Relationships between surface pollen and vegetation in the Meiling Mountains, southeast China: an aid to reconstruct past vegetation dynamics

Yiman Fang

University of Hull

School of Environmental Sciences

A thesis submitted for the degree of Doctor of Philosophy

10/2018

Acknowledgements

First of all, I would give my deepest thank to my supervisors M. Jane Bunting (University of Hull), Jeff Blackford (University of Hull), Anna Bird (University of Hull) and Michelle Farrell (Coventry University) for guidance through the whole project and comments on an earlier version of the thesis. In particular I want to thank my lead supervisor Jane for her professional guidance, kindness, encouragement, patience and timely response during the past three years. She introduced me to a new modelling world, which offers me a great chance to experience the joy of Palynology. As an international student, she also helped my English progress constantly. A special thank also goes to Chunmei Ma (Nanjing University) for her full support in both field work and lab work, and useful suggestions about the project which impressed me a lot.

Thanks are due to Rodney M Forster for his help with the remotely sensed data, Graham Ferrier for his guidance in classifying RS data and Yvonne C. Golding (British Pteridological Society) for her help in identifying the fern species during field work. My gratitude is also due to Lingyu Tang, Limi Mao and Chunhai Li for giving useful suggestions and guidance in identifying these pollen grains. Thanks also goes to Isabella Capellini for giving a chance to organize the Ecology Research Group every Friday, which is a wonderful opportunity to meet new people and talk about ecology stuff!

I appreciate the fieldwork guidance from Jane Bunting and Michelle Farrell. I also wish to thank Lei Yang (Wuhan University, China), Jue Sun, Liang Li, Lin Zhao (Nanjing University, China), Yulian Jia, Chaohao Ling, Jun Luo, Yuanhui Chen, Chuan Chen and Shiping Li (Jiangxi Normal University, China) for helping with fieldwork and plant identification. Happy memories with you all in this subtropical forest!

I am indebted to everyone at the Department of Geography for the various help I received from them. I am grateful to Brendan Murphy for his help with the insurance documents before the field work, to Mark Anderson, Kim Rosewell, Mike Dennett and Tim Bettley for their help I have received in carrying out palynological research, and Pauline Deutz & Andrew Jonas for inviting me to their house for a wonderful Christmas party. I also want to thank all the staff in the University Library, who always made me feel welcome and accompanied me throughout the writing process.

I would like to thank the co-ordinator Marie-José Gaillard and all PAGES Landcover6k members for valuable training and discussion in a supportive environment in Zaragoza, Spain. This project is a contribution to the PAGES LandCover6k working group (http://www.pastglobalchanges.org/ini/wg/landcover6k/intro).

This project was funded by the China Scholarship Council (Grant number 201506190128). Fieldwork and labwork were also supported by the Dudley Stamp Memorial Award from Royal Geographical Society (with IBG), UK (Grants DSMA 25/16), grants from the National Key Research and Development Program (2016YFA0600501) and the National Natural Science Foundation of China (NSFC, Grants 41671196), and Research Support Fund from the School of Environmental Science.

谨以此文献给亲爱的爸爸妈妈。28 岁，我活出了自己想要的样子。

Abstract

Understanding the relationships between modern pollen deposits and their surrounding vegetation is an important tool to improve the quality of reconstructions and interpretations of past vegetation changes from fossil pollen records. The overall aim of this research project is to validate and calibrate mathematical models of relationships for assemblages deposited in peatland and forest landscapes in southeast China, which will form an essential basis for quantitative reconstruction of past land cover from Quaternary peat deposits in the region. This field study area presents great challenges, being spatially large in extent compared to studies in northwestern Europe and of difficult and inaccessible mountainous terrain, which makes the vegetation survey (10m-100m) time-consuming.

Firstly, a study to decide whether to use moss or soil surface samples is presented. 42 paired moss and soils sample were collected in the five main forest types. Similar levels of variation in the pollen spectra are seen, but there are systematic differences in the mean values of key groups of taxa. Moss polsters are chosen since they record the most accurate representation of the contemporary vegetation.

Secondly, the behaviour of several pollen dispersal and deposition models is tested against a grassland-forest transect. The Prentice-Sugita model passed the test and is therefore considered suitable for use.

Thirdly, the first estimates of relative pollen productivity (RPP) for 9 key taxa (*Castanea*, *Cryptomeria*, *Cyclobalanopsis*, *Liquidambar*, *Pinus*, Poaceae, *Quercus*, Rosaceae and Theaceae) are presented. Two alternative methods (modified Davis method and iteration method) for estimating RPP are also developed, which have great potential for use in wider areas.

Fourthly, wetland herb taxa are important in the pollen spectra from mire records, therefore surface samples were taken from a mire surface. They show that there are inter-annual differences in pollen from wetland herbs.

Finally, the discussion presents suggestions for how these findings can be best applied to land-cover reconstruction, explores the strengths and limitations of the study and identifies future directions which such work could take.

Abbreviations

AP	arboreal pollen
a.s.l.	above sea level
CA	Correspondence Analysis
CANOCO	CANOnical Community Ordination
Comm.	communities
Cry. forest	*Cryptomeria japonica* var. *sinensis* forest
Cunn.-Pi. forest	*Cunninghamia lanceolata-Pinus massoniana* forest
d&d	dispersal and deposition
DEM	digital elevation model
DWPA	distance weighted plant abundance
EASM	East Asia Summer Monsoon
ERV	Extended R-value
ESA	European Space Agency
GPD	global pollen database
GPS	global positioning system
IM	Iteration method
LFs	likelihood function score
LGM	Last Glacial Maximum
LIA	Little Ice Age
LOVE	LOcal Vegetation Estimate
LRA	Landscape Reconstruction Algorithm
LSM	Lagrangian stochastic model
MAT	modern analogue technique
MDM	Modified Davis method
MSA	Multiple Scenario Approach
PAR	pollen accumulation rates
PCA	Principal Component Analysis
PL	pollen loading
P-S	Prentice-Sugita
REVEALS	Regional Estimates of VEgetation Abundance from Large Sites
RPP	relative pollen productivity
RSAP	relevant pollen source area of pollen
SCD	squared-chord distance
SD	standard deviations
SSCD	summed squared-chord distance
TWINSPAN	Two Way Indicator Species Analysis
UAV	unmanned aerial vehicle

Table of contents

List of figures

List of tables

Chapter 1. Introduction

The analysis of pollen records from sedimentary environments has proved of considerable value in palaeoecology and archaeology. From a botanical point of view, stratified sequences of pollen in lake sediments and peat bogs make it possible to document the past vegetation communities by considering the whole assemblage of pollen grains (Moore et al., 1991). However, due to differences in pollen productivity and in the dispersal and deposition (d&d) processes between species, and the effects of different types of sedimentary systems on the accumulating record, pollen spectra and the surrounding vegetation do not share a one-to-one mapping. As understanding of the complications caused by these factors improves, the quantitative reconstruction of past vegetation can become more realistic and robust.

There are multiple approaches to interpretation of pollen records from sediments. In recent years, models of pollen-vegetation relationship have been an important focus, and are being developed into systematic approaches not only for interpretation of pollen spectra, but also as part of reconstructing past landscapes and climates. The accuracy of these models depends on 1) the d&d model being suitable for the study area, 2) good estimates of model parameters being available (e.g. pollen productivity, fall speed) and 3) an understanding of the ways that the sedimentary system in which the sampled pollen assemblage accumulated affects the signal recorded.

The existing models have been tested and validated successfully, but work so far is geographically limited to temperate zones (mainly in northern Europe and northern America). Therefore in this project, I want to expand the geographic areas of the model-based approach by applying it to a subtropical mountainous area. Mountains are chosen because they are less influenced by human impact and are more likely to have semi-natural vegetation.

Southeast China has a long history of land use, characterised by cultivation in the floodplains and lowlands, which could have an effect on the abundances and distributions of vegetation (Normile, 1997; Zong et al., 2007; Li et al., 2009). The climate here is greatly influenced by the activities of the summer and winter monsoons (Ding, 1994). The summer monsoon brings necessary precipitation to this area, the variability and strength in summer monsoon are closely related to the country's economy and people's lives (Hu et al., 2003). This region is therefore of particular interest for studies in

historical climate variations, past land-cover changes and the related anthropogenic drivers.

Located in the subtropical area, the Meiling Mountains have a typical monsoon climate with four distinct seasons (Wang, 1974). There are several mires in small hollows where peat has accumulated. About 50 years ago, a peat profile derived from Xiyaowu mire indicated the history of forest went through 4 phases, i.e. forest mainly composed by *Castanopsis*, forest dominated by *Pinus*, and then *Alnus* coexisted with *Pinus*, and finally pine forest mixed with *Quercus* and *Castanopsis*. However, this record was not radiocarbon dated, so the timing of these changes is not well understood (Wang, 1974). In 2013, 11 cores in the Xiyaowu mire were collected by team members from Nanjing University and one of the cores was selected for pollen analysis with intervals of 2 cm. This radiocarbon dated core shows clear changes in pollen assemblages during the little ice age (LIA) (spanning from 1300 to 2010 A.D., Cui et al., 2018), and reconstructing land cover change for this period would help to test theories on human-ecosystem-climate interactions. The feasibility and reliability of past land cover reconstruction relies on an understanding of the modern pollen-vegetation relationship. However, there are no previous studies about applying modelling approaches to this type of area. In this protected forest area, there are still a range of common sub-tropical natural vegetation types present (for vegetation details, see chapter 2). This area is therefore selected as a suitable study site to investigate the use of models of pollen-vegetation relationship in a sub-tropical biome.

1.1 Aims and objectives

The overall aim of this project is to explore the relationships between pollen and vegetation for assemblages deposited in peatland and forest landscapes in southeast China, with the goal of validating mathematical models of pollen dispersal and deposition, thus providing an essential basis for quantitative reconstruction of past land cover from Quaternary peat deposits in the region.

This will be addressed by achieving the following specific objectives:

Objective 1:

In many terrestrial landscapes, surface pollen assemblages are collected from soil samples, but standard model calibration methods recommend use of moss or lake

sediment (Gaillard et al., 2010). By collecting paired samples of moss and soil, both linked to the same vegetation survey, this study will investigate whether calibration results are systematically affected by sample type.

Objective 2:

Through study of modern pollen distribution in relation to mapped vegetation and habitats in the mountains, assess the extent to which the widely-used Prentice-Sugita model, developed for use in northern hemisphere temperate zone deciduous forests in topographically subdued landscapes, is suitable for use in this mountain region and in mire environments, and determine any necessary adjustments.

Objective 3:

Simulation and empirical measurement of a key taxon-specific model parameters, relative pollen productivity (RPP), for the plant taxa which dominate pollen records from the region.

Objective 4:

Assess spatial sensitivity of pollen representation of dry land vegetation in mire pollen signals, through collecting pollen samples in the mire and surrounding slopes from two different years.

1.2 Outline of thesis

The thesis is divided into eight chapters each dealing with a particular aspect of the research project. The introductory chapter (chapter 1) outlines the aims and scope of the project. Chapter 2 provides context for the research by summarising and commenting on previous research, providing a general background on reconstruction of past land cover using pollen analysis and introducing the field area. Literature specifically relevant to each of the objectives is reviewed in more detail in the results chapters 4-7. Chapter 3 deals with the methodologies used, gives an account of the reasons for choosing them, and shows how they are applied to meet each objective. Chapter 4 presents results of investigation of differences between moss and adjacent soil samples in the forest (objective 1). In chapter 5, pollen spectra obtained from a grass-forest transect are compared with the simulation results from several commonly used pollen dispersal and

deposition models (objective 2). Chapter 6 presents novel approaches in estimating pollen productivity (RPP). The widely used Extended R-value (ERV) method is known to not work well on small datasets or those with non-random sample distribution, but in environments such as sub-tropical montane forest large, truly random datasets cannot always be obtained. I present two new methods of estimating RPP which are alternatives to the widely used ERV method, test their behaviour using a simulation study, and then apply them to an empirical dataset to give a first set of RPP estimates from sub-tropical China (objective 3). Chapter 7 investigates how the mire sample pollen assemblages record the dryland vegetation. The differences between two sets of mire samples, which were collected in two different years, are discussed in this chapter (objective 4). Chapter 8 brings together the findings from the four main objectives, which gave insight into the reconstruction of past landscape, to discuss their application in the study region and beyond.

Chapter 2. Literature review

2.1 Reconstruction approaches and modern samples

Pollen and spores preserved in lakes and mires record changes in land cover over time. But reconstructing past land cover from pollen data is not a simple process; it requires a knowledge of the relationship between fossil pollen and former vegetation communities, which is affected by pollen production, pollen preservation and pollen dispersal and deposition (d&d).

2.1.1 The problem of interpreting pollen records

Pollen and spores are disseminated through different agencies such as wind, animals (primarily insects) and water. Pollen grains and spores may be settle onto the ground, and in some environments like lakes, peat, soil, moss polsters and so on, they become embedded in the stratum and get preserved as microfossils (Nair, 1966). These preserved layers are of great importance in providing information regarding the vegetation history through geological time. One of the most important applications of pollen studies is tracing the history of vegetation communities (Godwin, 1975). Pollen-based quantitative estimates of vegetation abundances and how these change over time have aroused much concern among palynologists, but there are still some problems to be solved (Tauber, 1965; Andersen, 1970; Prentice, 1985, 1988; Sugita, 1993, 1994, 2007a, 2007b; Nielsen, 2005; Bunting & Middleton, 2009; Hjelle et al., 2015; Bunting et al., 2018; Bunting & Farrell, 2018).

For a better understanding of palaeoecology and palynology, modern pollen deposits and their surrounding vegetation can be studied. Understanding the relationships between modern pollen and the vegetation which produced it improves the quality of reconstructions of vegetation and interpretations of past vegetation changes from past pollen records. Developing a systematic approach using both descriptive and modelling approaches has enabled researchers to develop a better understanding of this complex relationship.

The primary aim of plant palaeoecologists has always been to reconstruct the situation of past vegetation as precisely as possible (van Post, 1916). Different approaches have been tried to (semi) quantitative reconstruction of past land cover. There are two main ways of reconstructing past environments from pollen assemblages. One is the data comparison

approach by combining the pollen assemblages with the modern vegetation distribution in a large scale (e.g. modern analogue technique (MAT, Overpeck et al., 1985), biomization (Prentice et al., 1992)) and the other is the algebraic equation method (e.g. Davis, 1963; Prentice & Parsons, 1983; Sugita, 2007a, 2007b).

2.1.1.1 MAT

MAT works by matching modern pollen spectra from investigated vegetation communities with fossil assemblages. When the modern and fossil assemblages have an agreed level of statistical similarity, the relevant modern vegetation community is considered to be an analogue for the past environment (Overpeck et al., 1985; Nakagawa et al., 2002).

Disadvantages of the approach are the lack of training and inspection assemblages (that is, of modern data), and that all pollen spectra include both local and background components. The background represents the wider landscape, and differs both in space and time. Therefore even if the vegetation sampled in a modern nature area is a perfect match for the past vegetation at the fossil site, the modern pollen assemblage will likely to be different from the fossil one because the background pollen component is different.

2.1.1.2 Biomization

The biomization technique is designed to aid evaluation of climate modelling by reconstructing the distribution of major vegetation biomes at a continental scale. Pollen data in China have been used in the global "BIOME 6000 project" (Prentice et al., 2000; Pickett et al., 2004). The obtained BIOME 6000 palaeo-biome map is important for model evaluations at the global scale.

The limitation of this method is that it overlooks background pollen and relative pollen productivity. And the scope for this method is too general, i.e. a rough estimate of the vegetation from tree and treeless areas, which is not suitable for many research questions.

2.1.2 Models of pollen-vegetation relationship

Algebraic modelling of pollen d&d, the processes which link the land cover to the pollen signal forming at one spot, has been a fruitful line of research and is the one this thesis focuses on.

2.1.2.1 Theory

Tauber (1965, 1967) illustrated that pollen deposition in lakes and bogs is considered to consist of three main components (pollen derived by the wind from above the canopy (Cc), from the trunk space (Ct), long-distance component by rain (Cr)) and two additional components (gravity component from overhanging plants (Cg) and runoff component from the catchment area of the basin (Cw)) (Figure 2.1).

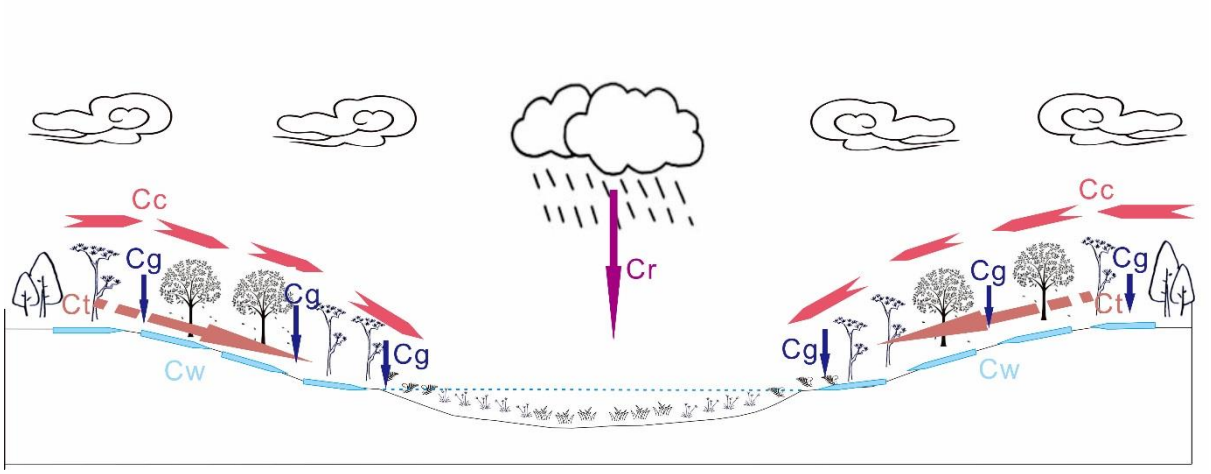


Figure 2. 1 Tauber model of pollen taphonomy into a surface sample in lakes and bogs. Cc: canopy component, Ct: trunk space component, Cr: washout component by precipitation, Cw: runoff component, Cg: gravity component (after Tauber, 1965, 1967; Jacobson & Bradshaw, 1981; redrawn from Bunting et al., 2013).

2.1.2.1.1 Davis – R-values

Davis (1963) proposed that a simple linear model could be used to "correct" pollen percentages to produce estimates of vegetation cover percentages. Vegetation and pollen values were linked by a taxon-specific ratio, the R-value, which gave a value for the amount of pollen produced per unit vegetation cover. She applied this method to show that the taxon with the highest pollen percentage in fossil pollen samples is not necessarily the most abundant plant, and argued that the interpretation that pine forests had once covered Vermont during the 'Pine Pollen Zone' time was incorrect. The model is of the form:

$$P_{ik} = R_{ik}V_{ik}$$

where R_{ik} is the pollen productivity of taxon i (assumed to be a constant for a given study region)

P_{ik} is the pollen percentage of taxon i at site k,

V_{ik} is the vegetation percentage of taxon i surrounding site k.

2.1.2.1.2 Andersen – adding the background term

In the discussion following Davis, Andersen (1970) found that some pollen from a tree can travel a very long distance, therefore if the vegetation is not measured to a very long distance, the R-value estimate will be wrong. However, long-travelled pollen should be well mixed in the atmosphere so should be the same in every lake in an area, so could be treated as a constant. He put forward the concept of background term (ω_i) which needs to be added in calculations of pollen deposition. The model now takes the form:

$$y_{ik} = R_{ik}X_{ik} + \omega_i$$

Where y_{ik} is pollen influx of taxon i at site k, X_{ik} is crown area of trees surrounding the site, ω_i is the background pollen contribution from taxon i in the study region.

2.1.2.1.3 Prentice – distance-weighting the vegetation

Prentice (1985) argues that vegetation needs to be distance-weighted – a tree close by makes a larger contribution to the pollen loading than one far off. He considers the Tauber model (Tauber, 1965) of taphonomy (Figure 2.1), argues that the canopy component (c_c) is most important for the mires he is interested in, and shows that the pollen signal has a distinctive shape when deposition is plotted against distance from the source – it is 'leptokurtic'. He suggests that Sutton's equation for a particle source at ground level under stable atmospheric conditions (Sutton, 1947a, 1953) is a suitable model to use, as it includes variation according to grain size/density (fall speed), variation in wind speed etc. He also argued that since pollen deposition from trees does not have a "skip distance" of low/no deposition close to the source, the elevated source or 'Smoke stack' model was less suitable. He presented an equation which is designed to model the pollen loading (PL) on a single point at the centre of a peatland basin, assuming no redeposition of pollen.

The basic form of the equation is:

$$y_{ik} = \alpha_i \psi_{ik} + \omega_i$$

where y_{ik} is the pollen loading of taxon i at site k, α_i is the pollen productivity of taxon i, and ψ_{ik} is distance-weighted plant abundance of taxon i in the vegetation within the source area around site k.

2.1.2.1.4 Sugita – a lake version of Prentice's model

In order to make the Prentice's model more appropriate for pollen deposited in a small lake (1-10 ha), Sugita (1993) extended the model to include mixing within a water body. His model integrates the pollen loading over the entire basin surface (effectively assuming total mixing).

Sugita (1993, 1994) also defined the boundary between local and background pollen components as the relevant pollen source area of pollen (RSAP). Determining this is an important step in both qualitative and quantitative vegetation reconstruction (Sugita, 1993, 1994). The 'relevant source area' of pollen is defined as the distance beyond which the pollen-vegetation relationship does not improve with additional vegetation data (Sugita, 1994). It is the distance where the correlation between pollen loading (PL) and distance-weighted plant abundance (DWPA) or likelihood function scores reaches an asymptote. The RSAP has been discussed and estimated from many landscapes and site types, including forest hollows in Michigan and Wisconsin (Calcote, 1995), cultural landscapes in southern Sweden (Broström et al., 2005) and Scotland (Bunting, 2003) and small lakes in Denmark (Nielsen & Sugita, 2005). Empirical estimation of RSAP for different site types vary markedly, for instance, studies in southern Sweden shows an RSAP of c.400m (Broström et al., 2005), whereas studies in Highland region of Scotland shows an RSAP on the order of 2 m or less (Bunting, 2003).

2.1.2.1.5 The generalised model of pollen dispersal and deposition

Their work combined is usually referred to as the Prentice-Sugita (P-S) model, which is a linear model based on the Sutton's distance-weighting algorithm.

The equation is:

$$y_{ik} = \alpha_i \int_R^{\zeta} \psi_{ik} + \omega_i$$

Where R is the basin radius, ζ is the pollen source area,

and $\int_R^\zeta \psi_{ik} = \int_R^\zeta g_i(z) x_{ikz} \mathrm{dz}$,

where $g_i(z)$ is the distance weighted term for taxon i, x_{ikz} is the plant abundance at distance z from the pollen deposition site k for taxon i.

2.1.2.2 Parameterising the models

A great number of studies have been carried out with the aim of applying the P-S model into modern pollen analysis in areas such as Minnesota (Sugita, 1994), Denmark (Nielsen & Odgaard, 2005) and southern Sweden (Broström et al., 2005). α_i and ω_i are two taxon specific parameters when using the P-S model for reconstruction of past vegetation. Works on estimations of relative pollen productivity and pollen fall speed have been made, such as those of Broström et al. (2004) in southern Sweden, Bunting et al. (2005) in UK, Duffin & Bunting (2008) in southern African and Hjelle & Sugita (2012) in Norway.

2.1.2.2.1 Pollen productivity – alpha (α_i)

It has been noted that not all plants produce the same amount of pollen. Previous studies have also stated that there are observed differences in estimation of pollen productivity for the same taxa between regions (Broström et al, 2008). For instance, entomophilous species usually produce less pollen than anemophilous plants; this led to variations in representation by comparison in the fossil record (Lowe & Walker, 2014).

Pollen productivity can be measured directly by measuring pollen grains produced per anther. But it is time consuming, hard to do and therefore rare (Pohl, 1937; Saito & Takeoka, 1985). For pollen diagram interpretation, the main interest concerns the number of grains produced per area unit of vegetation cover (Faegri et al., 1989). Therefore it is not necessary to estimate absolute pollen productivity for individual species.

Another possible way is to measure pollen productivity indirectly by comparing pollen assemblages and composition of surrounding vegetation. The slope of the linear function (α_i) represents the pollen productivity of an individual taxon. When pollen percentages instead of pollen loadings are used as pollen component, the linear relationship between pollen variables becomes non-linear, which is called the 'Fagerlind effect' (Fagerlind, 1952; Prentice & Webb, 1986). The Extended R-Value (ERV) models, which

incorporates a site-dependent weighting factor into R-Value model (Andersen, 1970), offer a good solution to handle the 'Fagerlind effect' problem.

A key issue here is to decide how far to collect the vegetation data and how much detail is needed. The minimum distance of vegetation survey needed to calculate x_{ik} for an individual site is termed the Relevant Source Area of Pollen (RSAP: Sugita, 1994, see section 2.1.2.1.4 for definition), and for empirical estimation of α_i using the ERV approach, vegetation survey is designed to cover an area typically around 2-5 times the anticipated RSAP (see e.g. Bunting et al., 2013) usually via a combination of direct field survey and remote sensing (e.g. aerial photograph interpretation). The empirical methods of estimating relative pollen productivity is considered in more detail in chapter 6.

2.1.2.2.2 Pollen fall speed - v_s

Deposition velocity of pollen is another important parameter to be included in using P-S model. Fall speed calculation is measured either directly (e.g. Eisenhut, 1961), or estimated based on the equation of Stoke's law (Gregory, 1973; Sugita et al., 1999), which takes particle size and density into consideration.

The settling velocity (v_s) for a spherical particle is expressed as

$$v_s = \frac{2r^2 g(\rho_0 - \rho)}{9\mu}$$

Where:

r = particle radius (cm)

g = the acceleration due to gravity (981 cm/s^2)

ρ_0 = particle density (g/ cm^3)

ρ = fluid density (1.27×10^{-3} g/cm^3 for air at 18°C),

μ = dynamic viscosity (1.8×10^{-4} g/ cm* s for air at 18°C).

For ellipsoid particles, the fall speed is expressed as

$$v_s = \frac{2\sqrt[2]{a/b}g(\rho_0 - \rho)}{9\mu}$$

Where:

a= major axes of the ellipsoid

b= minor axes of the ellipsoid (Gregory, 1973).

2.1.2.3 Reconstructing past land cover - Landscape Reconstruction Algorithm (LRA) and Multiple Scenario Approach (MSA)

Once calibrated, these algebraic descriptions of the pollen–vegetation relationships can be used to reconstruct past land cover. Two reconstruction methods based on the P-S model have been proposed in the last decade.

The 'Landscape Reconstruction Algorithm' (LRA), proposed by Sugita (2007a, 2007b), has been successfully empirically validated in southern Sweden (e.g. Hellman et al., 2008a, 2008b), central Europe (Soepboer et al., 2010), and the forest hollows in the northern US (Sugita et al., 2010). The LRA has two steps, each managed by a different model, the Regional Estimates of VEgetation Abundance from Large Sites (REVEALS) and the LOcal Vegetation Estimate (LOVE) model (Figure 2.2a). The REVEALS model uses pollen assemblages from large lakes (>=100ha) to reconstruct regional vegetation abundance within at least 50-100k radius of the lakeshore. From this, it is possible to estimate the past background (regional) pollen rain, and the LOVE model uses these background estimates along with pollen records from small sites (<100 ha) to work out the DWPA within the RSAP at these small sites (which ranges from 400-2000m radius).

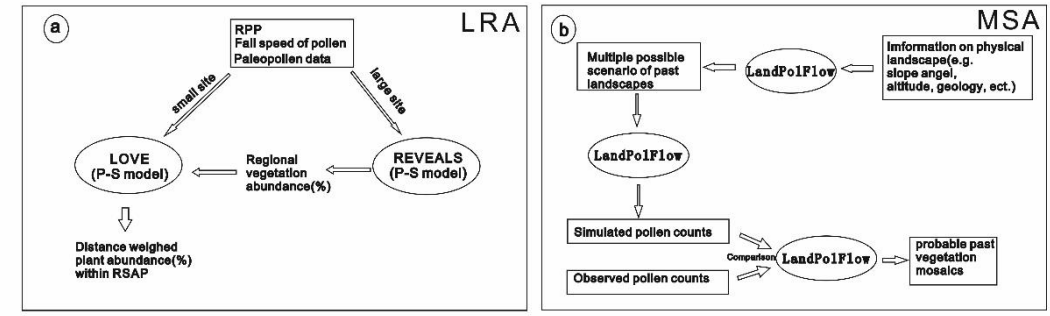


Figure 2. 2 Flow diagram of LRA and MSA (after Sugita, 2007a, 2007b and Bunting & Middleton, 2009; redrawn from Bunting & Middleton, 2009).

An alternative approach, the Multiple Scenario Approach (MSA) (Bunting & Middleton, 2009), begins by generating possible vegetation mosaics using GIS, and then uses the P-S model to simulate pollen spectra at known sampling points for each of these landscapes (Figure 2.2b). By comparing these simulated assemblages with real pollen assemblages from the past, the most likely reconstruction(s) of past land cover can be selected. This method has also been used to reconstruct land cover dynamics in recent years (e.g. Bunting et al., 2018).

A major advantage for MSA is that it is particularly suitable for landscapes with complex structure. Unlike the relatively uniform landscapes in temperate Europe, the vegetation mosaics in Meiling Mountains vary significantly over relatively short distances. Secondly, the mire available for reconstruction of the regional vegetation is too small for the REVEALS step of the LRA, and there are no suitable published regional pollen diagrams available.

2.2 Review of research in China and the case study region

2.2.1 Pollen records from China

Pollen records have long been used as proxies of past vegetation and climate. With the growing number of pollen records available, projects concerning regional and global pollen databases are being built up. In recent years, the global pollen database (GPD) and several regional pollen databases from Europe, North America, Latin America and Africa have been successfully established (Gajewski, 2008), aiming at traditional use of palaeoclimatic and palaeoecological analysis (Davis et al., 2003; Viau et al., 2006), testing ecological and biogeographical hypotheses (e.g. Mitchell et al., 2005; Foster et al.,

2006; Peros et al., 2008) and studying changes in the carbon cycle (e.g. Gajewski et al., 2001). Although China is not included in the GPD, there are numerous fossil pollen records available. The number of pre-Quaternary pollen records in China is much less than those covering the Quaternary period (Zhao, 2018).

The Holocene is the period since the end of the last glaciation, and when human impacts on global vegetation become significant. Researchers performed many investigations over the world on the aspect of reconstruction of changes during this period, particularly after the mid-Holocene (e.g. Ritchie et al., 1985; Waller & Hamilton, 2000; Yu et al., 2000; Fletcher et al., 2013; Quamar & Nautiyal, 2016). The Holocene is also a major focus for palaeoenvironmental research in China aimed at understanding both human activity and associated impacts on the natural environment.

During the early to mid-Holocene, the climate was warm and wet. The Holocene Optimum, which is often regarded as the time of maximum postglacial warmth (e.g. Winkler & Wang, 1993), was asynchronous in different regions of China (An et al., 2000; Guo et al., 2000; Chen et al., 2006). Recent research on climate change during Holocene is mostly concentrated on several critical periods, such as the Holocene Climatic Optimum, the Medieval Warm Period, the Little Ice Age (LIA) and the Late Holocene.

There are more than 200 pollen records starting from the Last Glacial Maximum (LGM) in China, which provide good record of the broad temporal and spatial patterns of climate and vegetation (summarised by Cao et al., 2013). Modern climatological research showed that the eastern Asia is currently affected by three relatively independent systems, i.e. the Indian monsoon, the East Asian monsoon and the westerlies (Hong et al., 2005; Wang et al., 2010). This area is therefore assumed to be particularly sensitive to monsoonal evolution and global climate change. Studies about the behaviour and dynamics of East Asia Summer Monsoon (EASM) since the last deglaciation is a key concern for many pollen analysts (e.g. Wen et al., 2010; Chen et al., 2015; Xu et al., 2017).

2.2.1.1 Holocene climate/vegetation patterns from selected pollen records in China

In the last two decades, numerous pollen-based studies in China have attempted to trace late Quaternary climate/vegetation histories by investigating fossil pollen profiles from lake records or by reviewing published pollen records. For example, a high-resolution pollen record over the past 18ka from Qinghai Lake, Tibetan Plateau (Shen et al., 2005)

indicated a cold and dry climate before the Holocene (steppe), a wet climate in the early and middle Holocene (steppe forest, 8.5-4 ka, 1ka means 1000 cal yr BP) and a dry climate in late Holocene (*Artemisia*-dominated steppe). Another high-resolution pollen record from the Zoige Basin on the eastern Tibetan Plateau (spanning 10ka from Zhao et al., 2011) suggest a warm and wet climate prevailed in the mid-Holocene (6.5-4.7 ka) and a cooling and drying trend in late Holocene. In the semi-arid Inner Mongolia area, pollen records (Jiang et al., 2006) over 12.5ka indicated a cold and dry climate in early Holocene (steppe, 12.5ka to 9.2ka), followed by a relativity humid climate in mid-Holocene (forest steppe 9.2ka to 6.7ka) and cooling again in late-Holocene (steppe). Analysis of pollen records covering the past 16000 years in central China shows that climate here has undergone significant change during the Holocene, divided broadly into a warm period after the Last Glaciation (from 16 to 12.7 ka, characterized by large amount of warm temperate conifers and broad-leaved taxa), the Holocene Climatic Optimum (from 11 to 6 ka, featured by substantial increasing amount of evergreen broad-leaved taxa), and a cooling period during late Holocene (from 4 ka characterized by a noticeable regression of arboreal plants) (Zhu et al., 2010).

There are also studies aiming at the general and regional differences in vegetation and climate based on published pollen records (e.g. Herzschuh, 2006; Zhao et al., 2009a, 2009b). In the arid and semi-arid northwestern China, Zhao et al. (2009a) summarized 30 published pollen records and found that although those records show a drying trend during the late Holocene, there are spatial difference in vegetation and climate during the Holocene. In the monsoonal region of China, Zhao et al. (2009b) synthesised 31 fossil pollen records and indicated a humid climate in the early and middle-Holocene and a drier climate during the late Holocene.

2.2.2 Modern pollen studies in China

As a basis for reconstruction, surface pollen studies have also aroused much interest in China. A variety of kinds of sampling materials (e.g. soil, moss, waterlogged surface sediments (sea bottom, mires, lake) and traps) have been used. Building on empirical field-based studies, more recent research has been extended to quantify pollen-vegetation-climate relationships. These studies have enabled the first analysis of pollen taxa relative to plant abundance at a sub-continental scale (Zheng et al., 2008).

For studies in pollen-climate relationships, in order to develop pollen-climate calibration or training set, sufficient surface samples analysed along long environmental gradients with modern climate data is an essential (Seppä et al., 2004). In China, many regional scale reconstructions or reconstructions spanning previous periods have used modern pollen-climate datasets as calibration sets. The main existing surface pollen database used incorporates the East Asian surface pollen database (consisted of more than 5000 surface samples, Zheng et al., 2014) and modern pollen on the Qinghai-Tibetan Plateau (1020 sites, Lu et al., 2011). Some site-specific climate reconstructions also use regionally restricted calibration pollen sites with less numbers, such as studies in the eastern Qinghai-Tibetan Plateau (227 modern pollen sites from Shen et al., 2006, 112 surface lake sites from Herzschuh et al., 2009a, 2009b, 2010).

Surface pollen-vegetation studies using modelling methods have begun in China in recent years. A key parameter in reconstruction, relative pollen productivity, has been produced in the northern China using ERV analysis (e.g. Wang & Herzschuh, 2011; Li et al., 2011, 2015; Wu et al., 2013; Xu et al., 2014; Ge et al., 2015; He et al., 2016; Han et al., 2017; Li et al., 2017, 2018). Availability of these values have enabled the application of REVEALS model to reconstruct regional land cover (Wang & Herzschuh, 2011; Xu et al., 2014). Land cover reconstructions over larger areas using LRA and MSA in other areas of China, i.e. subtropical areas, require RPP datasets across this region. However, there are no published RPPs in subtropical China, therefore this study is carried out. A global pollen-based reconstruction project, PAGES LandCover6k group (Gaillard, 2015; Gaillard et al., 2018), is currently being implemented. This study estimates RPPs in new areas of China, and will benefit the LandCover6k project.

2.2.3 Past vegetation and land cover history in southeast China

2.2.3.1 Past climate/vegetation in southeast China

In subtropical southeast China, there are pollen records available, which are mainly from lake, peat and fluvial sediments with abundant pollen types. Some geographic gaps still exist in the sampling locations, however. Pollen records in this region are mainly distributed in the Yangtze River Delta (e.g. Han et al., 2000; Shu et al., 2008; Zhao et al., 2007) and the southern coast (e.g. Chen et al., 1994; Huang et al., 1982; Zheng, 1990, 1991; Li et al., 1996). There are also some pollen studies using peatland profiles which are scattered in the hilly areas (e.g. Zhou et al., 2004; Dodson et al., 2006; Zhu et al.,

2006; Xiao et al., 2007; Zhao et al., 2016, 2017) or lakes (Jarvis, 1993; Chen et al., 2009). Based on these Holocene pollen records, the vegetation patterns can be summarised as from evergreen broadleaved forest during early and mid-Holocene to evergreen deciduous broadleaved/mixed forest during the late Holocene, but the dominant vegetation shows different patterns. A general synchronous wet-dry climate changes during the Holocene can be found from these records, characterised by a wet climate in the early and mid-Holocene (before 6-5 ka) and a dry climate during the late Holocene. Most of these studies take the assumption that pollen percentages are a direct reflection of plant abundance, but take little consideration about the pollen characteristics in reflecting the vegetation (i.e. differences in pollen productivity, the d&d properties between species and the effects of different types of sedimentary systems on the accumulating record).

There is limited research in climate reconstructions based on pollen records in southeast China. For example, Zhao et al. (2009b) synthesized 31 published Holocene pollen records in monsoonal region of China and reconstructed moisture histories using a 4-class ordinal wetness index.

About 50 years ago, a peat profile derived from Xiyaowu mire in the Meiling Mountains indicated that the history of forest went through 4 phases, i.e. forest mainly composed by *Castanopsis*, forest dominated by *Pinus*, and then *Alnus* coexisted with *Pinus*, and finally pine forest mixed with *Quercus* and *Castanopsis*. However, this record was not radiocarbon dated, so the timing of these changes is not well understood (Wang, 1974). In 2013, 11 cores in the same mire were collected by team members from Nanjing University and one of the cores was selected for pollen analysis with intervals of 1 cm. Pollen assemblages of a core in the mire were available (Cui et al., 2018). The ^{14}C AMS dating data from this core indicates a history during the Little Ice Age (LIA). The LIA is defined as a moderate cooling period which occurred from about the16th to the 19th centuries. It is the most recent period of dramatic mountain glacier expansion following the Medieval Warm Period (Mann, 2002). In recent years, based on evidence from historical literature (e.g. Zhu,1972; Yang et al., 2006; Ge et al., 2011; Yan et al., 2012), tree-ring (e.g. Liu et al., 2010; Zhu et al., 2011), lake sediments, stalagmites (e.g. Zeng et al., 2012) and cave records in China, significantly colder conditions were documented in most records. However, the temporal and spatial climate change patterns vary between different regions. Minor changes in temperature fluctuations is a key character of climate change during LIA. Records in Tibetan ice cores suggest a wet climate during the LIA (Yao et al., 1996, 1997). Chen et al. (2006) documented a humid LIA with oscillations

between warm-dry and cold-humid climate conditions in arid areas of northwest China. Zeng et al., (2012) report a wet LIA from lake sediments in tropical south China.

2.2.4 The gap identified and what is needed to address it

2.2.4.1 Environment and context of the Meiling Mountains

The Meiling Mountains (also known as the Xishan Mountains or Mount Hsishan, 28°31′N - 28°54′N, 115°34′E -115°53′E) are located northwest of Nanchang city, in southeast China (Figure 2.3). The mountains, which are aligned in a northeast-southwest direction, are mostly composed of porphyritic granite and gneiss. The mountains occupy an area of 150 km^2. Faulting led to the development of valleys which can support wetlands.

The highest peaks of the mountain reach about 950m a.s.l. (above sea level). The mean annual temperature and precipitation are about 18.8℃ and 1760mm, respectively (China Meteorological Administration Data Service Center, 1981-2010). Much of the land cover is secondary forest, with some tea plantations in valleys and grasslands on ridges and around small settlements (vegetation mapped in Figure 4.1).

Due to early tectonic denudation and late weathering erosion, the landscape in Meiling Mountains is characterised by a mosaic of hollows and low hills. This allows peat sediments to develop and survive. Pollen is well preserved in such environments because they provide anaerobic conditions which reduce oxidative stress (Knox, 1979). The Xiyaowu mire (28°44′N, 115°40′E) is a waterlogged area located in the ridge zone of the main peaks (Sun, 1980). The elevation of the mire is about 748-755 m. It covers an area of 13334 m^2 and contains a maximum depth of about 250 cm of peat. The basin is lined with a thick silty clay layer with high water retaining capacity which supports hygrophilous plants. The peats were formed gradually from the continuous deposition of plant residues and humus.

The vegetation communities in the Meiling Mountains are mainly composed of subtropical needleleaf forest (dominated by *Pinus massoniana* and *Cunninghamia lanceolata*), subtropical broadleaf deciduous forest (characterized by *Castanea sequinii*, *Quercus serrata* var. *breviptiolata* and *Platycarya strobilacea*), subtropical broadleaf evergreen forest (dominated by *Castanopsis sclerophylla* and *Cyclobalanopsis glauca*)

and subtropical bamboo forest (*Phyllostachys edulis*) and scrub. There are small areas of shrubs, farmland and grassland scattered in the valleys.

Many species of aquatic plants grow in the Xiyaowu mire, including shrubs (*Salix chaenomeloides*) and wetland herbs such as *Juncus effusus*, *Sphagnum palustre*, Poaceae spp., *Viola lactiflora*, *Polygonum thunbergii*, *Ligularia japonica* and *Rotala rotundifolia*. Belts of tea tree (*Camellia sinensis*) line three sides of the mire, with cedar (*Cryptomeria fortunei*) growing in higher altitudes. The southern slope is covered by shrubs such as *Camellia oleifera* (tea for oil) and *Lindera reflexa.*

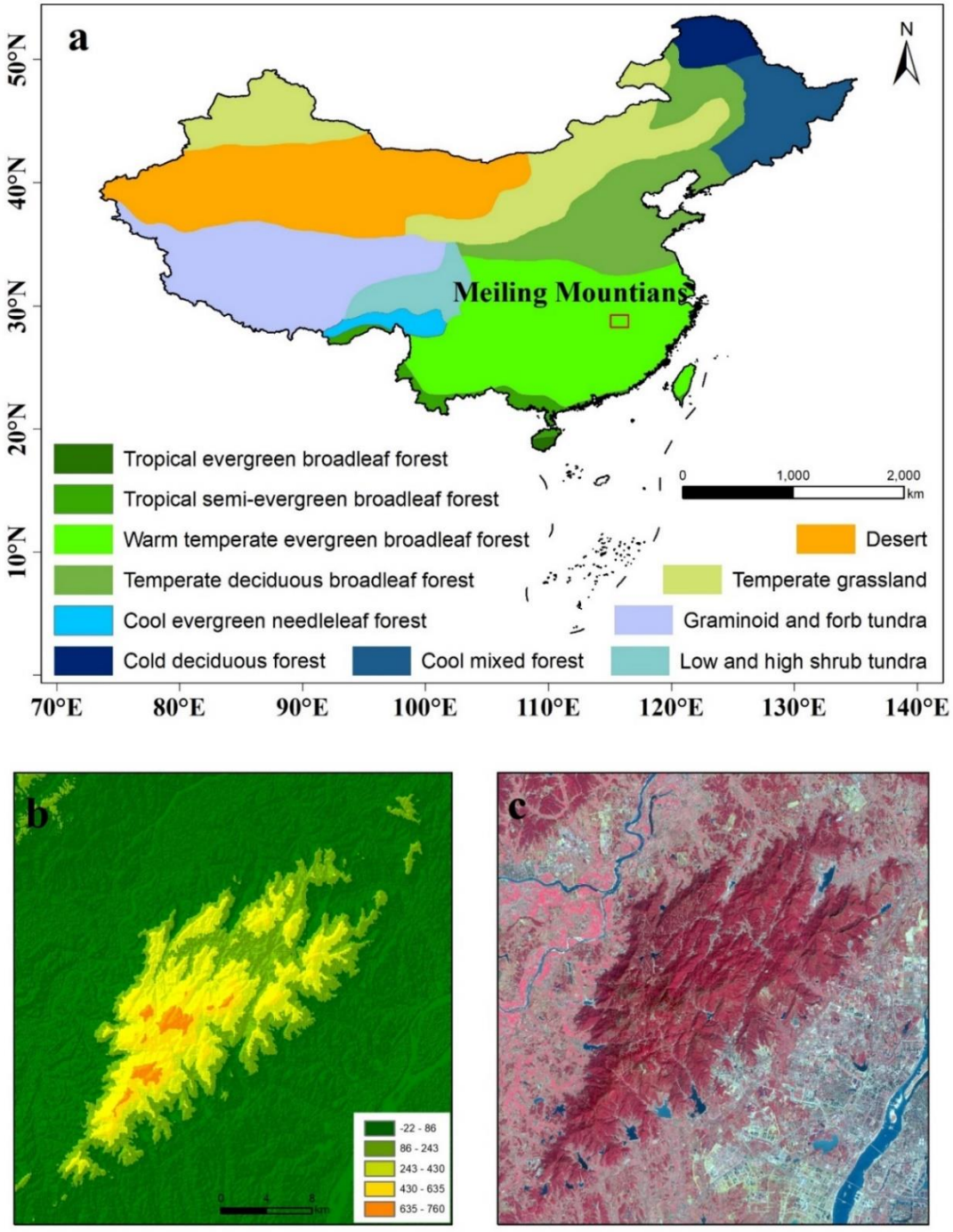


Figure 2. 3 Map showing the study area. a) The biome map of China (modified from Olson et al., 2001 and Zheng et al., 2014), b) topography of Meiling Mountains, c) satellite image of Meiling Mountains (from European Space Agency (ESA) data), dark red shows the forested areas and picks out the Scenic Area very effectively, blue shows the river and lakes, grey shows the roads and towns/city. Maps of sampling sites for each objectives are shown in chapters 4 to 7, respectively.

2.2.4.2 Other researchers' calls for further research in this area

The study of surface samples from modern vegetation types is a promising method of approaching the interpretation of past pollen spectra in terms of the vegetation which gave rise to them. The interpretation from a pollen spectrum to the vegetation is a process that requires understanding of all the relevant processes (i.e. pollination, sedimentation and sampling) involved. In the heavily populated regions of southeast China, there are pollen records from valley mires available (e.g. Zhou et al., 2004; Dodson et al., 2006; Zhao et al., 2016, 2017), which contain records of past land-cover, but few studies have conducted the surface study in the region.

With the developments of the theory in pollen-vegetation relationships (Andersen, 1970; Prentice, 1985; Sugita, 1994), new frameworks of vegetation reconstruction have been developed, i.e. the LRA and MSA. Both reconstruction methods are based on the assumption that the d&d model chosen is suitable for the system being studied. This is based on the 'pollen sample's eye of view' that vegetation growing closer to the sampling site contributes more than plants far away (Calcote, 1995). The most widely used DWPA term is a taxon-specific distance derived from Sutton's equation (Sutton, 1953), which assumes neutral atmospheric conditions and treats the atmospheric parameters as constants. However, pollen dispersal may occur under unstable conditions. Jackson & Lyford (1999) investigated the influences of three atmospheric parameters (i.e. wind speed, turbulence and vertical diffusion) under unstable atmospheric conditions, and indicated more widespread dispersal from individual sources. There are also more complex models (Lagrangian stochastic model (LSM), Andersen, 1991; Kuparinen et al., 2007) and less complex models (e.g. z^{-1}, z^{-2}, $\log_2(z)$, e^{-z}, $e^{-0.5z}$, Broström et al., 2004; Mazier et al., 2008; Bunting & Hjelle, 2010) which have been explored in the past decades. More details about these different distance weighting terms can be found in section 5.1. These models have never been tested in the hilly areas of south east China.

Application of these reconstruction methods (LRA and MSA) also require definition of taxon-specific parameters which need to be estimated via studies of modern pollen-vegetation relationship. Relative pollen productivity (RPP), a key taxon-specific model parameter, have been obtained mostly in temperate European areas (e.g. Soepboer et al., 2007; Mazier et al., 2008; Poska et al., 2011; Sjögren et al., 2008a, 2008b; Matthias et al., 2012; Theuerkauf et al., 2013). Another cluster of published values are available for

northern temperate China (reviewed in Li et al., 2017), but no values yet exist for southern parts of China.

Moss polsters and soil are considered as two main 'natural' traps of pollen. Standard model calibration methods recommend use of moss or lake sediment (Gaillard et al., 2010), but in areas such as arid and/or semi-arid environments, soil is the only easily accessed deposit of modern pollen. Soil samples are often biased by factors and should be interpreted with caution (Adam & Mehringer, 1975; Hill, 1996). Little research is available to state if there were systematic differences between moss and the adjacent soil samples, and their representative of the surrounding vegetation.

Pollen records preserved in mires are important in interpretation of past land cover, there are challenges, however, due to the influence of plants within a few meters of the sampling point (Bunting, 2003). How dry land vegetation is being recorded by surface mire samples is an important issue to address, therefore to aid the interpretation of pollen records in mire samples. A great amount of study have been put forward to explore the relationships of pollen and vegetation in forest hollows (e.g. Andersen, 1970; Bradshaw, 1981; Calcote, 1995) or small lake basins in forest (e.g. Prentice & Parsons, 1983; Bradshaw & Webb, 1985; Jackson, 1990), using both empirical and theoretical approaches, but there is no published study of modern pollen assemblages in a small mire in fully forested landscapes for subtropical southeast China.

Therefore, this project is planned and carried out aiming at filling the gaps listed above.

Chapter 3. Methods

3.1 Introduction

The aim of this thesis is to extend application of pollen modelling approaches to the subtropical climate zone and to mountainous terrain, seleting the Meiling Mountains of southeastern China as a suitable study region. The specific aims are to improve understanding of pollen-vegetation relationships in southeast China (explained in chapter 1), using the sampling of surface sediments (soil, moss or peat), investigation of the surrounding vegetation communities, with the goal of validating mathematical models of pollen dispersal and deposition, thus providing an essential basis for quantitative reconstruction of past land cover from Quaternary peat deposits in the region.

Four distinct objectives are identified in chapter 1, which will be achieved by using paleoecological methods to analyse different sets of surface samples. This chapter outlines these methods in general terms, referring for more detail to the relevant parts of chapters 4-7.

3.1.1 Breakdown of methods by objective

Table 3.1 summarises sampling strategies used for each of the objectives.

Table 3. 1 Objectives and the corresponding sampling strategies

Objective	Sampling strategies	Analysis included
1. By collecting paired samples of moss and soil, both linked to the same vegetation survey, investigate whether calibration results are systematically affected by sample type. (results presented in chapter 4)	42 paired soil and moss samples (both collected within 1m^2 area), distributed in five main forest zones (see section 3.2.1.1) Vegetation survey (see section 3.3)	Chemical treatment of surface samples Pollen counting Regional vegetation map Principal Component Analysis (PCA) Vegetation diversity analysis: Shannon index and Simpson index

		Counting error analysis Simulation of counting errors
2. Through study of modern pollen distribution in relation to mapped vegetation and habitats in the mountains, assess the extent to which the widely-used Prentice-Sugita model, developed for use in northern hemisphere temperate zone deciduous forests in topographically subdued landscapes, is suitable for use in this mountain region and in mire environments, and determine any necessary adjustments. (results are presented in chapter 5)	11 moss and adjacent soil samples collected along a 1000m transect (every 100m), perpendicular to a major vegetation boundary (see section 3.2.1.2) Vegetation survey (see section 3.3)	Chemical treatment of surface samples Pollen counting Regional vegetation map Comparison of simulated data using different distance weighting methods (taxon-specific weighting based on Sutton's equations; simple linear model; LSM) with empirical data
3. Simulation and empirical measurement of relative pollen productivities (RPPs) for taxa which dominate pollen records from the region, using 'Extended R-value' (ERV) analysis (Parsons & Prentice, 1981; Prentice & Parsons, 1983; Prentice, 1985; Prentice & Webb, 1986), modified Davis method (MDM) and iteration method (IM). (results are presented in chapter 6)	Simulation study to establish two alternative methods to estimate RPP • Modified Davis method (MDM) • Iteration method (IM) Empirical study: 10 moss samples with standardised vegetation survey (Bunting et al., 2013) (see section 3.2.1.3)	Simulation study • Landscape created in Mosaic • Pollen loading simulated in PolFlow • ERV, MDM and IM Empirical study: • Chemical treatment of surface samples • Pollen counting • Regional vegetation map • Fall speed • ERV, MDM and IM
4. Assess spatial sensitivity of pollen representation of dry land vegetation in mire pollen signals, through collecting pollen samples in the mire and surrounding slopes from two different years.	Samples collected in 2014 (7 surface sediment samples in the mire, and 14 moss polster samples from the surrounding slopes) Samples collected in 2016 (6 surface samples in the mire, and 3	Chemical treatment of surface samples Pollen counting Regional vegetation map

(results are presented in chapter 7)	moss polster samples from the surrounding slopes) (see section 3.2.1.4)	Principal Component Analysis (PCA)

3.1.2 Reasons for choosing different sampling methods for different objectives

Studies of surface sample and simulation approaches improve quantitative reconstruction of vegetation from pollen data. Modern pollen assemblages are usually obtained by collecting samples from natural pollen traps such as moss polsters, soil, peat and lake sediments. The type of surface material used for sampling must be selected with care, because the composition of pollen assemblages will generally vary with sedimentation conditions (Moore et al., 1991). Soil, moss and peat deposits are used in this thesis to address different objectives. The pros and cons of each sampling materials are fully discussed in chapter 4 (moss and soil) and chapter 7 (peat).

In this thesis, the purpose of modern samples is to aid the interpretation of fossil pollen assemblages from sediment cores, therefore it was decided to use moss polsters in most sub-studies as these have been found to be effective natural pollen traps (Berglund et al., 1986) and are regarded as the most appropriate analogue for the interpretation of peat cores (Prentice, 1986). Surface pollen studies in north and north-east China (e.g. Xu et al., 2007; Zhao et al., 2009c; Li et al., 2011), often have to use soil samples due to a lack of widely distributed mosses. Since soil and moss are both widely distributed in Meiling Mountains, this offers a useful opportunity to compare pollen signals from both, which will then make it easier to compare results with previous studies. Objective 1 was achieved by carrying out studies of paired moss and soil samples collected within 1 m^2 square area.

Model-based landcover reconstructions (LRA and MSA) are based on the assumption that the d&d model chosen is suitable for the system being studied. For objective 2, several different d&d models were tested against empirical data from a transect of samples across a vegetation boundary which is expected to have a clear pollen signal, a woodland to grassland transition.

Moss polsters are widely used in RPP estimation studies of northern Europe (e.g. Broström et al., 2004; Bunting et al., 2005; Baker et al., 2016). Previous studies have

identified some problems with soil surface sample pollen assemblages (Adam & Mehringer, 1975; Hill, 1996; Xu et al., 2016). Based on the results of the moss and soil comparison study (see detailed discussion in section 4.5.4 and 8.1), moss samples are used for RPP estimation for objective 3.

Pollen preserved in bogs provides a synthetic record of past vegetation (Tauber, 1965), including both a signal from the plants found within the bog and from the vegetation growing on the surrounding dry land. Surface peat samples were collected in Xiyaowu mire (Objective 4), where a sediment core was retrieved from the centre of the mire (Cui et al., 2018). This study sought to investigate the effects of different mire communities on the pollen signal of the dry land vegetation, in order to identify possible guidelines for reconstruction of past land cover from a mire pollen core (see e.g. Bunting et al., 1998a, 1998b).

3.2 Fieldwork for modern pollen sample collection

3.2.1 Site selection

The overall aim of the study is to aid the interpretation of pollen assemblages from a mire core (collected and counted before the start of this study) in the Meiling Mountains. The surface pollen sites were selected to address the four objectives above. The maps showing the sampling locations are given in each four main chapters (chapters 4-7).

3.2.1.1 Selection of paired moss and soil samples

This objective requires comparing soil and adjacent moss polster pollen assemblages (within 1 m distance) to see if there were systematic differences. The strategy for collecting samples was limited by the logistical constraints of a challenging field study area and time constraints of fieldwork. Both sample types were collected wherever present, in parallel with all other field-based sample collection when sampling for objectives 2-4.

3.2.1.2 Selection of transect samples

There is a sharp decrease in forest tree pollen incidence at the border between a forest and a non-forest area (Faegri et al., 1989). The aim for these samples is to see if the simulated pollen signal can replicate broad patterns of pollen assemblage. A 1000m transect which runs across major boundaries of grassland-forest, and across a range of altitude values

was selected. Moss polsters and the adjacent soil samples at 100m intervals along the transect were collected as surface samples. In order to make the vegetation information comparable, standardised vegetation survey using the Crackles Bequest Project methodology (Bunting et al., 2013) were conducted at locations 1, 6 and 8. For the other locations, the vegetation were recorded within a 1 m x 1 m square area by visual estimation of percentage cover.

3.2.1.3 Selection of RPPs samples

The following criteria was used to select RPP samples for empirical study: 1) random sample points covering the main species in the forest; 2) sample points must be at least 200m away from each other; 3) chosen so that overall there are samples with low, medium and high abundance in the plants of the main taxa

Criterion 1) is an attempt to ensure main taxa being included in the RPP estimation. As RPP estimation requires vegetation data at each sampling point to the 100m range, criterion 1) also ensures systematic and detailed information about the vegetation in regional areas. Criterion 2) is to avoid overlap of the vegetation survey area. Successful ERV analysis expects scatter plots of pollen proportion against vegetation abundance to show a generally monotonic relationship, that is, sufficient pollen samples within ranges of communities. Criterion 3) is to ensure the samples covering range of percentage cover of main pollen producing taxa, and thus adds the range of plant abundance in scatter plot.

3.2.1.4 Selection of peat samples

Pollen preserved in mires provides a synthetic record of past vegetation (Tauber, 1965), including both signal from the plants found within the bog and from the vegetation growing on the dry land. By recording the vegetation communities within typical mires, and taking multiple pollen samples, it should be possible to get better understanding of how the dryland signal is affected by variations in the local vegetation within the mire.

The boundaries of different vegetation communities were defined in and outside the mire. Samples in 2014 (pilot study before the start of this thesis) are mostly collected near the edge of the mire. In order to make the samples cover the inner area of the mire, we re-sampled in 2016. A central sample was first taken, then stratified random samples were taken ensuring samples are at a range of distances from the edge and a range of different vegetation communities.

3.2.2 Collection of samples

The coordinate of each sample site was recorded with a hand-held global positioning system (GPS). Altitude of each sample was also recorded. In order to make the samples taken of known volume, all samples were collected using inverted sample containers measuring 7.5 cm x 7.5 cm x 1 cm in size. Only the green parts of the moss polster were collected.

3.3 Vegetation map of field area

The Editorial Board of Vegetation Map of China (2007) provides land-cover data within major land types subdivided into several subclasses. Using this coarse classification of vegetation for the Meiling Mountains groups some samples which were collected in different communities into one community (e.g. only one mire community is shown). This is too coarse for the needs of this study. In addition, the publication year of the existing map is 2007, and the vegetation might have changed. Therefore in this study, a new regional vegetation map is obtained from interpretation of the remote sensed image coupled with detailed ground-based vegetation surveys to define the boundaries of the communities (Figure 3.1).

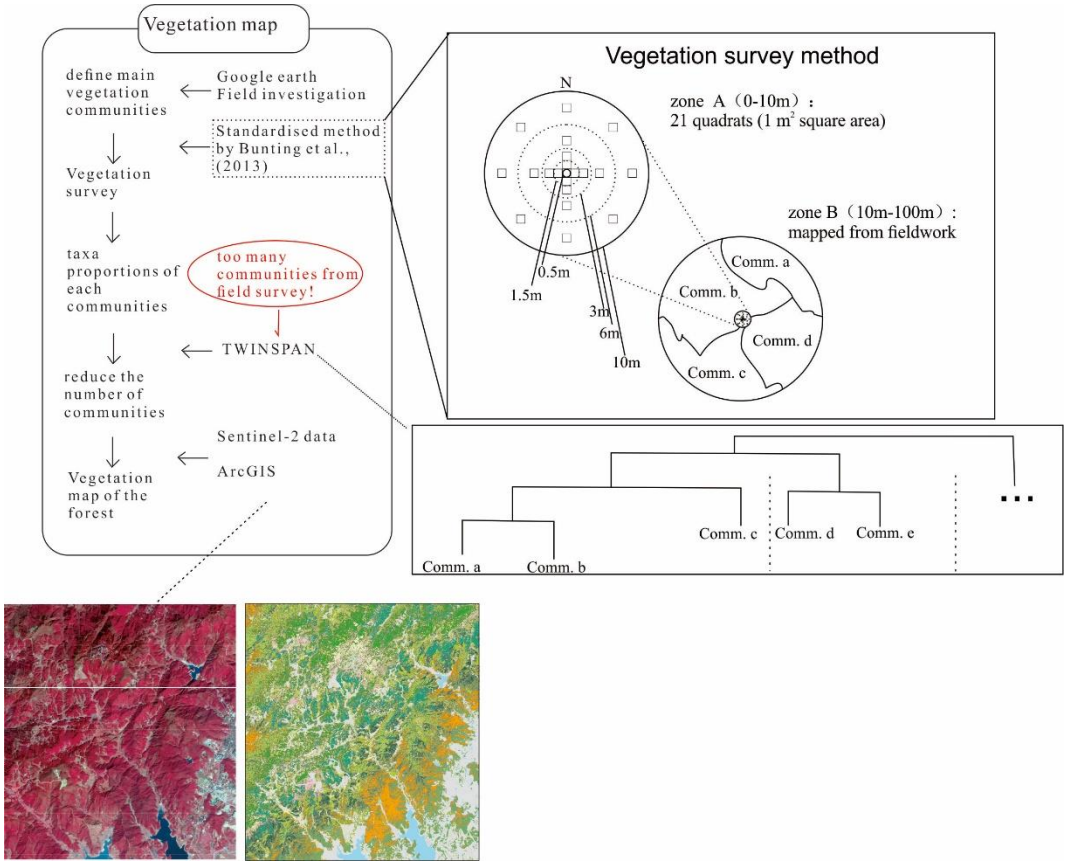


Figure 3. 1 Procedures of mapping the vegetation communities. This method was used for moss & soil comparison (see chapter 4), local map for transect (see chapter 5), RPPs in empirical study (see chapter 6) and vegetation map in and outside the mire (see chapter 7).

3.3.1 Ground-based vegetation survey in main forest types (0-100m around pollen sample points) for objective 3

10 random survey points were selected based on satellite maps from Google Earth and on-the-spot investigation. The criteria of the selection is discussed in detail in section 3.2.1.3.

In order to get systematic and detailed information about the vegetation at each sample site, vegetation survey was conducted using the Crackles Bequest Project methodology (Bunting et al., 2013). The vegetation survey strategy is shown in Figure 3.1. The community compositions in zone B (10-100m) were recorded as percentages.

3.3.2 Mapping communities in the wider landscape

Satellite remote sensing is a unique and useful source of data for the identification of different types of landscapes over large areas, as described in the last two decades in literature (Collado et al., 2002; Ozesmi & Bauer, 2002; Stow et al., 2004; Heumann, 2011). Generally, there are two major ways of image classification techniques, one is unsupervised classification, and the other is supervised classification. The outcomes of unsupervised classification are based on the software analysis of the image. The computer determines which pixel are related and grouped them into classes using the user defined algorithms and desired number of output classes. Supervised classification is a human-guided classification. The training sites are user-defined by selecting sample pixels in the image. And then the training sites are used as references for the classification of all other pixels in the image (Richards, 2013). In this study, supervised classification was used to classify the satellite image due to the availability of training sites from fieldwork (10m - 100m survey maps described above).

3.3.2.1 Defining communities to map – TWINSPAN

Based on fieldwork observations, 55 communities were identified, including 32 forests, 6 shrubs, 9 grasslands and 8 mires. This is too many for a successful supervised classification. There is a need to group the communities for further processing. This project uses the TWINSPAN (Two Way Indicator Species Analysis) method (Hill et al., 1975; Hill, 1979), which is popular among community ecologists and is useful for large datasets (Lepš & Šmilauer, 2003). This method uses indicators for the definition of vegetation types by translating quantitative species data into qualitative species data. In TWINSPAN, the division is built on the basis of a correspondence analysis (CA) ordination according to the sample score of the first axis. Therefore TWINSPAN was used to group the field data into a smaller set of communities which then could be used as training sites for supervised classification. The cut levels were selected using Domin cover scale (Kent, 2011), which is listed as follows:

1=present-1%, 2=2-5%, 3=6-10%, 4=11-20%, 5=21-50%, 6=>50%.

The 55 communities are classified into 13 groups using TWINSPAN software (Figure 3.2). Table 3.2 summaries the main forest types in each group.

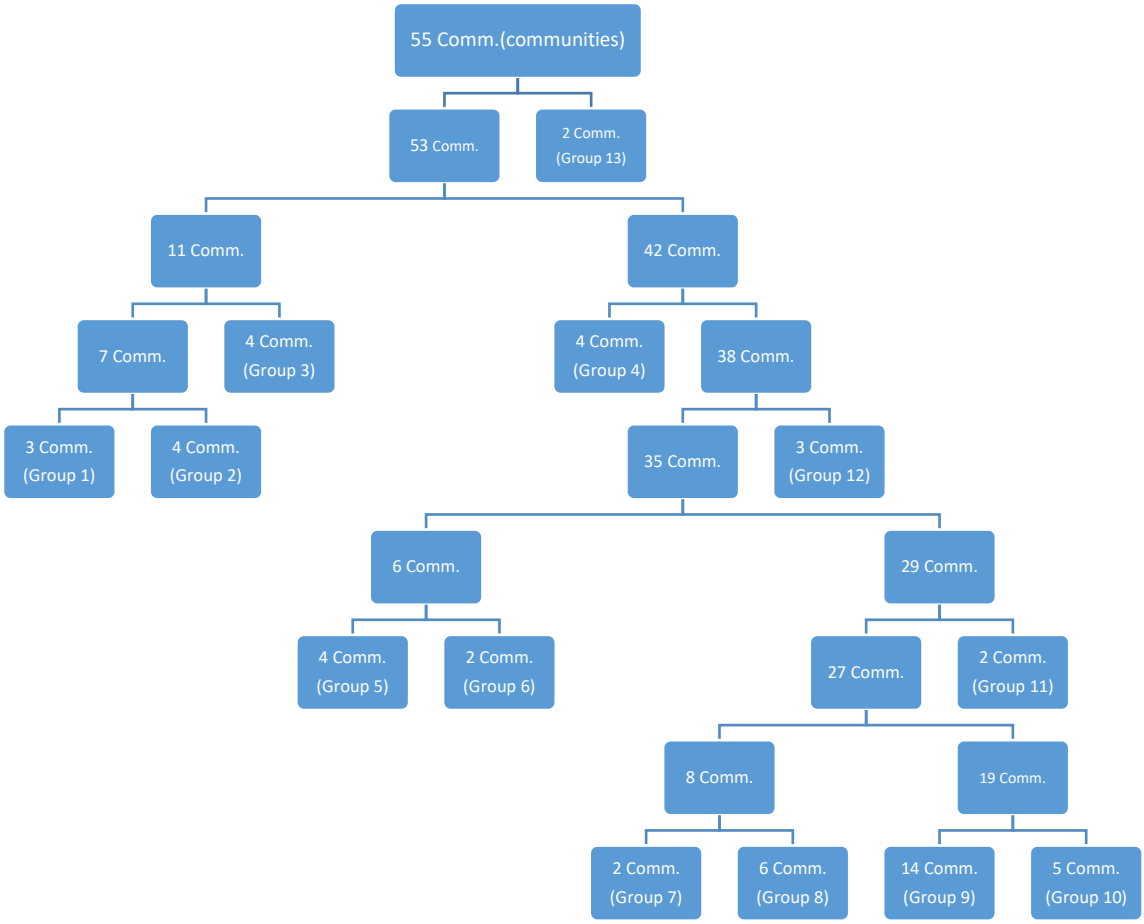


Figure 3. 2 The classification of 55 communities produced by TWINSPAN

Table 3. 2 Summary of the communities in each group (in column 2, mi=mire, g=grass, f=forest, sh=shrub)

Group number	Vegetation communities	Names of these communities	Main species
1	mi6, mi7, mi8	*Ligularia japonica-Viola lactiflora* mire	*Ligularia japonica*, *Viola lactiflora*
2	mi1, mi2, mi3, mi4	*Juncus effusus-Ligularia japonica* mire	*Juncus effusus*, *Ligularia japonica*, *Sphagnum palustre*
3	g2, g4, g8, g9	Poaceae spp. grassland	Poaceae spp.
4	g1, g5, g6, g7	*Astragalus sinicus*-Poaceae spp. grassland	*Astragalus sinicus*, *Poaceae* spp., *Lysimachia christiniae*
5	f4, f6, f7, f10	*Phyllostachys edulis* forest	*Phyllostachys edulis*, *Theaceae* sp.
6	f1, f8	*Cunninghamia lanceolata-Phyllostachys edulis* forest	*Cunninghamia lanceolata*, *Phyllostachys edulis*, *Lindera aggregate*
7	f5, f9	*Cunninghamia lanceolata-Pinus massoniana* forest	*Pinus massoniana*, *Cunninghamia lanceolata*

8	f11, f12, f13, f14, f15, f16	*Cyclobalanopsis glauca-Pinus massoniana-Loropetalum chinense* mixed forest	*Pinus massoniana, Loropetalum chinense, Cyclobalanopsis glauca, Liquidambar formosana*
9	f17, f18, f19, f20, f21, f22, f23, f24, f25, f26, f27, f28, f29, f30	*Pinus massoniana-Cyclobalanopsis glauca-Liquidambar formosana* mixed forest	*Cyclobalanopsis glauca, Quercus aliena, Platycarya strobilacea, Castanea* sp., *Pinus massoniana, Cunninghamia lanceolata*
10	f31, f32, sh4, sh5, sh6	*Theaceae* sp.*-Loropetalum chinense* shrub	*Liquidambar formosana, Pinus massoniana, Theaceae* sp., *Loropetalum chinense*
11	f2, f3	*Cryptomeria japonica* var. *sinensis* forest	*Cryptomeria japonica* var. *sinensis*
12	sh1, sh2, sh3	*Camellia oleifera* shrub	*Camellia oleifera*
13	g3, mi5	*Miscanthus sinensis-Arundinella anomala* grassland	*Miscanthus sinensis, Arundinella anomala, Themeda japonica*

3.3.2.2 Accessing and classifying remote sensed data

Sentinel-2 satellite images is important in determining various plant indices such as water content indexes and leaf area chlorophyll, mapping changes in land cover, monitoring plant growth and the world's forests (Drusch et al., 2012; Frampton et al., 2013). Since our focus is on the vegetation classification in wilder areas, a cloud-free clipped Sentinel-2 image is used as the original input data. The date of the image is 14th February, 2017. The regional vegetation map was then derived from the Sentinel-2 image using supervised classification in ArcGIS 10.5.1 (see Figure 3.1).

3.4 Pollen analysis

Chemical treatment of surface samples for pollen counting (Figure 3.3) is based upon the fact that pollen and spores are resistant to many corrosive chemical conditions (Moore et al., 1991), so other components of the sample can be removed whilst the pollen grains and spores survive. The greatest risk at this stage is the possibility of a sample being contaminated by atmospheric pollen. Ideally, the treatment of samples should be done during seasons with less pollen input.

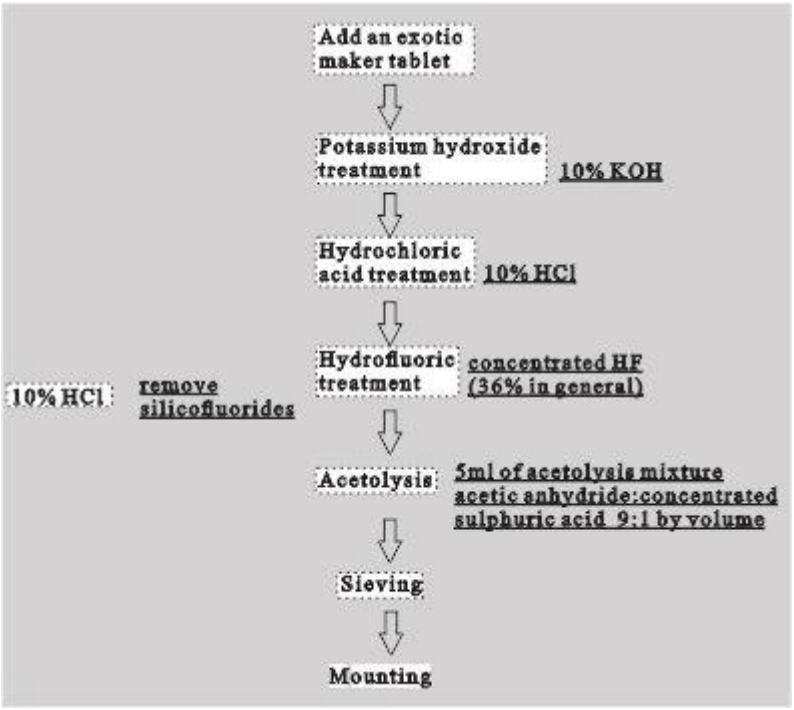


Figure 3. 3 Flow chart showing the processing procedures of surface samples in the laboratory

Each sample was treated with 1) 10% HCl, 2) 10% KOH, 3) 36% hot HF, and 4) hot acetolysis 5) sieving (with a mesh size of 7μm) following the standard procedures of Moore et al. (1991). A minimum of 400 grains (excluding fern spores and tracers) were counted for each sample sites. Identification of pollen types was made with reference to the Pollen Flora of China (Wang et al., 1995), An illustrated handbook of Quaternary pollen and Spores in China (Tang et al., 2016) and photographs of pollen grains from a herbarium collected in Wuhan Botanical Garden and Nanjing Botanical Garden. Pollen percentage diagrams were constructed in Tilia 2.0.4 (Grimm, 1991 and updated versions).

3.5 Data analysis methods

3.5.1 Vegetation data

3.5.1.1 Vegetation biodiversity indices

In forested landscapes, due to the dominant layers of trees and their impact on a variety of ecological gradients, the composition and characteristic of trees will have influences on plant biodiversity, i.e. understory vegetation diversity and composition (Palik & Engstrom, 1999). There were visible differences in the tree composition while collecting the paired samples, which might explain any observed differences between paired moss

and soil samples. The assumption of this test is to see if pairs were more similar (less Pythagorean distance) in lower diversity environments.

Quantitative definitions of vegetation diversity do exist. Two indices are selected to calculate the vegetation diversity, i.e. Shannon index (Shannon & Weaver, 1949) and Simpson index (Simpson, 1949) (see detailed definition in section 4.3.4.1). These two indices were selected for their widespread use in the ecological literature (e.g. Duelli & Obrist, 1998; Keylock, 2005; Allen et al., 2009). A key difference between the two indexes is their relative sensitivity to rare species (Gering et al., 2003). The Simpson index is a dominance measure weighted toward common species, while the Shannon index is weighted equally toward rare and common species (Magurran, 1988).

According to TWINSPAN results, the 42 paired samples are distributed in five main forest zones. The two vegetation diversity indices in the five forest zones were calculated using the proportional abundance of all species within each community from vegetation survey data (see details in chapter 4). In the same time, the Pythagorean distance between each paired soil and moss samples were calculated and averaged according to their location using sample scores from PCA.

3.5.1.2 Distance weighting of vegetation data

Plants are distance weighted based on the 'pollen sample's eye of view', that is, plants far from the sampling site has less impact on the pollen assemblage than plants nearer. The most widely used distance weighting term is the Prentice-Sugita model adjusted for neutral atmospheric conditions. There are also several other different weighting terms used in previous published works, such as simple equations (z^{-1}, z^{-2}, log (z), e^{-z}, $e^{-0.5z}$) (Broström et al., 2004; Mazier et al., 2008; Sjögren et al., 2008a; Bunting & Hjelle, 2010; Poska et al., 2011) and mechanistic models such as LSM (Kuparinen et al., 2007; Theuerkauf et al., 2013). Vegetation data needed to be distance weighted to address objectives 2 and 3.

The point of chapter 5 is to assess if the widely used P-S model captures the main patterns observed from pollen assemblages along a grass-forest transect. Other available models were also tested, including the more complex model (i.e. LSM) and less complex models (i.e. z^{-1}, z^{-2}, e^{-z}, $e^{-0.5z}$ and $\log_2(z)$). First of all, the vegetation map along the transect was created (see details in section 3.3). Percentage cover of the vegetation was extracted

in 500 concentric rings around each sample point each with width 10m from the regional map, and then distance weighted using each of the different distance-weighting models listed above. The pollen loadings using these weighting terms were simulated using PolFlow and then compared with the pollen assemblages of 11 real samples to assess their Squared-chord distance (SCD) and correlation coefficient. For objective 3, the Prentice-Sutton weighting is used (see details in chapter 6).

3.5.2 Simulation of pollen counts

3.5.2.1 Simulation of counts from a specified distribution

Each pollen sample contains a great amount of pollen grains, but counting all of them is impossible in most cases. Large counts is suggested to be sufficient to have confidence in the results, which mainly benefits rare taxa being present rather than greatly reduces the total amount of error. Data obtained by pollen counts are statistical estimates of the 'real' values, and therefore subject to a statistical uncertainty (Faegri et al., 1989). The influences of counting error (Maher, 1972) are being considered in this study.

For the moss and soil comparison study (objective 1, chapter 4), the point is to find to what extent can the difference between paired samples be explained by counting error. For the simulation study in RPP estimation (objective 3, chapter 6), the point is to make the simulated pollen counts more realistic. In both studies, counting errors were estimated from a known distribution and sum (500) using Excel.

3.5.2.2 Simulation of pollen assemblages from landcover grids

Models are always simplifications of the real world. In this study, HUMPOL0 software is used to simulate pollen assemblages. In order to enable the pollen loading to be simulated, three datasets are required, i.e. the landscape grid (made in Mosaic or mapping software such as ArcGIS), the HUMPOL data collection (made in PolSack) and the distance weighting look-up table (made in CreateLookup/Excel).

This suite is user-friendly because it allows more flexible land cover grids to be nested and additional weighting factor to be applied (Bunting & Middleton, 2005). For the transect study (chapter 5), the point is to compare the simulated pollen loadings with the real assemblages, in order to evaluate which pollen d&d model is suitable to be used in this area. For the simulation study in RPP estimation (chapter 6), HUMPOL 0 is used to

explore the possible effects of changes in vegetation homogeneity and number of data points on the RPP outputs.

3.5.3 Ordination analysis

Ordination analysis is widely used to explore patterns in ecological datasets. Because pollen samples are collected in different vegetation communities of the study area, there are visible variations among the pollen assemblages. PCA (principal component analysis) therefore is selected to explore possible environmental controls on sample variation in situations without explanatory variables, and to compare the distance of samples.

In this project, numerical comparison of pollen assemblages was carried out using principal component analysis (PCA) performed with CANOCO (CANOnical Community Ordination) 4.5 software (ter Braak & Verdonschot, 1995; ter Braak & Smilauer, 1997, 2002).

In chapter 4, PCA is used to see the distribution of paired soil and moss samples, and to explore differences between moss and soil within pairs. The Pythagorean distance between each pair and between the centroids of samples in each community are calculated from the sample scores of the first two axis of PCA ordination. In chapter 7, PCA is used to determine the distribution of samples collected in the mire and the surrounding slopes, and the main environmental controls (represented by taxa) driving them.

3.6 Summary

Methods selected for the four objectives are discussed in more detail in the relevant results chapters. The summarizing flow chat of each objective (Figure 3.4) shows how all the methods outlined above come together.

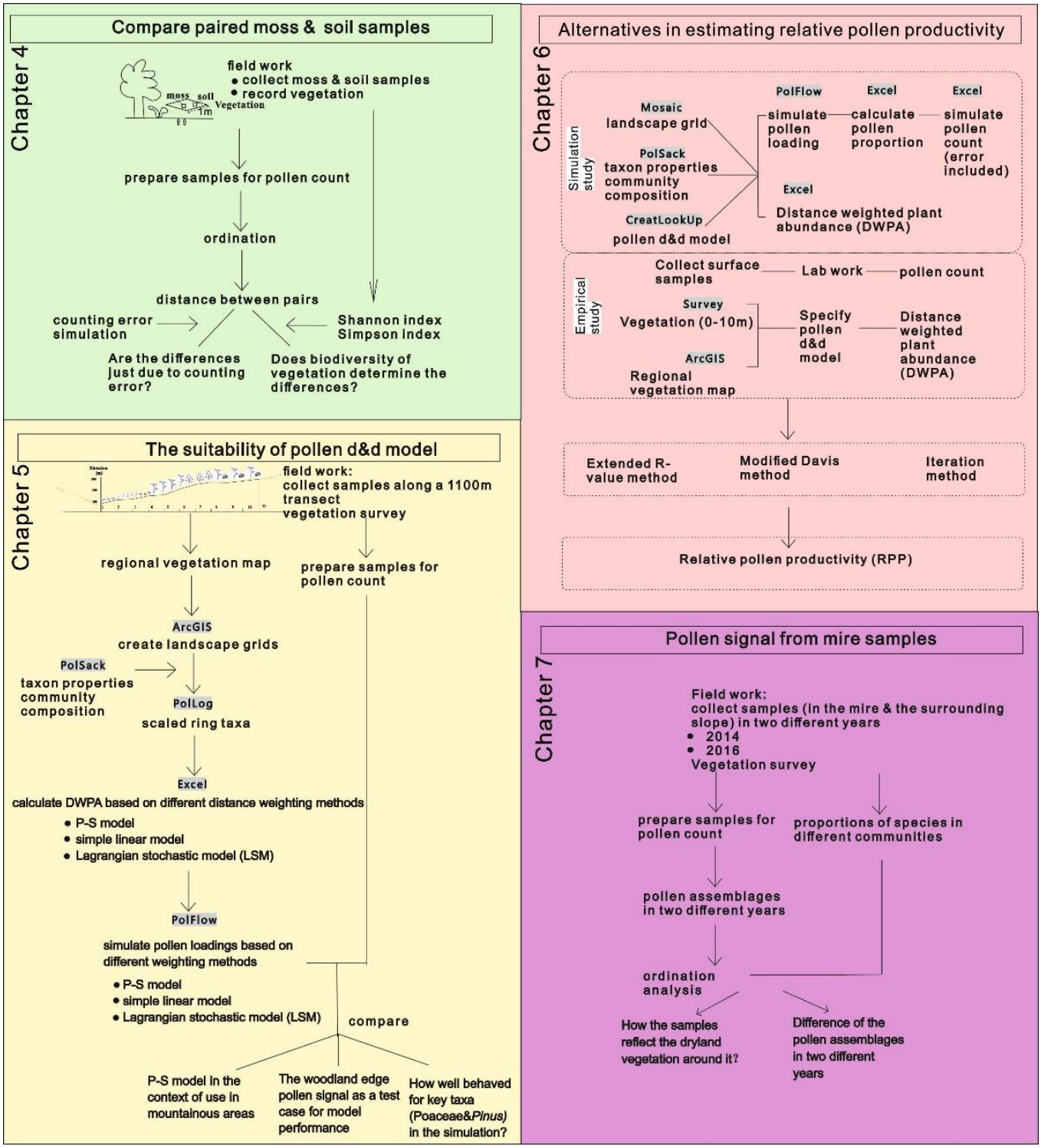


Figure 3. 4 Flow chart showing the procedures of each main chapters (chapters 4-7)

Chapter 4. Are modern pollen assemblages from soils and mosses the same? A comparison of natural pollen traps in the Meiling Mountains, southeast China

Abstract

Modern pollen assemblages collected from natural pollen traps such as moss polster, soil and peat/lake sediments are widely used to study the pollen-vegetation relationship, and to assist with the interpretation of sedimentary pollen records. However, there are few studies which directly compare the ways in which the pollen assemblages from different natural trap types record the contemporary vegetation, and whether choice of natural trap type alters the outcomes of investigations. This study compares the pollen spectra from 42 pairs of soil and moss samples (pairs collected within $1m^2$ squared area) sourced from five main forest zones in the Meiling Mountains, southeast China, with the vegetation proportions recorded at the collection point.

Soil samples and the adjacent moss samples show similar levels of variation in the pollen spectra records, but there are systematic differences in the mean values of key groups of taxa; there are lower percentages of arboreal types and higher percentage of herb taxa in soil samples than in moss samples. Simulation shows that in most cases intra-pair differences are greater than could be explained by counting errors. Ordination analysis is used to explore the differences between sample pairs and between samples from different forest types, finding that whilst intra-pair differences show no clear pattern, comparison between clusters of pollen signals from the different forest types do show systematic differences. Sample pairs from relatively homogenous vegetation areas are more similar than those from heterogenous areas. Comparison of local vegetation and pollen data suggests that pollen assemblages from moss samples provide a more accurate representation of the contemporary vegetation than those from soils. This study considers the implications of these findings for using studies of modern pollen assemblages from different natural traps as the basis for reconstructing past land cover.

4.1 Introduction

The use of surface samples in establishing the distribution of modern vegetation has becoming a standard interpretive technique for palaeopollen assemblages (Adam & Mehringer, 1975). Pollen spectra of surface samples used to improve interpretation of pollen diagrams come from multiple sampletypes, including waterlogged sediments (e.g. Faegri & Iversen, 1989; Wilmshurst & McGlone, 2005), artificial traps, ground surface samples such as moss polsters (e.g. Caseldine, 1981) and soil samples (e.g. Riding et al., 2007). In order to interpret the pollen diagrams, it is important to understand how the pollen assemblages present in surface samples form (Erdtman, 1969). Pollen assemblages can be recovered from a wide variety of contexts (e.g. forensic studies (Wiltshire, 2004; Mildenhall, 2006)), but caution needs to be taken in their interpretation, because the pollen assemblages vary under different sedimentation conditions (Potter, 1967).

4.1.1 Surface lake/peat sediment samples and artificial traps – pros and cons

Surface lake/peat sediment samples derive pollen and spores from many sources of the catchment and are considered to be most similar to sediment cores. However, limitation to the access of surface sediment samples do exist, especially in arid areas or heavily culturally modified landscapes. Artificial pollen traps has the advantage of overcoming the problems of extraneous material (e.g. soil which contains the older pollen assemblage) and of knowing the exact pollen accumulation rates (PAR; deposition of grains cm^{-2} $year^{-1}$) (Lisitsyna et al., 2012). There are different types of pollen traps designed for different aims (e.g. Petri dishes, Hesselman, 1919; cylindrical glass jars with a little glycerol, Lüdi et al., 1936; Tauber trap Tauber, 1974; Cundill trap, Cundill, 1986; partly cylindrical vessels with conical bottoms and a glycerol-drenched filter paper, Faegri et al., 1989; Bush trap, Bush, 1992). Due to differences in annual pollen productivity, for pollen traps, there is a requirement of consecutively sampling for multiple years. Pollen trap is also resource-expensive, which requires time-consuming setting and collecting. The survival of pollen trap is another issue, as it can be easily destroyed by human and animals, whether consciously or subconsciously. Ground surface 'natural traps' such as moss and soil are easy to find in most landscapes with sufficient quantity, but also have problems of their own.

4.1.2 Moss polsters - pros and cons

Although there are a variety of techniques available for sampling modern pollen assemblages, moss polsters are the most widely used material in terrestrial surface sampling (Moore et al., 1991; Bunting & Hjelle, 2010; Farrell et al., 2016). There are still some points that need to be considered for moss polsters as surface samples, and a lot has been learned about the way in which the pollen assemblages they contain reflect the surrounding vegetation.

4.1.2.1 Length of time represented by moss polsters

Unlike artificial pollen traps, controlling the time period for pollen deposition is not possible for moss polsters. Various views still exist about the length of time represented by moss samples. Crowder & Cuddy (1973) expect that moss may reflect the pollen deposition of 5-15 years, while Bradshaw (1981) proposed that the green parts of the moss represent the last 5 years of growth. Cundill (1985) indicated that there are variations regarding the exact period of pollen deposition represented by moss polsters collected from different months and years. A comparison study between modified Tauber trap and moss samples found that moss samples represent 1-2 years of pollen deposition (Räsänen et al., 2004). Determining the age of different moss types might be a solution for this problem, for the annual growing rate is identifiable for some species (e.g. *Sphagnum* and *Polytrichum*) (Räsänen et al., 2004).

4.1.2.2 Sampling - the part of moss being sampled

In order to get a reliable pollen assemblage, it is important to have an average of several years of pollen deposition. Previous study found variations on different visits while conducting a repeated sampling study on moss polsters, and the influences in different visits is greater than the sampling error associated with pollen counting (Farrell et al., 2016). The part of moss being sampled is therefore important to ensure containing the required accumulation of pollen.

In studies of moss polster samples, some researchers use the only green parts of the moss (e.g. Bradshaw, 1981; Prentice, 1986). However, some use large bulk samples of mixed moss, soil and litter to minimise factors such as the seasonal fluctuations and relative pollen quantities (e.g. Maher, 1963; Janssen, 1966; Birks, 1973; Caseldine & Gordon,

1978). The whole moss polster excluding the soil is collected by some researchers (e.g. Caseldine, 1981; Spieksma et al., 1994; Hjelle, 1998). Cundill (1991) conducted a comparison study of pollen deposition reflected by three parts of moss (green shoot, brown shoot and the moss humus), and indicated that the whole moss polster is the only reliable way of obtaining several years of pollen deposition.

4.1.3 Soil – pros and cons

Although soil is a widely distributed and easily obtained sampling material, it has received less attention from palynologists than waterlogged materials (Moore et al., 1991). Soil sequences are not considered as ideal pollen archives, partly because of the vertical mixing of the profile by invertebrate detritivores and poor preservation in aerobic environments (Andersen, 1986). Nevertheless, soil samples are still used to obtain modern pollen assemblages due to their easy accessibility, i.e. in areas where moss polsters may not available, such as arid and semi-arid areas of China (Li et al., 2007; Zhao et al., 2009c; Li et al., 2011).

Pollen assemblages from soils are often biased by several factors and should be interpreted with caution. One problem is the taphonomy of soil. Waterlogged sediments are accumulations of organic detritus, but soils are created by ongoing processes. There are potential risks of incorporation of older pollen by mixing, selective pollen decay and long-term pollen accumulation (Moore et al., 1991). Walch et al. (1970) demonstrated that soil fauna, such as earthworms, are responsible for producing considerable vertical movements of pollen in soils.

Another problem is that preservation of pollen in soil is also influenced by factors such as pH level and temperature. Havinga (1964) indicated that pollen degradation in soil always begins with oxidation, and different species have different oxidation susceptibility (Havinga, 1971, 1984). Lebreton et al. (2010) conducted an oxidation test on seven selected pollen taxa in soils, and indicated that short oxidation exposure times and low oxidant concentrations will make a fast decrease in pollen concentrations.

4.1.4 Comparing mosses and soils

In general, surface samples represent not merely a single year, but sample pollen accumulated over a longer period, which varies according to the pollen-trapping material. Moss polsters derive their pollen and spores mainly from canopy (Cc) and washout

component by precipitation (Cr), which make the modern spectra a high local pollen signal. Soil, on the other hand, comprise most resistant and corroded pollen and spores which has been stored for some time (Wilmshurst & McGlone, 2005).

The most commonly used material in terrestrial surface sampling is the moss polster, but in areas such as arid and/or semi-arid environments, soil is the only easily accessed deposit of modern pollen. A comparison study is therefore designed to see if there were systematic differences between moss and soil samples. In this chapter, paired surface soil and adjacent moss polster samples in range of forest types were collected. The aims of this study are to determine:

1) whether the pollen assemblages are actually different between the two types of traps in a sub-tropical environment,

2) if there are differences, are there systematic patterns which may indicate which processes are causing the difference?

4.2 Study area

4.2.1 Site description

This study was conducted in sub-tropical forest of Meiling National Scenic Area located in southeast China (28°31′N - 28°54′N, 115°34′E - 115°53′E; Figures 4.1a and 4.1b). The main axis of this mountain range is approximately 39 kilometres long and oriented in a northeast-southwest direction. The bedrock is granite which has experienced long-term weathering and denudation, with the highest point today being Laoluohan Peak at 950 m (Wang, 1974). The climate is mainly controlled by the southeast subtropical monsoon, with an annual average temperature of 18.8 ℃ and annual mean precipitation of 1760 mm (China Meteorological Administration Data Service Center, 1981-2010).

Forest coverage within the Scenic Area is around 89% (Ding et al., 1965). The vegetation communities present are mainly subtropical needleleaf forest (dominated by *Pinus massoniana* and *Cunninghamia lanceolata*), subtropical broadleaf deciduous forest (characterized by *Castanea sequinii*, *Quercus serrata* var. *breviptiolata* and *Platycarya strobilacea*), subtropical broadleaf evergreen forest (dominated *by Castanopsis sclerophylla* and *Cyclobalanopsis glauca*), subtropical bamboo forest (*Phyllostachys edulis*) and scrub. Several small mires occur, mainly located at an elevation of 700-800m

a.s.l., are home to aquatic plants, including *Juncus effusus*, *Sphagnum palustre*, *Ligularia japonica*, *Viola lactiflora*, *Polygonum thunbergii*, *Rotala rotundifolia* and *Salix chaenomeloides*. Scattered rice fields and vegetable occur in the valleys and foothill, which sometimes become grassland if the farmlands are empty.

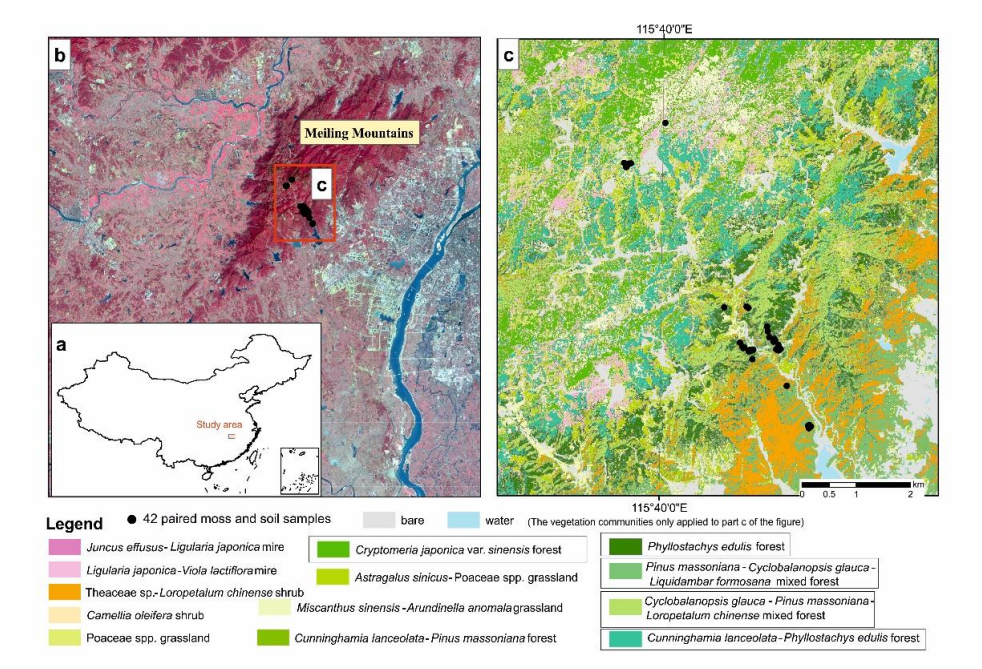


Figure 4. 1 Locations of the 42 paired sampling sites in Meiling Mountains, southeast China (a. location of the study area, b. satellite map showing the Meiling Mountains (from ESA data), c. location of 42 paired samples (the vegetation are divided into 13 biogeographic groups). The paired samples are located in five main communities, therefore their legends are highlighted with boxes.

Table 4. 1 Number of samples located in each communities

Vegetation communities	Main species	No. of samples in each vegetation communities
Phyllostachys edulis forest	*Phyllostachys edulis, Theaceae* sp.	10
Cunninghamia lanceolata-Pinus massoniana forest	*Pinus massoniana, Cunninghamia lanceolata*	7
Cyclobalanopsis glauca-Pinus massoniana-Loropetalum chinense mixed forest	*Pinus massoniana, Loropetalum chinense, Cyclobalanopsis glauca, Liquidambar formosana*	10
Pinus massoniana-Cyclobalanopsis glauca-Liquidambar formosana mixed forest	*Cyclobalanopsis glauca, Liquidambar formosana, Pinus massoniana, Platycarya strobilacea, Castanea* sp., *Cunninghamia lanceolata*	9
Cryptomeria japonica var. *sinensis* forest	*Cryptomeria japonica* var. *sinensis*	6

4.3 Methods

4.3.1 Site selection

During field excursions in March and April 2016, 42 paired soil and adjacent moss samples were collected from the 5 main forest communities of the region (Figure 4.1c, the legend of the five main forest are in box). Samples were only collected from sites that had not been visibly altered by human disturbance. The coordinate and altitude of each site were measured using a hand-held GPS.

4.3.2 Field methods

At each location, the single moss polster and adjacent soil samples were collected inside a 1*1 m^2 square area to ensure similar geographic distance. Vegetation cover (percentage of canopy species present) was recorded within a 100 m radius of the sampling point. In order to make the samples taken of known volume, all samples were collected using inverted sample containers measuring 7.5 cm x 7.5 cm x 1 cm in size. Only the green parts of the moss polster were collected.

4.3.3 Laboratory methods

The method of extracting pollen from surface samples followed standard preparation techniques (Faegri & Iversen, 1989; Moore et al., 1991). One *Lycopodium* tablet (27560

spores/tablet) was added to each sample as tracers. The chemical treatment procedures used include HCl (10%), KOH (10%), HF (40%), and samples were then acetolyzed with acetic anhydride and sulfuric acid (9:1), sieved and mounted in glycerine jelly. A minimum of 400 grains (excluding fern spores and tracers) were counted for each sample sites. Pollen were identified and counted under an optical microscope at magnification of 400×.

4.3.4 Data analysis methods

4.3.4.1 Vegetation data

The study of biodiversity between communities requires appropriate measurements of species richness and diversity (Keylock, 2005). Landscape richness is the number of species present in a community, while landscape evenness refers to the relative proportion of different species among these different cover types. To explore whether the landscape diversity affects the different distances of 42 paired surface soil & moss samples, the Shannon index (Shannon & Weaver, 1949) and the Simpson index (Simpson, 1949) were calculated on the percentage of species in each community, due to their extensive application in the ecological literature (e.g. Duelli & Obrist, 1998; Keylock, 2005; Allen et al., 2009).

The Shannon index of diversity is defined as:

$$S_{Shannon} = -\sum_{i=1}^{N} p_i \times lnp_i$$

Where N is the number of species in the community,

p_i is the percentage of species i

The Simpson index is calculated as follows:

$$S_{Simpson} = 1 - \sum_{i=1}^{N} p_i \times p_i$$

Where N is the number of species in the community,

p_i is the percentage of species i

4.3.4.2 Pollen data

In order to reduce the bias caused by abundantly produced but poorly dispersed spores (Wilmshurst & McGlone, 2005), pollen percentages of the terrestrial taxa (i.e., trees and shrubs, upland herbs) were calculated based on all terrestrial pollen grains. Percentages of wetland herbs' pollen and fern spores were calculated based on a sum of all pollen and spores. The pollen percentage diagram was plotted using Tilia 2.0.4 (Grimm, 1991 and updated versions).

For numerical analysis of pollen data, principal component analysis (PCA) is performed with CANOCO (CANOnical Community Ordination) 4.5 (ter Braak & Verdonschot, 1995; ter Braak & Smilauer, 1997, 2002). PCA is the linear method of indirect ordination, and it can display scores for samples, species and nominal dummy variables (Lepš & Šmilauer, 2003).

The PCA was carried out on a dataset including 49 pollen and spore types, all those with a value of over 2% in at least two samples. It is often expected that the samples are respective of their communities. A definition of the ordination is that it arranges sample points in such a way that distance between points corresponds as well as possible with the dissimilarity between sites (ter Braak, 1994). To find out if there were systematic changes in soil and moss samples, the individual combination and the centroids of each forest type combination are displayed with lines that connect corresponding sample types. The coordinates (X_i, Y_i) of centroids in each forest zone was calculated based on the formula below:

$$X_i = \frac{1}{n}(\sum_1^k x_k)$$

$$Y_i = \frac{1}{n}(\sum_1^k y_k)$$

Where X_i is x coordinate of a centroid, Y_i is y coordinate of a centroid,

x_k is the axis 1 value of sample core from PCA, y_k is the axis 2 value of sample core from PCA

n is the number of the forest zones, in this study, n=1,2,3,4,5

k is the number of the samples in each forest zones

The site scores in the ordination analysis are calculated in such a way that the close sites in the diagram are similar in species composition while far site are dissimilar (ter Braak, 1994). The changes in the Pythagorean distance between each pairs are calculated according to their sample scores from ordination. As the pairs are distributed in 5 forest zones, we grouped the pairs according to their locations and averaged the distance for each forest zones.

4.3.4.3 Counting error

Counting error is an important parameter which need to be taken into account. Maher (1972), for the first time, considered the results of possible measurement errors of pollen count, and assigned confidence limits to all pollen data. In order to test the pollen count error and see if the error is big enough to influence the distances of the paired soil and moss samples, we selected two random paired samples, namely S37 and M37, and simulated the pollen counts (as a total of 500 grains) five times for each sample. The simulation is done using the random number generation in Excel (see supplementary information for a simple example of simulating pollen count).

4.4 Results

42 paired soil and moss samples in the 5 main forest types were collected (see Table 4.1).

4.4.1 Landscape diversity of the five forest zones

For each of the five main forest zones, the Shannon index and the Simpson index were calculated and presented in Table 4.2.

Table 4. 2 Results of the Shannon index and the Simpson index, based on vegetation survey data recorded in the five forest zones of Meiling Mountains, southeast China.

	Shannon diversity	Simpson diversity
Pinus massoniana-Cyclobalanopsis glauca-Liquidambar formosana mixed forest	2.31	0.87
Cyclobalanopsis glauca-Pinus massoniana-Loropetalum chinense mixed forest	1.86	0.79
Cunninghamia lanceolata-Pinus massoniana mixed forest	1.46	0.69
Cryptomeria japonica var. *sinensis* forest	0.81	0.45
Phyllostachys edulis forest	0.57	0.32

The two diversity indexes show the same order among the five forest zones. Maximum diversity appears in the *Pinus massoniana-Cyclobalanopsis glauca-Liquidambar formosana* mixed forest, and minimum values were found in the *Phyllostachys edulis* forest. The *Phyllostachys edulis* forest (bamboo forest) is managed by the villagers living nearby. The strong roots of bamboos guaranteed their strong growth, even the main forest of the underground surface is covered by the bamboo root, therefore reduced the growth of other arboreal species.

There are significant changes in Shannon diversity values among the five forest zones, but gentle variations in Simpson diversity values. Shannon index is considered to emphasize the richness component of diversity and the latter is believed to show the evenness component (McGarigal & Marks, 1994; Nagendra, 2002).

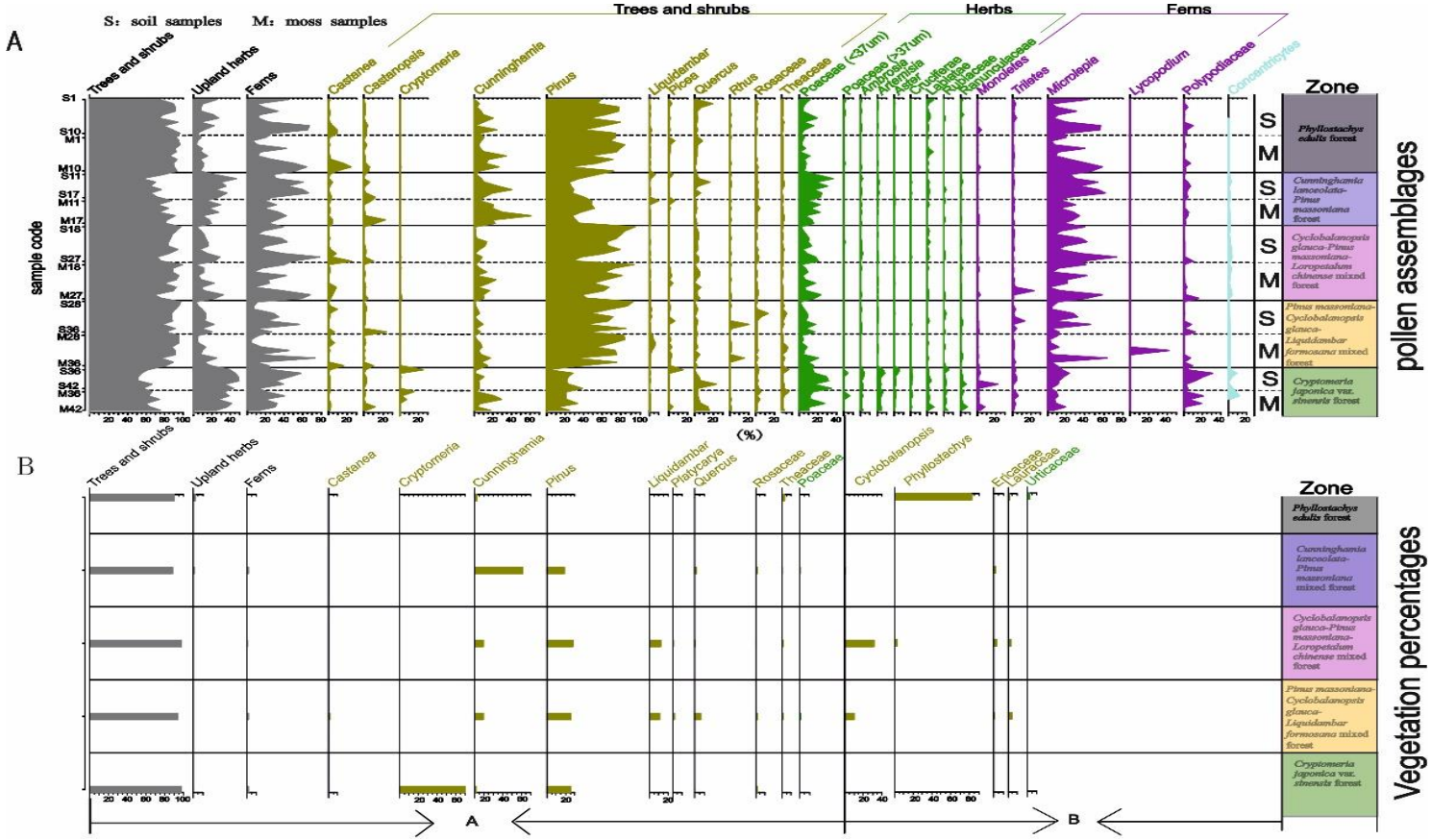


Figure 4. 2 Summary data collected from paired soil and moss samples collected in the five main forest types found in the Meiling Mountains, southeast China. Part A of the vegetation percentage diagram shows the species which are present both in the pollen spectra and the vegetation survey. Part B of the vegetation percentage diagram shows the species either found in the pollen spectra or in the vegetation.

4.4.2 Pollen assemblages

A total of 152 pollen and spore types were identified, including 93 tree and shrub taxa, 44 herb taxa and 15 spore taxa. Figure 4.2 shows the main taxa, and groups the samples according to the five main forest types. The vegetation composition in the five forest zones are presented for comparison as well. The differences between soil samples and moss samples are mainly shown as higher upland herbs percentages in soil samples. In general, *Pinus* and Poaceae are abundant in all forest types.

4.4.2.1 *Phyllostachys edulis* forest (10 pairs)

Pollen spectra in this group are characterised by high mean percentage of arboreal types, mainly from *Pinus* and *Cunninghamia*. Poaceae (<37μm) is the most abundant herb type, followed by Labiateae. However, *Pinus* and *Cunninghamia* rarely appear in the vegetation communities where these samples were collected. Soil and moss samples show similar characteristics and ranges, but the mean percentages of arboreal types are lower and of herb taxa are higher in soil samples than in moss samples.

4.4.2.2 *Cunninghamia lanceolata-Pinus massoniana* forest (7 pairs)

Pollen percentage of *Pinus* is markedly lower compared to the *Phyllostachys edulis* forest but it is still the main taxon in the pollen spectra. *Cunninghamia*, *Castanopsis* and Poaceae are the next three most abundant pollen types. The main species in the vegetation are *Cunninghamia* (51.6%) and *Pinus* (19%), which corresponds to the rise in *Cunninghamia* and decline in *Pinus* proportions in the pollen spectra. It is a similar case in the *Phyllostachys edulis* forest zone, soil samples have fewer arboreal types and more herb and fern types than the corresponding moss samples.

4.4.2.3 *Cyclobalanopsis glauca-Pinus massoniana-Loropetalum chinense* mixed forest (10 pairs)

Pollen spectra in this forest zone are dominated by high percentages of *Pinus*, Poaceae, *Cunninghamia* and *Castanopsis*. There is little visible difference in the percentages of arboreal pollen between soil and moss samples.

4.2.2.4 *Pinus massoniana-Cyclobalanopsis glauca-Liquidambar formosana* mixed forest (9 pairs)

Pollen assemblages from this forest feature high percentages of *Pinus*, Poaceae, *Cunninghamia* and Theaceae. *Liquidambar* percentages are slightly higher in moss samples than soil samples.

4.4.2.5 *Cryptomeria japonica* var. *sinensis* forest (6 pairs)

Pollen percentages of *Cryptomeria* are the highest compared to the above four forest zones, but are not high. The herb percentage increase compared to the other forest zones, including high proportions of Poaceae and Compositae types (e.g. *Artemisia*, *Ambrosia* and *Aster*).

4.4.3 Ordination analysis

The samples which were collected in the same forest zones were grouped in the ordination (Figure 4.3). Samples from the Cry. forest (*Cryptomeria japonica* var. *sinensis* forest) zone stays separately in the bottom right corner, while samples from the other four forest zones overlapped.

Ter Braak & Prentice (1988) indicated that the species arrow points in the direction of maximum variation in the species abundance, and the length is proportional to this maximum rate of change. An approximate ordering of the values of a species across the projected samples are obtained when projecting the sample point perpendicular to a species' vector. Therefore, species on the edge of the diagram are the most important for indicating site differences, while species near the centre are of minor importance. The biplot shows that herbaceous taxa (e.g. *Artemisia* and *Aster*) have greater impact on the green forest zone (*Cryptomeria japonica* var. *sinensis* forest). The species richness value of this forest zone ranked the second-lowest among the five zones, which dominated by *Cryptomeria japonica* var. *sinensis* (69.5%) and *Pinus* (25%).

In some cases, the soil sample envelope is larger than the moss sample envelope. Soil samples in Cunn.-Pi. forest zone (*Cunninghamia lanceolata-Pinus massoniana* forest) and Cry. forest zone have wider boundaries than the relevant moss samples. However, samples from the other three zones have relatively similar envelopes in both soil and moss samples.

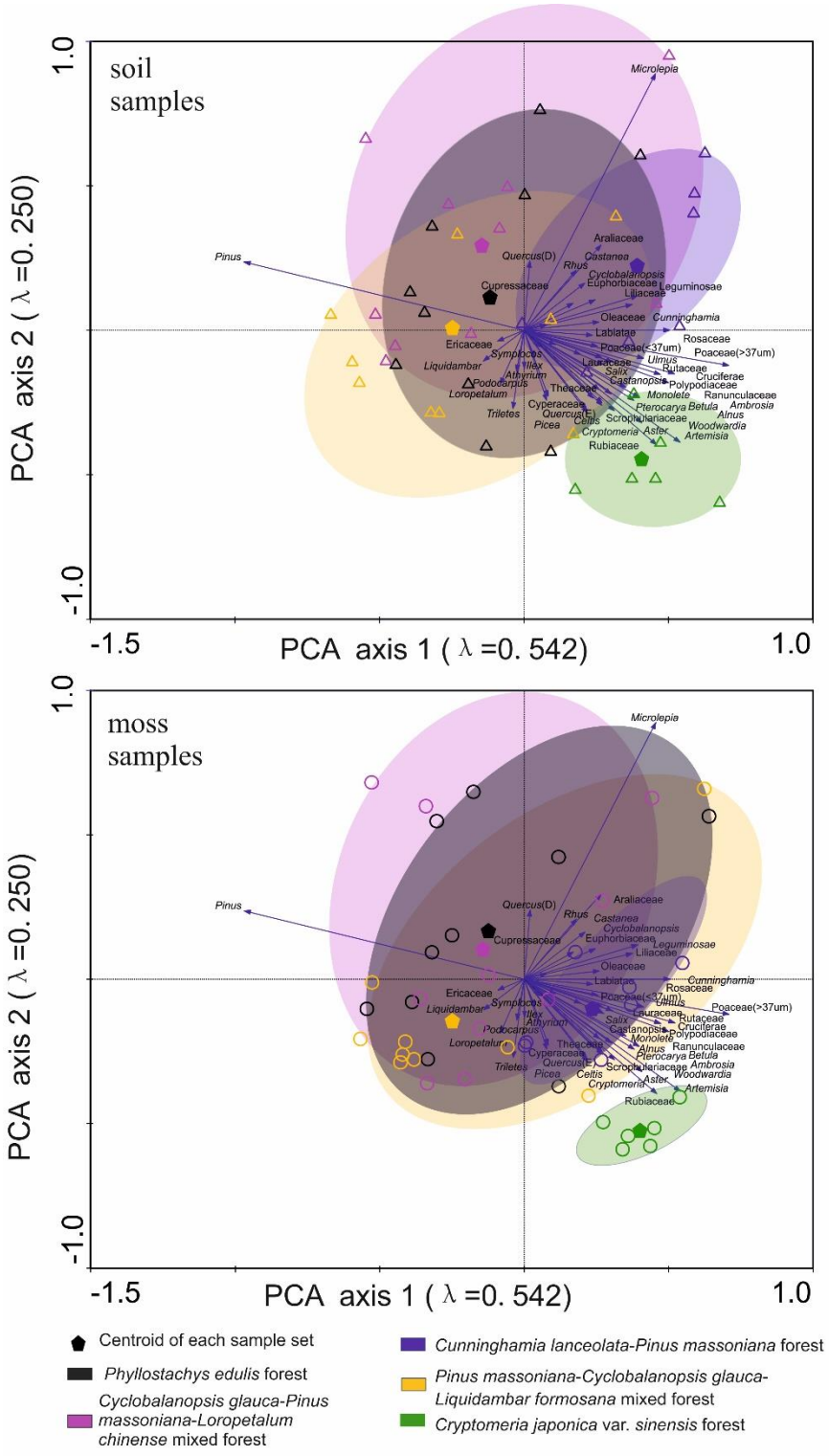


Figure 4. 3 Ordination based on Principle Components Analysis (PCA) diagram showing pollen assemblages in the 5 main forest zones. Ellipses are envelopes around samples from the five forest types. Centroids (multivariate centres of distribution) of each vegetation communities were plotted post hoc to visualize these relationships among sites. The eigenvalues of the two PCA axes are 0.542 and 0.250 respectively, all of which explain 79.2% of the total variance.

4.4.3.1 Paired sample distances

Direction and distance between paired samples are shown in Figure 4.4. There is no common direction of individual paired samples within all five forest zones. However, for the centroid of different samples in each forest zones (Figure 4.4 inset), the change is mainly in the direction 'north' to 'south', from the soil sample sets to the moss sample sets.

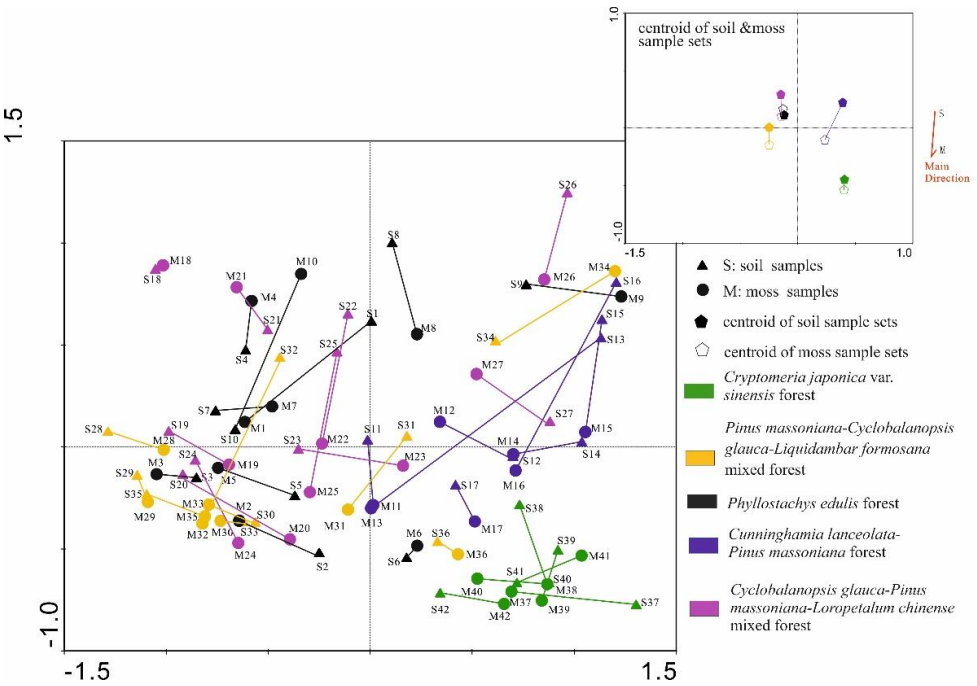


Figure 4. 4 Ordination (PCA) diagram showing the patterns of change on paired samples. Up-triangle symbols: soil samples; circle symbols: moss sample. Insert: plot of centroids of soil (filled pentagon) and moss (empty pentagon) samples for each forest zone.

4.4.4.2 Paired sample distances

Changes in the Pythagorean distance between each paired soil and moss sample scores are summarized in Figure 4.5a. The results show that sample pairs from Cry. forest zone have the shortest mean distances (mean = 0.59) and least variation (ranging from 0.43 to 0.84). Sample pairs from Cunn.-Pi. forest have the greatest separation (mean = 1.09), with a wide range of individual values (from 0.36 to 2.25).

In order to test the influence of counting error on the distance between the pairs, we simulated the pollen counts (as a total of 500 grains) five times for sample S37 and M37

individually, and calculated the Pythagorean distance between the 25 simulated pairs. Then, 25 absolute value of distance were calculated between each of the 25 simulated pairs and the baseline (i.e. real distance of S37 and M37) (Figure 4.5b). The distance between paired samples is generally larger than can be explained by sampling effects (counting error).

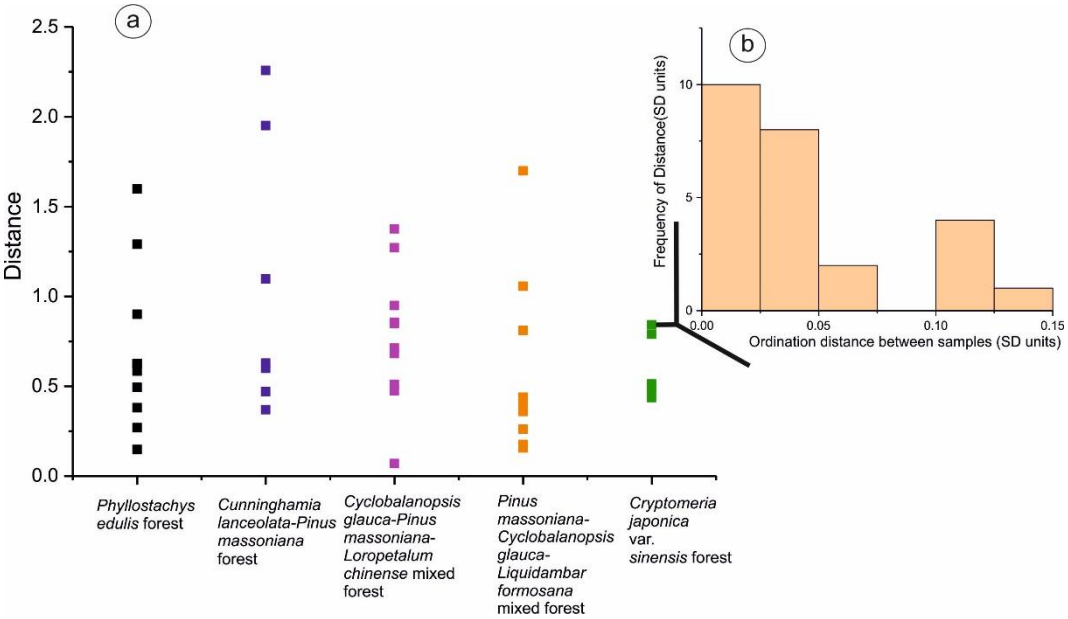


Figure 4. 5 Distance among paired samples in five forest zones (a. values of distance between 42 paired samples, the distance were calculated from sample scores of PCA output, b. frequency plot showing the distance between simulated repeat pollen counts. The distance is expressed as SD units of difference between simulated distance and real distance using the sample scores from PCA output. The simulation count is based on pollen portions of sample S37 and M37 using random number generation in Excel, for both samples, the simulation was performed five times).

4.5 Discussion

4.5.1 Surface pollen spectra from the five main forest types

The five forest types are hard to distinguish from the pollen spectra. According to the ordination, samples from the Cry. forest zone is clearly separated on the diagram, with high positive Axis 1 scores and high negative Axis 2 scores. The pollen assemblages in Cry. forest zone are characterised by lower percentages of trees and shrubs, and higher percentage of upland herbs. This forest zone has a second lowest diversity value among the five forest zones, which implies that forest with less diversity are more likely to be

distinguished from pollen assemblages. This might because in less diverse communities, the rarer taxa are less likely to be detected during pollen counting.

The bamboo forest can be considered as an invisible forest type. Pollen spectra in this forest zone is dominated by *Pinus* and *Cunninghamia*, both rarely present in the local communities but frequently appearing in neighbouring areas of the Meiling Mountains. *Phyllostachys edulis* (Poaceae) is the most important bamboo species and widely cultivated in southeast China (Zhang et al., 2016). The pollen shapes of *Phyllostachys* (Poaceae) were nearly spherical or sphere-like with monopolar hole (Lin & Ding, 2007), with an average diameter of 61.00μm (Zhang et al., 2016). However, the flowering cycle of bamboo can be as long as 3 to 120 years (Yuan et al., 2007) and the pollen germination percentage of *Phyllostachys* is quite low (ranging from 21.7% - 58.7%) (Lin et al., 2008). That might be the reason of little Poaceae presents in the pollen assemblage, therefore regional pollen dominates the pollen assemblage.

Pinus is the main taxon almost in all forest types, even in the bamboo forest where no pine trees were present. *Pinus* is found to be an over-represented pollen type and has been proved to travel thousands of kilometers into the Arctic (Campbell et al., 1999). Pine tree is widely spread in the Meiling mountains, which might explain this pattern.

4.5.2 Comparing pollen assemblages from paired soil and moss samples

As illustrated in Figure 4.2, soil samples and the adjacent moss samples show similar patterns in the pollen spectra. However, soil samples has less arboreal types but higher herbs and ferns than the corresponding moss samples. More upland herbs implies more background pollen component in soils. Pollen assemblages from moss polster are more influenced by local and extra-local vegetation than regional vegetation (Jacobson & Bradshaw, 1981; Prentice, 1985; Sugita, 1993). In this study, mosses tends to provide a more accurate representation of the contemporary local vegetation, which has also been shown by Wilmshurst & McGlone (2005).

4.5.3 Possible reasons for observed differences between paired moss and soil samples

The ordination diagram showed that there is no systematic differences in soil and adjacent moss samples. However, there are still observed difference between each pair. The

observed differences between paired moss and soil samples may be due to a mixture of reasons involved in pollen corrosion within the sediment.

4.5.3.1 Counting error

In pollen analysis, errors will be arise randomly. There are many reasons which might have an influence on the final pollen counts, such as the sample type, preparation and counting of the samples, data management (Galán et al., 2014). Comtois et al. (1999) indicated that in aerobiology, it is impossible to know the whole biological population of a certain volume of air. An estimation of the pollen component is tainted with error.

In this study, we estimate the influence of counting error using a simulation study. The Pythagorean distance between the real pairs (sample S37 and M37) is 0.84. The simulation study shows that 1 out of 25 simulated pairs (4%) makes a difference from 0.125 to 0.150, which is far less than the distance of real samples. Therefore, counting error is shown to generally not be able to account for the differences seen.

4.5.3.2 Differences in forest diversity

The diversity value is measured expecting that sample pairs from more diverse forest have greater distances between the pairs. The results corresponds to our expectations, which show that sample pairs from relatively homogenous vegetation areas are more similar than those from heterogenous areas. Sample pairs from Cunn.-Pi. forest, for instance, have the longest mean distance and high diversity values. Sample pairs from Cry. forest has the least mean distances and low diversity values.

There are exceptions, however. Sample pairs from *Phyllostachys edulis* forest, with the lowest diversity index, have relatively great distances. Although the pollen assemblages in this spectra are dominated by regional pollen types rather than local types, the opposite response possibly results from the fact that the bamboo forest, as a special economic crop managed by farmers, is more influenced by human disturbance.

4.5.3.3 Highly local taphonomic processes

Pollen taphonomy is defined as the whole process leading to the composition of pollen assemblages from living vegetation (West, 1973), which is usually summarized as pollen

production, pollen transport, sedimentation and post-depositional process (Coles et al., 1989; Campbell, 1999).

The surface sample is expected to be able to preserve the pollen grains from the surrounding vegetation. Soil samples are considered a problem because of the potential for systematic changes in pollen assemblage due to post-depositional differential decay (Havinga, 1964, 1967). For example, Havinga (1984) indicated that spores and Cichorioideae (*Taraxacum*) were more resistant to decay than the other taxa. A preservation study of *Pinus*, *Populus* and *Typha* pollen grains showed that *Pinus* is the best-preserved taxa (Sangster & Dale, 1961). Sangster & Dale (1964) sorted the pollen types according to their resistance to deterioration as '*Populus<Acer<Corylus<Alnus<Quercus<Fraxinus<Typha<Salix<Betula<Ulmus<Pinus'* . If this was the case, the expectation was that the direction of difference between paired samples always be in the same direction, that is, all samples lose the least decay-resistant type and therefore have relatively more of more resistant types. Figure 4.4 shows that there is no common direction of difference, therefore it appears that there is no substantial change in the soil pollen signal compared to the moss one as a result of consistent post-depositional losses. Figure 4.2 shows that the differences between soil samples and moss samples are mainly shown as higher upland herbs percentages in soil samples. However, the 'upland herbs' group contains both resistant and susceptible taxa, as does the 'arboreal pollen' group of taxa, therefore this difference is unlikely to be due to greater decay rates in soils.

Faunal activity and soil mixing have the potential risks to disturb the surface pollen spectra in soils (Moore et al., 1991; Walch et al., 1970). In addition, effects of water movement through soils and mosses may be differentially. Erosion by exogenic process such as raindrops and surface runoff are more likely to have on impact on soil than moss, especially in this hilly area. The relief in this mountainous area creates microtopographic variations in soils, which may mean the effects are not consistently seen in all the paired samples.

4.5.3.4 Very local factors (gravity, insects and erosion)

The distribution of species on the landscape will have a direct impact on the dispersal of their pollen to the sampling point. Obviously, in the real case, the landscape is never completely uniform (Davis, 1963). Flowering plants closely to the sample points will

affect the pollen assemblages by increasing the number of specific pollen types deposited due to gravity. In general, gravity should be same for both as the paired samples are close to each other, therefore I expect to see the same amount of variation for soil and moss polsters. However, as we can see in Figure 4.3, in some cases, the soil sample envelope is larger than moss sample envelope, which implies that the variation of soil samples in the same forest zone is larger than moss samples.

As for moss and soil, insects may have a preference to visit. Strong (1967) indicated that the significance of mosses to animals is mainly in providing shelter from the cold and wind. As indicated by Gerson (1969), moss may serve as food or shelter for certain athropods to survive than soil. This may lead to the difference between paired samples. The influence of insects is considered to be more local than regional, therefore more local pollen component in moss than soil. Single location samples is therefore a possible problem, as it may contain very local effects. Measuring of the local-scale component in a quantitative way (Farrell et al., 2016), and development of a standard sampling method which minimises its effect are needed for further study.

4.5.4 Selection of samples

This study shows that differences do exist in paired soil and adjacent moss samples. Comparison of local vegetation and pollen data suggests that pollen assemblages from moss samples provide a more accurate representation of the contemporary vegetation than those from soils.

In some cases, the soil sample envelopes are larger than the moss sample envelopes in the ordination diagram, which indicates that there is less variability among different samples collected in the same vegetation communities. Less variability suggests greater likelihood that small number of samples are “representative” of the community pollen signal, therefore less effort in collecting and processing samples. Although this study has not ruled out preservation and post-depositional reworking as factors altering soil samples, we do suggest that the former is not a dominant process modifying all paired samples in the same way.

All these factors discussed in section 4.5.3 are likely to be present and affecting pollen assemblages. Mosses grow in relatively stable locations, whereas soils may be more likely to be affected by runoff, especially in this mountain area. Factors such as faunal

activity and soil mixing are also more likely to affect soil samples. Therefore for this study, we choose moss polsters.

The future research directions and implications for other modern and palaeoecological studies of pollen are discussed further in chapter 8.

Supplementary information

How to use the random number generation in Excel

Assume we have a pollen assemblage of 3 taxa (1, 2 and 3), each taxa has a pollen percentage of 0.5, 0.25 and 0.25. (see Table 4.3. Note: the pollen percentage should have a sum of 1).

Table 4. 3 Example showing a pollen spectra of 3 taxa (A, B and C)

Pollen taxa	Pollen percentage
1	0.5
2	0.25
3	0.25

Open Excel. Select 'Data'-'Data Analysis'-'Random number generation'.

> If there is no 'Data Analysis' in the 'Data' menu, we should load it. Do it by right click under the Data menu, and select 'Customize the ribbon'-'Add-ins'-'Manage: Go'-'Analysis ToolPack'

Assume the total count of this simulation is 100 grains. Therefore, in the 'Random number generation' menu, type the Number of random numbers as 100. Select the input range and output range (Figure 4.6).

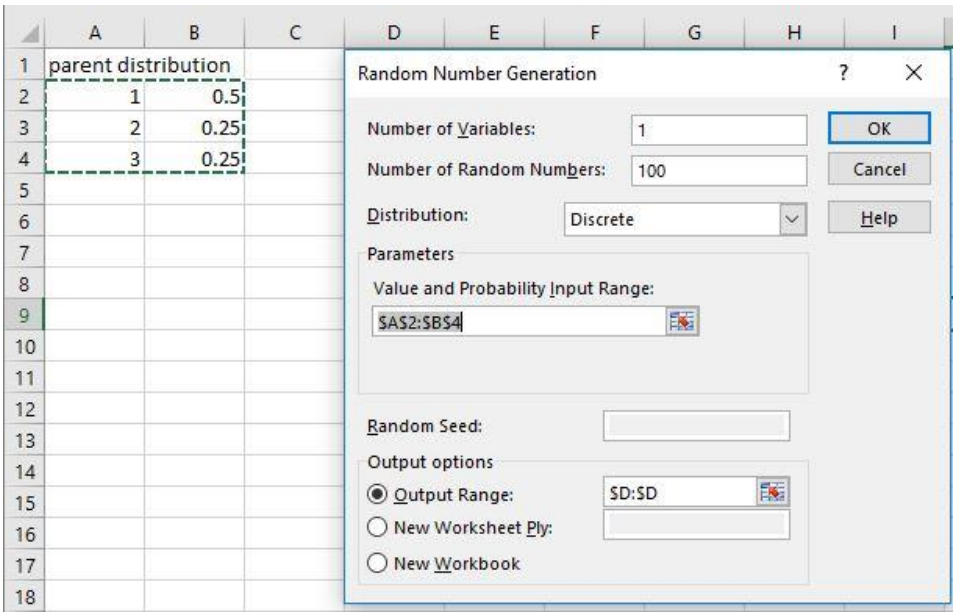


Figure 4. 6 Example showing how to simulate the pollen count as a sum of 100

Count the pollen grains of each taxa. For instance, if we are going to calculate the number of taxa 1, type '=COUNTIF(D$1:D$100,"=1")'.

Chapter 5. Assessing the suitability of pollen dispersal and deposition models for use in mountainous areas of southeast China

Abstract

In this chapter the choice of pollen dispersal and deposition model in mountain areas was considered. A detailed literature review is presented in order to explore the use of P-S model in mountainous areas. The woodland edge pollen signal is used as a test case for model performance. Pollen analysis has been carried out on moss polsters and adjacent soil samples from an 1100m transect across Poaceae grassland and *Pinus massoniana-Cunninghamia lanceolata* forest communities, and the results have been compared with simulation experiments employing different distance weighting terms. This study suggest that, the P-S model, the inverse distance squared (z^{-2}) test, the logarithm ($\log_2(z)$) test and the LSM test all performed well in terms of SCD and correlation between simulated and empirical data. The P-S model is a good predictor and is suitable for use to simulate arboreal pollen patterns in this study. The inverse distance squared (z^{-2}) test well reflect the Poaceae pollen, with low SCD and high correlation coefficient.

5.1 Introduction

Pollen analysis is widely used as a tool for the interpretation of qualitative changes in past land-cover. Pollen grains are produced by the anthers of flowering plants and disseminated through different agencies such as wind, animals (primarily insects) and water. Some pollen grains which dispersed through these mediums will finally settle on surfaces of lakes, peat, soil, moss polsters and so on, become embedded in the stratum and get preserved as microfossils (Nair, 1966). Therefore, the study of pollen is of great importance in providing valuable information regarding the vegetation history through geological time.

However, pollen assemblages are not direct reflections of contemporary vegetation. Factors such as pollen productivity, mode of transport and deposition will influence the formation of pollen assemblages. Simplified models which capture the main features of the system are a useful way to develop a better understanding of a complex system. Pollen dispersal and deposition (d&d) model is thus crucial for reconstructing past plant abundances quantitatively.

Pollen d&d models involve assumptions about system, which cannot always be tested for the past. In recent years, the most wildly used pollen d&d model is the P-S model. This model is developed from several previous studies. Tauber (1965) first made a conceptual evaluation of the aspects of pollen loading deriving from forest into lake basins, i.e. trunk space (C_t), above the canopy (C_c), rainout (C_r), gravity component from overhanging plants (Cg) and waterborne component (C_w) (Tauber, 1965, 1967; Bonny, 1980). Prentice (1985, 1988) developed a mathematical model of pollen loading deriving from above the canopy layer (C_c). His model was constructed by distance weighting the observed plant abundance relative to the point where the pollen assemblage accumulates. The Prentice's model was then modified further by Sugita (Sugita, 1994; Sugita et al., 1999) to include mixing of pollen within a water body after deposition to make it more suitable to predict the pollen loading of lakes. The P-S model incorporates aspects of pollen production, spatial distribution of vegetation and pollen dispersal, including both taxon specific properties (fall speed of pollen) and atmospheric conditions (e.g. wind velocity, stability). Sugita (1993) summarized the basic assumptions of the Prentice-Sutton model as follows (Table 5.1).

Table 5. 1 Basic assumptions of the Prentice-Sutton model

Assumptions		References	Notes
1.Assumptions about the model system	The sampling basin is a circular opening in a continuous flat vegetation canopy	Sugita, 1993	In most cases, the basin is not perfectly circular. Also vegetation is not flat, and may not be continuous.
	Wind is even in all directions	Sugita, 1993	This assumption seems inappropriate for areas with elevation changes.
	The dominant components of pollen transport are wind above the canopy	Sugita, 1993	In addition to wind dispersal, pollen can also be transported by other mechanisms such as water and insect.
2.Assumptions about the model parameters	Pollen productivity is constant for each taxon in space and time, and can be estimated from modern vegetation.	Sugita, 1993	Relative pollen productivity can be measured, but unlikely to be constant in space and time. RPP is affected by factors including vegetation structure, site position, regional climate and land management (Sugita et al., 1999; Waller et al., 2012).
	Sutton's ground-level model for particle deposition	Sutton, 1954	Treat the three atmospheric parameters (wind speed, turbulence, vertical diffusions) as constants by using predicted values under neutral atmospheric conditions. Jackson & Lyford (1999) investigated the influences of the three atmospheric parameters under unstable atmospheric conditions, and indicated more widespread dispersal from individual sources.
	Pollen deposition can be modelled using the diffusion model for small particles from a ground-level source	Sugita, 1993	'Skip distance' problem caused by elevated sources- 'Smoke stack' source (Prentice, 1985). But P-S uses ground level source.

The weighting term of the widely used P-S model is a taxon-specific distance weighting derived from Sutton's equation (Sutton, 1953) for particle from a continuously emitting point source at ground level. Application of Sutton's equation to dispersal of pollen and spores requires parameters including diffusion coefficient, turbulence factor, wind speed (Sutton, 1947b), deposition velocity of the particles (Chamberlain, 1953). The P-S model treat the first three atmospheric parameters as constants using prescribed values under neutral atmospheric conditions (Sutton, 1947a, 1947b; Chamberlain, 1953; Prentice, 1985). The deposition velocity (fall speed) is calculated based on Stokes' law, which

takes particle size and density into consideration. However, pollen dispersal may occur in unstable atmospheric conditions, and fall speed estimation might be 'wrong'. According to simulation studies, Jackson & Lyford (1999) indicated more wide spread pollen dispersal under unstable conditions than neutral conditions, and found substantial variability exist among tested datasets for sedimentation velocity predicted from Stokes' Law (Gregory, 1973). The P-S model assumes that pollen productivity is constant for each taxon in space and time. However, previous studies have shown that there are regional variations in previous reviews of RPP values.

For instance, simple models such as inverse distance (z^{-1}), inverse distance squared (z^{-2}) have been used in previous published works aiming at their influences on relative pollen productivity estimates (e.g. Broström et al., 2004; Mazier et al., 2008; Bunting & Hjelle, 2010; Poska et al., 2011; Theuerkauf et al., 2013). Small numbers of systematic empirical studies have been carried out to investigate how these differences affect the model outputs on showing the landscape changes. More recent work attempts to reduce the limitation of under-consideration of irregular and transitory airflow by using mechanistic models, such as the Lagrangian stochastic model (LSM) (Andersen, 1991; Kuparinen et al., 2007). The LS models assume multipule atmospheric layers with non-Gaussian turbulence in the upper layer. These models require the user to define various atmospheric parameters (e.g. the profile of wind velocity with increasing height). For this study, the LSM used is the same as Theuerkauf et al. (2013, 2016).

The P-S model was developed assuming a landscape which was forested, contained many small, roughly circular mires, lakes and ponds, and had relatively flat terrain. These assumptions are reasonable in the initial landscapes from where the model was developed and applied, e.g. in Denmark (Tauber, 1967; Andersen, 1970) or Minnesota (Sugita, 1993). This model has been used successfully in areas where the vegetation structure is more complex (such as fragmented cultural landscapes in southern Sweden (Sugita et al., 1999; Broström et al., 2004)). It has also been tested by inverse ERV (Nielsen & Odgaard, 2005) and used in reconstruction algorithms (LRA, e.g. Hellman et al., 2008a). Existing works illustrate the great potential of the model for reproducing the main features of the observed pollen patterns. It is in this area that P-S model may be able to make a significant contribution to studies in tracing the vegetation history. However, other d&d models might also be applicable.

There is often a distinctive change in pollen assemblages taken from two adjacent ecosystems, for example a woodland and a grassland or heathland. This "edge effect" has been studied empirically by many researchers (e.g. Table 5.2). A very rapid decline in arboreal pollen (AP) at the woodland edge was observed in the real surface samples (Faegri & Iversen, 1989), which is also found in woodland and moorland studies in West Yorkshire, UK (Tinsley & Smith, 1974). Different researchers fitted multiple different models (e.g. the exponential model $\left(e^{-bz}\right)$ by Turner (1964); the inverse distance model (z^{-1}) by Tinsley & Smith (1974)) to the data to explain the decline in tree pollen with distance – these studies were mostly conducted before P-S model being available. Here in this study, we collected surface samples along a grass - forest transect, assuming this declining pattern will exist in the pollen assemblages. If it is the case, can this pattern be simulated correctly by any of the different d&d models?

Table 5. 2 Selected previous studies on 'woodland edge effect'

Location	Approaches used	Species used	Reference
West Yorkshire, England	Empirical study with Mathematical model	Arboreal pollen types	Tinsley & Smith, 1974
North York Moors	Empirical study	Total tree pollen	Cundill, 1979
Fife, Scotland	Mathematical models	Arboreal pollen types, non- arboreal pollen types, Poaceae, *Betula*	Caseldine, 1981
Northern Minnesota, USA	Empirical study	Specific pollen and spore types (135 types)	Janssen, 1984
Outer Hebrides, Scotland	Empirical study	Tree pollen types	Gearey & Gilbertson, 1997
northwest Scotland	Empirical study and mathematical models	local arboreal (tree and shrub) pollen types	Bunting, 2002

Therefore, the aims of this chapter are to

1) review assumptions of P-S model (and allied models)
2) present results of a test of multiple pollen d&d models against a woodland edge case study in a mountainous region
3) determine which pollen d&d model is most suitable for the wider study.

5.2 Study area

Subtropical forest is the most extensive vegetation types in the Meiling Mountains (Ding et al., 1965). The main forest today are subtropical needleleaf forest (dominated by *Pinus massoniana* and *Cunninghamia lanceolata*), subtropical broadleaf deciduous forest (characterized by *Castanea sequinii*, *Quercus serrata* var. *breviptiolata* and *Platycarya strobilacea*), subtropical broadleaf evergreen forest (dominated by *Castanopsis sclerophylla* and *Cyclobalanopsis glauca*), subtropical bamboo forest (*Phyllostachys edulis*) and scrub. Several small mires occur, mainly located at an elevation of 700-800m a.s.l., are home to aquatic plants, including *Juncus effusus*, *Sphagnum palustre*, *Ligularia japonica*, *Viola lactiflora*, *Polygonum thunbergii*, *Rotala rotundifolia* and *Salix chaenomeloides*. Scattered rice fields and vegetable occur in the valleys and foothill, which sometimes become grassland if the farmlands are abandoned. The vegetation map can be found in Figure 5.1c.

5.3 Methods

5.3.1 Sample collection

The surface sampling was carried out in a pine-China fir forest in a valley within the Meiling Mountains, southeast China. Eleven paired moss polsters and adjacent soil samples (within 1 m^2 squared area) were collected along a 1000m transect runs across major vegetation boundaries from open grassland to forest, and across a range of altitude values (Figure 5.1d). The sampling interval along the transect is 100m in order to reduce the complications of the rapid decline at the boundary seen in studies which sample at intervals of around 10m (e.g. Turner, 1964). The major vegetation belts along the transect are illustrated in Figure 5.1d.

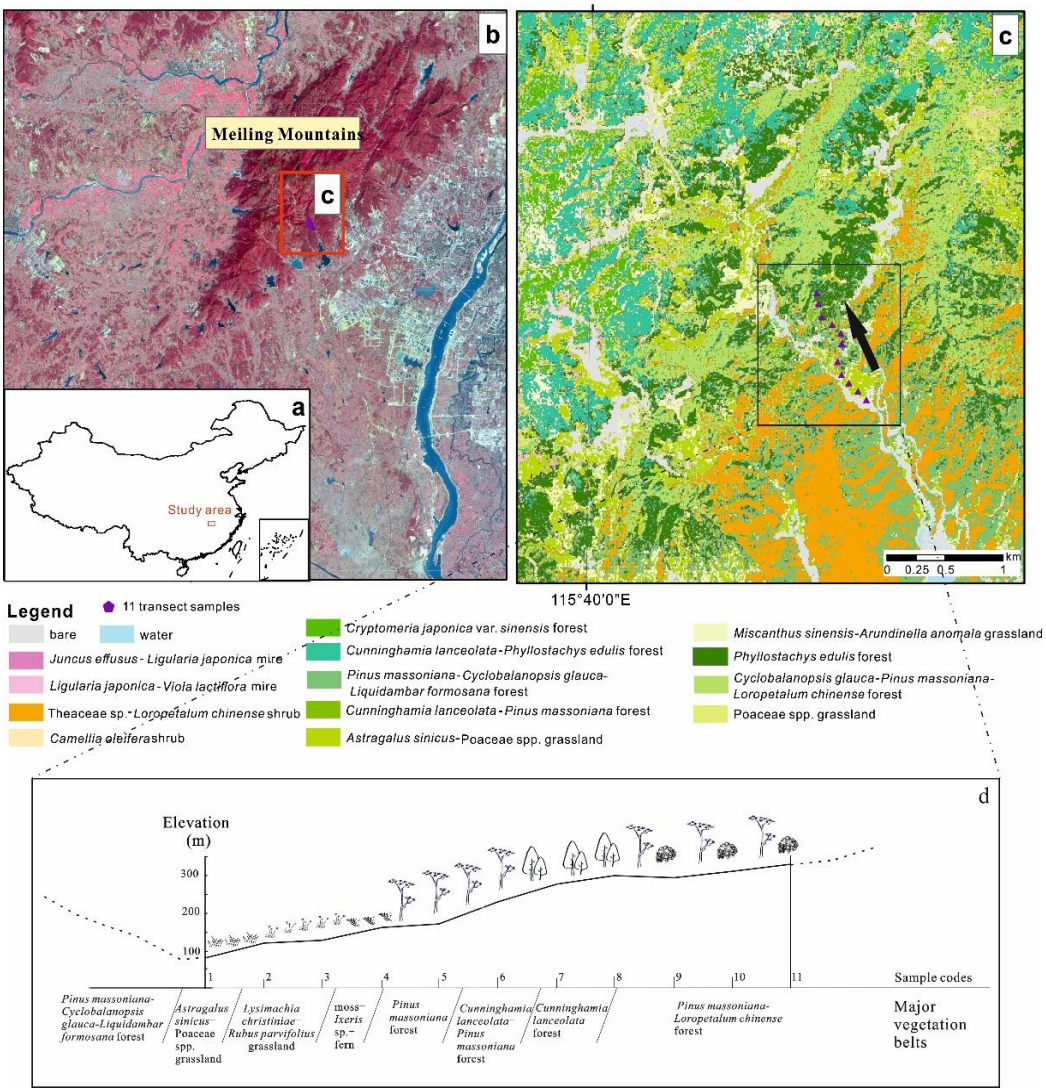


Figure 5. 1 Location map showing the 11 transect samples collected from the Meiling Mountains (a. location of the study area in China, b. Sentinel-2 data (ESA data) showing the Meiling Mountains, dark red (the forest areas) picks out the Scenic Area effectively, c. location of 11 transect samples on classified vegetation map. The vegetation communities were derived from Sentinel-2 data. d. sketch profile of the open-forest transition transect.).

5.3.2 Vegetation map

The composition of semi-open and forest communities were determined by counting all the mature trees above the canopy at 1m intervals along 30m transect. For each of these communities, at least 4 transects are surveyed. The composition of the ground flora in the open area was determined by four randomly selected quadrats (1m*1m) using visual estimates of cover. All of these compositions were recorded as percentages.

The vegetation map in the wider area was then created using supervised classification with ArcGIS 10.5.1 (Esri, Inc., 1999-2012, ArcGIS program version 10.5.1) from clipped Sentinel-2 remote-sensed data (see details in section 3.3).

5.3.3 Laboratory methods (see details in chapter 3)

Chemical treatment of surface samples for pollen counting is based upon the fact that pollen and spores are resistant to many corrosive chemical conditions (Moore et al., 1991), so other components of the sample can be removed whilst the pollen grains and spores survive. Pollen extraction was carried out using standard methods including hydrochloric acid treatment (10% HCl), potassium hydroxide treatment (10% KOH), hydrofluoric acid treatment (HF), acetolysis and sieving (with a mesh size of 7μm).

Pollen were mounted in glycerin and identified under an optical microscope at ×400 magnification. A minimum of 400 grains (excluding ferns and tracers) were counted for each sample sites. In order to reduce pollen representation bias produced by large amounts of spores, pollen percentages of terrestrial taxa (i.e., trees and shrubs, upland herbs) are calculated based on the total pollen of arboreal and upland herbs. Aquatic herbs and fern spores were calculated based on the sum of all pollen and spores.

5.3.4 Data analysis: vegetation data and distance weighting

The basic form of the linear model of pollen dispersal and deposition can be expressed mathematically as:

$$y_{ik} = \alpha_{\mathrm{i}} \mathrm{f}(x_{ik}) + \omega_{\mathrm{i}}$$

Where y_{ik} is the pollen loading of taxon i at site k, α_{i} is the relative pollen productivity of taxon i, $\mathrm{f}(x_{ik})$ is the distance weighted plant abundance of taxon i at site k, ω_{ik} is the background component of taxon i.

HUMPOL 0 allows the use of a variety of pollen dispersal and deposition models for simulation experiments. The input grid of the transect was derived from the classified Sentinel-2 remote-sensed data (see details in section 3.3). The composition of each community was defined using the percentage records from field survey.

Therefore, the pollen signals along the transect were simulated using ranges of weighting algorithms as follows:

1. Taxon-specific weighting based on Sutton's equations
2. Simple linear model:

Here we give some modifications of the linear model, basically based on different types of f(x_{ik}):

- inverse distance model :

$$f(x_{ik}) = z^{-1}$$

- inverse distance squared model:

$$f(x_{ik}) = z^{-2}$$

- exponential model (e^{-bz}):

$$f(x_{ik}) = e^{-z}$$

$$f(x_{ik}) = e^{-0.5z}$$

- logarithm model: It is achieved by giving equal weight to the vegetation data in each ring of a vegetation survey (Sjögren et al., 2008a). The area of each circle increases logarithmically compared to the previous ring, and the vegetation in each ring will be down-weighted compared to the next one.

$$f(x_{ik}) = \log_2(z)$$

3. Lagrangian stochastic model (LSM). This mechanistic model is considered more realistic than the P-S model as it predicts the trajectory of single pollen grains individually under turbulent conditions within the atmospheric boundary layer (Kuparinen et al., 2007; Theuerkauf et al., 2016), then sums across many thousand grains to develop a distribution model. It is suggested to be well suited to model pollen deposition in small and medium sized lakes (Kuparinen et al., 2007; Theuerkauf et al., 2013).

Percentage cover of the vegetation was extracted in 500 concentric rings each with width 10m and then distance weighted plant abundance was calculated using each of the different distance-weighting models listed above (except for the logarithm test). The

empirical pollen data are divided into two sections: one is the total arboreal taxa, the others are two of the most abundant taxa in the pollen assemblages, *Pinus* and Poaceae. All results are compared with actual data statistically to see which is most effective at capturing the main patterns observed.

5.3.5 Data analysis: comparing modelled and empirical data

The squared-chord distance measure is popular with palaeontologists and in studies on pollen. It is defined as:

$$D = \sum_{i=0}^{n} (\sqrt{x_i} - \sqrt{y_i})^2$$

where x_i is the number of observations for species i in sample x, y_i is the number of observations of species i in sample y.

In this study, the SCD is calculated based on the pollen percentages from the real samples and simulated percentage from the each of the d&d models. The lower the SCD, the better the simulation.

5.4 Results

5.4.1 Observed pollen spectra along the transect

Figure 5.2 shows the pollen percentages of selected types (AP, *Pinus* and Poaceae) from empirical dataset. In both moss and soil samples, a very rapid decline in arboreal pollen at the woodland edge was observed, the percentage of *Pinus* are less in the grassland than those in the forest. There is also a decline trend of Poaceae along the transect in both types of samples.

A very rapid decline of *Pinus* was observed in soil sample T4, which is not shown in moss samples.

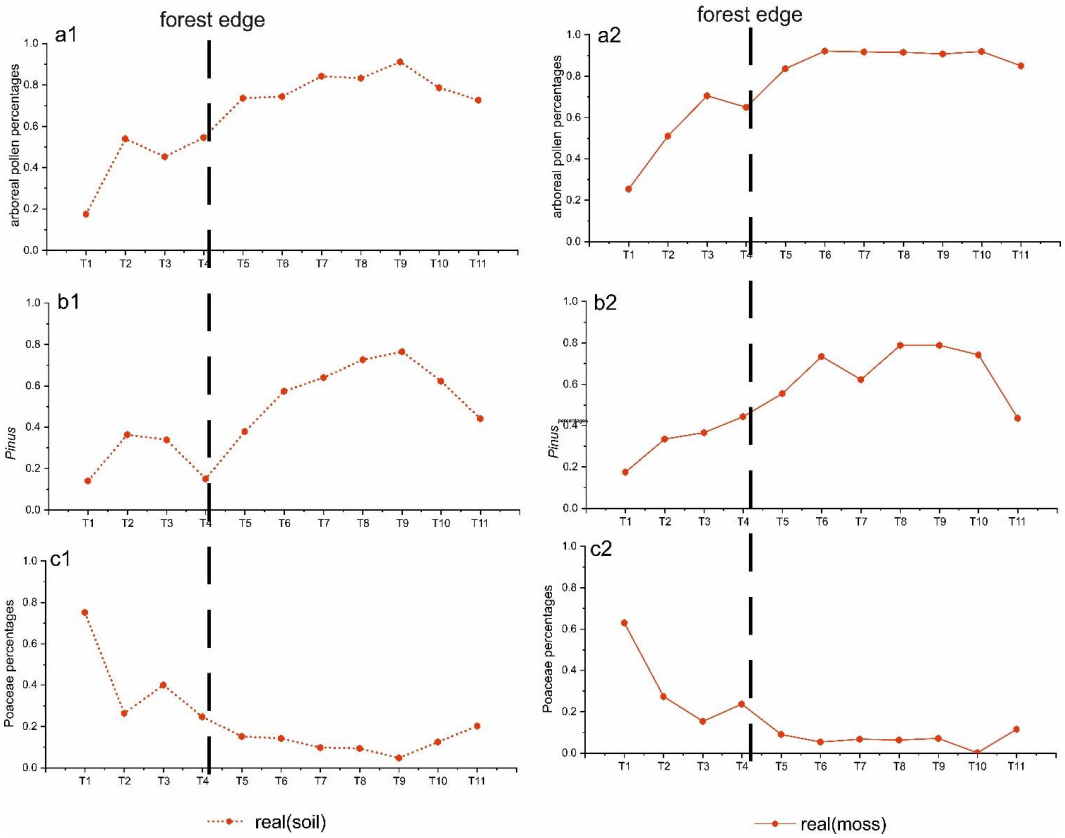


Figure 5. 2 Pollen percentage of selected types from empirical datasets (a1. arboreal pollen percentage (soil samples), a2. arboreal pollen percentage (moss samples); b1. *Pinus* pollen percentage (soil samples); b2. *Pinus* pollen percentage (moss samples); c1. Poaceae pollen percentage (soil samples); c2. Poaceae pollen percentage (moss samples))

5.4.2 Simulated pollen spectra under different distance weighting methods

5.4.1.1 Arboreal taxa

The simulated distributions of arboreal pollen at increasing distance along the open-forest transect have been examined in Figure 5.3a. The P-S model, the inverse distance squared (z^{-2}) test and logarithm ($\log_2(z)$) test show increasing patterns as the real samples. However, the inverse distance (z^{-1}) test are always flat in all samples, which is not similar to the empirical data. The LSM test show similar pattern for the samples in the forest (T5-T11), but show opposite trend simultaneously for samples in the grassland (T2 and T3).

5.4.1.2 *Pinus*

The simulated distribution of *Pinus* percentage at increasing distance along the transect is presented in Figure 5.3b. Moss samples show a good uptrend of *Pinus* for the samples in the grassland (Figure 5.2b), and this pattern is well described by the P-S model, inverse distance squared (z^{-2}) test and logarithm ($\log_2(z)$) test. The very rapid decline of *Pinus* in soil sample T4 is not been captured by any models. Again, the P-S model, the inverse distance squared (z^{-2}) test and logarithm ($\log_2(z)$) test have the most similar results as real samples.

5.4.1.3 Poaceae

Poaceae is an herbaceous taxon abundant in both the grassland and in the forest. Sample T1 is collected in an *Astragalus sinicus*-Poaceae spp. grassland, and here can be considered as a source of Poaceae. Similar as above, the P-S model, the inverse distance squared (z^{-2}) test, the logarithm ($\log_2(z)$) test have the most similar results as the real samples.

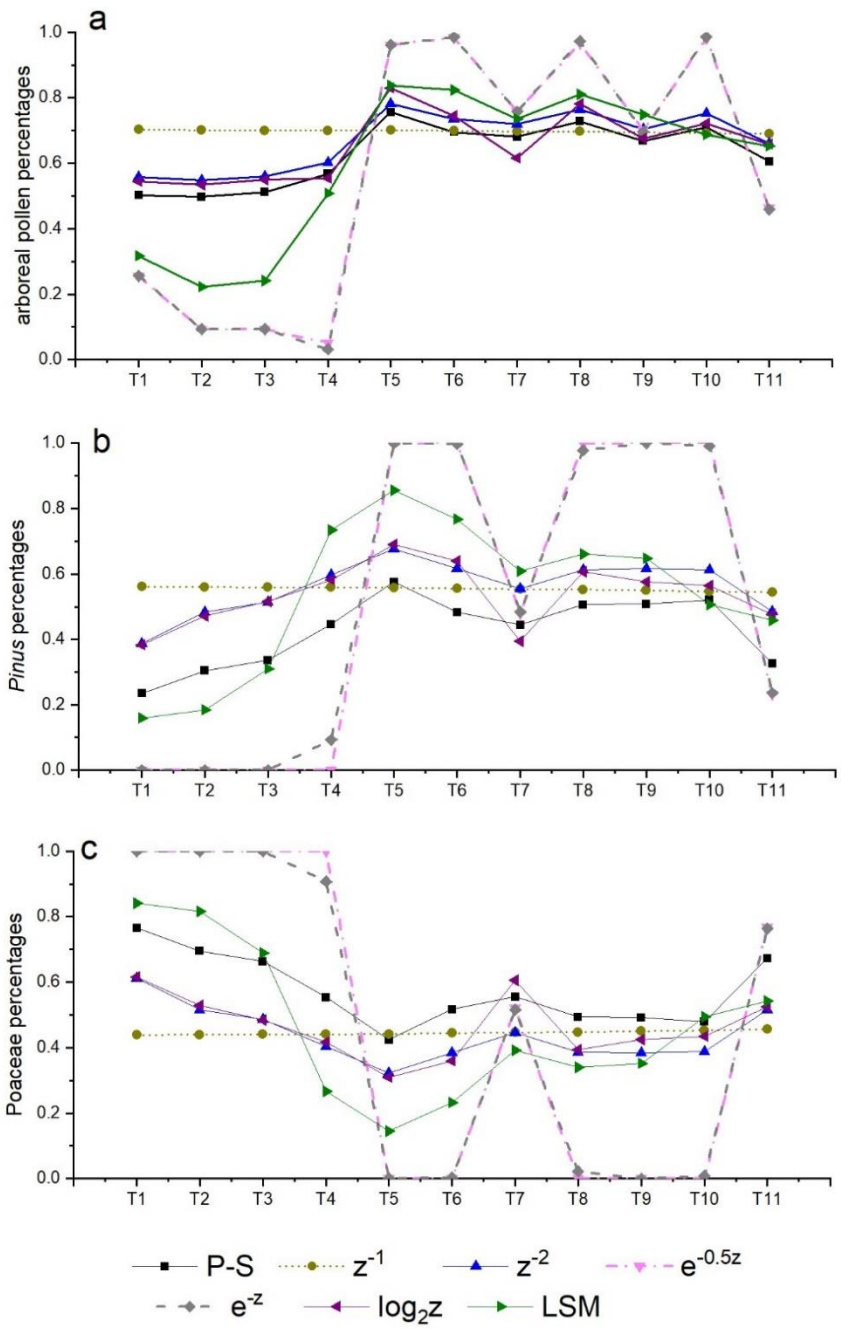


Figure 5. 3 Pollen percentage comparing empirial datasets and simulation from different weighting methods (a. arboreal pollen percentage; b. *Pinus* pollen percentage; c. Poaceae pollen percentage)

5.4.2 Comparison and classification of real and simulated pollen assemblages

Table 5.3 shows the squared-chord distance (SCD) and correlation coefficient (r) of pollen percentage between simulation and real samples.

Table 5. 3 Squared-chord distance (SCD) and correlation coefficient between simulations and real samples (left: soil samples, right: moss samples). The boxes show the tests which have relatively small SCD with the real samples, which is considered to be good simulations.

	Soil						Moss					
	Arboreal		*Pinus*		Poaceae		Arboreal		*Pinus*		Poaceae	
	SCD	r	SCD	r	SCD	r	SCD	r	SCD	r	SCD	r
P-S	0.13	0.81	0.19	0.59	1.21	0.79	0.16	0.82	0.14	0.85	1.98	0.76
$\log_2(z)$	0.15	0.67	0.32	0.25	0.83	0.54	0.15	0.70	0.15	0.58	1.44	0.56
z^{-2}	0.14	0.81	0.30	0.52	0.69	0.80	0.13	0.82	0.12	0.81	1.27	0.78
LSM	0.15	0.82	0.37	0.39	0.72	0.71	0.25	0.81	0.13	0.71	1.35	0.65
z^{-1}	0.26	-0.67	0.38	-0.66	0.86	-0.56	0.22	-0.68	0.24	-0.56	1.47	-0.59
e^{-z}	0.68	0.72	1.31	0.78	1.51	0.68	0.88	0.74	1.46	0.90	1.65	0.68
$e^{-0.5z}$	0.74	0.72	1.18	0.80	1.32	0.70	0.93	0.73	1.17	0.90	1.53	0.68

For the arboreal pollen types, the P-S model, the inverse distance squared (z^{-2}) test and the logarithm ($\log_2(z)$) test have lower SCDs with both soil and moss samples than the other four d&d models. The SCD between LSM test and soil pollen assemblages is also very small. For *Pinus* in the soil samples, the P-S model has the smallest SCDs with the real samples, but for *Pinus* in the moss samples, the P-S model, the inverse distance squared (z^{-2}) test, the logarithm ($\log_2(z)$) test and the LSM show lower SCD with the real samples than the other four d&d models. For Poaceae in both soil and moss samples, the SCDs are small in the inverse distance squared (z^{-2}) test and the LSM test.

Apart from the inverse distance (z^{-1}) test, all simulations has significant positive correlations with the real samples. In general, the P-S model, the inverse distance squared (z^{-2}) test, the logarithm ($\log_2(z)$) test and the LSM test have high relative coefficients, good stability and good predictability.

5.5 Discussion

All models compared used vegetation distance weighting, as Prentice (1985) argues - a tree close by a sampling point makes a larger contribution than one far off. In this study, two taxon-specific weightings (P-S model and LSM) and five non-taxon-specific weightings (inverse distance, inverse distance squared, exponential (e^{-bz}) with two values of b, and logarithmic ($\log_2(z)$) were compared against three sub-sets of empirical data. The empirical pollen samples were collected from a grass-forest transect in the mountainous area. Samples T1-T4 are collected from the grassland and samples T5-T11 are from the forest.

5.5.1 Simulation issues

5.5.1.1 Very local pollen signal

The basic mechanism of the distance models assume a monotonic decline of pollen with increasing distance. However, very local pollen signal, such as accidental falling of pollen anther and insect-borne pollen, might have an impact on the pollen assemblage (e.g. in the empirical dataset, an increase of arboreal percentages in soil sample T2 and an abrupt decrease of *Pinus* in soil sample T4). Insects, especially bees, carry pollen in fairly well recognized patterns and limited distances (Stanley & Linskens, 2012). Outliers could arise due to local deposition of entomophilous pollen close to the sampling point, or an anther above the sampling point. The very local component has not been considered into any of these d&d models, although it does exist. The local pollen signal might act as the result of the simulations do not perfectly reflect the pattern from empirical dataset.

5.5.2 Which model(s) are suitable for use in the Meiling Mountains?

In this study, we use SCD and correlation coefficient (r) to compare the simulated and empirical datasets. The two statistical methods used to compare simulated and real pollen assemblages show different results. In the comparison, pollen percentage was used. The correlation method may be affected by the interdependence of percentage values, the so-called 'Fagerlind Effect' (Fagerlind, 1952), therefore the SCD method is considered more reliable, and it has been recommended in previous works (Overpeck et al., 1985).

The results shows that the P-S model, the inverse distance squared (z^{-2}) test, the logarithm ($\log_2(z)$) test and the LSM all fit at least some of the empirical data strongly in terms of

SCD. In general, the correlation coefficient (r) is higher for P-S model than LSM. However, as in this study pollen percentage is used in coreelation calculations, the correlation method may be affected by the 'Fagerlind Effect' (Fagerlind, 1952), so it is considered not reliable.

It is therefore suggested that all of these four models could be used. But the inverse distance (z^{-1}) test, exponential model (e^{-z} and $e^{-0.5z}$) failed either with high SCD or with negative r, which suggested that these three methods were not suitable to use in this area.

5.5.2.1 Comparing the two 'process models'- P-S and LSM

The widely accepted P-S model is developed from an empirically based equation (Sutton, 1947a, 1953) for a particle source at ground level under stable atmospheric conditions, which includes parameters for the variation according to grain size/density (fall speed), variation in wind speed, etc.. The LSM simulates three dimensional turbulent airflows in the canopy and above the canopy in the surface layer and the convective layer, and the movement of pollen grains within them. Theuerkauf et al. (2016) argue that LSM describes pollen dispersal in a more realistic way particularly in terms of long distance transported pollen, whilst Sjögren et al. (2008a) state that the P-S model is a good approximation for pollen dispersal at medium distances.

In this study, both models were able to simulate the general patterns of arboreal pollen along the transect; it is not possible from these results to choose one or the other as 'better'.

The P-S model assume a flat landscape, which means that this model ignores the 'edge effect' between vegetation communities with height differences. Previous empirical studies show that the 'edge effect' at woodland edge, which may be in part due to height differences, extends less than 100m from the woodland edge (Tinsley & Smith, 1974; Gearey & Gilbertson, 1997; Bunting, 2002) and this study deliberately sampled at wider intervals (100m). In this study, the pollen assemblage is well modelled by the P-S model, which might because the elevation ranges along the transect is not high (less than 200m).

A few studies have been tried to use a combination of existing pollen d&d models, and test their behaviour. There is a wilder application to test these models. Previous studies have shown that different sampling materials (e.g. moss, soil, artificial pollen trap, ect.)

have different pollen signals (see details in section 4.1). Future studies can be done to test the model behaviour using different sampling contexts, for example, compare samples from moss, small mire, big mires, under canopy and in open areas.

5.5.3 Size range of pollen grains

Four models are considered to be able to reflect the general patterns of the empirical data (i.e. P-S model, LSM, z^{-2} and $\log_2(z)$). Among them, P-S and LSM use fall speed as an input parameter, while the two latter models ignore fall speed. Fall speed differences have greater effect in P-S than LSM. In this study, the fall speeds of *Pinus* and Poaceae are different by 0.033m/s (see appendix 7 in chapter 6), but in this test that did not seem to be enough to affect which model could replicate the patterns seen, possibly because the effect is less than other sources of variation such as counting effects within this test set-up.

Fall speed can be estimated from geometric measurements of grain size (see equation in section 2.1.2.2.2). It is generally accepted that grain size is affected both by chemical treatment (Christensen, 1946) and mounting media (Andersen, 1960). However, the impact of chemical analysis on pollen size is currently less well understood. Moore et al. (1991) pointed out that pollen treated with HF often appears smaller than untreated grains. Broström's work was based on measurements of single pollen grains from books (Punt et al., 1976-1995; Reille, 1992-1998) and papers (Eide, 1981). In this study, pollen grain measurements for fall speed were made on processed samples. Further investigations on how size measurements vary depending on different methods, and the feasibility of measuring fresh pollen grains, are needed, but do not appear to be vital since this test shows that the patterns can be modelled by different models.

5.5.4 Selection of distance weighting models

The choice of pollen dispersal model matters in LRA and MSA, both aiming at reconstructing vegetation composition on local and regional scale (Sugita, 2007a, 2007b; Bunting & Middleton, 2009).

Models should be tested before they can be used. In this study, 11 samples are collected at 100m intervals across two environmental gradients, which is aiming to test which d&d model fits the general pattern of the pollen signal. The P-S model requires fall speed measurements, and assumes constant pollen productivity for each taxon in space and

time. The atmospheric conditions in the past is not likely to be the same as today. Generally, the widely used P-S model fits well with the main features of the empirical pollen signal across the transect with low values of SCD. Therefore in the next two chapters, the P-S model is used.

Chapter 6. Novel approaches to estimating a key parameter for reconstruction of past land cover from pollen records in southeast China

This chapter has been submitted to Progress in Physical Geography, and is currently in review: Fang, Y., Ma, C., Bunting, M. J. (in review). Novel methods of estimating Relative Pollen Productivity, a key parameter for reconstruction of past land cover from pollen record. Progress in Physical Geography

Abstract

The analysis of pollen records from sedimentary environments has proved of considerable value for many aspects of Quaternary research from understanding past climates to environmental archaeology. One method of interpreting these records is to extract quantitative estimates of past land-cover using simplified models of the pollen-vegetation relationship, but these models require some parameters which need to be estimated via studying modern pollen-vegetation relationship.

A key model parameter, relative pollen productivity (RPP), is usually estimated from empirical data using the 'Extended R-value' (ERV) method, developed in the 1980s. This is an iterative model-fitting method which aims to estimate two parameters for each taxon, RPP and a background pollen term, and requires at least twice as many samples in the empirical dataset as the number of taxa for which parameters are sought. Collecting sufficient field data is time-consuming, and acts as a major barrier to wider use of reconstruction models. In this chapter, we present two alternative methods for estimating RPP which can be used with smaller data sets. Access to quicker computer processors and remote sensed data allow us to propose these alternatives, which we call the modified Davis method and the iteration method.

We demonstrate equivalence of all three methods in simulation, and apply them to a small dataset from southeast China, a region with no published pollen productivity estimates to date. These alternatives offer methods for RPP estimation from existing small sample datasets, such as pollen monitoring data, from areas where fieldwork is particularly difficult, or areas where no values exist and therefore no land cover reconstructions can be made.

6.1 Introduction

The analysis of pollen records from sedimentary environments has proved of considerable value for many aspects of Quaternary research from reconstructing past climates to environmental archaeology. The assemblages of pollen and spores contained in a peat deposits, lake or surface sediments are the product of a complex process of production, dispersal and deposition, which varies between sites depending on local characteristics and the wider catchment from which the grains and spores are derived. Quantitative estimations of past land-cover changes depend on a series of assumptions about the relationships between pollen deposited within the sedimentary system and the contemporary vegetation around the basin.

Land-cover is a fundamental variable which impacts on and links many aspects of the human and physical environments (Foody, 2002). Pollen records from sedimentary systems offer an important source of information for reconstructing past land cover (Poska et al., 2008). There are multiple methods of achieving this goal, such as the modern analogue technique (MAT, Overpeck et al., 1985; Calcote, 1998), biomisation (Prentice et al., 1996), or using mathematical models of the pollen–vegetation relationships (Davis, 1963; Prentice & Parsons, 1983; Sugita, 2007a, 2007b). There are two main reconstruction approaches developed from such mathematical models, namely the 'Landscape Reconstruction Algorithm' (LRA, Sugita, 2007a, 2007b) and the 'Multiple Scenario Approach' (MSA, Bunting & Middleton, 2009; Bunting et al., 2018). The LRA is currently being used to reconstruct global land cover through the PAGES Working Group LandCover6k (Trondman et al., 2015), but for many parts of the world this task is hampered by a lack of parameter estimates.

Application of these mathematical methods requires definition of taxon-specific parameters which need to be estimated via studies of modern pollen-vegetation relationship. Pollen productivity, a key taxon-specific model parameter, can be obtained in two main ways. Direct measurements (e.g. Erdtman, 1969) of pollen productivity are relatively few due to the limited reliability and laborious nature of available methods (Moore et al., 1991). Alternatively, surface samples of pollen assemblages collected from lake sediments, traps, soils or moss polsters can be combined with distance weighted plant abundances from the surrounding vegetation for each point, and this dataset can be used to develop estimates of the parameters. Most recent studies express pollen productivity as a dimensionless ratio relative to a reference taxon, called the relative

pollen productivity (e.g. Andersen, 1970; Broström et al., 2008). RPP values are widely assumed to be constant within a region through time (Sugita, 2007a, 2007b), but those values need to be empirically estimated for each new region of interest where different taxa are concerned.

RPP is usually estimated using the 'Extended R-value' (ERV) method (Parsons & Prentice, 1981; Prentice & Parsons, 1983). ERV analysis iteratively estimates two parameters, RPP and the background pollen component, for all pollen taxa in the empirical dataset simultaneously (Parsons & Prentice, 1981; Prentice & Parsons, 1983; Sugita, 1994); a detailed summary of the ERV method can be found in Bunting et al. (2004). In recent years, RPP estimates have been obtained for various taxa and localities by many researchers. Most of these estimates are from temperate areas in Europe such as Broström et al. (2004) in southern Sweden, Soepboer et al. (2007) in Swiss Plateau, Mazier et al. (2008) in Switzerland, Sjögren et al., (2008a) in Swiss Alps, Sjögren et al. (2008b) in SE Switzerland, Poska et al. (2011) in Estonia, Hjelle & Sugita (2012) in Norway, Matthias et al. (2012) in eastern Germany (reviewed in Broström et al. 2008; Mazier et al, 2012) and Theuerkauf et al. (2013) in northeast Germany. Another cluster of published values are available for northern temperate China (reviewed in Li et al., 2017), but no values yet exist for southern parts of China.

Successful ERV analysis depends on the empirical dataset available. One characteristic of a "good" dataset is that it contains sufficient samples. A "rule of thumb" of a minimum sample count of twice the number of taxa for which RPPs are required is widely used (Sugita, pers. comm.; Bunting et al. 2013), with more samples recommended. However, it is not always possible for researchers to collect enough samples, due to the time-intensive nature of the vegetation survey methods preferred (Bunting et al., 2013) especially in habitats with high biodiversity or where movement is challenging. Poor ERV analysis outcomes are attributed to having a small dataset in some studies e.g. Duffin & Bunting (2008; South Africa) and Li et al. (2015; northeast China). The RPP estimates derived from ERV analysis can also be sensitive to outliers within the dataset, and to variability in less well-represented taxa. Removing outliers from the dataset is often done on the basis of visual inspection, where it seems likely that some of the assumptions of the underlying mathematical model of the pollen-vegetation relationship are violated (see e.g. Bunting et al., 2016a, 2016b), but given the time-cost of each data point and lack of a formal objective basis for removal of outliers this is problematic. The ERV approach was developed almost 40 years ago, and developments in personal computer capability and

the resolution of and access to remotely sensed vegetation data since then have enabled us to explore alternative methods of estimating RPP which will have different limitations, and may be better suited to problematic and/or small datasets.

The aims of this paper are to:

1) present two alternatives to the ERV-analysis method of estimating RPP values,
2) show in simulation that these methods may be more effective than the widely-used ERV analysis method for small datasets, and
3) apply all three approaches to derive a first set of estimates of RPP for southeast China.

6.1.1 Theoretical explanation of RPP estimation methods

A key assumption of all these approaches is that the 'pollen sample's eye view' of land-cover around a sample point is weighted. Plants growing close to the sampling point contribute more to the pollen assemblage forming there than plants of the same species growing further away, and therefore the pollen data must be compared with distance-weighted plant abundance (DWPA) data. Multiple types of distance-weighting have been tried, and both different methods and different choices of vegetation survey increment to build the DWPA value will affect the results of data analysis (Bunting et al., 2013). In this paper we use a Gaussian Plume distance-weighting model, the Prentice-Sutton weighting, and empirical data were collected using the Crackles Bequest Project methodology (Bunting et al., 2013), whilst the modelled data used a simplified vegetation sampling approach due to model limitations (see methods section below for more details).

If both pollen loading and distance-weighted plant abundance are measured independently for a taxon, the relationship between them is rectilinear; the intercept with the y-axis indicates the background pollen component (pollen arriving from sources beyond the distance of the vegetation survey) and the slope of the line indicates the pollen productivity (the amount of pollen produced per unit of vegetation cover, expressed as distance-weighted plant abundance). However, in most cases one or both empirical values are only available as percentages, and the interdependence between the different taxa included in the base sum for percentage calculation means that the relationship between pollen and vegetation for a single taxon measured at multiple locations is no longer rectilinear (the "Fagerlind Effect"; Fagerlind, 1952). This non-linearity is more

pronounced at extreme values (i.e. close to 0% or 100%); Webb et al. (1981) suggested that linearity could be assumed for pollen percentage values between 20% and 30%.

6.1.2 ERV-analysis

ERV analysis assumes that the equation for the pollen vegetation relationship is:

$$y_{ik} = \alpha_i x_{ik} + \omega_i \qquad \text{(Equation 1)}$$

Where y_{ik} is the underlying 'absolute' pollen influx of taxon i at site k,

α_i is the pollen productivity for taxon i.

x_{ik} is the distance-weighted plant abundance of taxon i relative to site k summed over a "large" distance (a distance substantially greater than the relevant source area of pollen of site k, *sensu* Sugita, 1994).

ω_i is the background pollen contribution from taxon i in the study region, and is a constant for taxon *i* in a given study region assuming that the area of vegetation surveyed extends to the Relevant Source Area of Pollen (Sugita, 1994).

Therefore, the ERV method requires two parameters to be estimated for each taxon, which means there is no algebraic solution possible and an iteration method is required. ERV analysis seeks to relinearise the empirical dataset by calculating a site-specific correction factor based on the RPP and background pollen component. The RPP of one taxon is set to 1.00 (the reference taxon) and values of all other parameters allowed to vary. For each set of values, the method corrects the empirical dataset and calculates a likelihood function score based on the improvement in fit to a rectilinear model for each taxon of the corrected dataset compared to the raw dataset. Parameter values are then varied and the calculation repeated many times, to identify a solution with the lowest likelihood function score obtainable. Three different sub-models are available, representing three approaches to correcting the empirical dataset based on three different assumptions about the definition of the background pollen contribution, which should theoretically all return the same results. A detailed summary of the ERV method can be found in Bunting et al. (2004).

6.1.3 Modified Davis method (MDM)

Davis (1963) originally defined the "R-value", a measure of pollen productivity, as the ratio between a pollen assemblage and the vegetation community it represents, as a means of estimating vegetation cover from pollen proportions. The relationship assumed is:

$$R_{ik} = P_{ik}/V_{ik} \qquad \text{(Equation 2)}$$

Where P_{ik} is the pollen percentage of taxon i at site k, and V_{ik} is the vegetation percentage of taxon i surrounding site k. Davis (1963) argued that R_{ik} will not necessarily be constant between sites, since a change in pollen percentage of a taxon could occur without a corresponding change in the percentage cover of that taxon in the landscape because of the percentage interdependence problem, but that the ratio of two R-values for a pair of taxa, $\frac{R_{ik}}{R_{jk}}$, should be constant between sites. By using the same taxon as the denominator for all taxa in an assemblage, it should be possible to specify a set of relative pollen productivity values which are constant within a region (Andersen, 1970). In the original paper, vegetation is measured as area of cover with no distance weighting applied (Prentice, 1985), and it is assumed that there is negligible pollen input from outside the area of surveyed vegetation (no background pollen component).

The Modified Davis method proposed here uses the same approach to calculate RPP values for each pollen assemblage in turn, but uses distance weighted vegetation data collected to a distance two to five times larger than the "Relevant Source Area of Pollen" (*sensu* Sugita, 1994) around each point. This is made possible by the availability of remotely sensed data, processed through GIS software.

The mathematical model of the pollen-vegetation relationship underlying this method is simple:

$$y_{ik} = \alpha_i x_{ik} \qquad \text{(Equation 3)}^1$$

where y_{ik} is the underlying 'absolute' pollen influx of taxon i at site k,

[1] ERV analysis extended this R-Value model equation by adding in a term for pollen from sources located beyond the area of surveyed vegetation (see equation 1)

α_i is the pollen productivity for taxon i.

x_{ik} is the distance-weighted plant abundance of taxon i relative to site k summed over a "large" distance (a distance substantially greater than the relevant source area of pollen of site k, *sensu* Sugita, 1994).

Substituting equation 3 in equation 2 shows the equivalence of $\frac{R_{ik}}{R_{jk}}$ and $\frac{\alpha_i}{\alpha_j}$

$$\frac{R_{ik}}{R_{jk}} = \frac{\frac{P_{ik}}{V_{ik}}}{\frac{P_{jk}}{V_{jk}}} = \frac{P_{ik}}{V_{ik}} \cdot \frac{V_{jk}}{P_{jk}} = \frac{\frac{y_{ik}}{y_{\cdot k}}}{\frac{x_{ik}}{x_{\cdot k}}} \cdot \frac{\frac{x_{jk}}{x_{\cdot k}}}{\frac{y_{jk}}{y_{\cdot k}}} = \frac{y_{ik}}{x_{ik}} \cdot \frac{x_{jk}}{y_{jk}} = \frac{\alpha_i x_{ik}}{x_{ik}} \cdot \frac{x_{jk}}{\alpha_j x_{jk}} = \frac{\alpha_i}{\alpha_j} \quad \text{(Equation 4)}$$

By comparing all taxa (i) with a single reference taxon (*ref*), and setting α_{ref} to 1, we can calculate the relative pollen productivity of taxon i thus:

$$\alpha_{iref} = \frac{p_{ik}}{x_{ik}} \cdot \frac{x_{refk}}{p_{refk}} \quad \text{(Equation 5)}$$

Where p_{ik} is the proportion of pollen type i in the sample from site k, p_{refk} is the proportion of the reference pollen type ref in the sample from site k, and x_{ik} andx_{refk} are the DWPA of taxa i and ref calculated relative to site k.

A more detailed explanation of the modified Davis method can be found in appendix 1.

6.1.4 Iteration method (IM)

If the vegetation included in the DWPA term x_{ik} is surveyed over a large enough area, the background pollen term can be treated as negligible and we can revert to Equation 2. This permits us to estimate RPP using a simpler iteration approach, which we have simply called the iteration method. We take advantage of modern computer capacity to simply calculate the pollen assemblage that would be produced by given DWPA values under all possible combinations of a range of possible RPP values, calculate a similarity coefficient for the simulated pollen data and actual pollen data for each site, and use a maximum likelihood approach to decide which set of estimated parameters is the best estimate of the actual values. The detailed steps for the iteration method can be found in appendix 2.

6.2 Methods

In this paper, we first compare the three methods of RPP estimation (ERV, MDM and IM) in simulation to determine whether they produce similar results, and then apply them to an empirical dataset.

6.2.1 Simulation study

The simulation study was carried out using HUMPOL 0 software (Middleton, unpublished) and Excel. The purpose of simulation was to determine how well the three methods could return input RPP values.

Vegetation scenarios were created using Mosaic v3.2 (Middleton & Bunting, 2004). Two grids were created (Figure 6.1), an outer grid with 20m pixels and overall extent of 20 km x 20 km and an inner grid with 10m pixels and overall extent of 5 km x 5 km, centred within the outer grid. Five plant communities, four with a single dominant taxon, and one with a two-taxon mix, were included in the landscape.

It is widely accepted that "perfect" ERV analysis results cannot be expected when the landscape is not homogenous. In order to detect the influence of a homogenous vegetation structure upon the calculation of RPPs from ERV method, we created two kinds of vegetation structure, mainly distinguished from the outer grid (i.e. a homogeneous outer grid and non-homogeneous distributed outer grid, community properties are shown in appendix 3). The uneven landscape is conceptually based on the empirical case study area. This makes the test here more realistic since most real world landscapes do not meet the strict homogeneity assumption.

Nine moss samples were placed inside the inner grid, at least 1000m away from the edge and at least 1500m apart to minimize possible auto-corelation effects. Three replicate inner grids were created, giving a total of 27 pollen samples for analysis. 27 surface samples satisfies the "2n" rule of thumb for ERV analysis of a dataset with 6 pollen taxa.

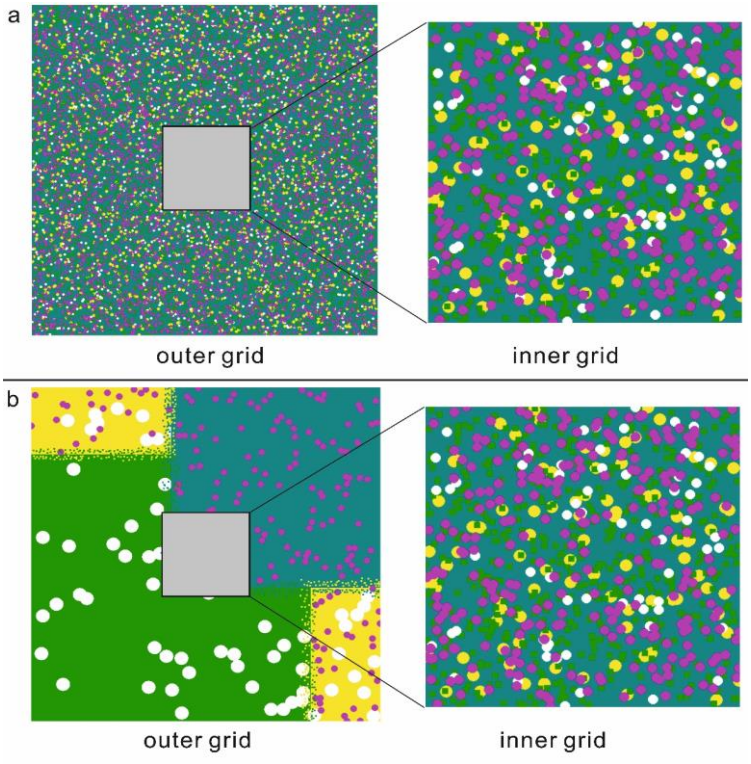


Figure 6. 1 Grids for simulation (outer grid: 20km x 20km; inner grid: 5km x 5km). a. homogeneous distribution of vegetation in the outer grid, b. non-homogeneous landscape in the outer grid

The properties of the communities and pollen taxa were defined in PolSack. These six taxa were selected because they were the main arboreal species in the empirical case study. PolFlow was then used to simulate pollen loadings and collect vegetation data in 400 concentric rings each of 20m width around the specified sampling points. Distance weighted plant abundance was then calculated from the vegetation composition of these rings using the Prentice weighting method (Sutton, 1953; Prentice, 1985, 1988) in Excel and expressed as percentages for analysis of each sample. Pollen assemblages are simulated as pollen loadings, but in order to incorporate the sampling errors inherent in actual pollen counts, the loadings were used as a probability distribution to simulate pollen counts of 500 grains for each sample. The counts were then expressed as percentages for analysis.

Taxa to include in ERV analysis and the reference taxon for all three analyses were selected using inspection of the scatter plots comparing pollen percentage with summed DWPA at 8000m; the taxon with the best range of values and scatter visually closest to rectilinear was selected as the reference taxon, and taxa with multiple outliers or very low

pollen or plant presence were excluded from ERV analysis. ERV analysis was then carried out using PolERV (Middleton, unpublished), RPPs were calculated using the MDM in Excel, and IM was applied to the dataset using R (see appendix 4 for the code).

Analysis was run both with 27 data points and with a subset of 9 samples to assess whether method performance changed with a small sample.

6.2.2 Empirical study

6.2.2.1 Study area

The Meiling Mountains are located in Jiangxi province, southeast China (Figure 6.2). The highest peaks reach about 950m a.s.l. (above sea level). The mountains lie along a southwest to northeast axis and occupy an area of 150 km^2 (28°31′N - 28°54′N, 115°34′E - 115°53′E). The mountains are mainly composed of granite and gneiss. The mean annual temperature and precipitation are about 18.8℃ and 1760mm, respectively.

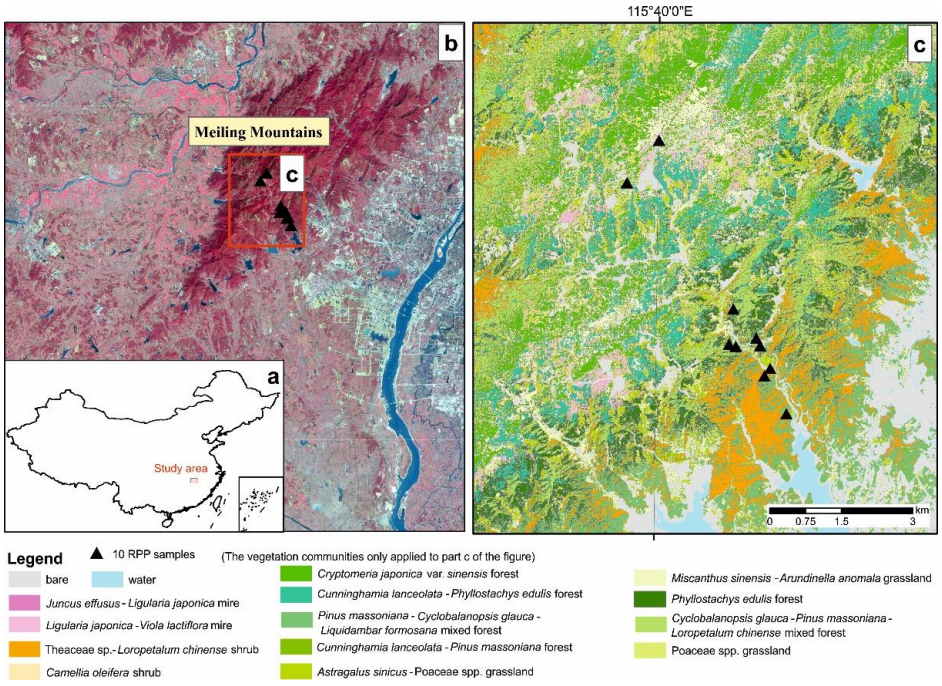


Figure 6. 2 Locations and regional vegetation communities of Meiling Mountains, southeast China (a. map showing the location of study area, b. map showing Meiling Mountains in their wider landscape setting, c. the inner study area and the locations of 10 RPPs samples in the forest (vegetation communities shown in the legend were mapped by classifying Sentinel-2 data (ESA data) using field data from 2016 for ground trothing.

Due to human impact, the original forest are mainly replaced by secondary forest, with some primary forest remnants scattered in the area. The main vegetation communities present are subtropical forests, including needleleaf forest (dominated by *Pinus massoniana*, *Cunninghamia lanceolata*), broadleaf deciduous forest (characterized by *Castanea sequinii*, *Quercus serrata* var. *breviptiolata* and *Platycarya strobilacea*), broadleaf evergreen forest (dominated by *Castanopsis sclerophylla* and *Cyclobalanopsis glauca*) and bamboo forest (*Phyllostachys edulis*). Shrubs are mainly distributed in the valleys, which are unstable communities caused by deforestation. Grassland are mainly found beside the small settlements of the mountains and ridges with barren lands, which are dominated by Poaceae spp..

Minor communities include small mires, plantations (tea tree and rice) and settlements. Mires are mainly located at an elevation of 700-800m a.s.l., support aquatic plants, including *Juncus effusus*, *Sphagnum palustre*, *Ligularia japonica*, *Viola lactiflora*, *Polygonum thunbergii*, *Rotala rotundifolia* and *Salix chaenomeloides*.

6.2.2.2 Vegetation data collection and processing

10 sample points covering the main forest regions described in Table 6.1 were selected for this study using a stratified random sampling method. Vegetation survey around each point was conducted using the Crackles Bequest Project methodology (Bunting et al., 2013). ArcGIS 10.5.1 (Esri, Inc., 1999-2012, ArcGIS program version 10.5.1) was used to digitize a map of communities in the wider area (>100m radius from the sample point) from clipped Sentinel-2 remote-sensed data.

Table 6. 1 Vegetation communities within 100m of the empirical sampling points (see text for details)

Sample Code	Vegetation types	Main species in the forest canopy	Main understory species
1	*Cryptomeria japonica* var. *sinensis* forest	*Cryptomeria japonica* var. *sinensis*	*Aster*, *Rhododendron simsii*, Theaceae spp.
2	*Pinus massoniana-Cyclobalanopsis glauca-Liquidambar formosana* mixed forest	*Pinus massoniana*, *Cyclobalanopsis glauca*, *Liquidambar formosana*, *Lithocarpus* sp.	Ferns, moss, *Buxus microphylla* subsp. *Sinica*, Theaceae spp.
3	*Cyclobalanopsis glauca-Pinus massoniana-Loropetalum chinense* mixed forest	*Cyclobalanopsis glauca*, *Pinus massoniana*	*Rubus corchorifolius*, *Loropetalum chinense*, *Rubus parvifolius*, Theaceae spp.
4	*Phyllostachys edulis* forest	*Phyllostachys edulis*	*Lindera aggregate*, Theaceae spp., *Adiantum capillus-veneris*, *Microlepia* sp.
5	*Astragalus sinicus*-Poaceae spp. grassland	*Astragalus sinicus*, Poaceae spp.	*Astragalus sinicus*, Poaceae spp., *Oxalis corniculata*, *Erigeron* sp.
6	*Cunninghamia lanceolata-Pinus massoniana* forest	*Cunninghamia lanceolata*, *Pinus massoniana*	*Loropetalum chinense*, moss, *Adiantum capillus-veneris*, *Microlepia* spp., *Buxus microphylla* subsp. *Sinica*
7	*Cunninghamia lanceolata-Pinus massoniana* forest	*Cunninghamia lanceolata*, *Pinus massoniana*	*Loropetalum chinense*, moss, *Dicranopteris pedata*, *Microlepia* spp., Theaceae spp.
8	*Cyclobalanopsis glauca-Pinus massoniana-Loropetalum chinense* mixed forest	*Pinus massoniana*, *Loropetalum chinense*, *Quercus* spp., *Castanea* sp., *Cyclobalanopsis glauca*, *Platycarya strobilacea*	Moss, *Dicranopteris pedata*, *Microlepia* spp., *Carex* spp., Theaceae spp.
9	*Pinus massoniana-Cyclobalanopsis glauca-Liquidambar formosana* mixed forest	*Pinus massoniana*, *Cyclobalanopsis glauca*, *Liquidambar formosana*, *Cunninghamia lanceolata*	Moss, *Microlepia* spp., *Dicranopteris pedata*, *Buxus microphylla* subsp. *Sinica*
10	*Cunninghamia lanceolata-Phyllostachys edulis* forest	*Cunninghamia lanceolata*, *Phyllostachys edulis*, *Cyclobalanopsis glauca*	*Microlepia* spp., *Schima argentea*, moss

6.2.2.3 Pollen analysis methods

Moss polsters were collected from each sampling point using inverted sample containers measuring 7.5 cm x 7.5 cm x 1 cm in size (Bunting et al., 2013). A *Lycopodium* spore tablet (concentration 27560 grains/tablet) was added in each sample as an exotic tracer. Pollen extraction was carried out using standard methods including hydrochloric acid treatment (10% HCl), potassium hydroxide treatment (10% KOH), hydrofluoric acid treatment (HF), acetolysis and sieving (with a mesh size of 7μm) (Moore et al., 1991). Residues were mounted in glycerin and identified under an optical microscope at ×400 magnification. A minimum of 400 grains (excluding fern spores and tracers) were counted for each sample sites. Identification of pollen types was made with reference to the Pollen Flora of China (Wang et al., 1995), An illustrated handbook of Quaternary pollen and Spores in China (Tang et al., 2016) and photographs of pollen grains from a herbarium collected in Wuhan Botanical Garden and Nanjing Botanical Garden.

6.2.2.4 Data analysis

Taxa to include in ERV analysis and the reference taxon for all three analyses were selected using inspection of the scatter plots comparing pollen percentage with summed DWPA at 2000m; the taxon with the best range of values and scatter visually closest to rectilinear was selected as the reference taxon, and taxa with multiple outliers or very low pollen or plant presence were excluded from ERV analysis. ERV analysis was then carried out using PolERV (Middleton, unpublished), RPPs were calculated using the MDM in Excel, and IM was applied to the dataset using R (see appendix 4 for the code).

6.3 Results

6.3.1 Simulation study

6.3.1.1 Big dataset (homogeneous vs non-homogeneous)

The scatter plots comparing summed distance-weighted plant abundance at 8000m and pollen proportions for the 6 taxa included are shown in Figure 6.3a (homogeneous) and Figure 6.3b (non-homogeneous). Most taxa show reasonably linear relationships, and following the criteria described in the methods section *Cunninghamia* was chosen as the reference taxon for both datasets.

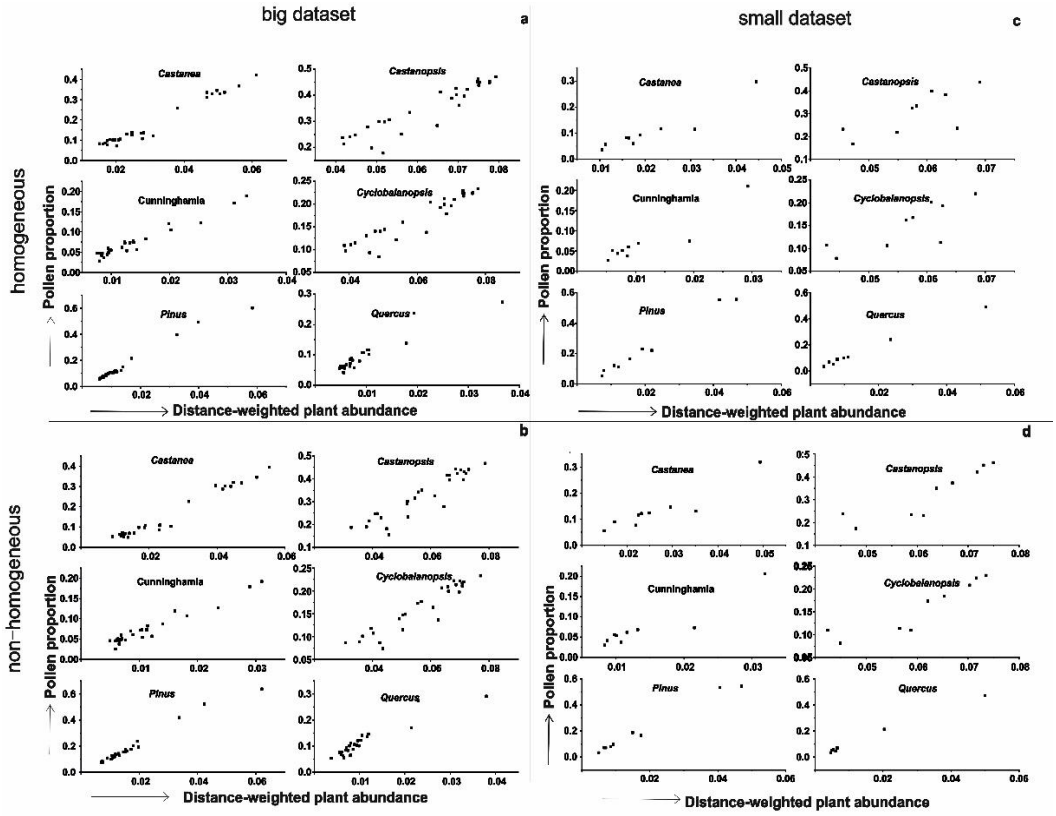


Figure 6. 3 Scatter plots of pollen and distanced weighted plant abundance at 8000m in simulation study (a. big dataset (27 samples) in homogeneous landscape, b. big dataset (27 samples) in non-homogeneous landscape, c. small dataset (9 samples) in homogeneous landscape, d. small dataset (9 samples) in non-homogeneous landscape)

6.3.1.1.1 ERV method

ERV analysis outputs are summarised via plots of likelihood function scores (LFs) against distance of DWPA included in each run. In PolERV analysis, lower values of the score shows improvement in the goodness-of-fit of the corrected data to the linear model, therefore better estimates of the model parameters (Prentice & Parsons, 1983; Sugita, 1994). For the big dataset in homogeneous landscape (Figure 6.4a), the results from all the three ERV sub-models show the same pattern, declining sharply from 0 to c. 1000m, stay moderate onto c. 2000m, and then gradually and monotonically increasing from c. 2000m to 8000m. As for the big dataset in non-homogeneous landscape (Figure 6.4b), the results from all the three ERV sub-models show the same pattern, declining sharply within 0-400m before rising to a secondary peak at about 1000m and falling to c. 2000m, then gradually and monotonically increasing from c. 4000m to 8000m.

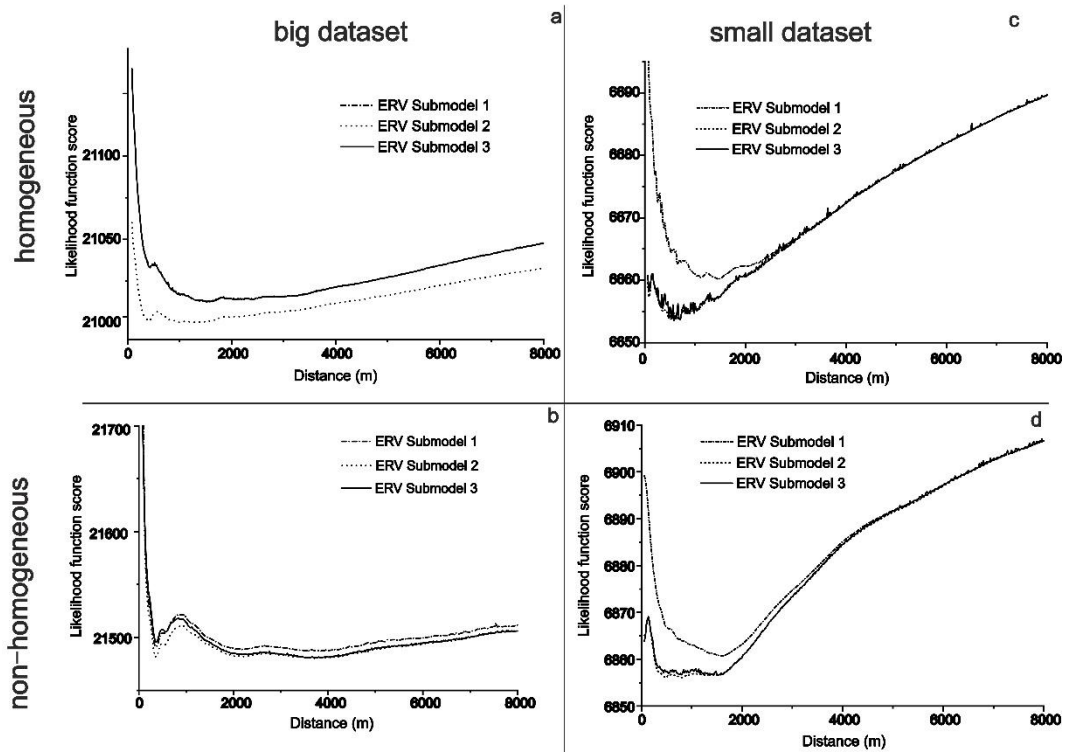


Figure 6. 4 Likelihood function score plots for the three ERV sub-models 1, 2 and 3 (a. big dataset (27 samples) in homogeneous landscape, b. big dataset (27 samples) in non-homogeneous landscape, c. small dataset (9 samples) in homogeneous landscape, d. small dataset (9 samples) in non-homogeneous landscape)

The estimated pollen productivity values relative to *Cunninghamia* ($RPP_{Cunninghamia}$, shortened to RPP here) for the 6 taxa using the three ERV sub-models are shown in Figures 6.5a and 6.5b. In the homogeneous landscape (Figure 6.5a), results from the three sub-models show that the RPPs of *Cyclobalanopsis* is lower than that for *Cunninghamia*. There is no significant difference between the values estimated from sub-model 1 and 3, and sub-model 2 tends to have higher estimation of RPPs values. There is little difference between the RPPs from the ERV methods and the input values, except for *Pinus.*

In the non-homogeneous landscape (Figure 6.5b), the standard deviation of the three methods are larger than results from the homogeneous landscape. The RPPs of *Cyclobalanopsis* and *Castanea* are lower than that for *Cunninghamia*, but there is no significant difference between the three sub-models. Similar to above, the difference between the RPPs from the ERV method and the input values is not obvious for most taxa, however *Pinus* and *Quercus* have higher variations than that in homogeneous landscape.

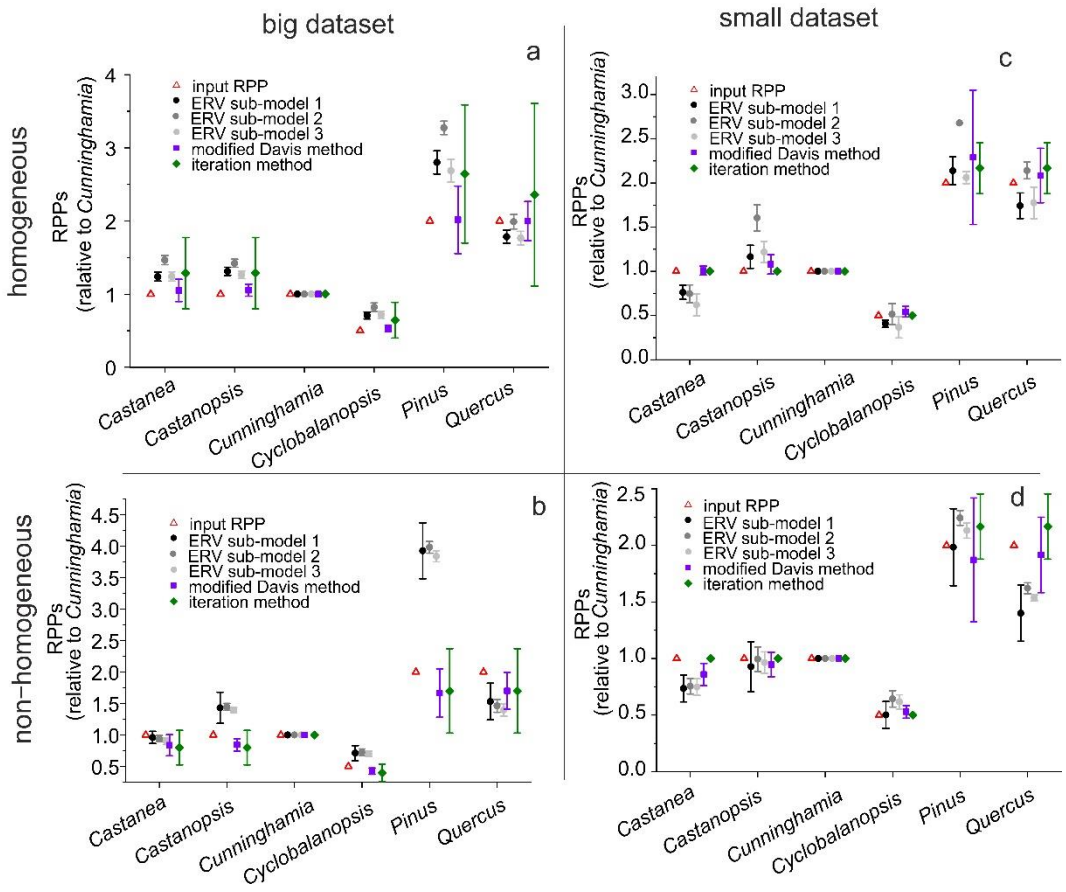


Figure 6. 5 Estimates of RPPs obtained from three methods (ERV, MDM and IM) with input values in Polsack (all three methods are considered with standard deviations (SD)) in simulation study. Note *Cunninghamia* as the reference taxon is set to 1. (a. big dataset (27 samples) in homogeneous landscape. b. big dataset (27 samples) in non-homogeneous landscape. c. small dataset (9 samples) in homogeneous landscape. d. small dataset (9 samples) in non-homogeneous landscape).The black, grey and light grey dots show the RPPs with standard deviation for six forest taxa estimated by 3 ERV sub-models. The purple square shows the RPPs estimated by modified Davis method. The green diamond presents the means of RPPs estimated from iteration method

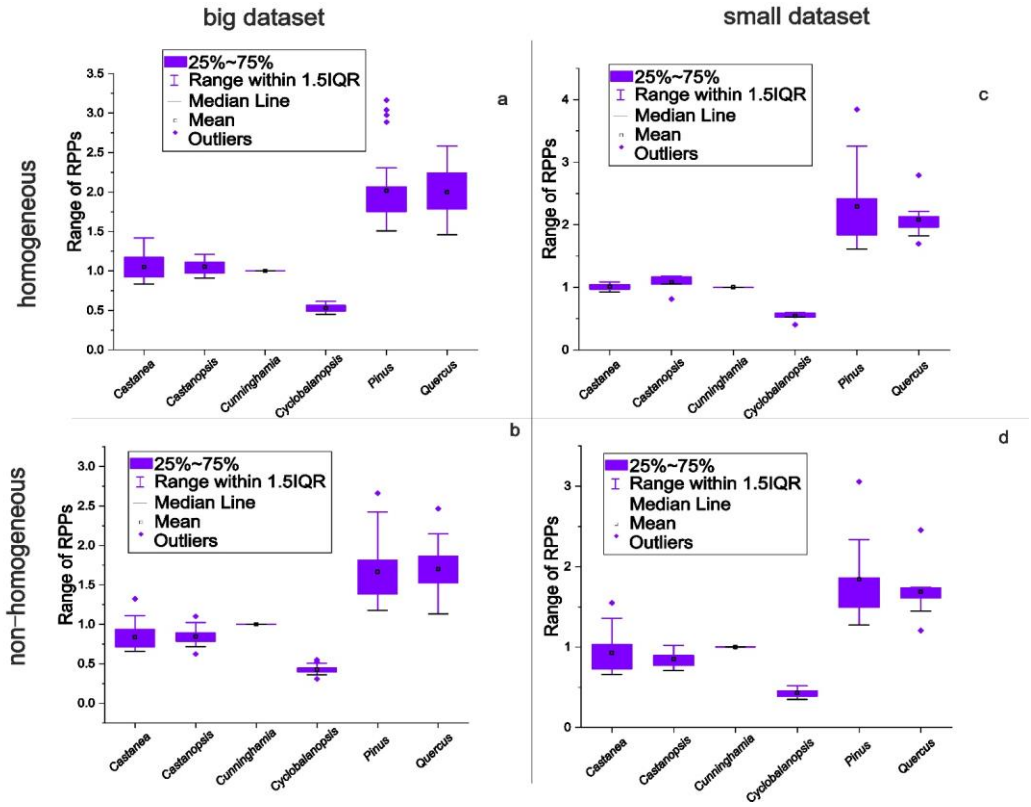


Figure 6. 6 Box plot showing RPPs derived from modified Davis method. Note *Cunninghamia* as the reference taxon is set to 1. (a. big dataset (27 samples) in homogeneous landscape, b. big dataset (27 samples) in non-homogeneous landscape, c. small dataset (9 samples) in homogeneous landscape, d. small dataset (9 samples) in non-homogeneous landscape)

6.3.1.1.2 Modified Davis method

Figures 6.6a and 6.6b present the box-plots of calculated pollen productivity relative to *Cunninghamia* obtained from each of the 27 sites in two types of landscape for all six taxa using modified Davis method.

In the homogeneous landscape, as Figure 6.6a shows, the amount of variation is intermediate for *Castanea* and *Castanopsis*, and smallest for *Cyclobalanopsis*. No outliers are defined for most taxa except for *Pinus* (4 outliers). In the non-homogeneous landscape (Figure 6.6b), the pattern of variations of each taxon are quite similar to that in the homogeneous landscape. One or two outlier values are identified for each taxon.

6.3.1.1.3 Iteration method

In this study, 6 possible values of RPP (0.1, 0.5, 1, 2, 4, and 10) were considered for each taxon (R code in appendix 4.1). In the homogeneous landscape, the results (Table 6.2a) illustrate that the summed squared-chord distance (SSCD) is lowest (0.08) when the pollen productivity of *Castanea* is 1, *Castanopsis* is 1, *Cunninghamia* is 1, *Cyclobalanopsis* is 0.5, *Pinus* is 2, and *Quercus* is 2. In the non-homogeneous landscape, the RPPs with the lowest SSCD (Table 6.2b) is the same as those in the homogeneous landscape. The small number of possible RPP values means that solutions with similarly low SSCD values have quite different RPP profiles. A better estimate of RPP might be obtained from taking the mean of values for multiple good-fit solutions, and in the final rows of each subset of Table 6.2 the mean RPPs from the best-fit profiles are shown, expressed with the value for *Cunninghamia* set to 1 for comparison with other methods. Histograms showing the top 200 smallest SSCD for each test are shown in appendix 5. The number of best-fit profiles for each test were selected individually based on their SSCD values. Selection is based on the visual inspection of the SSCD histograms with group boundaries defined by a small 'step' in the profile. Therefore, 7 profiles were seleted in the homogeneous landscape and 5 profiles in the non-homogeneous landscape.

Table 6. 2 The smallest SSCD obtained from iteration method for simulation study (A. *Castanea*, B. *Castanopsis*, C. *Cunninghamia*, D. *Cyclobalanopsis*, E. *Pinus*, F. *Quercus*). For each set of results from iteration methods, the cut-off value of SSCD is defined differently (homogeneous landscape with big dataset: 7 profiles with value less than 0.24, non-homogeneous landscape with big dataset: 5 profiles with value less than 0.3, homogeneous landscape with small dataset: 3 profiles with value less than 0.05, non-homogeneous landscape with small dataset: 3 profiles with value less than 0.1). The last main column shows the pollen productivity relative to the reference taxon (*Cunninghamia*).

		Pollen productivity estimated from IM							Pollen productivity relative to *Cunninghamia*					
a.	Taxa	A	B	C	D	E	F	SSCD	A	B	C	D	E	F
		1	1	1	0.5	2	2	0.08	1	1	1	0.5	2	2
Homogeneous landscape with big dataset		2	2	2	1	4	4	0.08	1	1	1	0.5	2	2
		4	4	4	2	10	10	0.16	1	1	1	0.5	2.5	2.5
		1	1	0.5	0.5	2	2	0.22	2	2	1	1	4	4
		2	2	1	1	4	4	0.22	2	2	1	1	4	4
		1	1	1	0.5	2	1	0.24	1	1	1	0.5	2	1
		2	2	2	1	4	2	0.24	1	1	1	0.5	2	1
	mean								1.29	1.29	1	0.64	2.64	2.36
	SD								0.49	0.49	0	0.24	0.94	1.25
b.	Taxa	A	B	C	D	E	F	SSCD	A	B	C	D	E	F
		1	1	1	0.5	2	2	0.10	1	1	1	0.5	2	2
Non-homogeneous landscape with big dataset		2	2	2	1	4	4	0.10	1	1	1	0.5	2	2
		4	4	4	2	10	10	0.19	1	1	1	0.5	2.5	2.5
		1	1	2	0.5	2	2	0.27	0.5	0.5	1	0.25	1	1
		2	2	4	1	4	4	0.27	0.5	0.5	1	0.25	1	1
	mean								0.8	0.8	1	0.4	1.7	1.7
	SD								0.27	0.27	0	0.14	0.67	0.67
c.	Taxa	A	B	C	D	E	F	SSCD	A	B	C	D	E	F
		1	1	1	0.5	2	2	0.0479	1	1	1	0.5	2	2
Homogeneous landscape with small dataset		2	2	2	1	4	4	0.0479	1	1	1	0.5	2	2
		4	4	4	2	10	10	0.048	1	1	1	0.5	2.5	2.5
	mean								1	1	1	0.5	2.17	2.17
	SD								0	0	0	0	0.29	0.29
	Taxa	A	B	C	D	E	F	SSCD	A	B	C	D	E	F
d.		1	1	1	0.5	2	2	0.06	1	1	1	0.5	2	2
		2	2	2	1	4	4	0.06	1	1	1	0.5	2	2
Non-homogeneous landscape with small dataset		4	4	4	2	10	10	0.06	1	1	1	0.5	2.5	2.5
	mean								1	1	1	0.5	2.17	2.17
	SD								0	0	0	0	0.29	0.29
Input values		**2**	**2**	**2**	**1**	**4**	**4**		**1**	**1**	**1**	**0.5**	**2**	**2**

6.3.1.1.4 Comparison of results across all three methods

Figure 6.5a compares the RPP values obtained from the three methods with the values input into the simulation, and shows that MDM and IM return values very close to the initial values, whilst ERV results are less accurate. This suggests that the two alternative methods are at least as effective as the ERV method at estimating RPP values.

6.3.1.2 Small dataset (homogeneous vs non-homogeneous)

We repeated these analyses for sub-sets of 9 samples in homogeneous and non-homogeneous landscape, which are considered too-small datasets for successful ERV analysis. *Cunninghamia* was again selected as the most suitable reference taxon (Figures 6.3c and 6.3d).

6.3.1.2.1 ERV method

For the small dataset in both landscapes, the likelihood function score (LFs) show less consistent profiles than those in the big datasets (Figures 6.4c and 6.4d). For samples in the homogeneous landscape, the LFs in sub-model 1 falls sharply from 0 to c.1000m, stay flat from c.100m to 1600m, and then gradually and monotonically increase from c. 1600m to 8000m. The LFs in sub-model 2 and 3 have the same pattern, falling from 0 to c.500m, keeping stable from c.500m to c. 1000m, and increasing from c.1000 to 8000m. As for the small dataset in non-homogeneous landscape, the LFs shows lows between 400m and 1600m for sub-model 2 and sub-model 3, around 1600m for sub-model 1, then increasing substantially to 8000m (Figure 6.4d). Therefore, RPP values with the lowest LFs were selected.

The estimated pollen productivity values relative to *Cunninghamia* for the 6 taxa in the two different types of landscapes from the three ERV sub-models are shown in Figures 6.5c and 6.5d. For the small dataset in the homogeneous landscape, the RPPs estimated from the three sub-models varies. ERV sub-model 2 again have higher values than sub-model 1 and 3. As for the small dataset in non-homogeneous landscape, results from the three sub-models show that the RPPs of *Cyclobalanopsis* and *Castanea* are lower than that for *Cunninghamia*. Comparison of the RPPs from three sub-models shows significant differences for *Pinus* and *Quercus*. Comparing the RPPs from the ERV methods and the input values, there is little difference in *Castanopsis* and *Cyclobalanopsis*, which are better behaved than those in the homogeneous landscape.

6.3.1.2.2 Modified Davis method

For the small dataset in the homogeneous landscape, most taxa show little ranges of pollen productivity except for *Pinus*, and most taxa have one outlier except for *Castanea* (Figure 6.6c). As for the small dataset in non-homogeneous landscape, *Pinus* and *Castanea* show the largest ranges of pollen productivity obtained from modified Davis method, and all taxa have one or two outliers (Figure 6.6d).

6.3.1.2.3 Iteration method

In this study, 6 possible values of RPP (0.1, 0.5, 1, 2, 4, and 10) were again considered for each taxon.

For the small dataset in the homogeneous landscape, the SSCD is lowest (0.479) when the pollen productivity of *Castanea* is 1, *Castanopsis* is 1, *Cunninghamia* is 1, *Cyclobalanopsis* is 0.5, *Pinus* is 2, and *Quercus* is 2 (Table 6.2c). As for the small dataset in non-homogeneous landscape, the SSCD is lowest (0.06) when the pollen productivity of *Castanea* is 1, *Castanopsis* is 1, *Cunninghamia* is 1, *Cyclobalanopsis* is 0.5, *Pinus* is 2, and *Quercus* is 2 (Table 6.2d). The average RPPs from the best-fit profiles (3 for both types of landscape) are shown with their standard deviations expressed with the value for *Cunninghamia* set to 1 for comparison with other methods (Figures 6.5c and 6.5d). *Pinus* and *Quercus* have larger differences compared to input data, and also relatively big standard deviations.

6.3.1.2.4 Comparison of methods

Considering the standard deviation of the RPP values obtained from MDM, there is no significant difference between the empirical RPP used as input and the RPP calculated from the simulated pollen and vegetation data, both in the homogeneous and non-homogeneous landscapes (Figures 6.5c and 6.5d). It is worth noting that *Pinus* and *Quercus* have large standard deviations from all three methods. The RPPs from iteration method coincide exactly with the results from input RPP values, which is the same case as the simulation study with big datasets (Figures 6.5a and 6.5b). This implies that the MDM and IM are suitable for small datasets, if outliers are removed from MDM.

ERV results are not worse than those with the larger dataset, probably due to that the inner grid (5km*5km), a big area much larger than RSAP, is homogeneous. But the likelihood function score is worse in small datasets than big datasets.

6.3.1.3 Summary of simulation study results

In simulation study, the two alternative methods (MDM and IM) perform at least as well as the ERV method. This implies that these two methods are suitable to be considered in further studies. *Pinus* and *Quercus*, two high pollen producers defined in input values, show large standard deviations from all three methods.

There are problems with the ERV method. Both datasets show worse LFs. The reasons are due to the size of dataset and the landscape structure. In simulation studies under non-homogeneous landscape, which do not perfectly meet the assumption of homogeneity, therefore background pollen component is not actually the same for every site.

6.3.2 Empirical study

6.3.2.1 Scatter plots showing the relationships between pollen proportions and DWPA

Figure 6.7a shows the relationship between standard pollen counts and distance weighted plant proportion for 9 selected taxa (*Castanea*, *Cryptomeria*, *Cyclobalanopsis*, *Liquidambar*, *Pinus*, Poaceae, *Quercus*, Rosaceae and Theaceae). *Pinus* was preferred as a reference taxon because it was present at all sites and had the largest variation in pollen and vegetation. It is worth noting that the number of outliers are especially large for *Castanea*, *Quercus* and Rosaceae, characterised by high percentage of pollen and low DWPA. The distance-weighting model chosen here is based on the assumption that canopy component is a major part of the pollen assemblage, therefore a single linear relationship between pollen and vegetation to be modelled. In the scatter plot here, there appeared to be two or more distinct relationship. The high pollen and low DWPA in the three taxa, may suggest other transport mechanisms such as gravity and insects make significant influences on the pollen assemblage (Bunting et al., 2016a). Another possible reason for this pattern could be major differences between forest types (Twiddle et al., 2012).

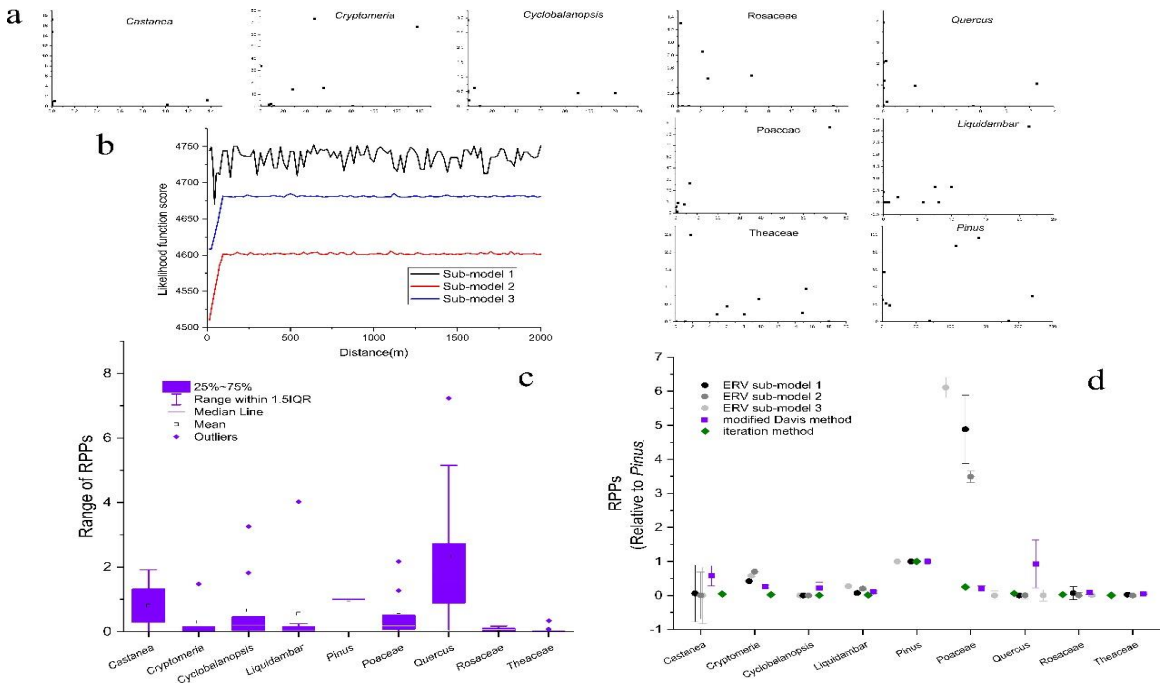


Figure 6. 7 Diagram showing estimates of relative pollen productivity (RPP) obtained from empirical study. Note *Pinus* is selected as a reference taxon. (a. scatter plots of standard pollen count of 9 key taxa (*Castanea*, *Cryptomeria*, *Cyclobalanopsis*, *Liquidambar*, *Pinus*, Poaceae, *Quercus*, Rosaceae and Theaceae) against distance weighted plant abundance (DWPA) to 2000m, b. Likelihood function scores of three sub-models, c. RPPs derived from modified Davis method, very large outliers (>>50) were removed from the figure (5 for *Castanea*, 2 for *Cryptomeria* and 1 for *Liquidambar*). d. RPPs obtained from three methods (ERV, MDM and IM), all three methods are considered with standard deviations (SD)). The black, grey and light grey dots show the RPPs with standard deviation for nine taxa estimated by 3 ERV sub-models at 55m, 120m, 120m, respectively. The purple square shows the RPPs estimated by modified Davis method. *Pinus* show relatively high standard deviations. The green diamond presents the means (the 4 best-fit profile with SSCD less than 2.615) of RPPs estimated from iteration method. All taxa show small standard deviations.

6.3.2.2 ERV analysis

Figure 6.7b shows the LFs obtained from the three ERV sub-models. Sub-model 2 produced the lowest LFs (better fit).The LFs was quite noisy in sub-model 1, while in sub-model 2 and sub-model 3 they were stable beyond 120m.

The black, grey and light grey dots in Figure 6.7d shows the RPP values with standard deviation obtained using ERV analysis. All three sub-models return similar results. The

standard deviations of *Castanea*, *Quercus* and Rosaceae are much higher than the RPP values the model estimates, meaning that the RPP for these three taxa is effectively zero.

6.3.2.3 Modified Davis method

The RPP values of each taxon relative to *Pinus* using modified Davis method are shown in Figure 6.7c. The range of values is small for *Cryptomeria*, *Liquidambar*, Rosaceae and Theaceae. *Castanea* and *Quercus*, both abundant in the samples but rarely seen in the vegetation survey, have much larger variation. RPP values are much smaller than *Pinus* for most taxa. Very large outliers (>>50) were removed from the figure (5 for *Castanea*, 2 for *Cryptomeria* and 1 for *Liquidambar*). Outliers are more common for taxa which were unevenly distributed in vegetation, such as *Castanea*, *Cryptomeria* and *Cyclobalanopsis*, and results are closely clustered for relatively rare taxa in the forest such as *Crytomeria*, *Liquidambar*, Rosaceae and Theaceae.

6.3.2.4 Iteration method

From the alpha ranges of ERV sub-models (appendix 6), we considered that some taxa are strongly under-represented in the pollen assemblages (e.g. Rosaceae, Theaceae and *Cyclobalanopsis*), we chose a different set of 6 possible RPP values for the iteration method (0.01, 0.1, 0.25, 0.5, 1 and 4), and iterated these over 9 taxa (*Castanea*, *Cryptomeria*, *Cyclobalanopsis*, *Liquidambar*, *Pinus*, Poaceae, *Quercus*, Rosaceae and Theaceae). The R code running this estimation can be found in appendix 4.2. The SSCDs are sorted from smallest to largest and a histogram showing the top 200 values are shown in appendix 5. Therefore, the outputs with the four lowest SSCD values (less than 2.615) are expressed relative to *Pinus* (Table 6.3). The SSCD is lowest (2.611) when the pollen productivity of *Castanea* is 0.063, *Cryptomeria* is 0.025, *Cyclobalanopsis* is 0.003, *Liquidambar* is 0.003, *Pinus* is 1.000, *Poaceae* is 0.250, *Quercus* is 0.063, *Rosaceae* is 0.025 and *Theaceae* is 0.003.

Table 6. 3 The smallest SSCD obtained from iteration method for empirical study. The cut-off value of SSCD is set to less than 2.615). The last main column shows the pollen productivity relative to the reference taxon *Pinus*. Symbol for the taxa (A: *Castanea*, B: *Cryptomeria*, C: *Cyclobalanopsis*, D: *Liquidambar*, E: *Pinus*, F: Poaceae, G: *Quercus*, H: Rosaceae and I: Theaceae).

	Pollen productivity estimated from IM									SSCD	Pollen productivity relative to *Pinus*								
Taxa	A	B	C	D	E	F	G	H	I		A	B	C	D	E	F	G	H	I
	0.25	0.1	0.01	0.01	4	1	0.25	0.1	0.01	2.611	0.063	0.025	0.003	0.003	1.000	0.250	0.063	0.025	0.003
	0.1	0.1	0.01	0.01	4	1	0.25	0.1	0.01	2.611	0.025	0.025	0.003	0.003	1.000	0.250	0.063	0.025	0.003
	0.25	0.1	0.01	0.1	4	1	0.25	0.1	0.01	2.615	0.063	0.025	0.003	0.025	1.000	0.250	0.063	0.025	0.003
	0.1	0.1	0.01	0.1	4	1	0.25	0.1	0.01	2.615	0.025	0.025	0.003	0.025	1.000	0.250	0.063	0.025	0.003
mean											0.044	0.025	0.003	0.014	1.000	0.250	0.063	0.025	0.003
SD											0.022	0	0	0.013	0	0	0	0	0

6.3.2.5 Comparison of results from all three methods

A comparative diagram of RPPs obtained from the three methods is shown in Figure 6.7d. RPPs estimated from the three methods are broadly similar except for Poaceae, *Castanea* and *Quercus*. Poaceae is the only herbaceous taxon included in the analysis, and the other two taxa have markedly clustered distribution in the vegetation. This indicates the accessibility of the two new methods in estimating the pollen productivity.

6.3.2.6 Summary of results of empirical study

In empirical study, the two alternative methods (MDM and IM) perform similar results as the ERV method, expect for Poaceae, *Castanea* and *Quercus*. This implies that these two methods are suitable to be considered in further studies.

In our study, the RPPs obtained from three ERV sub-models are quite similar except for Poaceae, the only herbaceous taxon. It is also worth noting that the standard deviations of *Castanea*, *Quercus* and Rosaceae are much higher than the RPP value the ERV method estimates, meaning that the RPP for these three taxa is not different from zero, and might be considered as unreliable. There are problems with the choice of *Pinus* as a reference taxon. *Pinus* is known as an over-represented taxon in previous studies (e.g. Campbell et al., 1999). If it is considered as a reference taxon, as this study shows, will drive down the values of other under-represented taxa.

6.4 Discussion

6.4.1 Simulation study

6.4.1.1 Likelihood function scores

Theoretically, the likelihood function score is expected to decrease quickly from the sampling point, slowing to an asymptote. The distance at which the asymptote is established is termed the relevant source area of pollen (RSAP) (Sugita, 1994; Sugita et al., 1999). Comparing the four tests in the simulation study, the likelihood function scores only present the excepted pattern from the big dataset in the homogeneous landscape. The LFs are much worse for the small sample tests in both types of landscapes, even though the sample locations were random relative to the vegetation patches (see e.g. Broström et al., 2004).

The primary cause of the observed likelihood function score trend is the lack of strict homogeneity in the wider landscape (the outer grid) (Prentice, 1985). Another possible reason is the number of data points. Sufficient data points, a minimum sample count of twice the number of taxa for which RPPs are required, is essential to present reliable estimation of RPPs (Sugita, pers. comm.; Bunting et al. 2013). 27 surface samples satisfies the "2n" rule of thumb for ERV analysis of a dataset with 6 pollen taxa, while 9 samples are quite insufficient.

6.4.1.2 RPP estimation

ERV analysis can be quite sensitive to the number of data points. It is widely accepted that a minimum sample count of twice the number of taxa for which RPPs are required is widely used (Sugita, pers. comm.; Bunting et al. 2013). In the simulation study, 6 taxa were included in the analysis, therefor 27 samples seems to be a suitable number for a successful ERV analysis, and 9 samples is really a small dataset. The comparison of the RPPs obtained from 27 and 9 samples suggest that the likelihood function score, as discussed above, are quite abnormal for the 9 sample test.

Comparison of the RPPs from ERV obtained from the four tests (Figure 6.5) in the simulation study reflect the influences of vegetation structure in the wider landscape and size of dataset. The results from homogeneous landscape with big datasets are more similar to those from the input data. Altering the homogeneity in the wider landscape appears to have more effect on pollen productivity estimates than changing the number of data points.

The modified Davis method works well despite outliers. The number of outliers are different for each taxon in different simulations (Figure 6.6). Because the samples are independent to each other, the outliers can be easily deleted. Even though the distance included in the four simulations studies, 8000m, is not full source area of pollen, and background is different for each sample, the RPP obtained overlap with the input data for all taxa when the standard deviations (SD) are taken into account. The variation for high pollen producers, *Pinus* and *Quercus*, is higher, as shown by the large SDs. The standard deviation of RPPs, which indicates that spread of values between random samples within the same homogenous inner grid, increases with smaller datasets.

From all four tests in the simulation study, when pinning the 'guessed value' from the input values in Polsack, the RPPs with the smallest SSCD turns out to be identical to the input values. When averaging the results with the best SSCDs and adding the standard deviations (the cut-off line of SSCDs in each test is shown in appendix 5), the RPPs obtained again overlap with the input data for all taxa, and the SDs are negligible for low pollen producers, but slightly big for high pollen producers. This suggest that the iteration method dose well for a successful estimation of RPPs, but care must be given to high pollen producers.

Our simulation study proved the applicability of the two alternative methods. The recommendations for future pollen productivity estimates of simulation studies are:

a. repeating with more simulated counts. As part of the estimation of pollen productivity, pollen counts for each sample point are simulated using a probability distribution based on the pollen loadings to include an element of counting error, which is done to accord with real-data systems. In this study, only one set of counts (500) are used in the simulations. Future work can be done with multiple running with different sets of pollen counts, to assess whether counting error will have an influence on the RPPs results.
b. including simulation errors in the vegetation survey. The vegetation component is another important parameter included in the analysis, and the survey area should be large enough to include the relevant source area of pollen (RSAP) of the sites, *sensu* Sugita (1994). However, the vegetation survey, in most cases, will have errors. Future work can be focused on adding the errors of vegetation survey.
c. including different distance weighting methods. The distance weighting method in this study is based on the widely used Prentice-Sugita model. There are also other weighting methods such as inverse distance (z^{-1}) , inverse distance squared (z^{-2}), LSM used in previous published works aiming at their influences on relative pollen productivity estimates (e.g. Broström et al., 2004; Mazier et al., 2008; Bunting & Hjelle, 2010; Poska et al., 2011; Theuerkauf et al., 2013). These model-dependent issues are worthy of considering while testing the applicability of the two methods.

6.4.2 Empirical dataset

Ideally, the sites for collection of the pollen-vegetation data should be randomly distributed in the study area (Broström et al. 2005; Mazier et al., 2008). If this is problematic, the sites should be selected so that the spread of values in vegetation cover is as large as possible (Broström et al., 2008).

The sample distribution in the empirical dataset does not fulfil the requirements for reliable RPP, as the variation in vegetation abundance is too low for several of the taxa of interest. The LFs from the three ERV sub-models are different. The background vegetation is closer to homogeneous therefore an asymptote is reached in sub-model 2 and 3. High degree of variability about that asymptote suggests that the algorithm is finding multiple stable solutions with differing LFs values. The reliability of the RPP estimates can be assessed by the standard deviation calculated in the ERV analysis, namely, if the RPP estimate is less than the SD, the estimated RPP can be considered unreliable (Li et al., 2017). The standard deviations of *Castanea*, *Quercus* and Rosaceae are much higher than the RPP value the ERV model estimates, meaning that the RPP for these three taxa is effectively zero and unreliable.

There are large numbers of outliers from the results of the modified Davis method (Figure 6.7c). The samples in the empirical study are from different forests (Table 6.1), therefore adds the variation of pollen-vegetation relationship. Other possible factor may be different pollen transport routes such as gravity or insects other than wind transport, or pollen transported in clumps other than single grains. We found *Castanea* sometimes appear as clumps under the microscope.

The values to iterate has an influence on the RPP results. In this study, the RPPs from the iteration method is similar to the results from the other two methods. The ‘guessed value’ of iteration method is based on the RPP ranges of ERV method (appendix 6) which shows that most taxa has little ranges except for Poaceae (ranging from 2.5-20 in sub-model 1, 2-4.5 in sub-model 2 and 3-9 in sub-model 3) and *Cunninghamia* (0.5 to 1 in three sub-models).

The empirical dataset estimates the pollen productivity relative to *Pinus*, a taxa widely accepted as a high pollen producer. In order to compare the RPPs obtained from this study with studies in other areas, which usually use Poaceae as a reference taxon, we transferred the pollen productivity relative to Poaceae for the results from modified Davis

method and iteration method (Table 6.4). Using a small dataset, as in the empirical study, the modified Davis performs the best answers when outliers are removed. Our RPP value for *Castanea* (1.14) is very low compared to that (11.49) of Li et al. (2017). This might be due to the poor representation of *Castanea* in the moss samples of this study area. Due to human impact, the natural growing forest are mainly replaced by trees with economic benefits such as *Pinus massoniana*, *Phyllostachys edulis* and *Cunninghamia lanceolata*. *Castanea* trees were scattered and seldom abundant in the study area. *Pinus* has a RPP of 5.93 in this study which is close to the mean value from Europe (6.38) but lower than published studies in northern and temperate China (Li et al., 2015; Li et al., 2017; Mu et al., in press). *Quercus* (9.14) is higher than the values from northern China. Studies in Northern Europe show that *Quercus* has a very high range of variation in RPP (e.g. 7.60 in open and semi-open woods in southern Sweden (Sugita et al., 1999), 2.56 in the lowland Swiss Plateau (Soepboer et al., 2007), 1.30 in open and semi-open Norway (Hjelle & Sugita, 2012), even between trees of the same species (*Quercus robur*) in topographically contrasting environments of the same region, so the difference simply reflects this variability. The large differences between RPP estimates might due to different species and vegetation types between regions. Our values were obtained from studies in subtropical mountainous areas, while the high value from northern China is from cultural and woodland landscape. Note that *Pinus* also varies in Europe between latitudes and landscape types (Broström et al., 2008). Again, the RPPs produced here can be improved with a larger number of sites from pollen traps as well as moss polsters. A comparision study using the same methods is required to better understand the controls on *Quercus* RPP variation, and therefore the implications for paleoenvironmental reconstruction.

Table 6. 4 Comparison of the relative pollen productivities (RPPs) in this study with RPPs obtained in two other Chinese studies, and RPPs from Europe.

Pollen type	This paper (modified Davis method)	Li et al. (2015) Woodland; Changbai Mountains	Li et al. (2017) cultural landscape, Shandong	Li et al., (2017)	Mu et al, in press	Mazier et al. (2012) Europe
Castanea	1.14			11.49±0.49		
Cryptomeria	1.85					
Cyclobalanopsis	2.29					
Liquidambar	2.45					
Pinus	5.93		8.96 ± 0.23		11.46±0.45	6.38 ± 0.45
Poaceae	1					
Quercus	9.14	5.19	4.89 ±0.16		5.17±0.06	5.83 ± 0.15
Rosaceae	0.41					
Theaceae	0.12					

6.4.3 Comparing the three methods

Table 6. 5 Comparison of the three methods

	Dataset	Vegetation area included	Background
ERV	Whole dataset	>RSAP	Calculated
Modified Davis method	Sample by sample	>>RSAP	Incorporated
Iteration method	Whole dataset	>>RSAP	Incorporated

Table 6.5 summarises the differences between the three methods. The ERV and iteration methods use the whole dataset in the analysis. For the modified Davis method, the outliers can be removed as the samples are independent of one another.

The ERV method needs vegetation data to the distance of the RSAP, and the background pollen component is calculated separately as a section of the pollen loading. And MDM and IM requires larger areas (far more beyond RSAP) in order to incorporate the background pollen component.

6.4.4 Implications for future research

The modified Davis method and iteration method show great potential in the simulation study, even when the dataset is small or the landscape is non-homogeneous, precluding a successful ERV analysis. The pollen productivity with the smallest summed squared-chord distance (SSCD) obtained from iteration method is identical to the input values in Polsack software. As in the empirical study, the modified Davis method performs the best answers when outliers are removed. This suggests that the two alternative methods are at least as effective as the ERV method at estimating RPP values, and are suitable to be considered in future studies.

6.4.5 The problem of vegetation data

The vegetation survey is the most time consuming part of model calibration, which costs large human and material resources in the fieldwork. All three methods require reliable vegetation data to run the analysis. In the empirical study, we used the Crackles Bequest Project methodology illustrated by Bunting et al. (2013), which seeks to record the vegetation information in a "pollen sample's point of view".

MDM and IM require vegetation information far beyond the RSAP. With the increasing development of remote sensing technology and availability of free or low-cost datasets, this is becoming easier to obtain, but still depends on ground survey for the composition of the vegetation communities. Potential problems of extrapolation may occur, such as errors in vegetation composition caused by inappropriate season to survey. The remote sensed imagery with the higher resolution is more beneficial for us to get even more abundant information about ground object. In this study, Sentinel-2 was used in combination with detailed field vegetation survey data. The advantage of using Sentinel-2 is that the mission provides a global coverage of the Earth's land surface every 10 days with one satellite and 5 days with two satellites, making the data of great use in multiple season and year studies.

6.4.6 Using the alternative methods

Cautions should be taken when choosing the suitable pollen productivity value to iterate. In the study of 10 moss samples, we selected six values and the iteration process takes about 3 hours. This is because the R code (appendix 4) is not so simple and succinct for datasets with so many taxa. Alternative softwares such as Matlab and C++ might speed

up the analysis of iteration method. In addition, the range between the smallest and biggest value should not be so huge, as to avoid enormous differences among taxa.

The alternative methods may be useful to help extend RPPs to more areas with available datasets such as existing Tauber trap data across Europe from Pollen Monitoring Programme (PMP) (Hicks et al., 2001) and use existing top samples from lakes across arid areas.

Appendix

1. Steps for the modified Davis method

1) Calculate the cumulated distance weighted vegetation data using Sutton term. Note that the distance should be much further than the relevant source area of pollen (RSAP) to ensure sufficient background component being considered.
2) Calculate R_{ik} for each taxon
3) Calculate ratios to a chosen reference taxon ($R_{ik}/R_{ref\ k}$)

2. Steps for the iteration method

Consider k pollen sites which receive i pollen types. For each taxon in each sample, the DWPA (x_{ik}) and pollen loading (y_{ik}) are given values. The calculation of DWPA is the same as above. This method is doing as follows:

1) 'Guess' the RPPs of each taxon. Set the ranges of simulated pollen productivity (α_i).
2) Create a matrix of iterated pollen productivity of each taxon using m different values. In this case, there are l ($l = i\text{^}m$) different combinations of pollen productivity.
3) Calculate the simulated pollen loading ($\hat{y}_{ik}$) using the background-less model. Pollen loading of taxon i is calculated as a result of cumulated distance weighted plant abundance (DWPA) multiplied by the guessed RPP value of taxon i. Here is the equation:

$$\hat{y}_{ik} = \alpha_{il} \times x_{ik}$$

4) Calculate the proportions of simulated pollen loading
5) Calculate the squared-chord distance (SCD) between the proportions of simulated pollen loading ($\hat{y}_{ik}$) and actual pollen loading (y_{ik}).

The squared-chord distance measure is popular with palaeontologists and in studies on pollen. It is defined as:

$$D = \sum_{i=0}^{n} (\sqrt{x_i} - \sqrt{y_i})^2$$

where x_i is the number of observations for species i in sample x, y_i is the number of observations of species i in sample y.

6) Calculate the summed squared-chord distance (SSCD) of all k pollen sites.
7) Repeat step 3-6 with many options using iterative guessed sets of RPPs.
8) Find the combination of pollen productivity with the shortest SSCD.

To make it more clearly, a simple hypothetical example is illustrated in Table 6.6. In this example, there are 2 pollen sites (S1 and S2, which means k is 2) which have 3 pollen types (A, B and C, so i is 3 here). The Guessed RPP values are 1, 2 and 4, which means that m is 3 and the combination of pollen productivity sets (l) is 27. The DWPA for the three taxa (A, B and C) in S1 are 0.06, 0.07 and 0.05, and in S2 they are 0.03, 0.09 and 0.01. The proportions of pollen loading for the three taxa (A, B and C) in S1 are 0.5, 0.2 and 0.3, and in S2 they are 0.2, 0.1 and 0.7.

Table 6. 6 Example showing the procedures of the iteration method.

		Guessed RPP (α_{il})			DWPA(x_{ik})			Simulated pollen loading ($\hat{y}_{ik}$)			Simulated pollen loading as proportions			Pollen loading as proportions (y_{ik})			SCD	SSCD
		A	B	C	A	B	C	A	B	C	A	B	C	A	B	C		
$l = 1$	S1	1	1	1	0.06	0.07	0.05	0.06	0.07	0.05	0.33	0.39	0.28	0.5	0.2	0.3	0.05	0.63
	S2				0.03	0.09	0.01	0.03	0.09	0.01	0.23	0.69	0.08	0.2	0.1	0.7	0.58	
$l = 2$	S1	1	1	2				0.06	0.07	0.1	0.26	0.30	0.43				0.06	0.51
	S2							0.03	0.09	0.02	0.21	0.64	0.14				0.45	
	S1	1	1	4				0.06	0.07	0.2	0.18	0.21	0.61				0.13	0.43
	S2							0.03	0.09	0.04	0.19	0.56	0.25				0.30	
…	…	…	…	…				…	…	…	…	…	…				…	…
$l = i\text{^}m$	S1	4	4	4				0.24	0.28	0.2	0.33	0.39	0.28				0.05	0.63
	S2							0.12	0.36	0.04	0.23	0.69	0.08				0.58	

3. Vegetation composition in simulation study

Table 6. 7 Vegetation composition of outer gird and inner grid

Landscape	Grid	Colour	Community	Proportion in the grid	Pollen taxon	Proportion in each community	RPP	Fall speed
Homogeneous	Outer grid	Yellow	*Cunninghamia lanceolata* forest	6.1%	*Cunninghamia*	100%	2	0.016
		Green	*Pinus massoniana* forest	15.2%	*Pinus*	100%	4	0.063
		Blue-green	*Castanopsis sclerophylla* - *Cyclobalanopsis glauca* forest	54.6%	*Castanopsis*	60%	2	0.006
							1	0.012
					Cyclobalanopsis	40%		
		White	*Quercus fabri* forest	4%	*Quercus*	100%	4	0.016
		Pink	*Castanea seguinii* forest	20%	*Castanea*	100%	2	0.004
	Inner grid	Yellow	*Cunninghamia lanceolata* forest	6.3%	*Cunninghamia*	100%	2	0.016
		Green	*Pinus massoniana* forest	15.4%	*Pinus*	100%	4	0.063
		Blue-green	*Castanopsis sclerophylla*- *Cyclobalanopsis glauca* forest	54.3%	*Castanopsis*	60%	2	0.006
							1	0.012
					Cyclobalanopsis	40%		
		White	*Quercus fabri* forest	4%	*Quercus*	100%	4	0.016
		Pink	*Castanea seguinii* forest	20%	*Castanea*	100%	2	0.004
Non-homogeneous	Outer grid	Yellow	*Cunninghamia lanceolata* forest	12.3%	*Cunninghamia*	100%	2	0.016
		Green	*Pinus massoniana* forest	43.4%	*Pinus*	100%	4	0.063
		Blue-green	*Castanopsis sclerophylla*- *Cyclobalanopsis glauca* forest	32.5%	*Castanopsis*	60%	2	0.006
							1	0.012
					Cyclobalanopsis	40%		
		White	*Quercus fabri* forest	6.9%	*Quercus*	100%	4	0.016
		Pink	*Castanea seguinii* forest	4.9%	*Castanea*	100%	2	0.004
	Inner grid	Yellow	*Cunninghamia lanceolata* forest	6.2%	*Cunninghamia*	100%	2	0.016
		Green	*Pinus massoniana* forest	14.7%	*Pinus*	100%	4	0.063
		Blue-green	*Castanopsis sclerophylla*- *Cyclobalanopsis glauca* forest	54.5%	*Castanopsis*	60%	2	0.006
							1	0.012
					Cyclobalanopsis	40%		
		White	*Quercus fabri* forest	4.3%	*Quercus*	100%	4	0.016
		Pink	*Castanea seguinii* forest	20%	*Castanea*	100%	2	0.004

4. R code for Iteration method

4.1 For simulation study

```
rm(list=ls())
gc()

test_x <- read.csv(file.choose())
head(test_x)

test_y <- read.csv(file.choose())
head(test_y)

library(tcltk)
pb <- tkProgressBar("progress","complete %", 0, 100)

startTime<-Sys.time()

nva=6
ntaxa=6
nsam=27

value<-c(0.1,0.5,1,2,4,10)
n=0
y_hat<-matrix(,nsam,ntaxa)
y_hat_pro<-matrix(,nsam,ntaxa)

maxRow = 19683
newFile = 1;
SSCDist<-matrix(,nrow=maxRow,ncol=ntaxa+1)
colnames(SSCDist)<-c("A","B","C","D","E","F","SSCD")
for (A in value){    ###the alpha of the first taxa changes according to value
 for (B in value){
  for (C in value){
   for (D in value){
    for (E in value){
     for (F in value){
                              alpha<-c(A,B,C,D,E,F)
                              y_hat<-t(alpha*t(test_x))
                              y_hat_pro<-y_hat/rowSums(y_hat)
                              distance<-sum((sqrt(y_hat_pro)-sqrt(test_y))^2)
                              SSCDist[n%%maxRow + 1,1:ntaxa]<-alpha
                              SSCDist[n%%maxRow + 1,ntaxa+1]<-distance
                              if (n%%maxRow == maxRow - 1 || n == nva^ntaxa - 1){
                                        if (newFile == 1){
```

```
                                        newFile = 0
                                        write.table(SSCDist[1:(n%%maxRow + 1),],file="E:\\output-
SSCD.csv",sep=",",row.names=FALSE)
                                } else {
                                        write.table(SSCDist[1:(n%%maxRow + 1),],file="E:\\output-
SSCD.csv",append=TRUE,sep=",",row.names=FALSE,col.names=FALSE)
                                }
                        }

                        if (n %% round(nva^ntaxa / 100) == 0){
                                info<- sprintf("complete %d%%", round(n*100/nva^ntaxa))
                                setTkProgressBar(pb, n*100/nva^ntaxa, sprintf("progress (%s)", info),info)
                        }
                        n<-n+1
    }
   }
  }
 }
}
}

close(pb)
endTime<-Sys.time()
endTime-startTime

a<- read.csv(file.choose())   ##read the output-SSCD.csv file
b<-a[order(a[,7]),]        ##order the file according to the value of SSCD,in this study, the SSCD is in colum 7
head(b)
write.csv(b,file="E:\\order output-SSCD.csv")
```

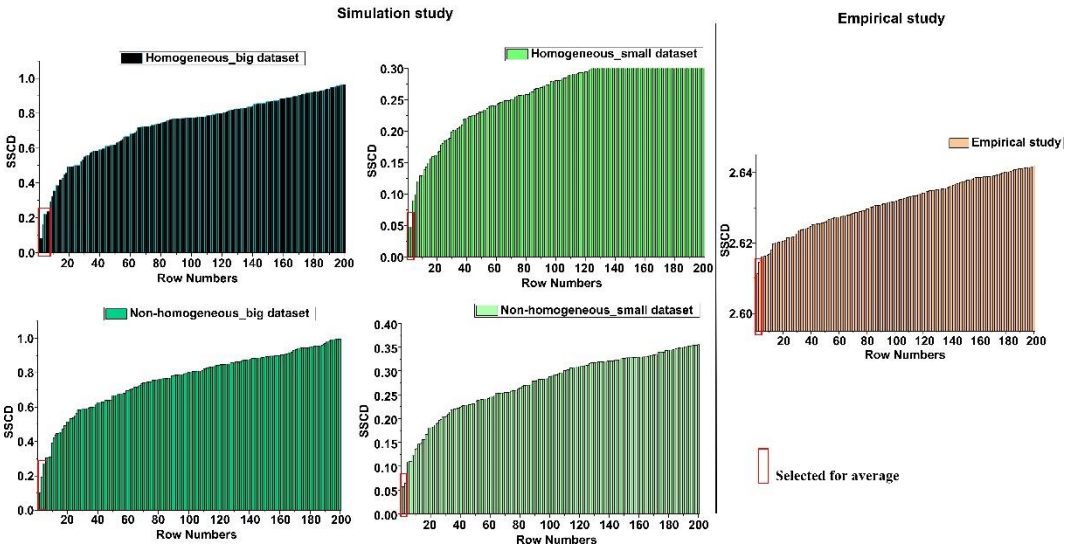
4.2 For empirical study

```
rm(list=ls())
gc()

test_x <- read.csv(file.choose())
head(test_x)

test_y <- read.csv(file.choose())
head(test_y)

library(tcltk)
pb <- tkProgressBar("progress","complete %", 0, 100)
```

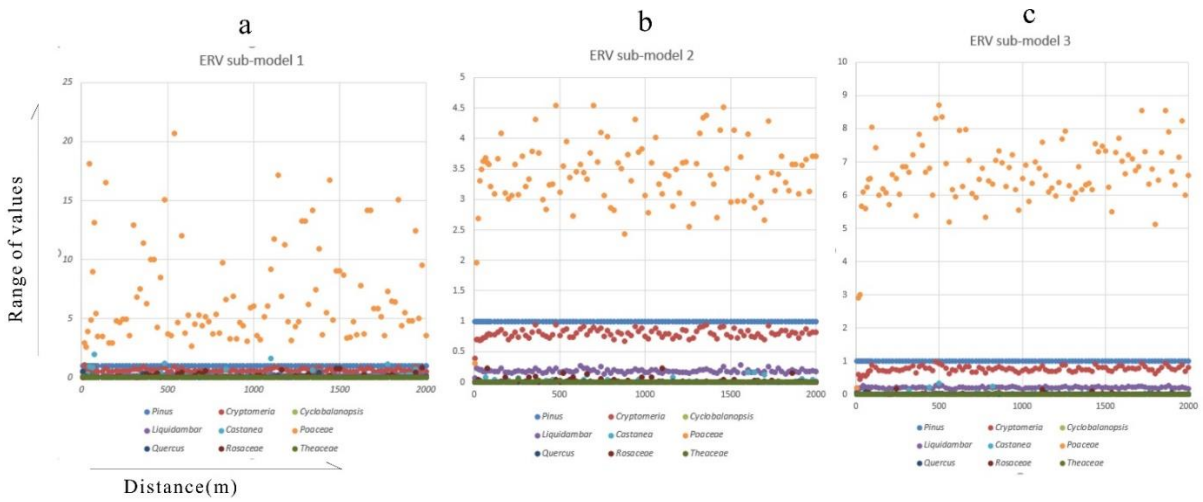
```
startTime<-Sys.time()

nva=6
ntaxa=9
nsam=10

value<-c(0.01,0.1,0.25,0.5,1,4)
n=0
y_hat<-matrix(,nsam,ntaxa)
y_hat_pro<-matrix(,nsam,ntaxa)

maxRow = 19683
newFile = 1;
SSCDist<-matrix(,nrow=maxRow,ncol=ntaxa+1)
colnames(SSCDist)<-c("A","B","C","D","E","F","G","H","I","SSCD")
for (A in value){    ###the alpha of the first taxa changes according to value
 for (B in value){
  for (C in value){
   for (D in value){
    for (E in value){
     for (F in value){
      for (G in value){
       for (H in value){
        for (I in value){
                        alpha<-c(A,B,C,D,E,F,G,H,I)
                        y_hat<-t(alpha*t(test_x))
                        y_hat_pro<-y_hat/rowSums(y_hat)
                        distance<-sum((sqrt(y_hat_pro)-sqrt(test_y))^2)
                        SSCDist[n%%maxRow + 1,1:ntaxa]<-alpha
                        SSCDist[n%%maxRow + 1,ntaxa+1]<-distance
                        if (n%%maxRow == maxRow - 1 || n == nva^ntaxa - 1){
                                if (newFile == 1){
                                        newFile = 0
                                        write.table(SSCDist[1:(n%%maxRow + 1),],file="E:\\output-
SSCD.csv",sep=",",row.names=FALSE)
                                } else {
                                        write.table(SSCDist[1:(n%%maxRow + 1),],file="E:\\output-
SSCD.csv",append=TRUE,sep=",",row.names=FALSE,col.names=FALSE)
                                }
                        }

                        if (n %% round(nva^ntaxa / 100) == 0){
                                info<- sprintf("complete %d%%", round(n*100/nva^ntaxa))
                                setTkProgressBar(pb, n*100/nva^ntaxa, sprintf("progress (%s)", info),info)
                        }
                        n<-n+1
```

```
      }
     }
    }
   }
  }
 }
}
}

close(pb)
endTime<-Sys.time()
endTime-startTime

a<- read.csv(file.choose())   ##read the output-SSCD.csv file
b<-a[order(a[,10]),]         ##order the file according to the value of SSCD,in this study, the SSCD is in colum 10
head(b)
write.csv(b,file="E:\\order output-SSCD.csv")
```

5. Histogram showing the top 200 smallest SSCD in simulation study and empirical study



6. Range of alpha values against distance from ERV sub-model1 (55m), sub-model2 (120m) and sub-model3 (120m) in empirical study



7. Fall speed

Table 6. 8 Fall speeds of the 9 selected main taxa from the Meiling Mountains

Pollen morphotype	Fall speed (m/s)	Published fall speed (m/s) and references
Castanea	0.004	
Cryptomeria	0.015	
Cyclobalanopsis	0.012	
Liquidambar	0.034	
Pinus	0.063	0.031 (Eisenhut, 1961)
Quercus	0.016	0.035 (Eisenhut, 1961)
Rosaceae	0.009	
Theaceae	0.025	
Poaceae	0.030	0.035 (Sugita et al, 1999)

Chapter 7. Effects of mire vegetation on pollen representation of dry land vegetation

Abstract

Surface sediment samples were collected in a mire located in a sub-tropical mountainous area during two different years to investigate the relationships between mire pollen assemblages and the surrounding vegetation, and to assess if there were systematic differences between two different years. Moss polster samples from the surrounding slopes were also collected for comparison.

Pollen assemblages from mire samples reflect the wider landscape to some extent (e.g. dominance of *Pinus*, *Cryptomeria*, *Castanea*, *Quercus* and *Castanopsis*), but they also include high proportions of herbaceous taxa (Poaceae, *Artemisia*, Cyperaceae, *Umbelliferae* and *Typha*) which mostly originate from taxa growing inside the mire. Wetland herbs play a significant role in the pollen spectra.

The PCA plot separated the mire samples collected from two different years, and the change is mainly driven by wetland herbs growing in the mire, but since samples were collected from different locations on different visits it is hard to disentangle the actual causes.

7.1 Introduction

The earth's land cover is an important part of the earth system (Trondman et al., 2015) and affect both ecological landscape functions and processes. In recent years, a great deal of attention has been devoted to understand the past and current land cover both in regional and global scale (e.g. Goldewijk et al., 2001, 2011; Bartholome & Belward, 2005; Gaillard et al., 2008). Pollen spectra from sedimentary records have proved to be an effective way to reconstruct past land cover. The outer layer (exine) of pollen grain is highly resistant to most decay processes (Moore et al., 1991), large numbers of pollen grains and spores become incorporated and fossilised in sediments.

Due to the good preservation properties of pollen and spores, wetland localities are more often used for pollen-analytical investigations (Jacobson & Bradshaw, 1981). Mires often preserve detailed record of environmental change (Lowe & Walker, 2014) and therefore important for reconstructing past land-cover. Studies focusing on the pollen record from valley mires, with the aim of reconstructing the vegetation dynamics and hydrology phases, informing conversation management and human activities have been put forwarded (e.g. Chambers et al., 1985; Hogg et al., 1995; Bunting et al., 1998a, 1998b; Southall et al., 2003; Pawłowski et al., 2012).

Although pollen records from mires are important in interpretation of past land cover, there are challenges, however, due to the additional complicating factor of wetland surface vegetation (Bunting, 2008). The presence of surface vegetation affects the forming pollen assemblage in many ways, including 1) confusion of pollen types which may originate from plants both within and beyond the wetland 2) dilution of the pollen assemblage component originating from the wider landscape by local pollen 3) filtration of pollen from the wider landscape by plant structures and 4) changes in the pollen assemblage caused by different seral stages and changes in the size of the basin rather than changes in the wider landscape. Also, at waterlogged areas where peat develops, solutes and particles deposited on the surface of a peat bog may subsequently move vertically and horizontally. Clymo (1987) indicated that downward, upward and alternating flows will move the median pollen position by about 1.5 cm. More detailed analysis of the complicating factors affecting the pollen signal in wetlands can be found in Jacobson & Bradshaw (1981) and Bunting (2008). Published works have found that the record of the surrounding dry land vegetation can be reflected by the pollen spectra in the mire, but can also be affected by pollen grains accumulating within peatland (Janssen,

1984). Through study of pollen transportation onto a bog surface, Caseldine (1981) proposed that pollen spectra in the central are of the bog is dominated by local pollen (*Calluna*) and taxa from the surrounding woodland. Binney et al., (2005) indicated that pollen assemblages from the fen carr can be used to infer the composition of adjacent dry woodland areas, but for taxa which are far from the edge are difficult to detect.

Despite the great value of the pollen deposition in the mire for reconstructing vegetation history, there is no published study of pollen assemblages in specified sampling points in a small mire within vegetation mosaics for southeast China yet. As the primary interest of the pollen deposit in the Xiyaowu mire is to reconstruct the dry land vegetation of this region, there are some issues to be resolved. Can the pollen deposits in the mire reflect the dry land vegetation? If so, is there any annual differences from the surface sediments collected in the mire? Which local input of pollen signal will make great differences to the pollen assemblages in the mire?

Here we collected surface samples from a mire and the surrounding slopes in two different years. The aims of this chapter are to

a) determine how the mire sample pollen assemblages record the dryland vegetation,
b) and to investigate whether the pollen assemblages vary between the two different years of sampling.

7.2 Study area

7.2.1 Site description

The Meiling Mountains (28°31′N - 28°54′N, 115°34′E - 115°53′E) are located northwest of Nanchang city, in southeast China (Figure 7.1a). The mountains, which are aligned in a northeast-southwest direction, are mostly composed of porphyritic granite and gneiss. Faulting led to the development of valleys and wetlands in the mountains.

The Xiyaowu mire is located in the ridge zone of the main peaks (Sun, 1980). The elevation of the mire is about 748-755 m. It covers an area of about 13334 m^2 and contains a maximum depth of about 250 cm of peat. The basin is lined with a thick silty clay layer with high water retaining capacity (Figure 7.1c) which supports hygrophilous plants.

7.2.2 Vegetation communities

The vegetation communities in the Meiling Mountains are mainly composed of subtropical needleleaf forest (dominated by *Pinus massoniana* and *Cunninghamia lanceolata*), subtropical broadleaf deciduous forest (characterized by *Castanea sequinii*, *Quercus serrata* var. *breviptiolata* and *Platycarya strobilacea*), subtropical broadleaf evergreen forest (dominated by *Castanopsis sclerophylla* and *Cyclobalanopsis glauca*) and subtropical bamboo forest (*Phyllostachys edulis*) and scrub. There are small areas of shrubs, farmland and grassland scattered in the valleys.

Many species of aquatic plants grow in the Xiyaowu mire, including shrubs (*Salix chaenomeloides*) and wetland herbs such as *Juncus effusus*, *Sphagnum palustre*, Poaceae spp., *Viola lactiflora*, *Polygonum thunbergii*, *Ligularia japonica*, and *Rotala rotundifolia*. Belts of tea tree (*Camellia sinensis*) line three sides of the mire, with cedar (*Cryptomeria fortunei*) growing in higher altitudes. The southern slope is covered by shrubs such as *Camellia oleifera* (tea for oil) and *Lindera reflexa* (Figures 7.1b and 7.1d).

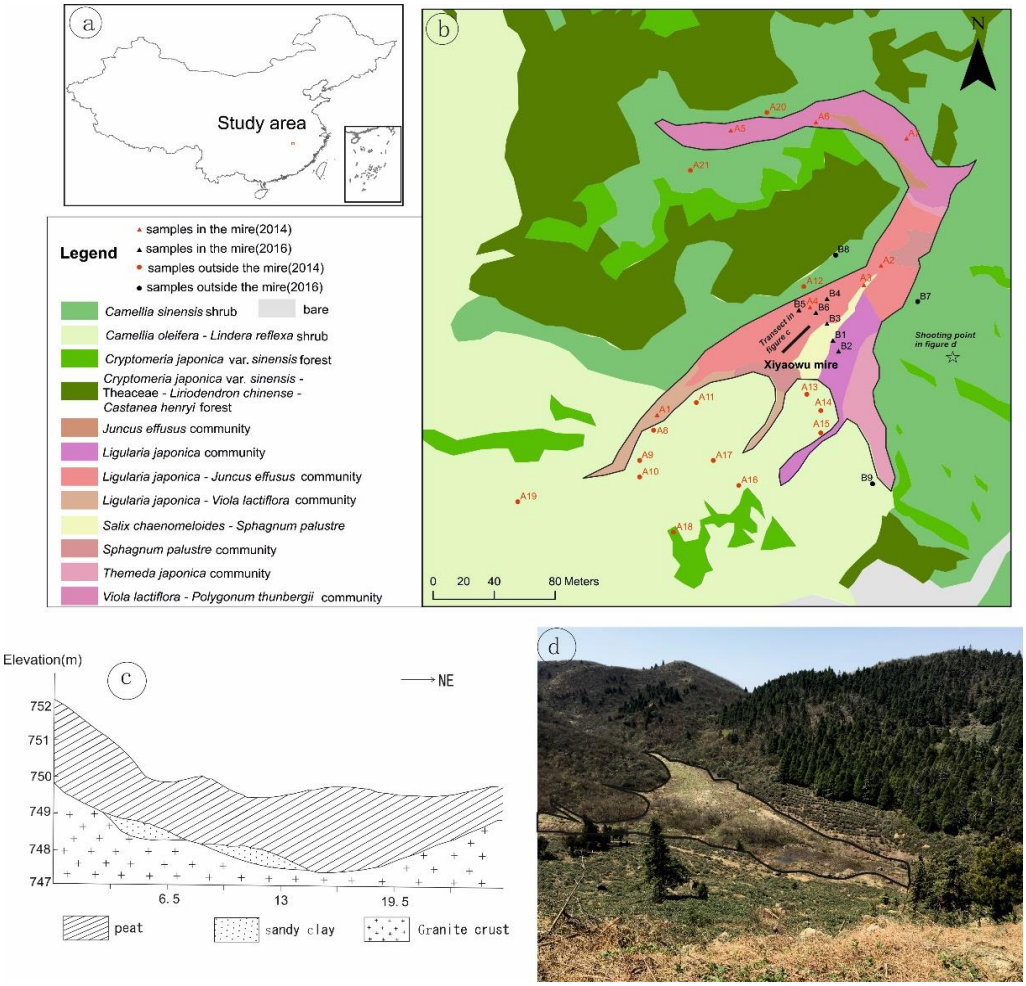


Figure 7. 1 Location of the study area and surface sediment samples collected in 2014 (red symbols) and 2016 (black symbols) (a. map showing the study are in China, b. map showing the sample locations of two years in and around the Xiyaowu mire, c. a SW-NE transect along the mire showing the depth of peat (redraw from Sun, 1980), d. photo showing the shape of the mire, taken from the black star in figure 7.1b whilst looking west-southwest).

7.3 Methods

7.3.1 Field methods

The site selection for collection of pollen samples was primarily designed to see how well the pollen assemblages from surface samples reflect the vegetation types around them. In April, 2014, 7 surface sediment samples (A1-A7) were collected near the edge of the mire, and 14 moss polster samples (A8-A21) were collected to the north and south on the surrounding slopes. In April, 2016, 6 surface samples (B1-B6) were collected inside the

mire, and 3 moss polster samples (B7-B9) were collected from the west and east slopes (Table 7.1).

Table 7. 1 Location, altitude and vegetation community of each sample sites

Year	Sample code	Latitude (°N)	Longitude (°E)	Site type	Landscape	Vegetation
2014	A1	28.74849448	115.6658278	surface sediment	wetland	*Ligularia japonica-Viola lactiflora* community
	A2	28.74939448	115.6671445	surface sediment	wetland	*Ligularia japonica-Juncus effusus* community
	A3	28.74927778	115.6670445	surface sediment	wetland	*Salix chaenomeloides-Sphagnum palustre* community
	A4	28.74914448	115.6667278	surface sediment	wetland	*Ligularia japonica-Juncus effusus* community
	A5	28.75021118	115.6662611	surface sediment	wetland	*Viola lactiflora-Polygonum thunbergii* community
	A6	28.75026118	115.6667611	surface sediment	wetland	*Viola lactiflora-Polygonum thunbergii* community
	A7	28.75016118	115.6672945	surface sediment	wetland	*Viola lactiflora-Polygonum thunbergii* community
	A8	28.74840274	115.6658084	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A9	28.74821944	115.6657251	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A10	28.74811944	115.6657251	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A11	28.74856944	115.6660584	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A12	28.74926944	115.6666917	moss polster	shrub	*Camellia sinensis* shrub
	A13	28.74861944	115.6667084	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A14	28.74851944	115.6667917	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A15	28.74838604	115.6667917	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A16	28.74806944	115.6663084	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A17	28.74821944	115.6661584	moss polster	shrub	*Camellia oleifera-Lindera reflexa* shrub
	A18	28.74778604	115.6659251	moss polster	forest	*Cryptomeria japonica* var. *sinensis* forest
	A19	28.74796944	115.6650084	moss polster	shrub	*Camellia oleifera–Lindera reflexa* shrub
	A20	28.75031944	115.6664751	moss polster	shrub	*Camellia sinensis* shrub
	A21	28.74996944	115.6660251	moss polster	shrub	*Camellia sinensis* shrub
2016	B1	28.74894448	115.6668611	surface sediment	wetland	*Salix chaenomeloides-Sphagnum palustre* community
	B2	28.74887781	115.6668944	surface sediment	wetland	*Ligularia japonica* community
	B3	28.74904448	115.6668278	surface sediment	wetland	*Salix chaenomeloides-Sphagnum palustre* community
	B4	28.74919448	115.6668278	surface sediment	wetland	*Ligularia japonica-Juncus effusus* community
	B5	*28.74912781*	*115.6666611*	*surface sediment*	wetland	*Ligularia japonica-Juncus effusus* community
	B6	28.74911114	115.6667611	surface sediment	wetland	*Ligularia japonica-Juncus effusus* community
	B7	28.74917781	115.6673611	moss polster	shrub	*Camellia sinensis* shrub
	B8	28.74946114	115.6668778	moss polster	shrub	*Camellia sinensis* shrub
	B9	28.748078	115.667094	moss polster	shrub	*Camellia oleifera - Lindera reflexa* shrub

7.3.2 Laboratory methods (see details in section 3.4)

Each sample was treated with (1) 10% HCl, (2) 10% KOH, (3) 36% hot HF, and (4) hot acetolysis 5) sieving (with a mesh size of 7μm) following the standard procedures of Moore et al. (1991). A minimum of 400 grains (excluding fern spores and tracers) were counted for each sample sites.

7.3.3 Data analysis methods

Pollen percentage data was plotted using Tilia 2.0.4 (Grimm, 1991 and updated versions). In order to reduce the impact of aquatic pollen and spores, the percentages of terrestrial pollen were calculated as a proportion of the total pollen originating from arboreal plants and upland herbs, and the proportion of aquatic pollen and spores were calculated based on total pollen and spores.

To explore the general patterns and to assess the similarities between samples, principle component analysis (PCA) was carried out using CANOCO (ter Braak & Smilauer, 1997, 2002). We used percentages of pollen types which reached over 2% in at least one sample for numerical analysis. The species growing within the Xiyaowu mire is floristically distinct from those growing in the surrounding landscape, which makes it possible to separate them from the pollen assemblage in the mire.

7.4 Results

7.4.1 Pollen assemblages from the mire

a) 2014: Pollen spectra from samples collected in the mire are dominated by high percentages of herb species such as Poaceae, *Artemisia*, Cyperaceae, *Umbelliferae* and *Typha* (Figure 7.2). Among the tree species, *Pinus*, *Cryptomeria*, *Castanea*, *Quercus* and *Castanopsis* are well represented. Spores of *Concentricyies*, an algae which grows in a shallow water environment, was present in almost all samples. There is not much visible variation between the samples.

b) 2016: Mire samples have higher proportion of arboreal pollen types than in 2014, especially of *Cryptomeria*. *Concentricyies* is again present in many samples (Figure 7.2). Sample B2 is characterised by less *Pinus* and *Cryptomeria* but more *Platycarya*, *Ambrosia*, Chenopodiaceae, *Humulus* and Liliaceae. Sample B6 also has a higher proportion of Poaceae (<37um) and less *Cryptomeria* than the other mire samples.

For the samples from the mire, the main species are similar in both years, including high percentages of herb species such as Poaceae, *Artemisia*, Cyperaceae, *Umberlliferae* and *Typha* and tree species such as *Pinus*, *Cryptomeria*, *Castanea*, *Quercus* and *Castanopsis*. The pollen spectra show distinct character of wetland communities, with higher amount

of upland herbs and wetland herbs. The proportion of wetland herbs is higher in 2014 than in 2016.

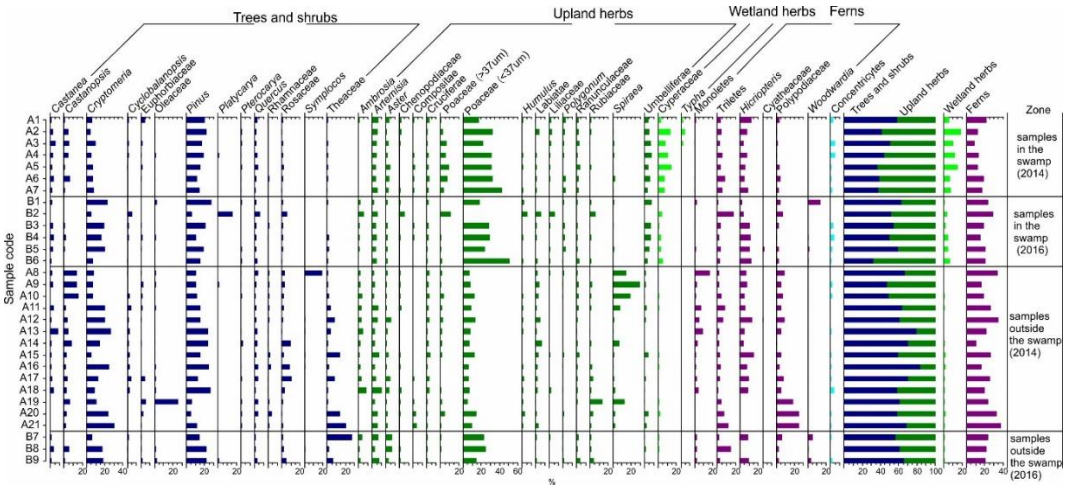


Figure 7. 2 Pollen assemblages in the mire and the surrounding slopes collected in 2014 and 2016

7.4.2 Pollen assemblages from the surrounding slope

a) 2014: Pollen assemblages from samples collected on the surrounding slopes are characterized by *Cryptomeria*, *Pinus*, *Quercus*, Rosaceae and Theaceae (with a maximum value of 20.1% in sample A21) (Figure 7.2). *Castanopsis* is abundant in A8, A9 and A10. The main herb types recorded are Poaceae (up to 42.7% in sample 30), *Artemisia* and *Ambrosia*. The wetland herb abundance dropped dramatically compared to the samples from the mire, while the amount of fern spores increases in general. Although A13, A14 and A15 are not far away from each other (20m in average), their pollen spectra show visible variations in percentage terms.

b) 2016: *Pinus*, *Cryptomeria*, Poaceae, Theaceae, and *Artemisia* are still well represented in the three samples (Figure 7.2). B7, a sample collected within the tea shrub and some distance from the *Cryptomeria fortunei* plantation (about 80 m), is characterised by high amount of Theaceae (26.8%) and less *Cryptomeria* (5.7%). B8 was collected at the boundary between the tea shrub and the *Cryptomeria fortunei* forest (Figure 7.1b), and Theaceae (1.9%) is less abundant, probably due to the high amount of *Cryptomeria* (17.2%).

For the samples in the surrounding slopes, there are much variations compared to samples in the mire. Arboreal pollen types show high amounts in this zone, especially for

Castanopsis, *Cryptomeria* and *Pinus*. Pollen abundance of some taxa are well presented in specified samples, such as Oleaceae in A19 (25.4%), *Symplocos* in A8 (18.7%), *Spireae* in A9 (29.4%). Theaceae remain present in almost all samples, and become important in sample A15 (13.7%), A20 (14.3%), A21 (20.1%) and B7 (26.8%).

7.4.3 Ordination of species and samples in two different years

30 samples and 45 species were included in the analysis. The eigenvalues of the two PCA axes are 0.356 and 0.158 respectively, and together explain 51.4% of the total variance (Table 7.2). As illustrated by the species and sample plots of axis 2 against axis 1 (Figure 7.3), distribution along the two axes is determined by different species groups which have distinct ecological significance, and there is a clear separation between samples from different vegetation communities. Figure 7.3a shows that species growing in the mire (Poaceae, Cyperaceae, *Polygonum* and Umbelliferae) are generally associated with positive scores of axis 1, and forest understory types (*Humulus*, *Artemisia* and Polypodiaceae) are associated with negative scores. Positive scores on axis 2 are driven by regional dryland taxa which are rare found in the surrounding slopes during vegetation survey such as *Castanopsis*, *Spiraea*, *Symplocos*, Poaceae (>37um), and negative scores are associated with the forest dominant species including *Cryptomeria*, Theaceae and *Pinus*.

Table 7. 2 Summary statistics for the first four axes of PCA

Axes	1	2	3	4	Total variance
Eigenvalues	0.356	0.158	0.132	0.079	1
Cumulative percentage variance of species data	35.6	51.4	64.6	72.5	
Sum of all eigenvalues					1

Mire samples in the two years are mostly distributed close to axis 1 (Figure 7.3b), reflecting the relatively high proportions of herbaceous taxa. Mire samples in 2014 (A1-A7) are distributed in right side while samples in 2016 (B1-B6), collected inside the mire, are distributed in the left side.

Figure 7.4 shows the sample scores of the two axes from PCA plotted onto the map of the wetland. For axis 1 (Figure 7.4a), samples in the mire have higher scores (larger markers), while samples on the surrounding slopes have lower scores, especially for those

with some distance from the mire (smaller markers). For axis 2 (Figure 7.4b), high scores mainly seen for samples which are close to the edge of the mire. The small circles, which are driven by the main arboreal types in the vegetation, are mainly distributed inside the surrounding slope. This shows that the regional dryland species have greater effects on samples in the edge of the mire than samples inside the forest.

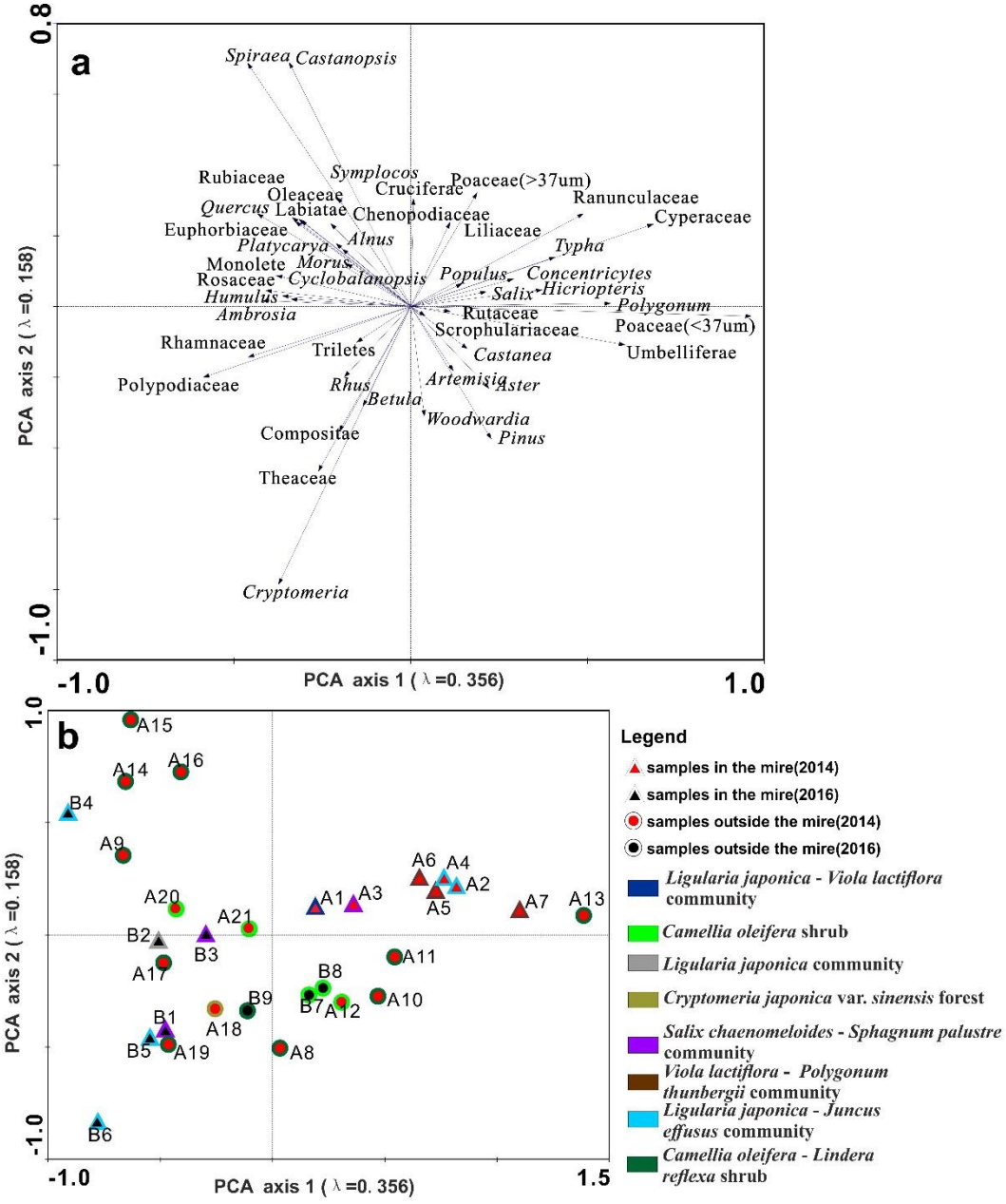


Figure 7. 3 Ordination based on principle components analysis (PCA) diagram showing pollen assemblages of the main vegetation zones in two years. The red symbol show the samples collected in 2014, and the black symbols show samples collected in 2016.

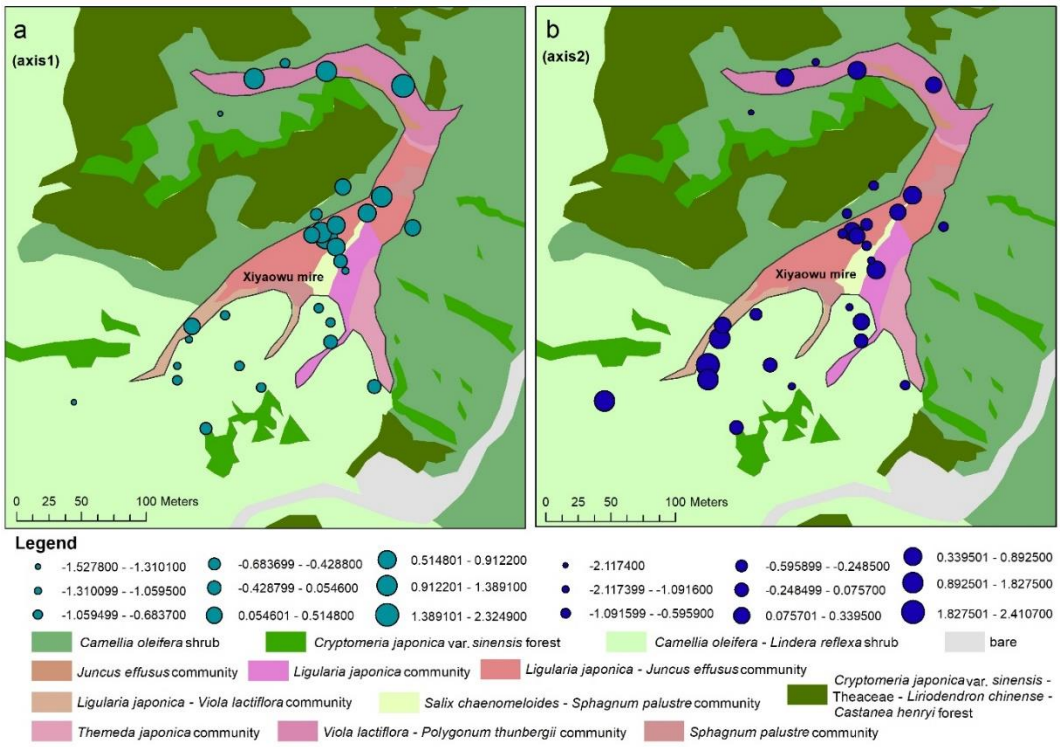


Figure 7. 4 The sample score of axis 1 (a) and axis 2 (b) from PCA. The scores are classified into 9 classes. The bigger the size of the circle, the higher the sample score.

7.5 Discussion

7.5.1 How the samples from the mire recorded the wider forest

Empirical studies have indicated the general relationship between basin size and pollen source area that larger sedimentary basins collect pollen signal from larger areas than smaller basins do (Janssen, 1973; Bradshaw & Webb, 1985; Jackson, 1990). A small mire, as in this study, receives pollen from surface offwash and streams in the drainage area (Peck, 1973; Bonny, 1976), therefore may provide a better chance for the under-represented species to be represented (Hjelle, 1997).

7.5.1.1 Local pollen signals reflected from surface sediment samples in the mire

In the surrounding slopes, belts of tea tree line three sides of the mire, with cedar (*Cryptomeria fortunei*) growing in higher altitudes. There are also small amount of trees (*Castanea henryi*, *Rubus parvifolius*, *Camellia oleifera* and *Liriodendron chinense*) scattered in the cedar forest. The southern slope of the mire is covered by shrubs such as *Camellia oleifera* and *Lindera reflexa* (Figures 7.1b and 7.1d). In the mire, the main

species are shrub (*Salix chaenomeloides*) and wetland herbs such as *Juncus effusus*, *Sphagnum palustre*, Poaceae spp., *Viola lactiflora*, *Polygonum thunbergii*, *Ligularia japonica* and *Rotala rotundifolia*.

The main arboreal species of the pollen assemblages in the mire are *Pinus*, *Cryptomeria*, *Castanea*, *Quercus* and *Castanopsis*. Pollen spectra in the mire reflected the local dryland vegetation to some extent, such as *Cryptomeria* (mean, 10.4%) and *Castanea* (mean, 2.2%). Tea is an important economic plant and there is a long history of tea planting in this area. Although the mire is surrounded with tea bushes, Theaceae pollen does not appear in all mire samples. There are two kinds of tea trees surrounding the mire, one is *Camellia sinensis* which lines three sides of the mire in large area; the other is *Camellia oleifera* mixed with other trees in the southern slope and in higher elevations. There is not sufficient references to distinguish the two species under the microscope, therefore in this study, these two taxa are grouped as Theaceae. *Camellia sinensis* is grown for drink and as flowering will consume the nutrients, farmers will take measures to stop it from flowering as much as possible. However, the number of flowers of *Camellia oleifera* is closely related to its production of fruits (Hu et al., 2015), which is used for extracting oil, therefore farmers will encourage the flowering. Theaceae is sometimes present with high abundance in the slope samples (e.g. A15, A20, A21 and B7). Very local component, such as insects and gravity, will have an effect in surface samples (see discussion in section 4.5.3.4). This pattern might due to anther falling due to gravity or occasional visit of insects.

Wetland species growing on the surface of a mire may be expected to deliver a large number of pollen grains to their immediate surroundings (Moore et al., 1991). The main herbaceous taxa in the mire are Poaceae, *Artemisia*, Cyperaceae, Umbelliferae and *Typha*, which are mostly the same as the wetland species found in the mire samples. The average percentage of wetland species in the mire samples is much higher than those in the surrounding slopes. Poaceae is a species both present in the bottom layer of the forest and in the mire, and the proportion of Poaceae is much higher in the mire samples (27.3% in average) than in the surrounding slopes (10.6% in average). Cyperaceae, a dominate wetland pollen taxon in the mire samples, is not recorded as a dominate wetland species in the vegetation survey. This might due to the seasonal variations in plant recordability. There are seasonal variations in species abundance, especially in grasslands caused by differences in the visibility and distinctiveness of taxa to the surveyor (Martínková et al., 2002). Determining the proportions of wetland species in early spring, as in this study,

can be influenced by the presence of identifiable flowers since the leaves and stems alone can be harder to distinguish with confidence. On the other hand, *Salix*, a dominant species growing in the southern part of the mire, appears with few grains in most mire samples. This might be due to the age of the plant, as most of *Salix chaenomeloides* in the mire are quite young with perhaps a relatively poor flowering.

The large numbers of fern spores in the mire may due to its good preservation capability. The chemical resistance of spore wall, a chemical composition comparable with that pollen exines, of some species are extremely resistant (Faegri et al, 1989). During the vegetation survey, ferns are dominant in the bottom layer of the surrounding forest. Spores in the mire samples are mainly transported by surface runoff and more local in character.

7.5.1.2 Regional vegetation signals reflected from surface sediment samples in the mire

A key regional taxon in the pollen assemblage is *Pinus*, which can be found in all surface samples in the mire. *Pinus massoniana* is absent within 200m radius around the mire but it is a major tree in the mountains. There is no distinct annual differences for *Pinus* in the mire samples, with 16.7% (average) in 2014 and 16.1% (average) in 2016. This suggests good transportation ability of *Pinus*.

7.5.2 Inter-annual differences of samples collected in the mire

Although the main types present in the two years in the mire samples are the same, the samples are grouped in two different parts of the ordination diagram (Figure 7.3a), and the change is mainly driven by wetland herbs growing in the mire. The possible reasons are listed below:

7.5.2.1 Different location

In this study, samples in 2016, which are mostly collected around the middle of the mire, have higher AP percentages than samples in 2014 (near the boundary). The location of samples will have an influence on the pollen assemblages. Both samples are collected in early-spring, when the arboreal trees and shrubs in the trunk space of the slope are blooming, therefore will increase the proportions of those taxa. The trunk space component is the pollen that falls from the tree canopy or produced by shrubs and herbs beneath the canopy (Tauber, 1965), and it is mostly carried by subcanopy air movements.

In pollen accumulation places (such as mire or lake) which borders a forest, the trunk space component is carried out for some distance on to the mire or lake surface, depending on the strength of subcanopy air currents (Andersen, 1974). This study suggests that samples in the middle of the mire are more likely to reflect the surrounding dryland vegetation. However, as the samples are not collected in the same year and there might be inter-annual differences among samples, this conclusion needs further evidence.

7.5.2.2 Different year

In this study, both samples were collected in the same month but in different years. Inter-annual meteorological factors such as precipitation and temperature might also have an effect on the richness and productivity of these annual wetland herbs.

According to the Tauber model (Tauber, 1965, 1967), precipitation will affect the pollen assemblages in the mire in two ways, one is the long-distance component by rain (Cr), and the other is the surface run-off component from the catchment area (Cw). The rainy season in southeast China is closely associated with the strength of the east summer monsoon, which is the main cause of precipitation variations in different years (Zhai et al., 2005).

A peatland is closely linked to its surrounding catchment area through water flow (Joosten & Clarke, 2002). In this area, the surface water flow in the surrounding slopes are influenced by seasonal precipitation, which is more focused in summer. Although runoff component (Cr) is not considered as a main component for pollen deposition in lakes and bogs (Tauber, 1965, 1967), it still makes a contribution.

7.5.3 Further work

The Xiyaowu mire receives drainage water from the surrounding catchment, and it is liable to receive pollen that has been deposited in the slope and subsequently mobilized and transported. This study can be strengthened by adding the elevation and aspect of the surrounding slopes using GIS tools, in order to assess the surface run-off component.

Chapter 8. Discussion and synthesis

The aim of this chapter is to discuss the four main objectives outlined in chapter 1 and presented in chapters 4 to 7, and relate them to the wider literature, to provide recommendations for sampling and coring, and to shape the way for future work.

8.1 Choice of moss or soil samples

8.1.1 Summary of results

Published comparison studies quantitatively assess the difference between different types of surface sampling materials (e.g. mosses and modified 'Tauber trap' by Räsänen et al., 2004; moss, soils and surface lake sediments by Wilmshurst & McGlone, 2005; Pardoe et al., 2010; Lisitsyna et al., 2012), but this study is the first to compare moss and soil samples. Chapter 4 gives the first comparison study of paired moss with adjacent soil collected inside a 1*1 m^2 square area. The results show that moss & soil reflect similar levels of variation in the pollen spectra records, but different mean values of key groups of taxa. There are lower percentages of arboreal types and higher percentage of herb taxa in soil samples than in moss samples. The simulation test shows that in most cases intra-pair differences are greater than could be explained by sampling effects. There might be several reasons for these non-counting effect differences.

8.1.2 Differences in the pollen signals

Pollen preservation conditions in soil samples are known to be problematic (Adam & Mehringer, 1975; Hill, 1996). Soil samples need to be selected carefully, as advised by Havinga (1971, 1984), since some soil types preserve pollen longer than other types. The depth of soil samples should also be chosen with care, since deeper soil may contain a more complex pollen signal due to the absence of the larger soil animals (Havinga, 1974). Figure 4.4 shows that there is no common direction of difference between individual paired samples, therefore it appears that there is no consistent change in the soil pollen signal compared to the moss one as a result of consistent post-depositional losses.

Soil samples contain more upland herb and less arboreal types than the adjacent moss samples. The upland herbs are mainly from areas outside the forest, therefore this suggests that soil has more regional pollen component than moss polster, and mosses have more local component. One possible reason is that surface inflow probably has a greater effect on soil samples than moss, especially in hilly areas with topographical

variation. Local pollen arriving into soil is then transported on, whereas that arriving in moss gets stuck. Another possible reason is the action of soil microfauna and macrofauna after pollen being deposited on the surface (Havinga, 1974), which will be greater in soil than moss. Since the samples are collected within a short distance of each other (within 1 m^2 squared area), the gravity component (Cg) can be seen as the same for both samples. The greater regional signal in soil samples suggests that soils have more long-distance component by rain (Cr) and less canopy component (Cc).

The distance between sample pairs is greatest in heterogenous areas, while for sample pairs in homogenous areas, their distance is shorter. This clearly indicates the effect of vegetation diversity on surface sample diversity. The background component which represents the wilder landscape is essentially uniform within a region (Prentice, 1985; Sugita, 1993, 2007a, 2007b; Bunting & Middleton, 2005, 2009), and variations between paired samples within communities are due to local variations in vegetation. As discussed above, moss better reflects the local vegetation, therefore the pollen assemblages are more variable in diverse communities.

The *Phyllostachys edulis* forest zone is a special case in this study, with the lowest diversity index but relatively large distances between pairs. This might due to the low pollen production by local vegetation (bamboo), which therefore increased the relative proportion of long distance pollen input and more of an even pollen rain. Further studies could be conducted in locations where the local dominant plant is over-represented (e.g. *Artemisia* grassland in northern China) and under-represented to see if this is true more widely.

8.1.3 Recommendations for future surface sample studies

In this study, all paired samples consist of one sample of each type. This is to make them more comparable with a sample from a peat core, because a peat core is from a single point. In order to smooth over spikes of local component (caused by pollen deposition through other routes, e.g. insects or anther fall) in the pollen assemblages, earlier studies have suggested that it is better to collect multiple subsamples (2-3) within the quadrats (Broström et al., 2004; Räsänen et al., 2004; Mazier et al., 2008). In order to minimise the effect of very local vegetation, it is advised to take samples from locations with few/no overhanging plants if possible.

In many cases, the procedures of collecting moss/soil samples, such as the amount/area of samples, the time of collection and the part of moss samples, is insufficiently described. In this study, the paired samples are single ones using inverted sample containers in mixed size. Further studies comparing different moss sampling strategies is clearly needed, such as comparing subsamples and single sample within a fixed area, to see if there are systematic differences. There is also work to do in sampling a wider variety of habitats. In this study, all paired samples were collected in closed forest. Further work can be carried out in open areas such as grasslands and farm lands. Given that pollen assemblages in soils are often biased by physical, chemical and biological factors (Xu et al., 2016), tracing the preservation and decay of pollen in soil requires a wider study. Further study by adding mixed ratios of exotic pollen types into soils under fixed time period (e.g. year/month) would be a valuable point to start.

On the basis of the diversity analysis presented here, it is usually advisable to collect moss samples because mosses better represent the contemporary vegetation. In some cases, the soil sample envelope is larger than the moss sample envelope in the ordination diagram, which suggest that there is less variability among the moss samples within the same vegetation communities. Less variability in samples means greater likelihood that small number of samples are “representative” of the community, which saves time and effort. From this point of view, moss samples is a better choice. As we can see from the direction test, preservation is not a dominant process modifying all paired samples in the same way. In this sub-tropical hilly areas, moss is growing in a more stable way than soil samples, and the latter will be more likely to be affected by factors such as soil fauna (Walch et al., 1970) and soil mixing. Therefore in my study, moss polsters are selected for RPP estimation.

8.2 Is the Prentice-Sugita model a good choice for pollen dispersal and deposition in sub-tropical mountains?

8.2.1 Summary of results

Chapter 5 presents pollen spectra obtained from a grass-forest transect, and the empirical data are compared with the simulation results from multiple pollen dispersal and deposition models. All these simulations here simply aim to find a simple and reasonably close approximation to real-world pollen dispersal across environmental gradients.

The test shows that the P-S model is one of the best performers, other models such as inverse distance squared (z^{-2}), the logarithm ($\log_2(z)$) and a Lagrangian Stochastic model (LSM) also performed as well as the P-S model. However, the inverse distance (z^{-1}) test failed to simulate the main patterns of pollen signal.

8.2.2 Comparison with other studies of woodland edge pollen signals

The 'edge effect' at the woodland edge has been studied, largely using total arboreal pollen rather than specific species, and in most studies empirical results are compared with one or a few mathematical models.

The 'edge effect' is important in reconstructions of past landscapes. Previous works show that the woodland edge effect is a distinctive signal, i.e. rapid decline of tree pollen within few tens of meters distance from the woodland edge (e.g. Tinsley & Smith, 1974; Cundill, 1979; Gearey & Gilbertson, 1997), and therefore can be used to test the behaviour of different models in terms of finer spatial scales. For small wetlands such as the Xiyaowu mire in the Meiling Mountains, which covers an area of about 13334 m^2, most of the sedimentary record in the mire is forming within the 'edge effect' zone.

In the transect pollen assemblages, percentages of AP fell from 89% (mean) in the forest to 54% (mean) by 100 meters, and reaches a minimum of 25% at 300m away from the edge. This conclusion can be compared with other similar studies, such as Tinsley & Smith (1974), which show rapid falls in tree pollen from source (woodland, mean 51%) to background levels (moorland, mean 14.5%) within distances of 100 metres, Gearey & Gilbertson (1997) indicates between 1 to 40m, and Bunting (2002) demonstrates less than 100m. The 'edge effect' is more spread out in this study area. Higher mean values of arboreal pollen in the grassland might be due to the extensive appearance of trees in the Meiling Mountains. The role of altitudinal position might be another factor. The transect

lies in a SW-NE direction, with grassland in lower elevation. The prevailing wind in this region is the same as the direction of this transect, i.e. from the southeast in summer and northeast in winter.

Previous studies tested fewer mixtures of simple algebraic models on 'edge effect', and these studies are from temperate zones. Turner (1964) found that the woodland edge effect was best modelled using the exponential model $\left(e^{-bz}\right)$. On the other hand, Tinsley & Smith (1974) proposed that the inverse distance model (z^{-1}) is the best. In this study, the P-S model is one of the best performers, other models such as the inverse distance squared (z^{-2}) test, the logarithm ($\log_2(z)$) test and the LSM also produce statistically good fits.

8.2.3 Choice of pollen d&d models

Other literature on choice of pollen d&d models looked at their influences on pollen productivity estimations by making assumptions about one model (generally the P-S model), or tested models by finding the best outputs (e.g. use z^{-1}, z^{-2} and P-S in ERV analysis and looking for the best solutions, i.e. those with the lowest likelihood function score). Although the most widely used pollen d&d model is the P-S model, some studies also argue that other d&d models are better performers. In particular, Sjögren et al. (2010) proposed that the composite dispersal function (CDF), which considers pollen dispersal on both local and regional scales, will better simulate the release from abundant pollen source such as trees. The Lagrangian stochastic model (LSM) (Kuparinen et al., 2007; Theuerkauf et al., 2013, 2016) is suggested to be well suited to model pollen deposition in small and medium sized lakes.

Different size of sampling sites and land-cover can be possible causes affecting the conclusions. For instance, the LSM studies (Theuerkauf et al., 2013, 2016) used surface sediment samples from small and medium sized lakes, and CDF studies (Sjögren et al., 2010, 2018) use pollen traps located in clearings in forests in the Swiss Alps. The bigger the sites, the more regional pollen component is likely to be included in the pollen signal, which needs special consideration. The size of those sampling sites are much bigger than the mire where the palaeo core of interest in the Meiling Mountains comes from. This transect test was therefore chosen for its relevance to the palaeo core to be used for reconstructing past land cover.

In the P-S model, fall speed is an important parameter which determines the distance pollen travels after release (Prentice et al., 1987). However, the simple models leave fall speed out of consideration. And for LSM, fall speed is also a less significant parameter compared with the magnitude of typical vertical fluctuations in the atmosphere (Kuparinen et al., 2007). Models of all types performed well in this test, which suggests that fall speed does not have a significant influence in this test.

The assumptions of the P-S model are better met in other different landscapes (e.g. flat areas in Denmark (Tauber, 1967; Andersen, 1970) or Minnesota (Sugita, 1993), fragmented cultural landscapes in southern Sweden (Sugita et al., 1999; Broström et al., 2004)) than in the hilly landscape of this study, but the performance of P-S model here suggests that it can at least capture the main characteristics of the pollen signal and therefore be used. All the models tested here are simply aiming to find a reasonably close approximation of the empirical datasets. Thus looking for more complex models, for example, adding the aspect/elevation into the models, seems unnecessary for this study.

The good behaviour of P-S model might be due to the relatively low altitudinal variation within the transect. Although the P-S model is previously successfully validated in many landscapes, there is still a need to test the behaviour of pollen d&d models before they can be applied in other new areas, especially areas with very huge elevation changes.

8.2.4 Implications of findings for palaeoecology and future research directions

A basic pre-requisite for interpreting palaeoecologial records using models is to test the model applicability to the site and landscape. The choice of model will strongly affect the estimated pollen productivity values (Theuerkauf et al., 2013, 2016), and hence the ensuing quantification of vegetation.

Prior to reconstruction, the selection of models can be informed by testing on a similar and suitable system within the area. In this study, the transect was selected because the scale of test case matches with the scale of target sedimentary system (Xiyaowu mire), both are small sized areas of grassland/mire and surrounded with forest. There are ways to do a similar study better in the future. In my study, there is only one transect with 100m intervals. The sampling internal to study the distribution of arboreal pollen at increasing distances from the woodland edge in other studies is finer, for example Tinsley & Smith (1974) collected samples every 20m along three transects, and increased to 10m

close to the woodland edge. A denser sampling, especially near the edge effect area (<100m), will be a good way to test what has been found from the results. This study uses an ecotone, i.e. from grassland to forest, which is available in this mountains. Future studies should also look at more environmental differences in mountains with more different ecotones.

In order to strengthen the discussion of the suitability of alternative models, future studies using several different kinds of sites in the same landscape could help to determine if pollen signals in different sites are best modelled using different d&d models. Simulation studies with the aim of identifying which sample location would most clearly distinguish between the different models is another option for future work. While adding counting errors, simulations can also be proposed with different counting sums (e.g. 500, 800 and 1000), to test how big the count in the empirical study is needed to be to confidently know if the models are different.

8.3 Obtaining estimates of Relative Pollen Productivity for the Meiling Mountains

8.3.1 Summary of results

Chapter 6 presents a first set of RPP values in southeast China derived using the alternative methods presented there, since the ERV analysis did not produce a useful solution. Two new methods (MDM and IM) are provided to estimate relative pollen productivity. These two alternatives use the same underlying model of the pollen-vegetation relationship as the ERV analysis approach to estimate RPP, but make different assumptions. The two methods open up chances to obtain values for existing pollen datasets whilst needing less field time than the standard means of using ERV analysis.

8.3.2 Comparing novel methods in simulation

The results from the simulation study showed that the two alternatives performed at least as well as the ERV method at returning the input values of relative pollen productivity, no matter the number of sites included in the analysis are big or small. For ERV analysis, there is a widely accepted ‘rule of thumb’ that the minimum number of samples should be 2n, twice the number of taxa for RPP estimation (Sugita, pers. comm.; Bunting et al., 2013). The ERV approach also assumes that a study region has homogenous vegetation composition at a large scale. Therefore in the simulation study, four datasets were

explored, combinations of homogeneous wider landscape vs heterogenous wider landscape and big dataset vs small dataset. Unexpectedly, the ERV results were not strongly affected by the homogeneity of the wider landscape or by the size of the dataset. This might be due to the design of the landscape. All samples were placed in the homogeneous inner grid, and they were at least 1000m away from the edge of the inner grid. The likelihood function scores for the four simulated datasets did not plot up with a perfect decline to an asymptote shape (e.g. Sugita, 1994). This is probably due to the inclusion of counting errors in pollen counts in the analysis.

8.3.3 Empirical estimates of RPP for the Meiling Mountains

The dataset used in the empirical study is not ideal for ERV analysis. The sample points are not randomly distributed (Broström et al., 2005; Mazier et al., 2008) and the range of vegetation cover values for some key taxa (e.g. *Castanea*, Rosaceae, *Quercus*) was very limited (Broström, 2008). Similar to Li's study in northern China (Li et al., 2017), these taxa have greater standard deviation than RPPs values, which mean these should be regarded as unreliable results. A wide range of values of RPP were found using the modified Davis method. This might be affected by the locations of the sample points which were collected from different vegetation communities. Another reason might be the influence of different transport mechanisms at some points such as insect and gravity as well as wind. *Castanea*, for example, sometimes appears as clumps on slides, which suggest that a whole anther fell into the sample from a nearby tree. This may actually be a strength of the Modified Davis method, since such individually anomalous samples can be identified and left out from the mean value for particular taxa. In the ERV method, all taxa must be included in the analysis or the whole sample has to be removed.

8.3.4 Recommendations for future RPP studies and applications of the alternative methods

Both alternative methods require vegetation data from a much larger area than the ERV analysis approach. For all three, a standard vegetation survey method, which is designed to capture the 'pollen's eye of view' of vegetation, is essential to get reliable RPP estimates (Bunting et al., 2013). This ideal vegetation method should reflect the P-S model assumption that the dominant pollen transportation is the air stream above the vegetation canopy. In the empirical study, the Crackles Bequest Project vegetation survey methodology was used (Bunting et al., 2013; see details in section 3.3). The vegetation survey for zone B (10m - 100m) is particularly time-consuming, especially in a dense

forest with variable slope gradients. In future studies, UAV (unmanned aerial vehicle) could be used to assist zone B vegetation survey. With the development of remote-sensed data, it is relatively straightforward to acquire vegetation information for zone C (>>100m). There is a wide range of spatial resolution available to choose from, depending on the user's requirements. A 10m resolution image was used in this study, which is a good choice for successful supervised classification in ArcGIS given the resolution of the available ground trothing data. Another important issue to be considered is the quality of the satellite image - images with no clouds are much easier to classify.

The two alternative methods presented here potentially allow analysts to extract estimates of RPP from small datasets, or datasets that were collected using sampling strategies which do not meet the assumptions of ERV analysis. The code for the iteration method became slow (3 hours) when there were 9 taxa and 6 possible RPP values to iterate. Though this is not a very long time for coding, larger datasets would take longer, and for future work faster code may be developed in R or in other languages such as C++ and Matlab.

The RPP values obtained for the Meiling Mountains represent the main taxa found in the pollen record of the Xiyaowu core (Cui et al., 2018), and therefore provide the necessary basis for reconstruction of past land cover using the MSA or other model-based methods. The intension as a next step is to achieve the application of pollen modelling to strengthen palaeoenvironmental interpretation of this published peat sequence during the LIA.

8.4 How does local mire surface vegetation affect the regional pollen signal?

8.4.1 Summary of results

The results in Chapter 7 show that pollen assemblages from mire samples reflect the wider landscape to some extent (e.g. dominance of *Pinus, Cryptomeria, Castanea, Quercus* and *Castanopsis*), but they also include high proportions of herbaceous taxa (Poaceae, *Artemisia*, Cyperaceae, *Umbelliferae* and *Typha*) which mostly originate from taxa growing inside the mire. The PCA results show that there are inter-annual differences between surface mire samples, and the change is mainly driven by wetland herbs growing in the mire, but since samples were collected from different locations on different visits it is hard to disentangle the actual causes.

8.4.2 Mire pollen assemblages and the dryland vegetation pollen signal

Wetland pollen assemblages do not always consistently record the dryland vegetation pollen signal due to four factors: 1) internal dynamics of the wetland basin itself such as changes in size and seasonal surface water depth, 2) the local wetland vegetation component, which can include the same pollen types as the dry land signal, 3) high pollen productivity of wetland species which can dilute the pollen signal from the dry land vegetation and 4) sampling error (Bunting, 2008). In this study the pollen spectra in the mire are clearly influenced by local herbaceous species such as Cyperaceae and Poaceae. In order to aid the reconstruction of the dryland land cover using pollen records from mires, caution must be taken when considering the wetland pollen taxa. Investigation of the vegetation is an initial and important step before coring.

As suggested by the PCA, the inter-annual differences between the surface mire samples is mainly caused by variation in the local vegetation components, but the sampling design does not allow a distinction to be made between differences in which plants are locally present and differences in the flower and pollen production of plants in different years. Samples in 2014 are located near the edge of the mire, while samples in 2016 are mainly around the middle of the mire, therefore given the zonation of the vegetation (Figure 7.1b) sampled different mixtures of vegetation. Inter-annual meteorological factors such as precipitation and temperature might also have an effect on the richness and productivity of these annual wetland herbs.

8.4.3 Recommendations for future studies

There is still need for further studies of the effects of local vegetation. The Xiyaowu mire is the only mire which can be found by the team members in the Meiling Mountains. It is quite small compared to other existing mires in southern China (Zhou et al., 2004; Dodson et al., 2006). The size of sedimentary basins is thought to influence the source area from which their pollen is derived (Sugita, 1994). The effects of mire size can be tested by creating mires with different areas using HUMPOL software in simulation, and by extending the study to other parts of southern China. Simulation studies could also help investigate differences caused by sample locations by placing samples in different locations of the mire.

The wetland pollen signal is not always easy to interpret. The dynamics of the wetland basin will profoundly affect the pollen record of the dry land (Bunting, 2008). The

reconstruction of wetland surface vegetation and its dynamics both contributes to the reconstruction and exploration of past environments and local archaeology and land use histories and also allied sciences such as ecology and wetland management. In this study, wetland herbs play an important role in the pollen spectra of the Xiyaowu mire. It is therefore recommended that for reconstruction using the core samples in this mire, comparable high values of some herbaceous taxa (Poaceae, Umbelliferae and Cyperaceae) as in this study should be taken with caution. The information of wetland vegetation can also be obtained by identifying other preserved plant fragments, such as diatoms, seeds, charcoal, organic detritus, monocot tissues, bryophyte fragments, aquatic plant tissues (e.g. Muller et al., 2003; Nicholson & Swinehart, 2005). The accuracy of identifying past wetland species will be strengthened by combination studies of these fragments.

8.5 Wider implications

The overall aim of this project was to explore the relationships between pollen and vegetation for assemblages deposited in peatland and forest landscapes in southeast China, with the goal of validating mathematical models of pollen dispersal and deposition, thus providing an essential basis for quantitative reconstruction of past land cover from Quaternary peat deposits in the region. This was addressed via the four objectives discussed above. This project compares the differences in moss and adjacent soil samples (chapter 4), tests the behaviour of pollen d&d model along a grass-forest transect (chapter 5), provides the first estimation of relative pollen productivity in southeast China (chapter 6) and studies the mire samples in reflecting the dryland signals (chapter 7). Surface sample studies and simulation approaches are improving the scientific basis of reconstruction of past landscapes.

The potential of using the P-S model in this mountains area was successfully demonstrated in chapter 5. Previous studies have reconstructed past land-cover from pollen records using this model in various ways and in different temperate regions (e.g. southern Sweden (Hellman et al., 2008a, 2008b); central Europe (Soepboer et al., 2010); the upper Great Lakes region of the US (Sugita et al., 2010); woodland in UK (Bunting & Farrell, 2018)). Since a core record is available from the mire studied here, there is potential to use one or more of these approaches along with the results of this study to reconstruct the past land cover change in the Meiling Mountains.

8.5.1 Recommendations for calibrating the models in other areas

In this study the model test was chosen to match the palaeoecological record available, that is, the grassland-forest transect situation is chosen to be comparable to the small mire in the mountains. Careful choice of model test, and of vegetation survey method so that it does not involve more detail than the pollen signal is able to resolve, is important. Future studies comparing the empirical-model fit for different types of site and different types of ecotone in the same region might help develop better guidelines for choosing models by improving our understanding of how the differences between them relate to the empirical data.

Including counting error in simulations and in investigations of model-data comparisons is also useful, as it helps determine where it is actually possible to see differences in empirical data, not just in the models.

8.5.2 Relative Pollen Productivity: two new methods, and suggestions for sample collection

Chapter 6 presents a first attempt to get RPPs from mountainous areas in southeast China, and also presented two new methods of estimating RPP from empirical datasets which are alternatives to the widely used ERV analysis approach, and demonstrated that, in simulation, these methods work at least as well as ERV analysis. The availability of these two methods enables extracting RPPs from other existing surface sample datasets which don't meet all the assumptions of ERV analysis. For example, Tauber trap samples from the INQUA work group Pollen Monitoring Programme (PMP) dataset (Hicks et al., 2001; http://www.pollentrapping.org/). The sampling strategy of PMP involves few samples at each location, and each dataset covers vegetation transitions so the background pollen rain cannot be assumed to be constant.

Quantitative reconstructions of past land cover using pollen data provide useful data to study the interactions among climate, land surface, and human activity. With the development of modelling methods, global land cover reconstruction projects have been put forward. One of the promising projects, the LandCover6k, is currently being implemented. This project aims to produce past land cover and land use maps using pollen, archaeological and historical data for key time windows, with the aim to aid the paleoclimate modelling (Gaillard et al., 2018). The reconstructions in LandCover6k based on models of pollen dispersal and deposition use the P-S model, and RPP is a key model parameter. This study opens up new ways to estimate RPP which can help obtain

estimates for more parts of the world for LandCover6k, since the existing ERV method can be time-consuming and problematic.

In this study, the TWINSPAN differentiates vegetation communities generally as grassland, mire and forest. However, there is a bias towards the clustering of 29 samples in very similar forest environments. In order to reduce the bias in sample types, there is a suggestion of using remote-sensed vegetation map with unsupervised classification before fieldwork, to provide a stronger basis from which to select a suitable range of sample types for collection in the field.

8.5.3 Fall speed measurements

In this study, fall speed is estimated based on the equation of Stoke's law (Gregory, 1973; Sugita et al., 1999), which takes particle size and density into consideration. This geometric method is more widely used by researchers (e.g. Gregory, 1973; Sugita et al., 1999; Broström et al., 2004; Li et al., 2015) for its convenience and efficiency. However, the impact of chemical analysis on pollen size is currently less well understood. Moore et al. (1991) pointed out that pollen treated with hydrofluoric acid often appears smaller than untreated grains. Broström's work was based on measurements of single pollen grains from books (Punt et al., 1976-1995; Reille, 1992-1998) and papers (Eide, 1981). In this study, pollen grain measurements for fall speed were made on processed samples, thus investigation of how size measurements vary depending on method, and the feasibility of measuring fresh pollen grains, are needed. Since there are several procedures of chemical analysis, the variation of pollen size after each procedure can also be analysed.

8.5.4 Conclusion

The aim of this project has been realised, the main finding of the four objectives are listed below:

Paired moss and soil samples show similar levels of variation in the pollen spectra records, but there are systematic differences in the mean values of key groups of taxa. There are lower percentages of arboreal types and higher percentage of herb taxa in soil samples than in moss samples. Moss polster are chosen since it provides a more accurate representation of the contemporary vegetation.

The P-S model, the inverse distance squared (z^{-2}) test, the logarithm ($\log_2(z)$) test and the LSM test all performed well in terms of SCD and correlation between simulated and empirical data. The Prentice-Sugita model passed the test and is therefore considered suitable for use.

The first estimates of RPPs of 9 key taxa (*Castanea*, *Cryptomeria*, *Cyclobalanopsis*, *Liquidambar*, *Pinus*, Poaceae, *Quercus*, Rosaceae and Theaceae) are provided in this subtropical region. Two alternative methods (modified Davis method and iteration method) for estimating RPP are also developed. These two new methods have great potential for use in areas with existing dataset where fieldwork is particularly difficult, and areas where no values exist and therefore no land cover reconstructions can be made.

The pollen spectra from the Xiyaowu mire reflect both regional and local pollen signal. Wetland herbs taxa play a significant role in the pollen spectra. There are inter-annual differences of the pollen spectra in the mire, the change is mainly driven by wetland herbs growing in the mire, but since samples were collected from different locations on different visits it is hard to disentangle the actual causes.

References

Adam, D.P., Mehringer, P.J., 1975. Modern pollen surface samples-an analysis of subsamples. Journal of Research of the US Geological Survey, 3(6), 733-736.

Allen, B., Kon, M., Bar-Yam, Y., 2009. A new phylogenetic diversity measure generalizing the Shannon index and its application to phyllostomid bats. The American Naturalist, 174(2), 236-243.

An, Z.S., Porter, S.C., Kutzbach, J.E., Wu, X.H., Wang, S.M., Liu, X.D., Li, X.Q., Zhou, W.J., 2000. Asynchronous Holocene optimum of the East Asian monsoon. Quaternary Science Reviews, 19(8), 743-762.

Andersen, M., 1991. Mechanistic models for the seed shadows of wind-dispersed plants. The American Naturalist, 137(4), 476-497.

Andersen, S.T., 1960. Silicone oil as mounting medium for pollen grains. Danmarks Geologiske Undersegelse. 4(1), 1-24.

Andersen, S.T., 1970. The relative pollen productivity and pollen representation of north European trees, and correction factors for tree pollen spectra. Danmarks Geologiske Undersøgelse, 2(96), 1–96.

Andersen, S.T., 1974. Wind Conditions and Pollen Deposition in a Mixed Deciduous Forest: II. Seasonal and Annual Pollen Deposition 1967–1972. Grana, 14(2-3), 64-77.

Andersen, S.T., 1986. Palaeoecological studies of terrestrial soils. Handbook of Holocene palaeoecology and palaeohydrology, 165-177.

Baker, A.G., Zimny, M., Keczyński, A., Bhagwat, S.A., Willis, K.J., Latałowa, M., 2016. Pollen productivity estimates from old-growth forest strongly differ from those obtained in cultural landscapes: evidence from the Białowieża National Park, Poland. The Holocene, 26(1), 80-92.

Bartholome, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing, 26(9), 1959-1977.

Berglund, B.E., Emanuelsson, U., Persson, S., Persson, T., 1986. Pollen/vegetation relationships in grazed and mowed plant communities of South Sweden. Anthropogenic indicators in pollen diagrams. Balkema, Rotterdam, 37-51.

Binney, H.A., Waller, M.P., Bunting, M.J., Armitage, R.A., 2005. The interpretation of fen carr pollen diagrams: the representation of the dry land vegetation. Review of Palaeobotany and Palynology, 134(3-4), 197-218.

Birks, H.J.B., 1973. Past and present vegetation of the Isle of Skye: a palaeoecological study (Vol. 1). Cambridge: University Press xii, 415.

Bonny, A.P., 1976. Recruitment of pollen to the seston and sediment of some Lake District lakes. The Journal of Ecology, 859-887.

Bonny, A.P., 1980. Seasonal and Annual Variation Over 5 Years in Contemporary Airborne Pollen Trapped at a Cumbrian Lake. The Journal of Ecology, 68, 421-441.

Bradshaw, R.H.W., 1981. Modern pollen-representation factors for woods in south-east England. Journal of Ecology, 45-70.

Bradshaw, R.H.W., Webb, T., 1985. Relationships between contemporary pollen and vegetation data from Wisconsin and Michigan, USA. Ecology, 66(3), 721-737.

Broström, A., Nielsen, A.B., Gaillard, M.J., Hjelle, K., Mazier, F., Binney, H., Bunting, J., Fyfe, R., Meltsov, V., Poska, A., Räsänen, S., 2008. Pollen productivity estimates of key European plant taxa for quantitative reconstruction of past vegetation: a review. Vegetation history and archaeobotany, 17(5), 461-478.

Broström, A., Sugita, S., Gaillard, M.J., 2004. Pollen productivity estimates for the reconstruction of past vegetation cover in the cultural landscape of southern Sweden. The Holocene, 14(3), 368-381.

Broström, A., Sugita, S., Gaillard, M.J., Pilesjö, P., 2005. Estimating the spatial scale of pollen dispersal in the cultural landscape of southern Sweden. The Holocene, 15(2), 252-262.

Bunting, M. J., Armitage, R., Binney, H. A., Waller, M., 2005. Estimates of 'relative pollen productivity' and 'relevant source area of pollen' for major tree taxa in two Norfolk (UK) woodlands. The Holocene, 15(3), 459-465.

Bunting, M.J., 2002. Detecting woodland remnants in cultural landscapes: modern pollen deposition around small woodlands in northwest Scotland. The Holocene, 12(3), 291-301.

Bunting, M.J., 2003. Pollen–vegetation relationships in non-arboreal moorland taxa. Review of Palaeobotany and Palynology, 125(3-4), 285-298.

Bunting, M.J., 2008. Pollen in wetlands: using simulations of pollen dispersal and deposition to better interpret the pollen signal. Biodiversity and conservation, 17(9), 2079-2096.

Bunting, M.J., Farrell, M., 2018. Seeing the wood for the trees: recent advances in the reconstruction of woodland in archaeological landscapes using pollen data. Environmental Archaeology, 23(3), 228-239.

Bunting, M.J., Farrell, M., Bayliss, A., Marshall, P., Whittle, A., 2018. Maps from mud—using the Multiple Scenario Approach to reconstruct land cover dynamics from pollen records: a case study of two Neolithic landscapes. Frontiers in Ecology and Evolution, 6, 1-20.

Bunting, M.J., Farrell, M., Broström, A., Hjelle, K.L., Mazier, F., Middleton, R., Nielsen, A.B., Rushton, E., Shaw, H., Twiddle, C.L., 2013. Palynological perspectives on

vegetation survey: a critical step for model-based reconstruction of Quaternary land cover. Quaternary Science Reviews, 82, 41-55.

Bunting, M.J., Gaillard, M.J., Sugita, S., Middleton, R., Broström, A., 2004. Vegetation structure and pollen source area. The Holocene, 14(5), 651-660.

Bunting, M.J., Grant, M.J., Waller, M., 2016a. Approaches to quantitative reconstruction of woody vegetation in managed woodlands from pollen records. Review of Palaeobotany and Palynology, 225, 53-66.

Bunting, M.J., Grant, M.J., Waller, M., 2016b. Pollen signals of ground flora in managed woodlands. Review of Palaeobotany and Palynology, 224, 121-133.

Bunting, M.J., Hjelle, K.L., 2010. Effect of vegetation data collection strategies on estimates of relevant source area of pollen (RSAP) and relative pollen productivity estimates (relative PPE) for non-arboreal taxa. Vegetation History and Archaeobotany, 19(4), 365-374.

Bunting, M.J., Middleton, D., 2005. Modelling pollen dispersal and deposition using HUMPOL software, including simulating windroses and irregular lakes. Review of Palaeobotany and Palynology, 134(3-4), 185-196.

Bunting, M.J., Middleton, R., 2009. Equifinality and uncertainty in the interpretation of pollen data: the Multiple Scenario Approach to reconstruction of past vegetation mosaics. The Holocene, 19(5), 799-803.

Bunting, M.J., Morgan, C.R., Bakel, M.V., Warner, B.G., 1998a. Pre-European settlement conditions and human disturbance of a coniferous swamp in southern Ontario. Canadian Journal of Botany, 76(10), 1770-1779.

Bunting, M.J., Warner, B.G., Morgan, C.R., 1998b. Interpreting pollen diagrams from wetlands: pollen representation in surface samples from Oil Well Bog, southern Ontario. Canadian Journal of Botany, 76(10), 1780-1797.

Bush, M.B., 1992. A simple yet efficient pollen trap for use in vegetation studies. Journal of Vegetation Science, 3(2), 275-276.

Calcote, R., 1995. Pollen Source Area and Pollen Productivity: Evidence from Forest Hollows. Journal of Ecology, 83, 591-602.

Calcote, R., 1998. Identifying forest stand types using pollen from forest hollows. The Holocene, 8(4), 423-432.

Campbell, I. D., McDonald, K., Flannigan, M. D., Kringayark, J., 1999. Long-distance transport of pollen into the Arctic, Nature, 399(6731), 29-30.

Campbell, I.D., 1999. Quaternary pollen taphonomy: examples of differential redeposition and differential preservation. Palaeogeography, Palaeoclimatology, Palaeoecology, 149(1-4), 245-256.

Cao, X.Y., Ni, J., Herzschuh, U., Wang, Y.B., Zhao, Y., 2013. A late Quaternary pollen dataset from eastern continental Asia for vegetation and climate reconstructions: set up and evaluation. Review of Palaeobotany and Palynology, 194, 21-37.

Caseldine, C.J., 1981. Surface pollen studies across Bankhead Moss, Fife, Scotland. Journal of Biogeography, 8, 7-25.

Caseldine, C.J., Gordon, A.D., 1978. Numerical analysis of surface pollen spectra from Bankhead Moss, Fife. New Phytologist, 80(2), 435-453.

Chamberlain, A.C., 1953. Aspects of travel and deposition of aerosol and vapour clouds (No. AERE-HP/R-1261). Atomic Energy Research Establishment, Harwell, Berks, England.

Chambers, F.M., Price, S.M., 1985. Palaeoecology of Alnus (alder): early post‐glacial rise in a valley mire, north‐west Wales. New Phytologist, 101(2), 333-344.

Chen, F., Huang, X., Zhang, J., Holmes, J.A., Chen, J., 2006. Humid little ice age in arid central Asia documented by Bosten Lake, Xinjiang, China. Science in China (Series D: Earth Sciences), 49(12), 1280-1290.

Chen, F., Xu, Q., Chen, J., Birks, H.J.B., Liu, J., Zhang, S., Jin, L., An, C., Telford, R.J., Cao, X., Wang, Z., 2015. East Asian summer monsoon precipitation variability since the last deglaciation. Scientific Reports, 5, 11186.

Chen, F.H., Cheng, B., Zhao, Y., Zhu, Y., Madsen, D.B., 2006. Holocene environmental change inferred from a high-resolution pollen record, Lake Zhuyeze, arid China. The Holocene, 16(5), 675-684.

Chen, M.H., Zhao, H.T., Wen, X.S., Zhang, Q.M., Song, C.J., 1994. Quaternary foraminiferal group and sporopollen zones in cores L2 and L16 in the Lingdingyang Estuary. Marine Geology and Quaternary Geology, 14, 11-22.

Chen, W., Wang, W.M., Dai, X.R., 2009. Holocene vegetation history with implications of human impact in the Lake Chaohu area, Anhui Province, East China. Vegetation History and Archaeobotany, 18(2), 137-146.

China Meteorological Administration Data Service Center, 1981-2010, http://data.cma.gov.cn/.

Christensen, B.B., 1946. Measurement as a means of identifying fossil pollen. Danmarks Geologiske Undersegelse, 4(3), 1-23.

Clymo, R.S., Mackay, D., 1987. Upwash and downwash of pollen and spores in the unsaturated surface layer of Sphangnum-dominated peat. New Phytologist, 105(1), 175-183.

Coles, G.M., Gilbertson, D.D., Hunt, C.O., Jenkinson, R.D.S., 1989. Taphonomy and the palynology of cave deposits. Cave Science, 16(3), 83-89.

Collado, A.D., Chuvieco, E., Camarasa, A., 2002. Satellite remote sensing analysis to monitor desertification processes in the crop-rangeland boundary of Argentina. Journal of Arid Environments, 52(1), 121-133.

Comtois, P., Alcazar, P., Néron, D., 1999. Pollen counts statistics and its relevance to precision. Aerobiologia, 15(1), 19-28.

Crowder, A.A., Cuddy, D.G., 1973. Pollen in a small river basin: Wilton Creek, Ontario. In Symposium of the British Ecological Society. Quaternary Plant Ecology, 61-78.

Cui, A.N., Ma, C.M., Zhao, L., Tang, L.Y., Jia, Y.L., 2018. Pollen records of the Little Ice Age humidity flip in the middle Yangtze River catchment. Quaternary Science Reviews, 193, 43-53.

Cui, A.N., Ma, C.M., Zhao, L., Tang, L.Y., Jia, Y.L., 2018. Pollen records of the Little Ice Age humidity flip in the middle Yangtze River catchment. Quaternary Science Reviews, 193, 43-53.

Cundill, P.R., 1979. Contemporary pollen spectra on the North York Moors. Journal of Biogeography, 127-131.

Cundill, P.R., 1985. The use of mosses in modern pollen studies at Morton Lochs, Fife. In Transactions of the Botanical Society of Edinburgh, 44(4), 375-383.

Cundill, P.R., 1986. A new design of pollen trap for modern pollen studies. Journal of biogeography, 83-98.

Cundill, P.R., 1991. Comparisons of moss polster and pollen trap data: a pilot study. Grana, 30, 301-308.

Davis, B.A., Brewer, S., Stevenson, A.C., Guiot, J., 2003. The temperature of Europe during the Holocene reconstructed from pollen data. Quaternary Science Reviews, 22(15-17), 1701-1716.

Davis, M.B., 1963. On the theory of pollen analysis. American journal of science, 261(10), 897-912.

Ding, J.H., Lu, S.Z., Lin, S.S., Lin, Q.H., 1965. Plant investigation report in Meiling area, Nanchang. Journal of Nanchang University, 1(7), 73-98.

Ding, Y.H., 1994. Monsoons over China. Springer Science & Business Media.

Dodson, J.R., Hickson, S., Khoo, R., Li, X.Q., Toia, J., Zhou, W.J., 2006. Vegetation and environment history for the past 14 000 yrBP from Dingnan, Jiangxi Province, South China. Journal of Integrative Plant Biology, 48, 1018-1027.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sensing of Environment, 120, 25-36.

Duelli, P., Obrist, M.K., 1998. In search of the best correlates for local organismal biodiversity in cultivated areas. Biodiversity and Conservation, 7(3), 297-309.

Duffin, K.I., Bunting, M.J., 2008. Relative pollen productivity and fall speed estimates for southern African savanna taxa. Vegetation History and Archaeobotany, 17(5), 507-525.

Editorial Board of Vegetation Map of China, Chinese Academy of Sciences, 2007. Vegetation Map of the People's Republic of China (1:1,000,000). The Geological Publishing House, Beijing, 192-257.

Eide, F.Y., 1981. Key for northwest European Rosaceae pollen. Grana, 20(2), 101-118.

Eisenhut, G., 1961. Untersuchungen über die Morphologie und Ökologie der Pollenkörner heimischer und fremdländischer Waldbäume. Forstwissenschaftliche Forschungen (translated into English by Jackson, S.T. and Jaumann, P., 1989)

Erdtman, G., 1969. Handbook of palynolgy: Morphology, taxonomy, ecology. An introduction to the study of pollen grains and spores. Hafner.

Faegri, K., Iversen, J., 1989. Textbook of pollen analysis. Chichester: John Wiley.

Faegri, K., Kaland, P.E., Krzywinski, K., 1989. Textbook of pollen analysis (No. Ed. 4). John Wiley & Sons Ltd.

Fagerlind, F., 1952. The real signification of pollen diagrams. Botaniska Notiser, 185-224.

Farrell, M., Bunting, M.J., Middleton, R., 2016. Replicability of data collected for empirical estimation of relative pollen productivity. Review of Palaeobotany and Palynology, 232, 1-13.

Fletcher, W.J., Debret, M., Goñi, M.F.S., 2013. Mid-Holocene emergence of a low-frequency millennial oscillation in western Mediterranean climate: Implications for past dynamics of the North Atlantic atmospheric westerlies. The Holocene, 23(2), 153-166.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote sensing of environment, 80(1), 185-201.

Foster, D.R., Oswald, W.W., Faison, E.K., Doughty, E.D., Hansen, B.C., 2006. A climatic driver for abrupt mid‐Holocene vegetation dynamics and the hemlock decline in New England. Ecology, 87(12), 2959-2966.

Frampton, W.J., Dash, J., Watmough, G., Milton, E.J., 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS journal of photogrammetry and remote sensing, 82, 83-92.

Gaillard, M.J., 2015. LandCover6k: Global anthropogenic land-cover change and its role in past climate. PAGES News, 23(1), 38-39.

Gaillard, M.J., Morrison, K.D., Madella, M., Whitehouse, N., 2018. Past land-use and land-cover change: the challenge of quantification at the subcontinental to global scales. Past Global Changes, 26(1), 2.

Gaillard, M.J., Sugita, S., Bunting, J., Dearing, J., Bittmann, F., 2008. Human impact on terrestrial ecosystems, pollen calibration and quantitative reconstruction of past land-cover. Vegetation History and Archaeobotany, 17, 415–418.

Gaillard, M.J., Sugita, S., Mazier, F., Trondman, A.K., Brostrom, A., Hickler, T., Kaplan, J.O., Kjellström, E., Kokfelt, U., Kunes, P., Lemmen, C., 2010. Holocene land-cover reconstructions for studies on land cover-climate feedbacks. Climate of the Past, 6, 483-499.

Gajewski, K., 2008. The Global Pollen Database in biogeographical and palaeoclimatic studies. Progress in Physical Geography, 32(4), 379-402.

Gajewski, K., Viau, A., Sawada, M., Atkinson, D., Wilson, S., 2001. Sphagnum peatland distribution in North America and Eurasia during the past 21,000 years. Global Biogeochemical Cycles, 15(2), 297-310.

Galán, C., Smith, M., Thibaudon, M., Frenguelli, G., Oteros, J., Gehrig, R., Berger, U., Clot, B., Brandao, R., EAS QC working group, 2014. Pollen monitoring: minimum requirements and reproducibility of analysis. Aerobiologia, 30(4), 385-395.

Ge, Q.S., Zhang, X.Z., Hao, Z.X., Zheng, J.Y., 2011. Rates of temperature change in China during the past 2000 years. Science China Earth Science, 41(9), 1233-1241.

Ge, Y., Li, Y., Li, Y., Yang, X., Zhang, R., Xu, Q., 2015. Relevant source area of pollen and relative pollen productivity estimates in Bashang steppe. Quaternary Sciences, 35, 934-945.

Gearey, B., Gilbertson, D., 1997. Pollen taphonomy of trees in a windy climate: Northbay Plantation, Barra, Outer Hebrides. Scottish Geographical Magazine, 113(2), 113-120.

Gering, J.C., Crist, T.O., Veech, J.A., 2003. Additive partitioning of species diversity across multiple spatial scales: implications for regional conservation of biodiversity. Conservation biology, 17(2), 488-499.

Gerson, U., 1969. Moss-arthropod associations. Bryologist, 495-500.

Godwin, H., 1975. The history of the British flora, 2nd edition. Cambridge University Press.

Goldewijk, K. K., Beusen, A., Van Drecht, G., De Vos, M., 2011. The HYDE 3.1 spatially explicit database of human‐induced global land‐use change over the past 12,000 years. Global Ecology and Biogeography, 20(1), 73-86.

Goldewijk, K.K., 2001. Estimating global land use change over the past 300 years: the HYDE database. Global biogeochemical cycles, 15(2), 417-433.

Gregory, 1973. The microbiology of the atmosphere, 2nd edition. New York. J. Wiley: & Sons

Grimm, E.C., 1991. TILIA v2.0.4 (computer software). Illinois State Museum, Research and Collections Centre, Springfield, IL.

Guo, Z., Biscaye, P., Wei, L., Chen, X., Peng, S., Liu, T., 2000. Summer monsoon variations over the last 1.2 Ma from the weathering of loess‐soil sequences in China. Geophysical Research Letters, 27(12), 1751-1754.

Han, H., Li, S., Zhang, L., Zhou, S., Zhang, J., Wu, S., Wang, F., 2000. Study on pollen analysis and palaeoenvironments in Baohuashan Mt. of Jurong county, Jiangsu Province. Acta Palaeontologica Sinica, 39(2), 295-300.

Han, Y., Liu, H., Hao, Q., Liu, X., Guo, W., Shangguan, H., 2017. More reliable pollen productivity estimates and relative source area of pollen in a forest-steppe ecotone with improved vegetation survey. The Holocene, 27(10), 1567-1577.

Havinga, A.J., 1964. Investigation into the differential corrosion susceptibility of pollen and spores. Pollen et spores, 6(2), 621-635.

Havinga, A.J., 1967. Palynology and pollen preservation. Review of Palaeobotany and Palynology, 2(1-4), 81-98.

Havinga, A.J., 1971. An experimental investigation into the decay of pollen and spores in various soil types. Sporopollenin, 446-479.

Havinga, A.J., 1974. Problems in the interpretation of pollen diagram of mineral soils. Geologie Mijnbouw, 53, 449-453.

Havinga, A.J., 1984. A 20-year experimental investigation into the differential corrosion susceptibility of pollen and spores in various soil types. Pollen et spores, 26 (3-4), 541-558.

He, F., Li, Y., Wu, J., Xu, Y., 2016. A comparison of relative pollen productivity from forest steppe, typical steppe and desert steppe in Inner Mongolia. Chinese Science Bulletin, 61(31), 3388-3400.

Hellman, S.E., Gaillard, M.J., Broström, A., Sugita, S., 2008a. The REVEALS model, a new tool to estimate past regional plant abundance from pollen data in large lakes: validation in southern Sweden. Journal of Quaternary Science: Published for the Quaternary Research Association, 23(1), 21-42.

Hellman, S.E., Gaillard, M.J., Broström, A., Sugita, S., 2008b. Effects of the sampling design and selection of parameter values on pollen-based quantitative reconstructions of regional vegetation: a case study in southern Sweden using the REVEALS model. Vegetation History and Archaeobotany, 17(5), 445-459.

Herzschuh, U., 2006. Palaeo-moisture evolution in monsoonal Central Asia during the last 50,000 years. Quaternary Science Reviews, 25(1-2), 163-178.

Herzschuh, U., Birks, H.J.B., Liu, X.Q., Kubatzki, C., Lohmann, G., 2009a. What caused the mid-Holocene forest decline on the eastern Tibet-Qinghai Plateau? Global Ecology and Biogeography 19, 278-286.

Herzschuh, U., Birks, H.J.B., Mischke, S., Zhang, C., Böhner, J., 2010. A modern pollen–climate calibration set based on lake sediments from the Tibetan Plateau and its application to a Late Quaternary pollen record from the Qilian Mountains. Journal of Biogeography, 37(4), 752-766.

Herzschuh, U., Kramer, A., Mischke, S., Zhang, C., 2009b. Quantitative climate and vegetation trends since the late glacial on the northeastern Tibetan Plateau deduced from Koucha Lake pollen spectra. Quaternary Research, 71(2), 162-171.

Hesselman, H., 1919. Om pollenregn på hafvet och fjärrtransport af barrträdspollen. Geologiska Föreningens i Stockholm Förhandlingar, 16:89-99,105,107

Heumann, B.W., 2011. Satellite remote sensing of mangrove forests: Recent advances and future opportunities. Progress in Physical Geography, 35(1), 87-108.

Hicks, S., Tinsley, H., Huusko, A., Jensen, C., Hättestrand, M., Gerasimides, A., Kvavadze, E., 2001. Some comments on spatial variation in arboreal pollen deposition: first records from the Pollen Monitoring Programme (PMP). Review of Palaeobotany and Palynology, 117(1-3), 183-194.

Hill, M.O., 1979. TWINSPAN: A FORTRAN Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of Individual and Attributes. Cornell University Press.

Hill, M.O., Bunce, R.G.H., Shaw, M.W., 1975. Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland. The Journal of Ecology, 597-613.

Hill, T.R., 1996. Statistical determination of sample size and contemporary pollen counts, Natal Drakensberg, South Africa. Grana, 35(2), 119-124.

Hjelle, K. L., Sugita, S., 2012. Estimating pollen productivity and relevant source area of pollen using lake sediments in Norway: How does lake size variation affect the estimates?. The Holocene, 22(3), 313-324.

Hjelle, K.L., 1997. Relationships between pollen and plants in human-influenced vegetation types using presence-absence data in western Norway. Review of Palaeobotany and Palynology, 99(1), 1-16.

Hjelle, K.L., 1998. Herb pollen representation in surface moss samples from mown meadows and pastures in western Norway. Vegetation History and Archaeobotany, 7(2), 79-96.

Hjelle, K.L., Mehl, I.K., Sugita, S., Andersen, G.L., 2015. From pollen percentage to vegetation cover: evaluation of the Landscape Reconstruction Algorithm in western Norway. Journal of Quaternary Science, 30(4), 312-324.

Hogg, P., Squires, P., Fitter, A.H., 1995. Acidification, nitrogen deposition and rapid vegetational change in a small valley mire in Yorkshire. Biological Conservation, 71(2), 143-153.

Hong, Y.T., Hong, B., Lin, Q.A., Shibata, Y., Hirota, M., Zhu, Y.X., Leng, X.T., Wang, Y., Wang, H., Yi, L., 2005. Inverse phase oscillations between the East Asian and Indian Ocean summer monsoons during the last 12 000 years and paleo-El Niño. Earth and Planetary Science Letters, 231(3-4), 337-346.

Hu, Y.L., Yao, X.H., Ren, D.H., Wang, K.L., Cao, Y.Q., 2015. Effects of Main Environment Factors on Flowering in Camellia oleifera. Journal of tropical and subtropical botany, 23(2), 211-217.

Hu, Z.Z., Yang, S., Wu, R., 2003. Long‐term climate variations in China and global warming signals. Journal of Geophysical Research: Atmospheres, 108(D19), 4614.

Huang, Z.G., Li, P.R., Zhang, Z.Y., Li, K.H., Qiao, P.N., 1982. Generation, Development and Evolution of the Zhujiang River Delta. Guangzhou Popular Science Press, Guangzhou.

Jackson, S.T., 1990. Pollen source area and representation in small lakes of the northeastern United States. Review of Palaeobotany and Palynology, 63(1-2), 53-76.

Jackson, S.T., Lyford, M.E., 1999. Pollen dispersal models in Quaternary plant ecology: assumptions, parameters, and prescriptions. The Botanical Review, 65(1), 39-75.

Jacobson, G.L., Bradshaw, R.H., 1981. The Selection of Sites for Paleovegetational Studies. Quaternary research, 16(1), 80-96.

Janssen, C.R., 1966. Recent pollen spectra from the deciduous and coniferous‐deciduous forests of northeastern Minnesota: A study in pollen dispersal. Ecology, 47(5), 804-825.

Janssen, C.R., 1973. Local and regional pollen deposition. Symposium of the British Ecological Society.

Janssen, C.R., 1984. Modern pollen assemblages and vegetation in the Myrtle Lake peatland, Minnesota. Ecological Monographs, 54(2), 213-252.

Jarvis, D.I., 1993. Pollen evidence of changing Holocene monsoon climate in Sichuan Province, China. Quaternary Research, 39(3), 325-337.

Jiang, W., Guo, Z., Sun, X., Wu, H., Chu, G., Yuan, B., Hatté, C., Guiot, J., 2006. Reconstruction of climate and vegetation changes of Lake Bayanchagan (Inner Mongolia): Holocene variability of the East Asian monsoon. Quaternary Research, 65(3), 411-420.

Joosten, H., Clarke, D., 2002. Wise use of mires and peatlands. International Mire Conservation Group and International Peat Society, 304.

Kent, M., 2011. Vegetation description and data analysis: a practical approach. John Wiley & Sons.

Keylock, C.J., 2005. Simpson diversity and the Shannon-Wiener index as special cases of a generalized entropy. Oikos, 109(1), 203-207.

Knox, R.B., 1979. Pollen and allergy. Edward Arnold.

Kuparinen, A., Markkanen, T., Riikonen, H., Vesala, T., 2007. Modelling air-mediated dispersal of spores, pollen and seeds in forested areas. Ecological modelling, 208(2-4), 177-188.

Lebreton, V., Messager, E., Marquer, L., Renault-Miskovsky, J., 2010. A neotaphonomic experiment in pollen oxidation and its implications for archaeopalynology. Review of Palaeobotany and Palynology, 162(1), 29-38.

Lepš, J., Šmilauer, P., 2003. Multivariate analysis of ecological data using CANOCO. Cambridge university press.

Li, F., Gaillard, M. J., Xu, Q., Bunting, M. J., Li, Y., Li, J., Mu, H., Lu, J., Zhang, P., Zhang, S., Cui, Q., 2018. A review of relative pollen productivity estimates from temperate China for pollen-based quantitative reconstruction of past plant cover. Frontiers in Plant Science, 9, 1214.

Li, F., Gaillard, M.J., Sugita, S., Mazier, F., Xu, Q., Zhou, Z., Zhang, Y., Li, Y., Laffly, D., 2017. Relative pollen productivity estimates for major plant taxa of cultural landscapes in central eastern China. Vegetation History and Archaeobotany, 26(6), 587-605.

Li, G.Z., Liang, W., Liao, S.M., Fang, G.X., 1996. Climatic changes since Holocene along Guangxi coast. Marine Geology and Quaternary Geology, 16, 49-60.

Li, X., Dodson, J., Zhou, J., Zhou, X., 2009. Increases of population and expansion of rice agriculture in Asia, and anthropogenic methane emissions since 5000 BP. Quaternary International, 202(1-2), 41-50.

Li, Y., Bunting, M.J., Xu, Q., Jiang, S., Ding, W., Hun, L., 2011. Pollen–vegetation–climate relationships in some desert and desert-steppe communities in northern China. The Holocene, 21(6), 997-1010.

Li, Y., Nielsen, A. B., Zhao, X., Shan, L., Wang, S., Wu, J., Zhou, L., 2015. Pollen production estimates (PPEs) and fall speeds for major tree taxa and relevant source areas of pollen (RSAP) in Changbai Mountain, northeastern China. Review of Palaeobotany and Palynology, 216, 92-100.

Li, Y., Xu, Q., Liu, J., Yang, X., Nakagawa, T., 2007. A transfer-function model developed from an extensive surface-pollen data set in northern China and its potential for palaeoclimate reconstructions. The Holocene, 17(7), 897-905.

Lin, S.Y., Ding, Y.L., 2007. Studies on the floral biological characteristics of three bamboo species of *Phyllostachys*. Science China Technological Sciences, 21(5), 52-55.

Lin, S.Y., Ding, Y.L., Zhang, H., 2008. Pollen Germination percentage and the floral character of five bamboo species. Scientia Silvae Sinicae, 44(10), 159-164.

Lisitsyna, O.V., Hicks, S., Huusko, A., 2012. Do moss samples, pollen traps and modern lake sediments all collect pollen in the same way? A comparison from the forest limit area of northernmost Europe. Vegetation history and Archaeobotany, 21(3), 187-199.

Liu, Y., Sun, J., Song, H., Cai, Q., Bao, G., Li, X., 2010. Tree-ring hydrologic reconstructions for the Heihe River watershed, western China since AD 1430. Water research, 44(9), 2781-2792.

Lowe, J.J., Walker, M.J., 2014. Reconstructing quaternary environments. 3rd Edition, Routledge. London.

Lu, H., Wu, N., Liu, K.B., Zhu, L., Yang, X., Yao, T., Wang, L., Li, Q., Liu, X., Shen, C., Li, X., 2011. Modern pollen distributions in Qinghai-Tibetan Plateau and the development of transfer functions for reconstructing Holocene environmental changes. Quaternary Science Reviews, 30(7-8), 947-966.

Lüdi, W., Rübel, E., Vareschi, V., 1936. die Verbreitung, das Blühen und der Pollenniederschlag der Heufieberpflanzen im Hochtale von Davos. Geobotan. Forschungsinst.,47-111.

Magurran, A.E., 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey, 1-5.

Maher, L.J., 1963. Pollen analyses of surface materials from the southern San Juan Mountains, Colorado. Geological Society of America Bulletin, 74(12), 1485-1503.

Maher, L.J., 1972. Absolute pollen diagram of Redrock Lake, Boulder County, Colorado. Quaternary Research, 2(4), 531-553.

Mann, M.E., 2002. Medieval Climatic Optimum. Encyclopedia of global environmental change, 1, 514-516.

Martínková, J., Smilauer, P., Mihulka, S., 2002. Phenological pattern of grassland species: relation to the ecological and morphological traits. Flora-Morphology, Distribution, Functional Ecology of Plants, 197(4), 290-302.

Matthias, I., Nielsen, A.B., Giesecke, T., 2012. Evaluating the effect of flowering age and forest structure on pollen productivity estimates. Vegetation history and archaeobotany, 21(6), 471-484.

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., Gaillard, M.J., Kuneš, P., Sugita, S., Trondman, A.K., Broström, A., 2012. Testing the effect of site selection and parameter setting on REVEALS-model

estimates of plant abundance using the Czech Quaternary Palynological Database. Review of Palaeobotany and Palynology, 187, 38-49.

McGarigal, K., Marks, B. J., 1994. FRAGSTATS: spatial analysis program for quantifying landscape structure. Unpublished report, Oregon State University.

Middleton, R., Bunting, M.J., 2004. Mosaic v1.1: landscape scenario creation software for simulation of pollen dispersal and deposition. Review of Palaeobotany and Palynology, 132(1-2), 61-66.

Mildenhall, D.C., 2006. *Hypericum* pollen determines the presence of burglars at the scene of a crime: an example of forensic palynology. Forensic Science International, 163(3), 231-235.

Mitchell, F.J., 2005. How open were European primeval forests? Hypothesis testing using palaeoecological data. Journal of Ecology, 93(1), 168-177.

Moore, P. D., Webb, J. A., Collison, M. E., 1991. Pollen analysis.2nd edition, Blackwell scientific publications.

Mu et al., in prep. This paper is in preparation - reference to be added depending on its progress.

Muller, S.D., Richard, P.J., Larouchel, A.C., 2003. Holocene development of a peatland (southern Québec): a spatio-temporal reconstruction based on pachymetry, sedimentology, microfossils and macrofossils. The Holocene, 13(5), 649-664.

Nagendra, H., 2002. Opposite trends in response for the Shannon and Simpson indices of landscape diversity. Applied Geography, 22(2), 175-186.

Nair, P.K.K., 1966. Trends in the morphological evolution of pollen and spores. Journal of the Indian Botanical Society, 44, 468-78.

Nakagawa, T., Tarasov, P.E., Nishida, K., Gotanda, K., Yasuda, Y., 2002. Quantitative pollen-based climate reconstruction in central Japan: application to surface and Late Quaternary spectra. Quaternary Science Reviews, 21(18-19), 2099-2113.

Nicholson, B. J., Swinehart, J. B., 2005. Evidence of Holocene climate change in a Nebraska Sandhills wetland. Great Plains Research, 45-67.

Nielsen, A. B., 2005. Quantifying the relationship between pollen sedimentation in lakes and land cover using historical maps. Geological Survey of Denmark and Greenland Bulletin, 7, 49-52.

Nielsen, A. B., Sugita, S., 2005. Estimating relevant source area of pollen for small Danish lakes around AD 1800. The Holocene, 15(7), 1006-1020.

Nielsen, A.B., Odgaard, B.V., 2005. Reconstructing land cover from pollen assemblages from small lakes in Denmark. Review of Palaeobotany and Palynology, 133(1-2), 1-21.

Normile, D., 1997. Yangtze seen as earliest rice site. Science, 275(5298), 309-309.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., 2001. Terrestrial Ecoregions of the World: A New Map of Life on EarthA new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience, 51(11), 933-938.

Overpeck, J.T., Webb, T. III, Prentice, I.C., 1985. Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. Quaternary Research, 23(1), 87-108.

Ozesmi, S.L., Bauer, M.E., 2002. Satellite remote sensing of wetlands. Wetlands ecology and management, 10(5), 381-402.

Palik B, Engstrom RT, 1999. Species composition. Maintaining Biodiversity in Forest Ecosystems, Cambridge University Press, Cambridge, UK, 65-94.

Pardoe, H.S., Giesecke, T., van der Knaap, W.O., Svitavská-Svobodová, H., Kvavadze, E.V., Panajiotidis, S., Gerasimidis, A., Pidek, I.A., Zimny, M., Święta-Musznicka, J., Latałowa, M., 2010. Comparing pollen spectra from modified Tauber traps and moss samples: examples from a selection of woodlands across Europe. Vegetation History and Archaeobotany, 19(4), 271-283.

Parsons, R. W., Prentice, I. C., 1981. Statistical approaches to R-values and the pollen-vegetation relationship. Review of Palaeobotany and Palynology, 32(2), 127-152.

Pawłowski, D., Kloss, M., Obremska, M., Szymanowski, M., Żurek, S., 2012. Evolution of small valley mire in central Poland as a result of hydroclimatic oscillations. Geochronometria, 39(2), 133-148.

Peck, R.M., 1973. Pollen budget studies in a small Yorkshire catchment. Quaternary Plant Ecology, 43-60.

Peros, M.C., Gajewski, K., Viau, A.E., 2008. Continental‐scale tree population response to rapid climate change, competition and disturbance. Global Ecology and Biogeography, 17(5), 658-669.

Pickett, E.J., Harrison, S.P., Hope, G., Harle, K., Dodson, J.R., Kershaw, A.P., Prentice, I.C., Backhouse, J., Colhoun, E.A., D'Costa, D., Flenley, J., 2004. Pollen‐based reconstructions of biome distributions for Australia, Southeast Asia and the Pacific (SEAPAC region) at 0, 6000 and 18,000 ^{14}C yr BP. Journal of Biogeography, 31(9), 1381-1444.

Pohl F, 1937. Die Pollenerzeugung der Windblütler. Beihefte zum Botanischen Centralblatt 56, 365-470.

Poska, A., Meltsov, V., Sugita, S., Vassiljev, J., 2011. Relative pollen productivity estimates of major anemophilous taxa and relevant source area of pollen in a cultural landscape of the hemi-boreal forest zone (Estonia). Review of Palaeobotany and Palynology, 167(1-2), 30-39.

Poska, A., Sepp, E., Veski, S., Koppel, K., 2008. Using quantitative pollen-based land-cover estimations and a spatial CA_Markov model to reconstruct the development of cultural landscape at Rouge, South Estonia. Vegetation History and Archaeobotany, 17(5), 527-541.

Potter, L.D., 1967. Differential Pollen Accumulation in Water‐Tank Sediments and Adjacent Soils. Ecology, 48(6), 1041-1043.

Prentice, C., 1988. Records of vegetation in time and space: the principles of pollen analysis. Vegetation history, 17-42.

Prentice, C., Guiot, J., Huntley, B., Jolly, D., Cheddadi, R., 1996. Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka. Climate Dynamics, 12(3), 185-194.

Prentice, I. C., Parsons, R. W, 1983. Maximum likelihood linear calibration of pollen spectra in terms of forest composition. Biometrics, 1051-1057.

Prentice, I. C., Webb, T. III, 1986. Pollen percentages, tree abundances and the Fagerlind effect. Journal of Quaternary Science, 1(1), 35-43.

Prentice, I.C. 1986. Forest-composition calibration of pollen data. In: Berglund, B.E. (ed.), Handbook of Holocene Palaeoecology and Palaeohydrology. John Wiley and Sons, Chichester, 799-816.

Prentice, I.C., 1985. Pollen Representation, Source Area, and Basin Size: Toward a Unified Theory of Pollen Analysis. Quaternary Research, 23, 76-86.

Prentice, I.C., Berglund, B.E., Olsson, T., 1987. Quantitative forest‐composition sensing characteristics of pollen samples from Swedish lakes. Boreas, 16(1), 43-54.

Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A., Solomon, A.M., 1992. A global biome model based on plant physiology and dominance, soil properties and climate. Journal of biogeography, 117-134.

Prentice, I.C., Jolly, D., Biome 6000 Participants, 2000. Mid‐Holocene and glacial‐maximum vegetation geography of the northern continents and Africa. Journal of biogeography, 27(3), 507-519.

Punt, W., Blackmoie, S., Clarke, G.C.S., Hoen, P.P., 1976-1995, The Northwest European pollen flora, 1 (1976), 2 (1980), 3 (1981),4 (1984), 5 (1988), 6 (1991), 7 (1995). Amsterdam: Elsevier Science.

Quamar, M.F., Nautiyal, C.M., 2017. Mid-Holocene pollen records from southwestern Madhya Pradesh, central India, and their palaeoclimatic significance. Palynology, 41(3), 401-411.

Räsänen, S., Hicks, S., Odgaard, B.V., 2004. Pollen deposition in mosses and in a modified ‘Tauber trap’ from Hailuoto, Finland: what exactly do the mosses record?. Review of Palaeobotany and Palynology, 129(1-2), 103-116.

Reille, M., 1992-1998. Pollen et spores d'Europe et d'Afrique du nord: Supplement 1 and 2. Laboratoire de Botanique hostorique et Palynologie: Marseille.

Richards, J.A., 2013. Supervised classification techniques. Remote Sensing Digital Image Analysis, 247-318.

Riding, J.B., Rawlins, B.G., Coley, K.H., 2007. Changes in soil pollen assemblages on footwear worn at different sites. Palynology, 31(1), 135-151.

Ritchie, J.C., Eyles, C.H., Haynes, C.V., 1985. Sediment and pollen evidence for an early to mid-Holocene humid period in the eastern Sahara. Nature, 314(6009), 352-355.

Saito, H., Takeoka, M., 1985. Pollen production rates in a young Japanese red pine forest. Japanese Journal of Ecology, 35(1), 67-76.

Sangster, A.G., Dale, H.M., 1961. A preliminary study of differential pollen grain preservation. Canadian Journal of Botany, 39(1), 35-43.

Sangster, A.G., Dale, H.M., 1964. Pollen grain preservation of underrepresented species in fossil spectra. Canadian Journal of Botany, 42(4), 437-449.

Sentinel-2 data, European Space Agency (ESA), http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2, https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/sentinel-2, https://sentinel.esa.int/web/sentinel/sentinel-data-access.

Seppä, H., Birks, H.J.B., Odland, A., Poska, A., Veski, S., 2004. A modern pollen–climate calibration set from northern Europe: developing and testing a tool for palaeoclimatological reconstructions. Journal of Biogeography, 31(2), 251-267.

Shannon, C.E., Weaver, W., 1949. The mathematical theory of communication. Illinois: University of Illinois Press.

Shen, C., Liu, K.B., Tang, L., Overpeck, J.T., 2006. Quantitative relationships between modern pollen rain and climate in the Tibetan Plateau. Review of Palaeobotany and Palynology, 140(1-2), 61-77.

Shen, J., Liu, X.Q., Wang, S.M., Ryo, M., 2005. Palaeoclimatic changes in the Qinghai Lake area during the last 18,000 years. Quaternary International, 136(1), 131-140.

Shu, Q., Xiao, J.Y., Zhang, M.X., Zhao, Z.J., Chen, Y., Li, J.J., 2008. Climate change in northern Jiangsu Basin since the last interglacial. Geological Science and Technology Information 27, 59-64.

Simpson, E.H, 1949. Measurement of Diversity. Nature, 163, 688.

Sjögren, P., Connor, S.E., van der Knaap, W.O., 2010. The development of composite dispersal functions for estimating absolute pollen productivity in the Swiss Alps. Vegetation History and Archaeobotany, 19(4), 341-349.

Sjögren, P., van der Knaap, W.O., Huusko, A., van Leeuwen, J.F., 2008a. Pollen productivity, dispersal, and correction factors for major tree taxa in the Swiss Alps based on pollen-trap results. Review of Palaeobotany and Palynology, 152(3-4), 200-210.

Sjögren, P., van der Knaap, W.O., Kaplan, J.O., van Leeuwen, J.F., Ammann, B., 2008b. A pilot study on pollen representation of mountain valley vegetation in the central Alps. Review of Palaeobotany and Palynology, 149(3-4), 208-218.

Sjögren, P., van der Knapp, W.O., Van Leeuwen, J.F., 2018. A practical approximation of pollen dispersal–deposition for calculation of pollen productivity and quantification of vegetation. Grana, 1-7.

Soepboer, W., Sugita, S., Lotter, A.F., 2010. Regional vegetation-cover changes on the Swiss Plateau during the past two millennia: a pollen-based reconstruction using the REVEALS model. Quaternary Science Reviews, 29(3-4), 472-483.

Soepboer, W., Sugita, S., Lotter, A.F., van Leeuwen, J.F., van der Knaap, W.O., 2007. Pollen productivity estimates for quantitative reconstruction of vegetation cover on the Swiss Plateau. The Holocene, 17(1), 65-77.

Southall, E.J., Dale, M.P., Kent, M., 2003. Spatial and temporal analysis of vegetation mosaics for conservation: poor fen communities in a Cornish valley mire. Journal of Biogeography, 30(9), 1427-1443.

Spieksma, F.T.M., Nikkels, B.H., Bottema, S., 1994. Relationship between recent pollen deposition and airborne pollen concentration. Review of palaeobotany and palynology, 82(1-2), 141-145.

Stanley, R.G., Linskens, H.F., 2012. Pollen: biology biochemistry management. Springer Science & Business Media

Stow, D.A., Hope, A., McGuire, D., Verbyla, D., Gamon, J., Huemmrich, F., Houston, S., Racine, C., Sturm, M., Tape, K., Hinzman, L., 2004. Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. Remote sensing of environment, 89(3), 281-308.

Strong, J., 1967. Ecology of terrestrial arthropods at Palmer station, Antarctic Peninsula. Antarctic Research Series, 10, 357-371.

Sugita, S., 1993. A model of pollen source area for an entire lake surface. Quaternary research, 39, 239-244.

Sugita, S., 1994. Pollen representation of vegetation in Quaternary sediments: theory and method in patchy vegetation. Journal of Ecology, 881-897.

Sugita, S., 2007a. Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. The Holocene, 17(2), 229-241

Sugita, S., 2007b. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. The Holocene, 17(2), 243-257.

Sugita, S., Gaillard, M.J., Broström, A, 1999. Landscape openness and pollen records: a simulation approach. The Holocene, 9(4), 409-421.

Sugita, S., Parshall, T., Calcote, R., Walker, K., 2010. Testing the Landscape Reconstruction Algorithm for spatially explicit reconstruction of vegetation in northern Michigan and Wisconsin. Quaternary Research, 74(2), 289-300.

Sun S.Y., 1980. The distribution and cause of formation of peatland in Xishan Mountain, Nanchang, southeast China. Journal of Nanjing Normal University (engineering and technology edition), 2, 43- 51.

Sutton, O.G., 1947a. The problem of diffusion in the lower atmosphere. Quarterly Journal of the Royal Meteorological Society, 73, 257-281.

Sutton, O.G., 1947b. The theoretical distribution of airborne pollution from factory chimneys. Quarterly Journal of the Royal Meteorological Society, 73, 426-436.

Sutton, O.G., 1953. Micrometeorology: a study of physical processes in the lowest layers of the Earth's atmosphere. New York: McGraw-Hill Book Conpany, Inc.

Tang, L. Y., Mao, L. M., Shu, J. W., Li, C. H., Shen, C. M., Zhou, Z. Z., 2016. An Illustrated Handbook of Quaternary Pollen and Spores of China. Beijing: China Scientific Book Services.

Tauber, H., 1965. Differential pollen dispersal and the interpretation of pollen diagrams, with a contribution to the interpretation of the elm fall. Danmarks Geologiske Undersogelse, 2(89), 1-69.

Tauber, H., 1967. Investigations of the mode of pollen transfer in forested areas. Review of Palaeobotany and Palynology, 3(1), 277-286.

Tauber, H., 1974. A static non‐overload pollen collector. New Phytologist, 73(2), 359-369.

Ter Braak, C.J., 1994. Canonical community ordination. Part I: Basic theory and linear methods. Ecoscience, 1(2), 127-140.

Ter Braak, C.J., Prentice, I.C., 1988. A theory of gradient analysis. Advances in ecological research, 18, 271-317.

Ter Braak, C.J., Verdonschot, P.F., 1995. Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquatic sciences, 57(3), 255-289.

Ter Braak, C.J.F., Smilauer, P., 1997. Canocofor Windows. Centre for Biometry Wageningen, CPRO-DLO, Wage-ningen, NL.

Ter Braak, C.J.F., Smilauer, P., 2002. Canoco for Windows version 4.5. Biometris-Plant Research International, Wageningen.

Theuerkauf, M., Couwenberg, J., Kuparinen, A., Liebscher, V., 2016. A matter of dispersal: REVEALSinR introduces state-of-the-art dispersal models to quantitative vegetation reconstruction. Vegetation history and archaeobotany, 25(6), 541-553.

Theuerkauf, M., Kuparinen, A., Joosten, H., 2013. Pollen productivity estimates strongly depend on assumed pollen dispersal. The Holocene, 23(1), 14-24.

Tinsley, H.M., Smith, R.T., 1974. Surface pollen studies across a woodland/heath transition and their application to the interpretation of pollen diagrams. New Phytologist, 73(3), 547-565.

Trondman, A.K., Gaillard, M.J., Mazier, F., Sugita, S., Fyfe, R., Nielsen, A.B., Twiddle, C., Barratt, P., Birks, H.J.B., Bjune, A.E., Björkman, L., Broström, A., Caseldine, C., David, R., Dodson, J., Dörfler, W., Fischer, E., van Geel, B., Giesecke, T., Hultberg, T., Kalnina, L., Kangur, M., van der Knaap, P., Koff, T., Kuneš, P., Lagerås, P., Latałowa, M., Lechterbeck, J., Leroyer, C., Leydet, M., Lindbladh, M., Marquer, L., Mitchell, F.J.G., Odgaard, B.V., Peglar, S.M., Persson, T., Poska, A., Rösch, M., Seppä, H., Veski, S. Wick, L, 2015. Pollen-based quantitative reconstructions of Holocene regional vegetation cover (plant-functional types and land-cover types) in Europe suitable for climate modelling. Global Change Biology, 21(2), 676-697.

Turner, J., 1964. Surface Sample Analyses from Ayrshire, Scottland. Pollen et spores, 6, 583-592.

Twiddle, C.L., Jones, R.T., Caseldine, C.J., Sugita, S, 2012. Pollen productivity estimates for a pine woodland in eastern Scotland: the influence of sampling design and vegetation patterning. Review of palaeobotany and palynology, 174, 67-78.

van Post, L., 1916. Om skogstradspollen i sydsvenska torfmosselagerfoljder (foredragsreferat). Geologisk Forening, Stockholm, 38: 384-394.

Viau, A., Gajewski, K., Sawada, M., Fines, P., 2006. Mean-continental July temperature variability in North America during the past 14,000 years. Journal of Geophysical Research:Atmospheres, 111, D09102.

Walch, K.M., Rowley, J.R., Norton, N.J., 1970. Displacement of pollen grains by earth-worms. Pollen et spores, 12, 39-44.

Waller, M., Grant, M.J., Bunting, M.J., 2012. Modern pollen studies from coppiced woodlands and their implications for the detection of woodland management in Holocene pollen records. Review of Palaeobotany and Palynology, 187, 11-28.

Waller, M.P., Hamilton, S., 2000. Vegetation history of the English chalklands: a mid‐Holocene pollen sequence from the Caburn, East Sussex. Journal of Quaternary Science, 15(3), 253-272.

Wang, F.X., Qian, N. F., Zhang Y. L., 1995. Pollen flora of China. Science Press, Beijing.

Wang, K.F., 1974. Palynological analysis peat-bogs of siyao lake Mount Hsishan, Nanchang, Kangxi Province. Atca Botanica Sinica, 16(1), 83-93.

Wang, Y., Herzschuh, U., 2011. Reassessment of Holocene vegetation change on the upper Tibetan Plateau using the pollen-based REVEALS model. Review of Palaeobotany and Palynology, 168(1), 31-40.

Wang, Y., Liu, X., Herzschuh, U., 2010. Asynchronous evolution of the Indian and East Asian Summer Monsoon indicated by Holocene moisture patterns in monsoonal central Asia. Earth Science Reviews, 103(3-4), 135-153.

Webb, T. Ⅲ, Howe, S.E., Bradshaw, R.H.W., Heide, K.M., 1981. Estimating plant abundances from pollen percentages: the use of regression analysis. Review of Palaeobotany and Palynology, 34(3-4), 269-300.

Wen, R., Xiao, J., Chang, Z., Zhai, D., Xu, Q., Li, Y., Itoh, S., 2010. Holocene precipitation and temperature variations in the East Asian monsoonal margin from pollen data from Hulun Lake in northeastern Inner Mongolia, China. Boreas, 39(2), 262-272.

West, R.G., 1973. Introduction. In: Birks, H.J.B., West, R.G. (Eds.), Quaternary Plant Ecology. Blackwell, Oxford, 1-3.

Wilmshurst, J.M., McGlone, M.S., 2005. Origin of pollen and spores in surface lake sediments: comparison of modern palynomorph assemblages in moss cushions, surface soils and surface lake sediments. Review of Palaeobotany and Palynology, 136(1-2), 1-15.

Wiltshire, P.E.J., 2004. Current applications of environmental profiling and forensic palynology in the United Kingdom. Program, Challenges & Changes 17th International Symposium on the Forensic Sciences. The Australian and New Zealand Forensic Science Society, Wellington, 28, 202.

Winkler, M.G., Wang, P.K., 1993. The late‐Quaternary vegetation and climate of China. Global climates since the last glacial maximum (ed. by H. E. Wright Jr, J. E. Kutzbach, T. Webb III, W. F. Ruddiman, F. A. Street‐Perrott and P. J. Bartlein), University of Minnesota Press, Minneapolis, 221-264.

Wu, J., Ma, Y. Z., Sang, Y.L., Meng, H.W., Hu, C.L., 2013. Quantitative reconstruction of palaeovegetation and development of the R-value model: an application of R-value and ERV model in Xinglong Mountain natural protection region. Quaternary Sciences, 33, 554-564.

Xiao, J., Lu, H., Zhou, W., Zhao, Z., Hao, R., 2007. Evolution of vegetation and climate since the last glacial maximum recorded at Dahu peat site, South China. Science in China Series D: Earth Sciences, 50(8), 1209-1217.

Xu, Q., Cao, X., Tian, F., Zhang, S., Li, Y., Li, M., Li, J., Liu, Y., Liang, J., 2014. Relative pollen productivities of typical steppe species in northern China and their potential in past vegetation reconstruction. Science China Earth Sciences, 57(6), 1254-1266.

Xu, Q., Chen, F., Zhang, S., Cao, X., Li, J., Li, Y., Li, M., Chen, J., Liu, J., Wang, Z., 2017. Vegetation succession and East Asian Summer Monsoon Changes since the last

deglaciation inferred from high-resolution pollen record in Gonghai Lake, Shanxi Province, China. The Holocene, 27(6), 835-846.

Xu, Q., Li, Y., Yang, X., Zheng, Z., 2007. Quantitative relationship between pollen and vegetation in northern China. Science in China Series D: Earth Sciences, 50(4), 582-599.

Xu, Q., Zhang, S., Gaillard, M.J., Li, M., Cao, X., Tian, F., Li, F., 2016. Studies of modern pollen assemblages for pollen dispersal-deposition-preservation process understanding and for pollen-based reconstructions of past vegetation, climate, and human impact: A review based on case studies in China. Quaternary Science Reviews, 149, 151-166.

Yan, J. H., Ge, Q. S., Zheng, J. Y., 2012. Reconstruction and analysis on the series of winter-half-year temperature change during the Qing Dynasty in the Northern China Region. Progress in Geography, 31(11), 1426-1432.

Yang, Y.D., Man, Z.M., Zheng, J.Y., 2006. Reconstruction of series in later or earlier starting date of rainy season in Yunnan Province and evolvement of summer monsoon in Qing period. Acta Geographica Sinica, 61(7), 705-712.

Yao, T.D., Jiao, K.Q., Tian, L.D., Thompson, L.G., 1996. Climatic variations since the Little Ice Age recorded in the Guliya Ice Core. Science in China Series D: Earth Sciences, 39(6), 587-596.

Yao, T.D., Shi, Y.F., Thompson, L.G., 1997. High resolution record of paleoclimate since the Little Ice Age from the Tibetan ice cores. Quaternary International, 37, 19-23.

Yu, G., Chen, X., Ni, J., Cheddadi, R., Guiot, J., Han, H., Harrison, S.P., Huang, C., Ke, M., Kong, Z., Li, S., 2000. Palaeovegetation of China: a pollen data‐based synthesis for the mid‐Holocene and last glacial maximum. Journal of Biogeography, 27(3), 635-664.

Yuan X.L., Lin X.C., Lin X., Fang W., 2007. Advances in the studies of bamboo flowering. Journal of bamboo research, 26(1), 6-14.

Zeng, Y., Chen, J., Zhu, Z., Li, J., Wang, J., Wan, G., 2012. The wet Little Ice Age recorded by sediments in Huguangyan Lake, tropical south China. Quaternary International, 263, 55-62.

Zhai, P., Zhang, X., Wan, H., Pan, X., 2005. Trends in total precipitation and frequency of daily precipitation extremes over China. Journal of climate, 18(7), 1096-1108.

Zhang, Y., Sun, L. F., Ran, H., Feng, Y., Zhang, Y., Guo, Q. R., 2016. Pollen morphology and double fertilization of Phyllostachys edulis. Guihaia, 11, 1325-1329.

Zhao, B.C., Wang, Z.H., Li, X., 2007. Characteristics and palaeogeographic significance of paleo-incised valley sediments in southern Changjiang Delta Plain. Journal of Palaeogeography， 9, 217-226.

Zhao, L., Ma, C., Leipe, C., Long, T., Liu, K.B., Lu, H., Tang, L., Zhang, Y., Wagner, M., Tarasov, P.E., 2017. Holocene vegetation dynamics in response to climate change and human activities derived from pollen and charcoal records from southeastern China. Palaeogeography, Palaeoclimatology, Palaeoecology, 485, 644-660.

Zhao, L., Ma, C., Tang, L., Liu, K.B., Mao, L., Zhang, Y., Lu, H., Wu, S., Tu, Q., 2016. Investigation of peat sediments from Daiyun Mountain in southeast China: late Holocene vegetation, climate and human impact. Vegetation History and Archaeobotany, 25(4), 359-373.

Zhao, Y., 2018. Vegetation and climate reconstructions on different time scales in China: a review of Chinese palynological research. Vegetation History and Archaeobotany, 1-12.

Zhao, Y., Xu, Q., Huang, X., Guo, X., Tao, S., 2009c. Differences of modern pollen assemblages from lake sediments and surface soils in arid and semi-arid China and their significance for pollen-based quantitative climate reconstruction. Review of Palaeobotany and Palynology, 156(3-4), 519-524.

Zhao, Y., Yu, Z., Chen, F., 2009a. Spatial and temporal patterns of Holocene vegetation and climate changes in arid and semi-arid China. Quaternary International, 194(1-2), 6-18.

Zhao, Y., Yu, Z., Chen, F., Zhang, J., Yang, B., 2009b. Vegetation response to Holocene climate change in monsoon-influenced region of China. Earth-Science Reviews, 97(1-4), 242-256.

Zhao, Y., Yu, Z., Zhao, W., 2011. Holocene vegetation and climate histories in the eastern Tibetan Plateau: controls by insolation-driven temperature or monsoon-derived precipitation changes?. Quaternary Science Reviews, 30(9-10), 1173-1184.

Zheng, Z., 1990. Holocene pollen flora and paleoenvironment in Chaoshan Plain. Tropical Oceanography 9, 31-38.

Zheng, Z., 1991. Pollen flora and paleoclimate of the Chao-Shan Plains during the last 50000 years. Acta Micropalaeontologica Sinica, 8, 461-480.

Zheng, Z., Huang, K., Xu, Q., Lu, H., Cheddadi, R., Luo, Y., Beaudouin, C., Luo, C., Zheng, Y., Li, C., Wei, J., 2008. Comparison of climatic threshold of geographical distribution between dominant plants and surface pollen in China. Science in China Series D: Earth Sciences, 51(8), 1107-1120.

Zheng, Z., Wei, J., Huang, K., Xu, Q., Lu, H., Tarasov, P., Luo, C., Beaudouin, C., Deng, Y., Pan, A., Zheng, Y., 2014. East Asian pollen database: modern pollen distribution and its quantitative relationship with vegetation and climate. Journal of biogeography, 41(10), 1819-1832.

Zhou, W.J., Yu, X.F., Jull, A.J.T., Burr, G., Xiao, J.Y., Lu, X.F., Xian, F., 2004. High-resolution evidence from southern China of an early Holocene optimum and a mid-Holocene dry event during the past 18,000 years. Quaternary Research, 62, 39-48.

Zhu, C., Ma, C., Yu, S.Y., Tang, L., Zhang, W., Lu, X., 2010. A detailed pollen record of vegetation and climate changes in Central China during the past 16 000 years. Boreas, 39(1), 69-76.

Zhu, C., Ma, C.M., Zhang, W.Q., Zheng, C.G., Tang, L.Y., Lu, X.F., Liu, K.X., Chen, H.Z., 2006. Pollen record from Dajiuhu Basin of Shennongjia and environmental changes since 15.753 kaB. P. Quaternary Sciences, 26(5), 814-826.

Zhu, H. F., Shao, X. M., Yin, Z. Y., Xu, P., Xu, Y., Tian, H., 2011. August temperature variability in the southeastern Tibetan Plateau since AD 1385 inferred from tree rings. Palaeogeography, Palaeoclimatology, Palaeoecology, 305(1), 84-92.

Zhu, K.Z., 1972. A preliminary study of nearly 5000 years Chinese to climate change. Science in China: Series. A, 4(2), 168-189.

Zong, Y., Chen, Z., Innes, J.B., Chen, C., Wang, Z., Wang, H., 2007. Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China. Nature, 449(7161), 459.