and has its lower value
Reconstruction Algorithm (LRA, Sugita, 2007a and b). For such techniques quantification, other methods are required such as the Landscape percentage cover of the vegetation type or the individual plant taxa. For vegetation in the past, i.e. they do not provide estimates of the abundance or this study shows that numerical analyses are very valuable for qualitative of using pollen assemblages to reconstruct climate in the past.

Vegetation and climate together explain 32.9% of the variation in pollen assemblages, while The results of the variation partitioning analysis show that climate alone explains only 2.8% of the variation in the pollen assemblages, while found to be an important climate variable that accounts for high proportion of variation in pollen data (Zhao and Herzschuh, 2009; Herzschuh, 2007; Li et al., 2011. Zhao et al., 2012a; Shen et al., 2006) in northern China and T ann is in influencing pollen assemblages in north China (Luo et al., 2010; Shen et al., 2006; Li et al., 2015), which is confirmed in our study. Except Pann, TJul is also vegetation explains 13%. Vegetation and climate together explain 32.9% of the variation in the pollen assemblages, i.e. 21.3% and 10.7% respectively, followed by T Jul and PJul are the climate variables that explain the larger part of the variation in 6.2.2 Relative Pollen Productivity estimates (RPPs)
The studies within this thesis have provided RPPs for 18 taxa from a case study in the cultural landscapes of Shandong (PAPER II) and 13 additional taxa from other available RPP studies in China (PAPER III). The estimation of RPP requires significant gradients in the values of pollen and vegetation percentages. Among the 18 taxa, six taxa (Gramineae, Artemisia, Cyperaceae, Compositae SF Cichoriodae, Pinus, and Quercus) show good linear relationships (PAPER II Fig. 5). For the remaining twelve taxa, the individual values are either irregularly distributed or have strongly deviating high values with either high pollen values corresponding to low vegetation values or the inverse. It is common in RPP studies based on field data, in particular in open and semi-open landscapes, that only a few taxa have their pollen-vegetation data closely fitting to the ideal ERV-model linear relationship (e.g. Mazier et al., 2008 in Europe; Xu et al., 2014 and Ge et al., 2015 in China). Further, the reliability of the RPPs will depend on the RSAP estimate obtained from the curve of the log likelihood score at increasing distances from the pollen sample. In our study, the RSAP could be identified with good precision from the log likelihood curve obtained with the ERV sub-model 3.

Moreover, the cultural landscape in Shandong is particular compared to the semi-open to open landscapes studied earlier in Europe or China. It is characterized by a much higher cover of barren soils than any landscape formerly used for RPP studies. In addition, we assume that the cover of weed taxa such as Artemisia, Humulus, Gramineae, Aster, Compositae SF Cichoriodae, etc. in the terraces at one point of time (period of survey) is a fair approximation of the mean cover of these taxa over the 1-3 years of pollen deposition in the moss sample, even though the weeds may have changed their spatial distribution through time due to the shifting of field under cultivation or fallow, and/or the change of cultivated crops (and therefore weed cover) in some fields.

Finally, the reliability of the RPP estimates is described by their relative standard deviations (SDs). RPPs with SDs > RPP mean are considered as unreliable and should not be used for pollen-based quantitative vegetation reconstructions. If we consider the estimates based on ERV sub model 3 and the taxon-specific distance weighting method, there is one such RPP estimate

at a distance of 2200 m which is also the radius of the area from which vegetation data was available, i.e. these lake basins may have had a larger RSAP.

In Europe, a RSAP of 400 m was estimated for moss polster samples in the open and semi-open cultural landscapes of southern Sweden (Broström et al 2005), of ca. 300 m in the pasture woodland landscape of the Jura Mountains (Mazier et al., 2008), and of 50-150 m in woodlands in Norfolk (Bunting et al., 2005).

Moreover, the cultural landscape in Shandong is particular compared to the semi-open to open landscapes studied earlier in Europe or China. It is characterized by a much higher cover of barren soils than any landscape formerly used for RPP studies. In addition, we assume that the cover of weed taxa such as Artemisia, Humulus, Gramineae, Aster, Compositae SF Cichoriodae, etc. in the terraces at one point of time (period of survey) is a fair approximation of the mean cover of these taxa over the 1-3 years of pollen deposition in the moss sample, even though the weeds may have changed their spatial distribution through time due to the shifting of field under cultivation or fallow, and/or the change of cultivated crops (and therefore weed cover) in some fields.

Finally, the reliability of the RPP estimates is described by their relative standard deviations (SDs). RPPs with SDs > RPP mean are considered as unreliable and should not be used for pollen-based quantitative vegetation reconstructions. If we consider the estimates based on ERV sub model 3 and the taxon-specific distance weighting method, there is one such RPP estimate

6.2.2 Relative Pollen Productivity estimates (RPPs)
The studies within this thesis have provided RPPs for 18 taxa from a case study in the cultural landscapes of Shandong (PAPER II) and 13 additional taxa from other available RPP studies in China (PAPER III). The estimation of RPP requires significant gradients in the values of pollen and vegetation percentages. Among the 18 taxa, six taxa (Gramineae, Artemisia, Cyperaceae, Compositae SF Cichoriodae, Pinus, and Quercus) show good linear relationships (PAPER II Fig. 5). For the remaining twelve taxa, the individual values are either irregularly distributed or have strongly deviating high values with either high pollen values corresponding to low vegetation values or the inverse. It is common in RPP studies based on field data, in particular in open and semi-open landscapes, that only a few taxa have their pollen-vegetation data closely fitting to the ideal ERV-model linear relationship (e.g. Mazier et al., 2008 in Europe; Xu et al., 2014 and Ge et al., 2015 in China). Further, the reliability of the RPPs will depend on the RSAP estimate obtained from the curve of the log likelihood score at increasing distances from the pollen sample. In our study, the RSAP could be identified with good precision from the log likelihood curve obtained with the ERV sub-model 3.

Moreover, the cultural landscape in Shandong is particular compared to the semi-open to open landscapes studied earlier in Europe or China. It is characterized by a much higher cover of barren soils than any landscape formerly used for RPP studies. In addition, we assume that the cover of weed taxa such as Artemisia, Humulus, Gramineae, Aster, Compositae SF Cichoriodae, etc. in the terraces at one point of time (period of survey) is a fair approximation of the mean cover of these taxa over the 1-3 years of pollen deposition in the moss sample, even though the weeds may have changed their spatial distribution through time due to the shifting of field under cultivation or fallow, and/or the change of cultivated crops (and therefore weed cover) in some fields.

Finally, the reliability of the RPP estimates is described by their relative standard deviations (SDs). RPPs with SDs > RPP mean are considered as unreliable and should not be used for pollen-based quantitative vegetation reconstructions. If we consider the estimates based on ERV sub model 3 and the taxon-specific distance weighting method, there is one such RPP estimate
in our results, i.e. that of *Vitex negundo* (0 ± 0.02). The value for Amaranth./Chenop. also has a very large SD (0.18 ± 0.16).

### 6.2.3 The possible factors behind between-study differences in the estimates of relevant source area of pollen (RSAP) and relative pollen productivity (RPP)

The factors accounting for the between-study discrepancies observed in RPPs and RSAP from studies can be summed up into two aspects, each of which gives a unique effect on the results of RPP: 1. Methodological differences including site selection, pollen and vegetation data collection, choice of reference taxon, method used to estimate the RSAP; 2. differences in climate, landscape and vegetation spatial structure, and involved plant taxa between the study regions (e.g. Broström et al., 2008; Bunting et al., 2013).

#### Methodological factors

Within the group of methodological factors, the differences between the methods used to assemble the vegetation dataset are probably those that influence the RSAP and RPP values most, together with the distribution of sites in a study, i.e. random or random stratified distribution, or selected sites according to various criteria (vegetation or land-use types, occurrence of particular plant taxa). Simulation studies using hypothetical landscape/vegetation structures or mimicking actual vegetation have shown that the smaller the patches and the more homogenous the patches distribution and the taxa composition and distribution, the smaller the RSAP (e.g. Bunting et al., 2004; Gaillard et al., 2008; Hellman et al., 2009 a and 2009 b). The fall speed of pollen (FSP) used for the taxa involved in the analysis also has an effect on the obtained RSAP distance and RPP estimates. FSP is usually calculated using the Stokes’s law and measurements of the size of the pollen grains. The value of FSP for same taxa differs between studies (PAPER III, Table 2). The type of pollen sample may also have an effect on the RPP values obtained because preservation of pollen and taphonomic processes differ between sample types i.e., in the case of the ten studies reviewed, between surface soil, lake sediment, moss polster, and pollen trap. It is known that there are differences between sediment types, for instance trap (Hicks, 2001) and moss polsters (Räsänen et al., 2004). Comparison of differences from pollen assemblages in soil samples, pollen traps and moss polsters were discussed by e.g. Xu et al. (2016), and differences in pollen assemblages from surface soils and lake-surface sediments by Zhao et al. (2009c). The RPP estimates of *Artemisia* and Chenopodiaceae using pollen data from lake surface sediments (Wang and Herzschuh, 2011) are much lower than those obtained from soil and moss samples. Simulation studies show that, for a given vegetation/landscape, the RSAP becomes larger with increasing basin size, i.e. pollen traps and moss polsters have a smaller RSAP than small lakes,
and small lakes have a smaller RSAP than middle-sized to large lakes (e.g. Sugita, 1994; Broström et al., 2005; Hellman 2009a and b).

Finally, the choice of ERV submodel and method to distance weight vegetation data influence significantly the values of RSAP and RPP. The choice of submodel varies between the ten studies reviewed in PAPER III.

Environmental factors
The three major environmental factors are i) the species included in the pollen taxa involved in each study, ii) the climate characterizing the study region, and iii) the vegetation/landscape structure in the study region. The species included may have significant effects on the RPP estimates obtained in the ten studies reviewed. For instance Artemisia, Chenopodiaceae and Pinus involved different species in the different regions. Differences in climate between the study regions do not seem to have a major effect on the RPP estimates obtained so far, as the RPP estimates of Artemisia in northern China (meadows and steppes of Inner Mongolia) and the Shandong province are comparable, as well as the RPP estimates of Pinus in the Changbai Mountains (Li et al., 2015) and the Shandong province (PAPER II). Finally, differences in vegetation/landscape structure have an effect on the RSAP and RPP, which was already discussed in the section on methodological factors above.

6.2.4 Selection of RPP values for the standardized RPP datasets
We adjusted the rules used in Mazier et al. (2012) to establish the standard RPP dataset for northern and temperate China. The RPPs selected for the dataset standard (std) 3 are the values obtained by removing the most deviating ones when more than three values were available and using expert knowledge when values were very variable for a single taxon. Although the proposed standard RPP dataset includes a large number of taxa for which only 1-2 estimates are available, it can already be used i) to test the RPP values by using them with modern pollen data and the REVEALS model (Sugita, 2007a) and compare the obtained estimates of plant cover with modern vegetation data, and ii) to achieve REVEALS-based quantitative reconstructions of past plant cover using fossil pollen records and evaluate the results by comparison with other palaeoecological information, such as climate reconstructions, knowledge on vegetation/climate relationships, and archaeological/historical data on human activity. The latter will provide first insights on past changes in plant cover and highlight the taxa for which results appear to be problematic, which might be due to uncertain estimates of relative pollen productivity.
6.3 Holocene vegetation abundance (PAPER IV)

6.3.1 Differences between pollen percentages and pollen-based REVEALS estimates of plant cover

The reconstructed vegetation abundances from different regions show different patterns during the Holocene (figures 2-6 and S1-14) that agree with former syntheses based on pollen records from multiple sites (Ren et al., 2007; Zhao et al., 2009a, 2009b; Zhao and Yu, 2012; Tian et al., 2016). The most important new insights provided by the REVEALS models are the differences between REVEALS reconstruction and pollen percentages. These differences have been discussed in several studies in Europe (Hellman 2008a, 2008b; Gaillard et al., 2010; Soepboer et al., 2010; Marquer et al., 2014; Trondman et al., 2015) and in North America (Sugita et al., 2010). The first published pollen-based REVEALS reconstructions of land cover (including the four taxa Artemisia, Gramineae, Cyperaceae and Chenopodiaceae) was performed by Wang and Herzschuh (2011) using pollen data from 11 lakes on the Tibetan Plateau. The authors found that the REVEALS-based changes in vegetation abundance over the Holocene were characterized by larger vegetation turnover than changes in pollen percentages. The comparison between REVEALS reconstruction and pollen percentages from 18 large lakes in northern Europe showed that there was an earlier increase in Ulmus and Corylus and a larger increase in grassland and deforested areas from mid-Holocene (Marquer et al., 2014). Trondman et al. (2015) found that the REVEALS-based proportions of open land and forested land are very different from the pollen percentages, whereas the general trends are similar to those observed in pollen percentages, although the age of the establishment of taxa and their maximum extent can also differ between REVEALS estimates of plant cover and pollen percentages (Trondman et al., in press).

This study confirms the earlier observation that REVEALS estimates of vegetation abundance are very different from pollen percentages. The general differences are as follows: the taxa that have high relative pollen productivities (e.g. Pinus, Betula, Quercus and Artemisia) show lower REVEALS-based vegetation abundances than pollen percentages, while the taxa that have low relative pollen productivities (e.g. Ulmus and Tilia among trees, and Gramineae, Cyperaceae, Polygonaceae, Rosaceae, Ranunculaceae, Compositae, Fabaceae, Brassicaceae, and several other herb taxa,) have higher REVEALS-based estimates of vegetation abundances than pollen percentages. Therefore, the REVEALS estimates show a consistently lower percentage cover of tree taxa (Pinus, Betula) and higher percentage cover of broad leaved trees (Ulmus, Tilia) and herbaceous taxa (particularly Gramineae and Cyperaceae), hence larger land-cover openness than previously suggested (Zhao et al., 2009a, 2009b; Ren et al., 2007; Zhao and Yu, 2012; Cao et al., 2015). For instance, when Artemisia is dominant over Cyperaceae as pollen
percentages, Cyperaceae is dominant over *Artemisia* in vegetation abundance. The inverted relationship between *Artemisia* and Cyperaceae is due to the very high pollen productivity of *Artemisia*, e.g. ca. 40 times larger than that of Cyperaceae (Paper III). If we compare the REVEALS results with pollen percentages of *Pinus*, *Betula*, and Gramineae, similar inverted relationships are observed, since the RPP of *Pinus* is 10 times that of Gramineae, and the RPP of *Betula* is 18 times that of Gramineae (Paper III).

Below we take two vegetation zones in China as examples to compare more specifically the differences between pollen percentages (Zhao et al., 2009a and 2009b) and REVEAL reconstruction in vegetation zone II (site group 1-Qindeli bog) and VII (site groups 20 – Sanjiaocheng and Hurleg lake, 21 – Qinghai Lake, and 24 – Mannas lake).

In the coniferous deciduous mixed forest vegetation zone, trees (especially *Betula* and *Pinus*) are over estimated while herbs (especially Cyperaceae and Gramineae) are underestimated. At Qindeli bog pollen percentage was dominated by *Betula* (10-40%) from 11-10 ka, shifted to *Ulmus* (15-25%), *Quercus* (15-25%), and *Abies* (10-20%) at 10-5.5 ka, and to *Ulmus* (15-35%), *Pinus* (5-35%), *Betula* (5-20%), *Quercus* (10-30%) after 5 ka (Zhao et al., 2009b). In contrast, the REVEALS results (site group 1) show that the landscapes are characterized by a large degree of land-cover openness (mainly Cyperaceae) throughout the Holocene, where the openness was 65-85% during 11-10 ka, 57%-93% during 10-5.5, and 78-85% after 5.5 ka. The trees are always represented by low cover, *Quercus* (ca. 1-3%), *Ulmus* (<2%) and *Pinus* (lower <3%) and ca. 8% during the last 0.5 ka. until recent time.

In the temperate desert zone, *Artemisia* and Amaranthaceae/Chenopodiaceae are generally underrepresented by pollen percentages while Gramineae and Cyperaceae are underestimated. At Sanjiaocheng pollen percentages of *Artemisia* range between 10 and 70%, at Hurleg lake between 10 and 40%, and at Mannas lake around 80%, while the REVEALS estimates of *Artemisia* plant cover are around 15% for site group 20, 2-4% for site group 21, and 15% for site group 24. At Qinghai Lake for instance, *Artemisia* is represented by 40-80%, Gramineae and Cyperaceae by ca. 10% each, and Amaranthaceae/Chenopodiaceae by ca. 5-10%. The corresponding REVEALS cover for site group 21 are 2-4% for *Artemisia*, 20-30% for Gramineae, 5-50% for Cyperaceae, and 5-10% for Amaranthaceae/Chenopodiaceae.

### 6.3.2 Spatio-temporal Holocene forest dynamics and landscape openness: climate or human induced?

This first large-scale pollen-based REVEALS reconstruction shows a much larger degree of land-cover openness than pollen percentages suggest. The REVEALS model corrects biases caused by inter-taxonomic differences in pollen productivity and characteristics of dispersal and deposition, basin size and wind speed, therefore, provides better estimates of regional vegetation/land-cover changes, and in particular for temperate steppe in central northeast
China, temperate desert in northwest China and alpine vegetation in Tibetan Plateau, than the traditional use of pollen percentages provides. REVEALS thus allows a more robust assessment of human-induced land-cover at the regional scale.

As seen from the REVEALS reconstructions, large parts of northwestern China, central eastern China and the Tibetan Plateau were covered by sparse woodland or open vegetation throughout the Holocene, while northeastern China, central southern China and southwestern China were covered by woodland. For most of the study region woodlands were expanding in early Holocene, reached a maximum abundance at around 6 ka, and steadily declined until recent 0.5 ka. Trees reached a maximum at around 8 ka and decreased slightly around 5.5 ka and expanded again after 5.5 ka for conifer-broadleaved deciduous forests in northeastern China. There is archaeological evidence of significant human disturbances in all vegetation zones, except in northeastern China. Hosner et al. (2016) reviewed the spatio-temporal distribution of Neolithic and Bronze Age archaeological sites in China. There is a noticeable increase in the concentration of Neolithic sites between around 7 and 6 ka and highest site concentrations were reached between 4 and 2.5 ka, which is in good agreement with the REVEALS reconstruction of decreasing forest cover. Nevertheless, climate change at the regional scale is still the major driver of vegetation changes.

The attempts at creating scenarios of past anthropogenic land-cover dynamics in HYDE (Kelin Godelwijk et al., 2011) and KK10 (Kaplan et al., 2009) suggest a noticeable opening of the landscapes since 6 ka. However, based on the numerous archaeological findings, the increase of land-cover openness from HYDE and KK seems to be largely underestimated (e.g. Gaillard et al., 2010). These scenarios are uncertain due to the population data used and the method for translating the population numbers into fractions of deforested land.

Archaeological studies provide great information on the start, and development of Neolithic agriculture and later cultures. However, it is a challenge to draw any firm conclusion on the factors behind land-cover/vegetation abundance changes since it is difficult to distinguish human-induced from climate-induced vegetation change (e.g. Zhao and Yu, 2012). This work needs to be performed in collaboration with archaeologists in order to further evaluate the results of this first REVEALS reconstruction and its interpretation.
ACKNOWLEDGEMENTS

Now! At the end of my PhD studies, I would like to express my sincere appreciation to all the colleagues and friends who helped me, took care of me and encouraged me during these four years!

My gratitude from deep heart goes first to my supervisor Marie-José for her wonderful support, care, supervision for every moment of my study. Thank you for providing me so many chances and opportunities to attend workshops and conferences and the two invaluable study stays in France and Germany, from which I have learned a lot from colleagues there. Thank you for editing all my applications, posters and power point presentations whenever I needed to submit those or present my results. Thank you for editing all manuscripts. I am touched when I think about so many moments that I have spent together with you, e.g. when we were seating together in the computer room to check the land-use maps site by site for 37 sites, when we were identifying and defining pollen types for more than 400 species; you made drawings to teach me the structure and ornamentation of pollen walls. I was almost depressed when my ERV runs failed so many times, but your patience and strength gave me energy to try again and again. I am sure your way of working and your style of always making things perfect will influence my whole life in future.

My sincere appreciation goes to my co-supervisor Shinya Sugita at Tallin University, Estonia, for supporting me with knowledge regarding pollen dispersal and deposition models and answering all the questions I had on the models. Thank you also for analysing my data and solving problems when I was running ERV models and for helping me to design the new PFT program to handle the output data from REVEALS model. Thank you for your kind help and patience even when you were very busy with your own work. Your detailed description and clear examples in the e-mails made the complicated issues easier for me to understand.
My thanks go also to my co-supervisor professor Xu Qinghai at Hebei Normal University for all his help during my stay in Shijiazhuang. Thank you for providing all facilities and help for my two field work seasons. My thanks go also to Carole, Laurent, Elodie, Mu Huishuang, Zhang Shengrui, Li Jie, Zhang Panpan, Hu Yanan, and Li Yang for your help during fieldwork. Thank you also Professor Li Yuecong for identifying plants during my first year fieldwork. My appreciation goes to Professor Zhou Zhongze and Zhang Yuyun at Anhui University; without your knowledges in identifying plants, the fieldwork would not have been so successful.

I am very grateful to Florence at Toulouse University for all your support from the beginning of my PhD, thank you for helping in designing the fieldwork strategy for the first time, for joining my second fieldwork, for presenting me to your colleagues in Toulouse University who helped me a lot in the later stages of my work. Thank you and Stephane for all your support during my stay in Toulouse and helping me with data handling for running ERV models. Special thanks go to Dominique Laffly for his invaluable teaching in creating the land-use maps for Paper II.

My appreciation goes to my paleoecology group, Qiao, Anna-Kari, Laurent, Ganna. We have had very nice discussion and nice time together at the beginning of my PhD and later even though you had moved to other places. Special thanks to Anna-Kari for answering all my questions regarding data preparation and REVEALS model runs. Thank you also Anna-Kari and Qiao for your comments on the first draft of this thesis. Thank you very much Geoffrey for translating the Swedish abstract.

To my co-supervisor Mats, thank you for good suggestions in designing the project.

To Ulrike Herzschuh and Xianyong Cao at Alfred-Wegener-Institute for hosting my one month research stay and providing the data for me to run REVEALS model. Thank you very much Xianyong for writing code in R on data handling whenever I asked for help, which saved a lot of my time when I was very busy with large amounts of data.

To Jane Bunting from Hull University! Thank you for providing the survey program which allowed me to extract vegetation data for within 100 m, and also Michelle to answer my questions regarding the fieldwork protocol for pollen productivity estimates at the beginning of my PhD.

To my master’s supervisor Professor Yan Zhao at the Institute of Chinese Academy, thank for your help with my master study. My thanks goes to my
good friends Guo Peng, Chen Guan and Li Huan at Lanzhou University for always willing to help me whenever I needed for ArcGIS.

To my Chinese friends, without you, the life in Kalmar wouldn’t have been so interesting! Xiaofen and Shan, it is so nice to have you here, we share so many things in work and life, since we have had so similar experiences. Thank you Qiao, you act like my elder sister, I like to talk with you when I am worrying about something and you always give me good advises. Thank you Xiaofen, Stephan, Shan, Zhichao, Jingyang, Elias for nice food and celebrating almost all Chinese and Swedish festivals together if we are all in Kalmar, the time with you is always full of happiness. Thank Gaofeng and Ke, for the delicious dinners and nice time together. Thank you Daniel and Qing, it is always nice to talk with you.

To colleagues and friends in the MERGE group and ClimBEC research school, thank you for the nice workshops and courses! Thank you Annika for booking all my travels for the courses and programs!

To Anders, Henrik and Sara, thank you for being always helpful whenever I need you! To the administrator people, Ymke, Christina, Ulrika, Per, thanks for your support! To the colleagues and friends around me in the department Lee, Nayani, Eva, Devi, Margarita, Frédéric, Mireia, Alex, Olga, Karin, Fabio, Emelie, Henrik, Carina and many more…, thank you for nice chat over lunches and fika!

To my mentor Yael, thank you for all your support and spending time with me.

最后，感谢爸爸妈妈一直以来的支持与疼爱！感谢姐姐姐夫的关心与帮助以及照顾爸妈！感谢弟弟越来越懂事！我爱你们！不管我走到哪里，家永远是最温暖的港湾！

Thank you!
Tack!
谢谢!
LITERATURE CITED


Barton, L. 2016. The cultural context of biological adaptation to high elevation Tibet. Archaeological Research in Asia, 5: 4-11.


LITERATURE CITED


Barton, L. 2016. The cultural context of biological adaptation to high elevation Tibet. Archaeological Research in Asia, 5: 4-11.


Fagerlind, F. 1952. The real significance of pollen diagrams, Botaniska


Bunting, M. J., & Middleton, R. 2009. Equifinality and uncertainty in the

Bunting, M. J., Gaillard, M. J., Sugita, S., Middleton, R., & Broström, A.


El-Moslimany AP 1990. Ecological significance of common nonarboreal


Cao, X., Herzschuh, U., Ni, J., Zhao, Y., & Böhmer, T. 2015. Spatial and


Chen, F., Dong, G., Zhang, D., Liu, X., Jia, X., An, C., Ma, M., Xie, Y.,


Munksgaard, Copenhagen.


9}


beginning c. 9000 cal. BP. of rice domestication in China: a response. The

Liu, M., & Tian, H. 2010. China's land cover and land use change from 1700
to 2005: Estimations from high-resolution satellite data and historical

Climate Simulation of the Evolution of Global Monsoons in the Holocene.
Journal of Climate, 1615: 2472-2490.

Lu, H., Zhang, J., Liu, K., Wu, N., Li, Y., Zhou, K., Ye, M., Zhang, T., Zhang,
H., Yang, X., Shen, L., Xu, D., & Li, Q. 2009. Earliest domestication of
common millet Panicum miliacuem. in East Asia extended to 10,000 years

Shen, C., Li, X., Tong, G., Jiang, H., 2011. Modern pollen distributions in
Qinghai-Tibetan Plateau and the development of transfer functions for
reconstructing Holocene environmental changes. Quaternary Science
Reviews. 30, 947-966.

Characteristics of the modern pollen distribution and their relationship to
vegetation in the Xinjiang region, northern and western China. Review of

Luo, C., Zheng, Z., Tarasov, Nakagawa, T., Pan, A., Xu, Q., Lu, H., Huang,
K. 2010. A potential of pollen-based climate reconstruction using a modern
pollen-climate dataset from arid northern and western China. Review of

Marquer, L., Gaillard, M.-J., Sugita, S., Trondman, A.-K., Mazier, F., Nielsen,
A. B., Fyfe, R. M., Odgaard, B. V., Alenius, T., Birks, H. J. B., Bjune, A.
E., Christiansen, J., Dodson, J., Edwards, K. J., Giesecke, T., Herzschuh,
Holocene changes in vegetation composition in northern Europe: why
quantitative pollen-based vegetation reconstructions matter. Quaternary
Science Reviews, 90: 199-216.

Mazier, F., Broström, A., Gaillard, M-J, Sugita, S., Vittoz, P., Buttler, A.
2008. Pollen productivity estimates and relevant source area of pollen for
selected plant taxa in a pasture woodland landscape of the Jura Mountains
Switzerland.. Vegetation History and Archaeobotany 17 5: 479-495.

Mazier, F., Nielsen, A. B., Broström, A., Sugita, S., & Hicks, S. 2012. Signals
of tree volume and temperature in a high-resolution record of pollen
accumulation rates in northern Finland. Journal of Quaternary Science,
276: 564-574.


