Modern pollen assemblages as climate indicators in southern Europe

Walter Finsinger[1,2]*, Oliver Heiri[1], Verushka Valsecchi[2], Willy Tinner[2] and André F. Lotter[1]

[1]*Palaeoecology, Institute of Environmental Biology, Utrecht University, Laboratory of Palaeobotany and Palynology, Budapestlaan 4, NL-3584 CD Utrecht, The Netherlands and* [2]*Section of Palaeoecology, Institute of Plant Sciences, University of Bern, Altenbergrain 21, CH-3013 Bern, Switzerland*

ABSTRACT

Aim and Location Our aim is to develop pollen–climate inference models for southern Europe and to test their performance and inference power by cross-validation with modern climate data. Surface sediments collected from lakes along a climate gradient from the winter-cold/summer-wet Alps to winter-wet/summer-dry Sicily were analysed for modern pollen assemblages.

Methods For each lake, mean monthly temperatures, seasonal precipitation and site-specific climate uncertainties have been estimated. Pollen–climate relationships were studied using numerical analyses, and inference models were derived by partial least squares (PLS) and weighted-averaging PLS (WA-PLS) regressions for January and July temperatures (T), and for winter, spring and summer precipitation (P). In order to assess whether these variables are also of ecological importance for vegetation in the subregions, we split the data set into an Alpine and a Mediterranean subset.

Results Low bootstrap cross-validated root mean square errors of prediction (RMSEP) for January T (1.7 °C), July T (2.1 °C) and summer P (38 mm), as well as low RMSEPs expressed as a percentage of the gradient length (8–9%), indicate a good inference power. Models revealed excellent to good performance statistics for January T, July T and summer P ($r^2 = 0.8$), and for winter and spring P ($r^2 = c.$ 0.5). We show that the variables with the highest explanatory power differ between the two subregions. These are summer T and P for the Alpine set, and January T, winter P and July T for the Mediterranean set.

Main conclusions The study reveals the influence of climatic conditions during the growing season on modern pollen assemblages and indicates the potential of pollen data for long-term climate reconstructions of parameters such as winter precipitation and temperature, which seem to be the main factors having an influence on the variability of Mediterranean climate. These models may therefore provide important information on past regional climate variability in southern Europe.

Keywords
Climate, Italy, pollen, transfer function, validation, vegetation.

*Correspondence: Walter Finsinger, Palaeoecology, Institute of Environmental Biology, Utrecht University, Laboratory of Palaeobotany and Palynology, Budapestlaan 4, NL-3584 CD Utrecht, The Netherlands. E-mail: w.finsinger@bio.uu.nl

INTRODUCTION

An understanding of climate change in the Mediterranean region is highly important with regard to present and future environmental and socio-economic issues. Yet our ability to predict the effects of the present global warming trend on the Mediterranean climate is hampered by a lack of long-term information on climate variability (Luterbacher *et al.*, 2006). Since instrumental meteorological records generally cover time-scales that are too short (<250 years) to reconstruct natural centennial-scale

climatic variability and to provide estimates of the likely magnitude and rate of climatic change, documentary sources and palaeoclimatic proxy data are needed to fill this gap (e.g. Luterbacher *et al.*, 2004).

Amongst the most abundant microfossils preserved in sediment archives are pollen grains, whose sedimentary assemblages are related to regional and local vegetation (Sugita, 1994). Hence, if vegetation assemblages are related to climate, sedimentary pollen assemblages may be used for reconstructing past climate variability that affected vegetation changes (e.g. Guiot, 1990;

Lotter *et al.*, 2000; Davis *et al.*, 2003; Peyron *et al.*, 2005). The success of pollen in quantifying past climate is strongly dependent on high-quality modern calibration data for establishing the relationship between vegetation, pollen assemblages and climatic parameters (e.g. Birks, 2004). Modern samples used for pollen–climate calibration currently have a low spatial resolution on the Italian Peninsula, i.e. < 20 samples (e.g. Guiot, 1990; Prentice *et al.*, 2000; Davis *et al.*, 2003). These samples were generally used in the context of continent-wide, broad-scale pollen calibration sets (e.g. Peyron *et al.*, 1998; Prentice *et al.*, 2000; Davis *et al.*, 2003). The use of such geographically widespread calibration data sets is associated with a number of problems. They are commonly analysed by a number of different analysts, and taxonomic consistency in the identifications may become an issue when quantitative palaeoenvironmental reconstructions are implemented (Birks, 1994). Moreover, due to the low taxonomic resolution that can be achieved for some pollen types (e.g. *Pinus* or Gramineae pollen), these data sets can often feature multiple modern analogues for past pollen assemblage composition (Birks, 1998). These multiple analogues, however, will originate from areas with different climatic conditions, complicating the use of calibration techniques. To avoid these problems, the use of regional calibration data sets has been suggested (Birks, 1998). For pollen-based palaeoenvironmental reconstruction in Holocene sediments it can be expected that such regional data sets provide reasonable analogues for most of the fossil pollen assemblages without the problems associated with multi-analyst, continent-wide data sets, since they are often analysed by one or a few analysts using consistent nomenclature.

The aim of the present study was to create a modern pollen–climate calibration data set for the Italian Peninsula at a high spatial resolution and covering the full range of modern climatic conditions in this region. We developed and validated pollen–climate inference models based on numerical techniques, such as partial least squares (PLS) and weighted-averaging PLS (WA-PLS) regression, that have previously been used successfully for developing pollen transfer functions in Fennoscandia (e.g. Birks, 2004; Seppä *et al.*, 2004) and Central Europe (e.g. Lotter *et al.*, 2000). These methods perform well in 'no analogue' situations, they operate consistently well with a range of data and do not involve an excessive number of parameters to be estimated and fitted, and are therefore relatively statistically robust and computationally economical (Birks, 2004). We avoided the modern analogue technique (MAT) (e.g. Guiot, 1990), because: (1) estimates of the predictive power of MAT models for pollen–climate inference models may be over-optimistic and misleading (e.g. Telford & Birks, 2005); and (2) species may behave individualistically and therefore associations have very likely not been constant over time (e.g. Gleason, 1939; Jackson & Williams, 2004).

Our modern pollen–climate inference model spans large gradients of temperature and precipitation regimes from the Alps to Sicily, i.e. from a temperate alpine climate (with cold winters and warm humid summers) to a Mediterranean climate (with mild winters and hot dry summers). This implies that plants are subject to different stressors along this long climatic gradient (Box & Fujiwara, 2005). Hence, in an attempt to estimate the performances of subregional inference models and to test whether environmental factors at a subregional scale are comparable with those operative at a regional scale, we developed 'Alpine' and 'Mediterranean' pollen–climate inference models and compared these with the regional model.

METHODS

Field and laboratory analyses

Surface sediment samples were collected with a Kayak corer (Renberg, 1991) in the centre of 92 lake basins. Eighty per cent of these lakes have a lake basin area of < 1 km^2, and only 5% of lakes have a basin area > 10 km^2. The top 2 cm of sediment were sampled and stored in plastic bags prior to subsampling for pollen analysis, which was carried out by W.F. in 2005. The top 1–2 cm of sediment, indicative of 5–10 years of time, were treated with standard physical (sieving with 500 μm mesh and decanting) and chemical methods (HCL, KOH, HF, acetolysis, KOH), embedded in glycerine, and stained with basic fuchsin. At least 300 pollen grains originating from terrestrial plants were counted from each slide at 400× magnification under a light microscope. Pollen grains were identified using the following pollen keys and atlases: Punt *et al.* (1976–1996), Moore *et al.* (1991) and Reille (1992). Pollen nomenclature of *Quercus ilex*-Typ, *Quercus cerris*-Typ, *Quercus robur*-Typ, *Plantago coronopus*, *Elaeagnus*, *Liquidambar* and *Cryptomeria*-Typ follows Beug (2004).

In total, 184 pollen and spore types of terrestrial plants were identified. Percentages were calculated on the basis of the sum of trees, shrubs, herbs and terrestrial ferns, whereas aquatics, mosses and other palynomorphs were excluded from the pollen sum (> 300 terrestrial pollen). The initial screening of pollen percentages revealed the presence of abundant *Humulus/Cannabis* type pollen in 13 samples, probably reflecting either a contamination from reworked sediment or sediments older than the 20th century (Finsinger *et al.*, 2006). Hence, samples with a *Humulus/Cannabis* pollen percentage exceeding 3% were excluded from the training set. The resulting full training set (FTS) consists of 79 sites (Fig. 1).

Environmental data

Geographical coordinates (datum: WGS84; projection: NUTM32) and elevations were obtained for each site (Fig. 1 & Appendix S1 in Supplementary Material). In order to estimate climatic variables at the sites, mean monthly and annual temperature (358 stations) as well as mean monthly and annual precipitation data (115 stations) were used (Petrarca *et al.*, 1999). A series of algorithms were tested for predicting mean monthly and mean annual temperatures and precipitation values for localities on the Italian Peninsula based on the climate station data. Temperatures were corrected to sea-level elevation using regional monthly and annual lapse rates calculated from the climate station data separately for different Italian regions using linear regression (see Appendix S2 in Supplementary Material). The calculated regional lapse rates
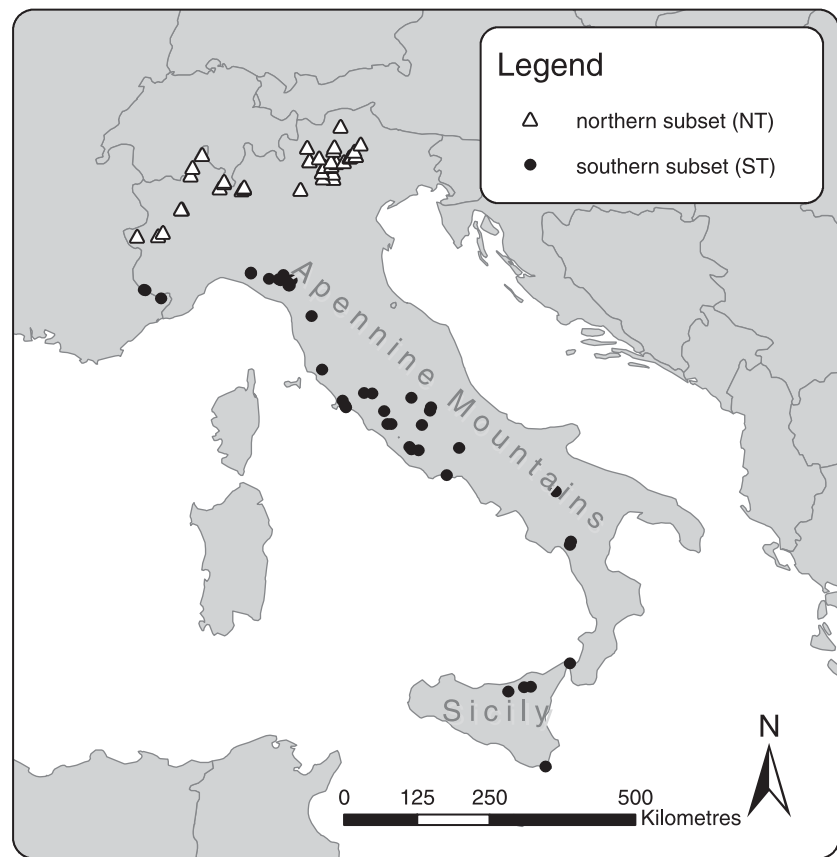
Figure 1 Map illustrating the position of the 79 sites used to develop a pollen transfer function for the Italian Peninsula. Open triangles, sites contained in the northern subset (NT); full circles, sites contained in the southern subset (ST).

show a remarkable range from −0.27 to −0.81 for monthly temperatures and a smaller range for annual temperatures (−0.47 to −0.67). In the Alpine regions (Trentino Alto Adige, Lombardia and Piemonte), summer and spring topographical temperature gradients are generally stronger than in winter or autumn as a consequence of thermal inversions during the cold season (Agustí-Panareda & Thompson, 2002). In contrast, temperature gradients are stronger in winter than in summer in Sicily, probably as a consequence of the subdued winter temperatures along the coast in winter and free circulation during summer, which tends to dampen vertical temperature gradients. In a second step, the *k*-nearest neighbour method was used to predict temperature values for the individual climate stations. For each station, the mean of the geographically closest *k* climate stations was calculated and compared with the measured temperatures at this station. The performance of this approach was assessed by calculating the root mean square error (RMSE), which approximates the standard deviation of the prediction residuals for all climate stations in the data set. RMSEs were calculated for increasing numbers of *k* nearest neighbours (starting with $k = 1$) until the performance of the method in predicting monthly and mean annual temperatures for the meteorological station data declined for increasing values of *k*. This *k*-nearest neighbour approach was repeated using a mean weighted by the inverse geographical distance of the *k*-nearest neighbours for calculating the prediction residuals. Of all these calculated nearest neighbour approaches the one with the best (i.e. smallest) RMSE was

selected individually for each environmental variable leading to a series of 13 optimal temperature prediction functions for inferring mean monthly and mean annual temperature on the Italian Peninsula (see Appendix S2). These functions were then used to estimate the 13 temperature parameters at the studied lakes. The temperatures, which were estimated for sea-surface level, were subsequently corrected to the elevation of the study sites using the regional temperature lapse rates (see Appendix S2). Mean monthly and mean annual precipitation values were calculated in the same way with the exception that the calculations did not include any corrections for elevation. For numerical analyses, the monthly precipitation values were combined into sums of spring (March, April, May), summer (June, July, August), autumn (September, October, November) and winter precipitation (December, January, February).

Numerical analyses

We used numerical analyses to select the minimal subset of environmental variables that best explains, in a statistical sense, the pollen assemblages in the training set, and to estimate the relationships between these environmental variables and individual pollen types. We subsequently derived pollen-based inference models and tested their ability to predict accurately the present-day climate at different sites along the climate gradient of the FTS (intraset cross-validation, following Bigler & Hall, 2002). Finally, we split the FTS into two smaller, climatically different,

training sets, derived pollen-inference models for these training sets, and compared these models with the models based on the FTS with regard to the influence of seasonal variables on biomes and statistical performance of the models.

In an attempt to optimize the 'signal' to 'noise' ratio, pollen percentages were square-root transformed before numerical treatment (e.g. Seppä *et al.*, 2004). To determine whether to use linear or unimodal-based numerical techniques, the FTS was first analysed by detrended correspondence analysis (DCA) to estimate the length of the compositional gradient (ter Braak & Prentice, 1988). The gradient length of the pollen assemblage in the FTS is 1.84 standard deviation (SD) units. ter Braak and Prentice (1988) indicate that linear ordination methods perform better for short gradients (< 1.5 SD units) and unimodal methods are more suitable for long gradients (> 3 SD units), whereas both methods may be appropriate for intermediate gradient lengths. With a gradient length > 1.5 SD units, we have chosen to use unimodal-based numerical methods, namely canonical correspondence analysis (CCA) and detrended CCA (DCCA) as implemented in CANOCO version 4.5 (ter Braak & Šmilauer, 2003).

In order to select a subset of environmental variables, the explanatory power of each of the 20 variables (latitude, longitude, altitude, average monthly temperatures and the sum of seasonal precipitation) was first assessed individually for the FTS, using a series of CCA ordinations in which each environmental variable was the sole constraining parameter. The percentage of variance in the pollen data explained by each variable was calculated, and the statistical significance of each model was assessed by Monte Carlo permutation tests involving 999 unrestricted permutations. Then, an attempt was made to find the minimal set of environmental variables that explained the highest significant amounts of variance in the biological data. Since all environmental variables have individually statistical significant marginal relationships with the pollen data set (see Table 1), a CCA with forward selection was used for variable selection, and the statistical significance of each variable added was assessed by a Monte Carlo permutation test with a Bonferroni-type adjustment for significance level (see Lotter *et al.*, 1997).

This numerical analysis identified nine environmental variables: March T, April T, July T, September T, November T, winter P, spring P, summer P, autumn P. However, since high variance inflation factors (VIFs; i.e. VIFs > 20) indicate high collinearity between mean monthly temperature variables (ter Braak & Šmilauer, 2003), only the mean temperatures of the warmest and coldest months (July and January T; see also Table 1) were chosen. For seasonal precipitation variables, we selected the sums of winter, spring and summer precipitations because: (1) percentage variance explained by CCA with autumn precipitation as the sole constraining variable is lowest among the precipitation variables; (2) autumn precipitation is highly correlated with spring precipitation ($r = 0.8$); and (3) because in a CCA with January T, July T and all four seasonal precipitation variables, autumn precipitation was the sole variable with an inflation factor greater than 20. The first axis of this reduced CCA (with five environmental variables) explains 52.7% of the variance in the pollen data.

Table 1 Percentage variance explained by each environmental variable in a CCA (999 unrestricted Monte Carlo permutations) with this variable as the sole constraining variable.

Variables	Percentage variance
Geography	
Latitude	5.7**
Longitude	2.5*
Altitude (m)	5.8**
Temperature (°C)	
January	13.9**
February	13.5**
March	13.3**
April	12.8**
May	12.6**
June	12.9**
July	13.3**
August	13.5**
September	13.3**
October	13.3**
November	14.3**
December	14**
Precipitation (mm)	
Winter	6.8**
Spring	5.4**
Summer	10**
Autumn	3.3*

$^*P \leq 0.05$; $^{**}P \leq 0.01$.

The statistical relationships between all pollen taxa occurring in more than 10% of the samples and the five environmental variables were tested statistically using a hierarchical series of response models fitted by generalized linear modelling. This taxon response modelling was implemented using the program HOF version 2.3 with square-root transformed percentages and assuming a Poisson error distribution (Oksanen & Minchin, 2002).

Prior to the development of quantitative climate inference models, the gradient length of each individual environmental variable with respect to the pollen data was estimated using a DCCA (Table 1). In principle, if the gradient length is short (< 1.5 SDs), the linear-based partial least squares (PLS) regression would be expected to outperform the weighted average PLS (WA-PLS) regression. However, since this is not always the case (Birks, 1998, 2004) we used both WA-PLS and PLS regression and chose the best model with regards to r^2_{boot}, lowest root mean square error of prediction (RMSEP), average and maximum bias, number of outliers and lowest number of components. The RMSEP and the coefficient of determination (r^2_{boot}) were assessed by bootstrapping (bootstrap sample size = 10) and including all taxa, following Birks (1994). PLS components were added until a reduction in cross-validated prediction error of < 5% with respect to the RMSEP for the simplest one-component model was reached (Birks, 1998). Outliers were

identified as samples having: (1) a very high leverage in a CCA using the environmental variable as the sole constraining variable; and (2) a bootstrap residual (expected – bootstrap estimated) higher than 1 SD of the environmental variable of interest. Transfer functions were developed using the computer program C^2 (version 1.4.2, Juggins, 2003). In addition, to evaluate the training set in terms of r^2_{boot}, RMSEP, average and maximum bias, we used an independent method of intraset cross-validation. For this purpose, the full 79-lake training set was split into two subsets, a 'reduced training set' (59 lakes; RTS) and a 'test training set' (20 lakes; TST). Samples for the TST were chosen by selecting every fourth sample along the sample classification of the two-way indicator species analysis (TWINSPAN, calculated with the WINTWINS program; Hill & Šmilauer, 2005, see Fig. 2). Climate inference models (PLS and WA-PLS) for the RTS were developed for the same environmental variables that were used with the FTS after outlier screening and applied to the TST. Pollen-inferred environmental variables and their sample-specific errors were subsequently compared with the expected climate data for each TST site and their relationship evaluated in terms of r^2.

Finally, the 79-lake training set was divided into two subsets according to the relative amounts of summer and winter precipitation, which reflect opposite seasonal precipitation regimes in Italy. The resulting division resulted in a separation into a northern 40-lake training set (NT), where the amount of summer precipitation is higher, and a southern 39-lake training set (ST), where the amount of winter precipitation is higher (Fig. 3). For each subset, pollen-inference models (PLS and WA-PLS) have been developed using the same methods as for the full training set. As the DCA gradient length of the two training sets was close to the 1.5 SD limit (i.e. 1.73 and 1.56 SD units, respectively), both unimodal and linear ordination methods have been used (i.e. CCA and redundancy analysis (RDA); ter Braak & Prentice, 1988).

All variables had a statistically significant relationship with the NT pollen assemblage in the CCAs and RDAs except for summer P ($P > 0.05$ with the CCA). The same environmental variables that were selected for the FTS (except summer P) have subsequently been tested using forward selection procedures, which led to the exclusion of winter P from further analysis ($P > 0.05$). January T and autumn P were also excluded due to their high inflation factors. In contrast, the RDA with forward selection retained summer P. The final models for the NT explain 84.5% (CCA with July T and spring P) and 60.4% of variance of the species–environment relationship (RDA with July T, spring P and summer P). DCCA gradient lengths do not exceed 1.3 SD units. In contrast to the NT, all variables had a statistically significant relationship with the pollen assemblages in the ST. With forward selection procedures (CCAs and RDAs), spring P and summer P were excluded at the $P = 0.05$ level. The final models for the ST explain 40.7% (CCA with January T, July T, autumn P and winter P) and 37.6 % of the variance of the species–environment relationship (RDA with January T, July T, autumn P and winter P). A partial CCA was used to determine if the two T and P variables have a significant independent effect on

assemblage composition (following Telford, 2006). DCCA gradient lengths for January T, July T, autumn P and winter P are < 2.

RESULTS

Individual responses of pollen taxa to climatic variables

The most abundant pollen types are shown in the pollen diagram (Fig. 2). The order of the samples, as implemented with TWINSPAN (Hill & Šmilauer, 2005), is mainly related to January T and July T ($r^2 \sim 0.6$), while correlations to winter and summer P are weaker ($r^2 = 0.3$ and 0.4, respectively). As can be seen in the diagram, only Gramineae pollen is abundant across the entire gradient, while most taxa are limited to a smaller portion of the entire gradient.

Of the 74 taxa with more than 10 occurrences in the FTS, none showed a skewed response model to any environmental variables (see Appendix S3 in Supplementary Material). Overall, symmetric responses of pollen taxa to single environmental variables are less common than flat or monotone models (Table 2). Between 30% and 40% of pollen taxa showed no statistically significant response to individual variables. However, tree pollen showed fewer taxa with no statistically significant response models than herb pollen (e.g. 8% vs. 47% for July T; 13% vs. 47% for January T, respectively). In comparison with herb pollen, a higher percentage of tree pollen showed symmetric response models to all environmental variables (e.g. 28% vs. 3% for July T, respectively). Only four herb and fern taxa (*Artemisia*, *Plantago lanceolata*, *Pteridium* and *Ranunculus acris*-type) showed no statistically significant relationship to any environmental variable included in this study. Their continuous presence at low abundance across the whole gradient suggests that major climatic thresholds for these taxa are not present within the sampled environmental gradient. In fact, *P. lanceolata* and *Pteridium aquilinum*, as well as *Artemisia* and *Ranunculus* species, are present in all Italian regions (Pignatti, 1982). The first two pollen and spore types are frequently associated with human impact (Behre, 1981); *Pteridium* spores often increase in response to forest fires (Moore *et al.*, 1991; Tinner *et al.*, 1999); *Ranunculus acris*-type includes a high number of different species from different environments (Punt *et al.*, 1976–1996). The higher incidence of symmetric distributions for trees and shrubs, on the other hand, reveals that distributions of woody plants are climatically limited along at least one of the environmental gradients (Table 2).

Modern pollen–climate inference models

Transfer functions were developed for January T, July T, winter P, spring P and summer P, as these parameters were selected with the CCA forward selection procedure and because they are of greatest palaeoecological and palaeoclimatic interest. The first CCA axis separates sites with high July and January T from sites in the Alpine region (lower July and January T), and is significantly stronger than the second axis $\lambda_1 = 0.19$, $\lambda_2 = 0.09$; Fig. 4). The
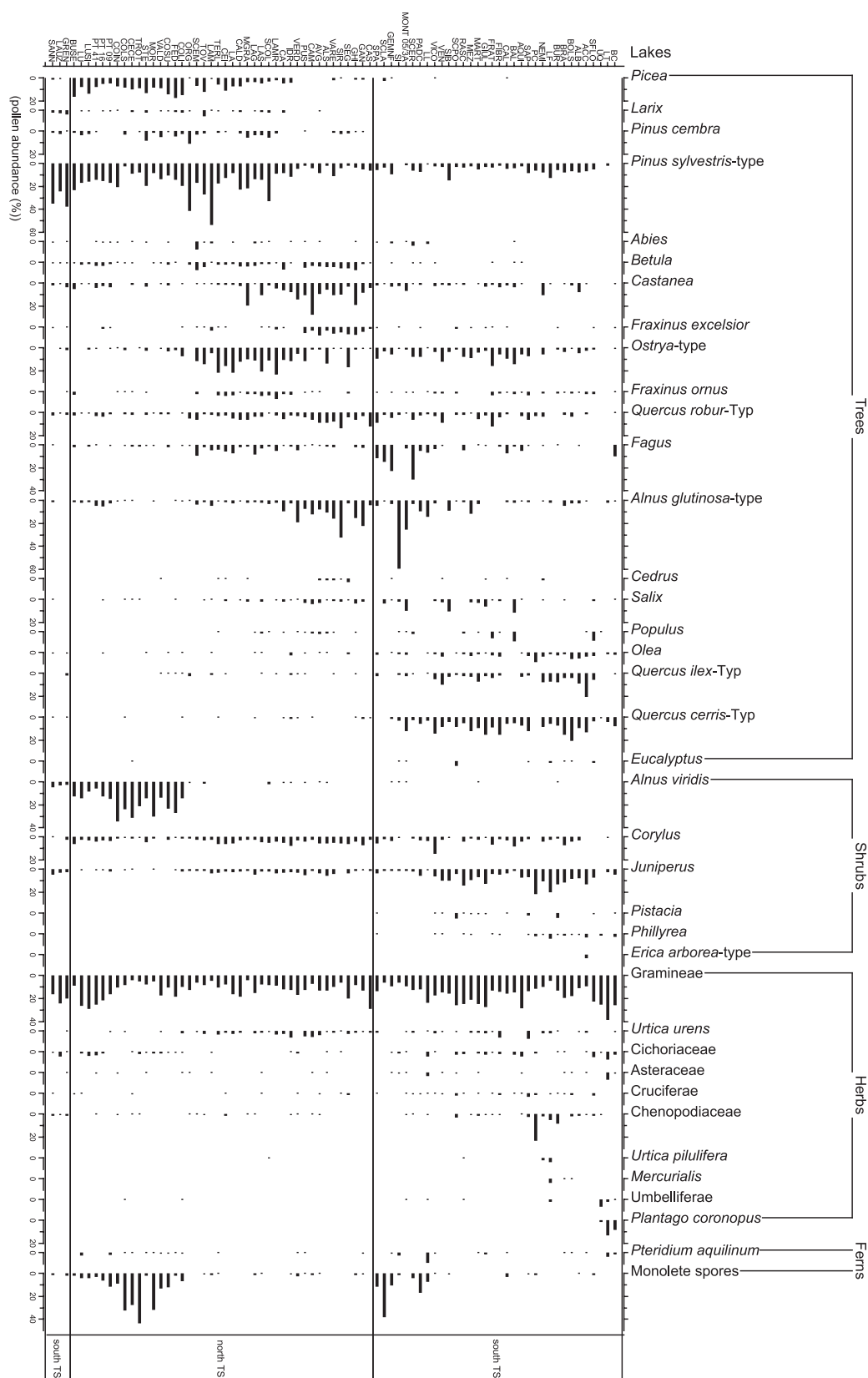
Figure 2 Simplified pollen diagram showing pollen and spore types with more than one occurrence and > 2% maximum abundance. Samples are ordered according to a two-way indicator approach implemented with wintwins (Hill & Šmilauer, 2005). Pollen and spore types are arranged according to their weighted averages on the *y*-axis. The names and locations of the lakes are presented in Appendix S1 (see Supplementary Material).
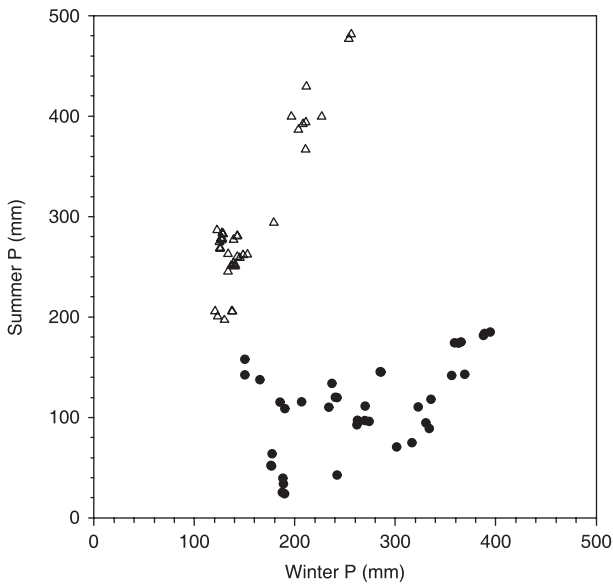
Figure 3 Scatter-plot showing estimated winter versus summer precipitation (P) for the 79 sites used in the full training set (FTS). Empty triangles, sites contained in the northern subset (NT); full circles, sites contained in the southern subset (ST).

Figure 4 Canonical correspondence analysis (CCA) ordination biplot with site scores and major environmental variables for the full 79-lake training set (FTS). Lake symbols are chosen according to the relative amount of summer and winter precipitation at the sites (see Methods and Figs 1 & 3 for details). Open triangles, sites contained in the northern subset (NT); full circles, sites contained in the southern subset (ST).

second axis separates sites on the basis of spring and summer P. Winter P mainly separates sites on the third axis (not shown).

Although the gradient length is only 1.84 SD units for DCA axis 1 and the symmetric model is not the most common statistically significant individual response model (Table 2), the unimodal-based WA-PLS models appear to be the most robust, as they are based on a smaller number of components, produced fewer outliers and showed lower RMSEP than the corresponding linear-based PLS models (Table 3). The reasons for this paradox (see discussions in Seppä *et al.*, 2004) may be related to the 'two faces of compositional data' (ter Braak & Šmilauer, 2003).

While only three outliers (LF, CUBA and RASC) were identified for January T and only two outliers for July T (LAUZ and CALD), a number of outliers were identified for the seasonal precipitation variables (Table 3). These lakes were removed from the training set prior to performing PLS and WA-PLS regression and calibration. For all variables (except for spring P) the best transfer
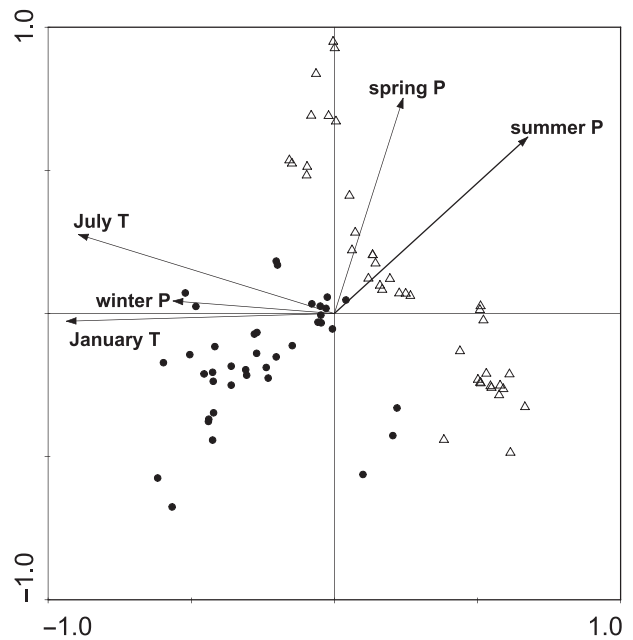
function or the 'minimal adequate model' (*sensu* Birks, 1998) was a two-component WA-PLS model (Table 3).

For most environmental variables, a strong relationship could be established between pollen-inferred variables and the values estimated for the study sites based on instrumental data (Fig. 5). Highest values of r^2_{boot} were found for January T, July T and summer P (r^2_{boot} ~ 0.85; see Table 3). The RMSEP of January T, July T, spring P and summer P amount to *c.* 10% or less of the total gradient lengths (after deletion of outliers). The RMSEP covers a higher fraction of the total gradient length for winter P (14.9 %; Table 3). Due to the deletion of outliers from the models, the upper range limits of prediction for January T, winter P, spring P and summer P are strongly reduced. Maximum January T was

Table 2 Frequencies of the different types of response models for trees and shrubs (T/S) and for herb (H) pollen with more than 10 occurrences in the full training set (FTS; 79 lakes).

	July T		January T		Winter P		Spring P		Summer P	
	T/S	H	T/S	H	T/S	H	T/S	H	T/S	H
None	8	47	13	47	33	56	23	56	26	35
Flat model	31	35	23	26	33	18	41	32	31	44
Monotone model	36	15	46	24	26	23	23	9	28	9
Symmetric (plateau) model	25	3	18	3	8	3	13	3	15	12

T, temperature (°C); P, precipitation (mm).

Table 3 Model performance of the full training set (79 lakes).

Variable	January T	July T	Winter P	Spring P	Summer P	January T	July T	Winter P	Spring P	Summer P
DCCA gradient length (SD units)	1.84	1.52	0.99	1.34	1.51	1.84	1.52	0.99	1.34	1.51
Model type	WA-PLS	WA-PLS	WA-PLS	WA-PLS	WA-PLS	PLS	PLS	PLS	PLS	PLS
No. of components	2	2	2	1	2	2	3	2	2	2
No. outliers	3	2	17	10	10	4	3	16	10	8
Apparent										
RMSE	0.92	1.36	21.31	27.34	21.07	1.38	1.42	30.55	31.07	29.32
r^2	0.96	0.93	0.90	0.70	0.94	0.91	0.92	0.77	0.65	0.88
Ave. bias	0.04	0.06	0.25	−0.62	0.32	0.00	0.00	0.00	0.00	0.00
Max. bias	0.76	2.03	16.23	74.49	37.50	1.44	1.80	49.69	76.42	133.88
Cross-validation (bootstrap 999 iterations)										
r^2_{boot}	0.89	0.84	0.69	0.51	0.84	0.86	0.84	0.66	0.50	0.84
Ave. bias	0.13	0.13	0.36	−1.73	−3.08	0.03	0.07	0.00	−0.13	−2.39
Max. bias	2.75	3.48	94.96	110.49	156.98	2.75	2.70	69.85	96.18	148.35
RMSEP	1.7	2.1	41.0	36.9	38.0	1.8	2.1	39.2	38.4	35.7
RMSEP % range	8.5	11.9	14.9	8.9	8.3	9.0	11.8	14.3	9.3	7.8

T, temperature (°C); P, precipitation (mm); PLS, partial least squares; WA-PLS, weighted-averaging-PLS; RMSEP, root mean square error of prediction.

reduced by 15% (to 9.5 °C), maximum winter P by 20% (335 mm), maximum spring P by 30% (367 mm) and maximum summer P by 40% (294 mm).

Intraset cross-validation

The intraset cross-validation was performed by reconstructing the environmental variables for 20 test sites using inference models developed with the 59-lake RTS. The RTS is comparable to the FTS with regard to minimum, maximum, mean, median and standard deviation of the variables (Table 4), and the length of the gradient (as assessed by a DCA), therefore providing an appropriate replicate for the intraset cross-validation. Three sites were detected as outliers for January T (LF, CUBA and RASC), two lakes for July T (CUBA and LAUZ) and a number of sites for winter P, spring P and summer P. These lakes were removed from the RTS prior to performing PLS and WA-PLS regression and prior to applying the transfer functions to the test set. Performance statistics and 'minimal adequate' number of components of the reduced set transfer functions were almost identical to those of the FTS (Table 5). We used the WA-PLS models to implement the validation because of their better performance with regard to RMSEP, average and maximum bias, number of outliers, and r^2_{boot}.

Overall, the intraset cross-validation indicates that the pollen-based temperature and precipitation inference models of the RTS provide, within the limits of the bootstrap-estimated standard error of prediction, accurate estimates of the expected values for the test sites (Fig. 6). The good performance of the pollen–climate inference models for January T, July T and summer P is indicated by the high r^2 (> 0.80; Table 6). However, pollen-inferred estimates of winter P and spring P were lower ($r^2 \leq 0.50$). Most of the under- or overestimated values in the test set originate from outliers in the FTS (Fig. 5).

Table 4 (a) Descriptive statistics of the full training set (79 lakes) and (b) of the reduced training set (59 lakes).

(a)

Variable	Minimum	Maximum	Mean	Median	SD
January T	−8.1	12.1	0.4	−0.3	4.89
July T	8.1	26.1	18.0	19.0	5.08
Winter P	120.7	394.5	209.3	188.0	81.24
Spring P	82.7	495.4	233.7	222.7	76.18
Summer P	24.1	481.6	202.9	197.0	109.64

(b)

Variable	Minimum	Maximum	Mean	Median	SD
January T	−8.1	12.1	0.2	−0.3	4.96
July T	8.1	26.1	17.9	19.0	5.11
Winter P	120.7	389.1	205.6	183.2	80.91
Spring P	82.7	488.8	231.5	222.4	71.66
Summer P	24.1	476.7	203.9	205.6	106.71

T, temperature (°C); P, precipitation (mm).

Subregional pollen–climate models

To address the importance of different environmental variables for the vegetation in the Mediterranean area and in the Alps, the FTS was split into two parts according to the relative amount of mean winter vs. mean summer precipitation (Figs 1 & 3). The minimal subsets of variables explaining the pollen assemblages of the two data sets differ between those two regions: for the Alpine training set July T, spring P and summer P were retained, whereas January T, July T, winter P and autumn P were retained
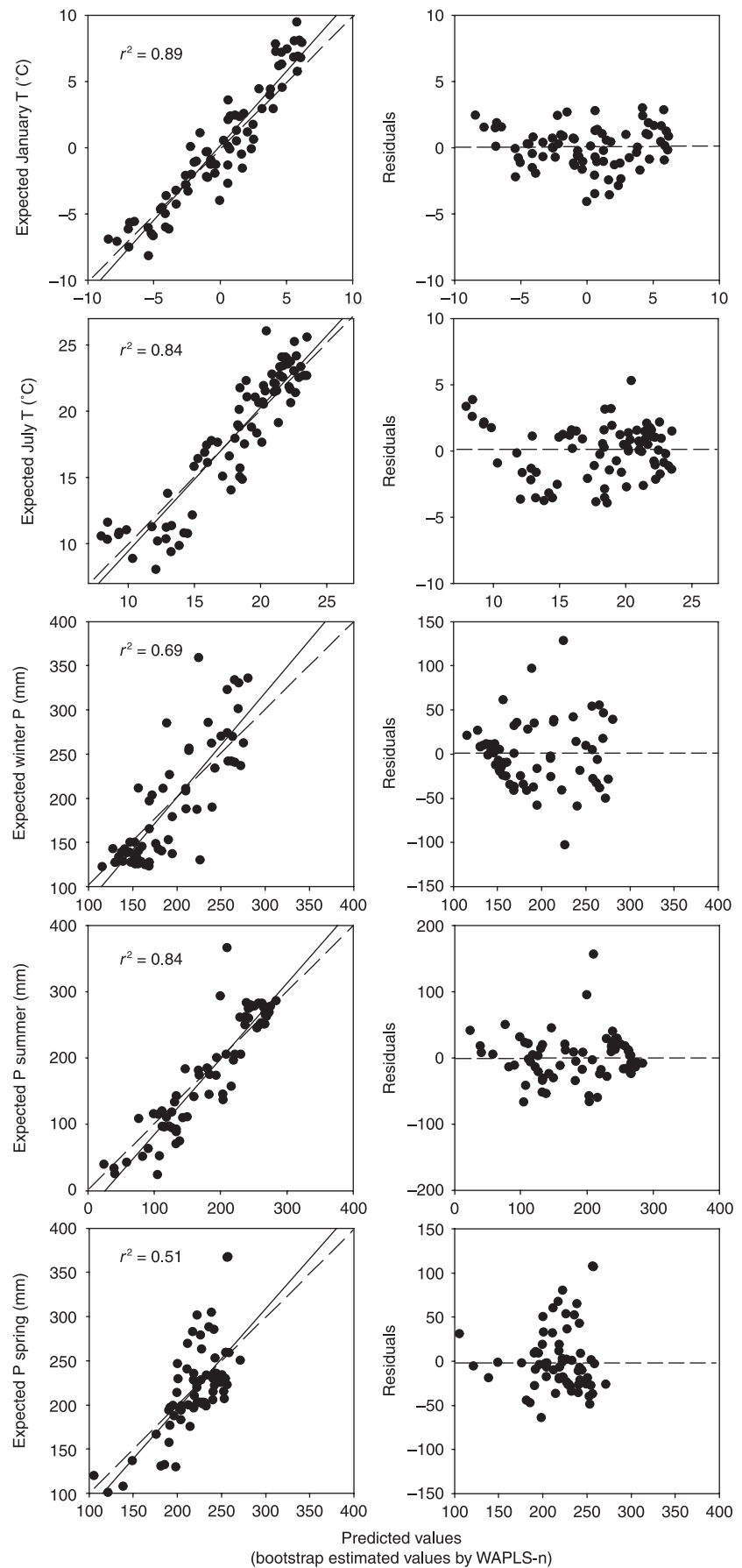
Figure 5 Scatter plots illustrating the relationships between pollen-inferred estimates in bootstrap cross-validation vs. expected temperature (T) and precipitation (P) (left) and between bootstrap residuals vs. expected temperature (T) and precipitation (P) (right) for January T, July T, winter P, summer P and spring P of the full 79-lake training set (FTS).

Table 5 Performance of the reduced training set (RTS; 59 lakes) used for the intraset cross-validation.

Variable	January T	July T	Winter P	Spring P	Summer P	January T	July T	Winter P	Spring P	Summer P
Model type	WA-PLS	WA-PLS	WA-PLS	WA-PLS	WA-PLS	PLS	PLS	PLS	PLS	PLS
No. of components	2	2	1	1	2	2	1	1	1	2
No. outliers	3	2	11	6	8	4	2	10	9	6
Apparent										
RMSE	0.86	1.21	30.50	26.85	18.18	1.35	2.20	34.20	29.53	24.38
r^2	0.96	0.94	0.74	0.68	0.95	0.91	0.80	0.69	0.51	0.92
Ave. bias	0.02	0.04	−0.13	−0.24	0.32	0.00	0.00	0.00	0.00	0.00
Max. bias	0.81	2.90	59.55	83.60	13.94	1.21	3.17	71.12	58.42	23.88
Cross-validation (bootstrap 999 iterations)										
r^2_{boot}	0.86	0.84	0.60	0.42	0.87	0.85	0.77	0.62	0.33	0.87
Ave. bias	0.04	0.08	0.11	−3.22	−5.28	0.05	0.09	0.79	−1.04	−2.99
Max. bias	2.57	3.38	94.43	119.20	54.32	2.24	3.74	86.60	68.05	43.15
RMSEP	1.8	2.1	38.8	37.8	34.5	1.9	2.5	39.3	35.8	32.1

T, temperature (°C); P, precipitation (mm); PLS, partial-least squares; WA-PLS, weighted-averaging PLS; RMSEP, root mean square error of prediction.

Table 6 Performance of the intraset cross-validation and edge values (maximum and minimum) of the reduced training set (RTS; 59 lakes) after the exclusion of outliers.

Variable	r^2 incl. outliers	r^2 excl. outliers	Min. RTS excl. outliers	Max. RTS excl. outliers
January T	0.90	No outlier	−8.1	9.5
July T	0.83	0.85	8.1	26.1
Winter P	0.20	0.54	122.4	335.7
Spring P	0.44	0.51	101.0	367.5
Summer P	0.48	0.87	24.1	294.0

T, temperature (°C); P, precipitation (mm).

for the Mediterranean training set. This indicates that for pollen assemblages in the Alps, winter climate conditions are less important explanatory variables than the climate conditions of the spring and summer seasons. Pollen assemblages along the southern training set (ST) gradient are not related to summer and spring precipitation as much as they are to other environmental variables. The analysis of the two subsets indicates therefore that pollen assemblages of the northern training set (NT) are mainly related to environmental conditions of the growing season (i.e. mainly spring–summer), while, rather surprisingly, summer P was not considered as important as winter P in the ST. This might be related to the fact that variations in summer P are not as strong as for winter P along the ST gradient.

Inference models have been developed for July T, spring P and summer P using both PLS and WA-PLS. While no outliers were detected for July T, a number of outliers that showed residuals higher than 1 SD were detected for spring P and summer P and were therefore excluded. WA-PLS models performed better than PLS models in having higher r^2_{boot}, lower RMSEP, lower average and maximum bias and fewer outliers (Table 7).

DISCUSSION

Individual responses of pollen taxa to climatic variables

Interannual climate variability may affect pollen production in single years, as shown by analyses of peat-sequences and pollen traps (e.g. Autio & Hicks, 2004). However, since our sediment-surface samples cover a period of at least 5–10 years, the influence of interannual variability is of minor importance when compared with the longer-term pollen–vegetation relationship. In effect, numerical analyses of paired pollen and tree-inventory data reveal good relationships between pollen assemblages in lake sediments and tree percentages (Sugita, 1994).

Overall, the distribution of tree pollen taxa (Fig. 2) mainly reflects the major biomes of the study area. *Pinus* (*Pinus mugo* and *Pinus cembra*), *Alnus viridis*, *Picea abies* and *Larix* occur at higher elevation in the Alpine area. Dispersal bias, however, may also be responsible for high values of the wind-pollinated *Pinus sylvestris*-type pollen at sites above the timberline, and for the low abundances of heavy *Larix* pollen grains (Prentice, 1985). *Castanea sativa*, *Fraxinus excelsior*, *Q. robur*-Typ (deciduous oaks such as *Q. robur*, *Quercus petraea*, *Quercus pubescens*, *Quercus frainetto*), *Fagus sylvatica* and *Alnus glutinosa* occur mainly at lower elevations in northern Italy and in the Apennine Mountains (Pignatti, 1997). *Ostrya*-type pollen shows a bimodal distribution, probably because it combines two species with overlapping biogeographical areas: *Ostrya carpinifolia* (northern and southern Italy) and *Carpinus orientalis* (southern Italy) (Pignatti, 1982). At the warmer end of the gradient (i.e. the upper part of the diagram), evergreen trees and shrubs dominate the pollen assemblages, notably *Olea*, *Pistacia*, *Phyllirea*, *Eucalyptus*, *Q. ilex*-Typ, and *Juniperus*. The higher occurrence of herb pollen in these landscapes (see Fig. 2) might be attributed to a more pronounced impact of widespread deforestations, agriculture and to the drier summer climate (Pignatti, 1997). *Alnus glutinosa* spans both the
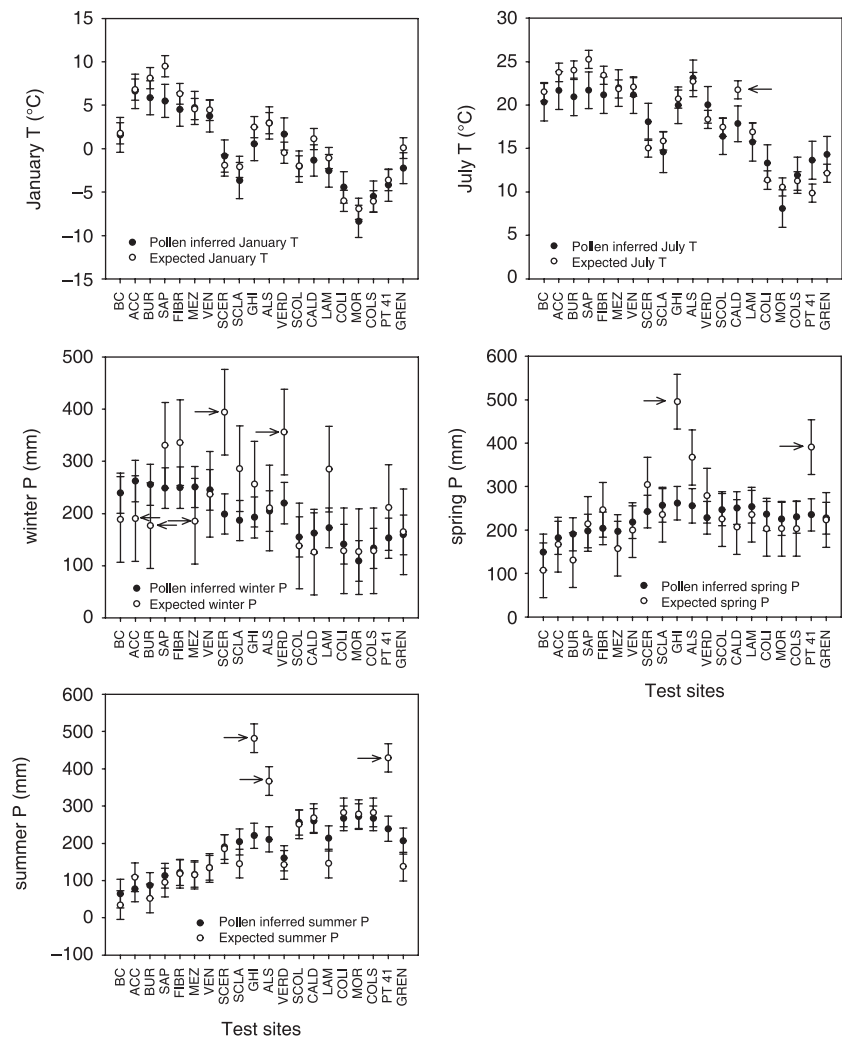
Figure 6 Results of the intraset cross-validation. Bootstrap pollen-inferred values (full circles) and associated bootstrap estimated standard error of prediction. Expected temperature (T) and precipitation (P) (empty circles) and associated standard deviation of the prediction errors. Arrows indicate sites that were identified as outliers in the full training set.

southern and the northern training set, whereas most other taxa show marked changes at the boundary between the two sub-regional data sets (Fig. 2). *Alnus glutinosa* is a riparian species and thus potentially less sensitive to climate variables than to landscape or ecosystem processes (Bennett & Birks, 1990). Nevertheless, it appears that its pollen abundance in this data set is limited to low-elevation sites in northern Italy and to moister sites in southern Italy.

Modern pollen–climate inference models

Strong relationships could be established between pollen-inferred variables and values estimated for the study sites based on climate station data. RMSEP for the models are fully comparable with those obtained for models based on similar numerical techniques from Norway and Estonia for annual temperature, T_{ann} (8.8%) (Seppä *et al.*, 2004) and for July T in Switzerland (8.7%) (Lotter *et al.*, 2000), whereas the RMSEP for eight different diatom-based calibration sets range between 10.7% and 25.2% (Rosen *et al.*, 2000) and for nine chironomid-based calibration sets between 7.9% and 21.4% (Larocque *et al.*, 2001; Heiri *et al.*, 2003).

In comparison, for a continent-wide, broad-scale pollen-inference model (2636 surface samples) based on the MAT, RMSEP of 2.58 °C ($r^2 = 0.83$) and of 2.25 °C ($r^2 = 0.75$) for mean temperatures of the coldest and warmest months, respectively, were reported (Davis *et al.*, 2003). A slightly lower RMSEP was obtained for a smaller model also based on the MAT (868 surface samples). In the latter case, a RMSEP of 2.0 °C ($r^2 = 0.94$) for the mean temperature of the warmest month was obtained (Peyron *et al.*, 2005) after deletion of analogues that were not compatible with inferred lake levels. The spatial extent of the latter data set is large, including surface samples from the Tibetan Plateau, Kazakhstan, the Near East, Siberia, the Scandes Mountains and Spain (Peyron *et al.*, 1998, 2005; Gachet *et al.*, 2003), while the spatial density on the Italian Peninsula is low. The higher spatial density of samples on the Italian Peninsula, however, is likely to make our 79-lake model more suitable for pollen-inferred quantitative palaeoenvironmental reconstructions on this regional scale. In addition, Davis *et al.* (2003) as well as Peyron *et al.* (2005) use the MAT, which provides over-optimistic error statistics in the presence of spatial autocorrelation (including terrestrial pollen and climate; see, e.g., Telford & Birks, 2005). Hence, our 79-sample inference model might provide reconstructions for temperatures

Table 7 Model performance of northern subset (NT, 40 lakes).

Variable	July T	Spring P	Summer P	July T	Spring P	Summer P
Model type	WA-PLS	WA-PLS	WA-PLS	PLS	PLS	PLS
No. of components	1	2	2	1	2	2
No. outliers	0	8	11	1	9	9
Apparent						
RMSE	1.7	10.6	6.7	1.9	21.6	18.8
r^2	0.89	0.95	0.96	0.85	0.85	0.82
Ave. bias	0.05	−0.02	−0.23	0.00	0.00	0.00
Max. bias	1.86	6.83	15.00	3.25	66.45	24.51
Cross-validation (bootstrap 999 iterations)						
r^2_{boot}	0.84	0.63	0.52	0.82	0.68	0.57
Ave. bias	0.07	1.97	0.87	0.11	1.83	0.13
Max. bias	3.49	83.08	104.90	3.66	113.97	73.87
RMSEP	2.2	31.9	29.8	2.2	34.3	32.0

T, temperature (°C); P, precipitation (mm); PLS, partial least squares; WA-PLS, weighted-averaging PLS; RMSEP, root mean square error of prediction.

Table 8 Model performance of southern subset (ST, 39 lakes).

Variable	January T	July T	Autumn P	Winter P	January T	July T	Autumn P	Winter P
Model type	WA-PLS	WA-PLS	WA-PLS	WA-PLS	PLS	PLS	PLS	PLS
No. of components	2	2	1	1	2	1	2	2
No. outliers	2	0	6	6	3	3	8	3
Apparent								
RMSE	0.8	0.7	24.0	28.0	1.1	1.5	26.7	21.6
r^2	0.96	0.97	0.87	0.83	0.92	0.83	0.87	0.90
Ave. bias	0.03	0.02	0.08	0.07	0.00	0.00	0.00	0.00
Max. bias	1.19	0.63	36.34	41.79	1.98	4.60	50.05	33.91
Cross-validation (bootstrap 999 iterations)								
r^2_{boot}	0.79	0.86	0.68	0.62	0.79	0.71	0.72	0.71
Ave. bias	0.07	0.03	−4.03	−4.85	0.02	−0.03	−3.71	−0.79
Max. bias	4.05	2.73	72.49	93.62	4.26	7.93	73.15	52.99
RMSEP	2.3	2.0	44.5	50.0	2.0	2.2	42.8	39.8

T, temperature (°C); P, precipitation (mm); PLS, partial least squares; WA-PLS, weighted-averaging PLS; RMSEP, root mean square error of prediction.

of the warmest month in this region with a prediction error that is at least comparable to the model by Peyron *et al.* (2005).

The performance of the pollen–climate inference models was increased by deletion of outliers from the FTS. Some sites identified as outliers for January T (RASC) and July T (LAUZ and CALD) are located in mountain regions. Due to pollen transport across altitudinal zones, the temperature of CALD, which is a low-elevation site in the south-eastern Alps, was underestimated, while the temperature of RASC and LAUZ, which are high-altitude sites, was overestimated by the models. In contrast, January T at LF and at CUBA, which have the highest values of the FTS, was underestimated. The high number of outliers with regard to precipitation reflects the heterogeneous spatial distribution of precipitation and the inherent difficulties in estimating modern precipitation values in mountain areas. Nevertheless, as shown by the intraset cross-validation, the pollen-based temperature

and precipitation inference models provide, within the limits of the bootstrap-estimated standard error of prediction, accurate estimates of the expected values (Fig. 6).

Subregional pollen–climate models

The environmental variables that were used in the FTS might not be the primary stressors for plant growth across the entire climatic and ecological gradient. In fact, it is likely that different ecological mechanisms may act in different subregions. For example, Pigott & Pigott (1993) observed that where Mediterranean scrub and temperate forests intergrade in southern France, evergreen sclerophyll shrubs dominate south-facing slopes while summergreen trees dominate northern slopes. They concluded that this is because summer drought is the proximal factor excluding soft-leaved deciduous trees from the sun-baked south-facing

slopes. Hence, in the Mediterranean, summergreen plants are at a disadvantage in comparison to evergreen trees and shrubs, as the period of physiological activity for summergreen species coincides with the most unfavourable season with regard to water availability (Givnish, 2002). In addition to precipitation during the warmest months, minimum and maximum temperatures also play a role if plants have a low tolerance to absolute or cumulative low temperatures (winter) or if gross primary production does not exceed respiration (summer) (Box & Fujiwara, 2005). In northern Italy, evergreen species are limited by winter cold (i.e. frost causing lethal damages) in the Alpine forelands (Walther *et al.*, 2002), while the performance of deciduous species is reduced by drought (Givnish, 2002). Seasonal factors, however, not only influence different vegetation types within the same environment, but can also influence a species in different ways along an environmental gradient. This is, for example, the case for *F. sylvatica* (beech), which is limited by late frosts in regions where winters are cold (Prentice & Helmisaari, 1991), whereas it is limited by drought where summers are warm and dry (Piovesan & Schirone, 2000).

The performances of the regional training sets are comparable to or lower than the performances of the FTS. Moreover, even when reducing the environmental gradient, sites identified as outliers in the FTS also had to be excluded from the regional subsets. Hence, the use of the subsets did not enhance the model performances. On the other hand, our results imply that care must be taken with the interpretation of environmental variables that are not ecologically meaningful for pollen–climate relationships in specific biomes, e.g. January T for the Alpine subset. The two training sets have only July T as a common influential environmental variable.

Influence of human impact on vegetation

Despite a history of long-term human impact that strongly disrupted the forested areas on the Italian Peninsula by deforestation, cultivation and forest fires (e.g. Sadori *et al.*, 2004; Finsinger & Tinner, 2006), the observed pollen–climate relationships reconstruct modern climatic variables within the site-specific climate data uncertainties (Fig. 6). This unambiguously reveals that vegetation can be related to present climate conditions and is not the result of landscape management alone. On the other hand, since the influence of human activity on vegetation and landscape depends on scale (Birks & Seppä, 2004), it may be envisaged that our training set covers a sufficiently large gradient to overarch the anthropogenic effect on most taxa.

We are, however, aware that climatic change may not have been the only factor influencing past changes in vegetation. It has, for example, been suggested that the major vegetation change from broad-leaf deciduous forests to evergreen trees and shrubs that occurred in the Mediterranean during the Holocene was caused by factors such as forest burning, wildfires and pasturing (Pons & Quézel, 1998). Hence, in future applications of these inference models to downcore pollen records it is advisable to cross-validate pollen-inferred climate changes by means of other, independent climate proxy indicators.

CONCLUSIONS

Our analysis of pollen assemblages in the surface sediments of lakes across the large temperature and precipitation gradients of the Italian Peninsula indicates that pollen can provide useful and accurate quantitative estimates for January T, July T, winter P, spring P and summer P. This result is based upon two lines of evidence: first and foremost, the overall inference models have provided reasonably low bootstrap cross-validated RMSEPs. The environmental variables with the highest explanatory power are January T, July T and summer P, with RMSEP of 1.8 °C, 2.1 °C and 35.7 mm, respectively. Second, the bootstrap cross-validation of the FTS and the intraset cross-validation indicate similar performance in terms of error statistics (Tables 3, 5 & 6). For January T, July T and summer P, r^2 values are > 0.8, while pollen-inferred winter and spring P show a somewhat weaker relationship with the observed variables ($r^2 \sim 0.50$).

In an attempt to further optimize the inference models, we split the 79-lake data set into two subsets and developed new inference models for each subset. Our results indicate that the prediction errors are not lower than those of the full data set but reveal that pollen–climate relationships differ within the two subsets. Modern pollen assemblages (and thus vegetation composition) are mainly related to climate conditions during the growing seasons: summer temperature and precipitation for the 'Alpine', and winter temperature and precipitation for the 'Mediterranean' subregions.

In the Mediterranean region, recent climate variability mainly had an impact on winter precipitation and temperature (Luterbacher *et al.*, 2006) and changes in these environmental variables are considered as causes for vegetation shifts during the past century (Walther *et al.*, 2002). Our pollen-inference models can provide quantitative estimates of seasonal climatic variables, including winter precipitation and temperature, and are therefore promising tools for the reconstruction of low-frequency past climate variability in southern Europe.

REFERENCES

Agustí-Panareda, A. & Thompson, R. (2002) Reconstructing air temperature at eleven remote alpine and arctic lakes in Europe from 1781 to 1997 AD. *Journal of Paleolimnology*, **28**, 7–23.

Autio, J. & Hicks, S. (2004) Annual variations in pollen deposition and meteorological conditions on the fell Aakenustunturi in northern Finland: Potential for using fossil pollen as a climate proxy. *Grana*, **43**, 31–47.

Behre, K.-E. (1981) The interpretation of anthropogenic indicators in pollen diagrams. *Pollen et Spores*, **23**, 225–245.

Bennett, K.D. & Birks, H.J.B. (1990) Postglacial history of alder (*Alnus glutinosa* (L.) Gaertn.) in the British Isles. *Journal of Quaternary Science*, **5**, 123–133.

Beug, H.-J. (2004) *Leitfaden der Pollenbestimmung für Mitteleuropa und angrenzende Gebiete*. Verlag Dr. Friedrich Pfeil, München.

Bigler, C. & Hall, R.I. (2002) Diatoms as indicators of climatic and limnological change in Swedish Lapland: a 100-lake calibration set and its validation for paleoecological reconstructions. *Journal of Paleolimnology*, **27**, 97–115.

Birks, H.J.B. (1994) The importance of pollen and diatom taxonomic precision in quantitative palaeoenvironmental reconstructions. *Review of Palaeobotany and Palynology*, **83**, 107–117.

Birks, H.J.B. (1998) Numerical tools in palaeolimnology — progress, potentialities, and problems. *Journal of Paleolimnology*, **20**, 307–332.

Birks, H.J.B. (2004). Quantitative palaeoenvironmental reconstructions from Holocene biological data. *Global change in the Holocene* (ed. by A.W. Mackay, R.W. Battarbee, H.J.B. Birks and F. Oldfield), pp. 107–123. Arnold, London.

Birks, H.J.B. & Seppä, H. (2004) Pollen-based reconstructions of late-Quaternary climate in Europe — progress, problems, and pitfalls. *Acta Palaeobotanica*, **44**, 317–334.

Box, E.O. & Fujiwara, K. (2005). Vegetation types and their broad-scale distribution. *Vegetation ecology* (ed. by E. van der Maarel), pp. 106–128. Blackwell Publishing, Oxford.

Davis, B.A.S., Brewer, S., Stevenson, A.C., Guiot, J. & Data Contributors (2003) The temperature of Europe during the Holocene reconstructed from pollen data. *Quaternary Science Review*, **22**, 1701–1716.

Finsinger, W. & Tinner, W. (2006) Holocene vegetation and land-use changes in the forelands of the southwestern Alps, Italy. *Journal of Quaternary Science*, **21**, 243–258.

Finsinger, W., Bigler, C., Krähenbühl, U., Lotter, A.F. & Ammann, B. (2006) Human impact and eutrophication patterns during the past ~200 yrs at Lago Grande di Avigliana (N. Italy). *Journal of Paleolimnology*, **36**, 55–67.

Gachet, S., Brewer, S., Cheddadi, R., Davis, B., Gritti, E. & Guiot, J. (2003) A probabilistic approach to the use of pollen indicators for plant attributes and biomes: an application to European vegetation at 0 and 6000 ka. *Global Ecology and Biogeography*, **12**, 103–118.

Givnish, T.J. (2002) Adaptive significance of evergreen vs. deciduous leaves: solving the triple paradox. *Silva Fennica*, **36**, 703–743.

Gleason, H.A. (1939) The individualistic concept of the plant association. *American Midland Naturalist*, **21**, 92–110.

Guiot, J. (1990) Methodology of the last climatic cycle reconstruction in France from pollen data. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **80**, 49–69.

Heiri, O., Lotter, A.F., Hausmann, S. & Kienast, F. (2003) A chironomid-based Holocene summer air temperature reconstruction from the Swiss Alps. *The Holocene*, **13**, 477–484.

Hill, M.O. & Šmilauer, P. (2005) *TWINSPAN for Windows, version 2.3*. Centre for Ecology and Hydrology, Huntingdon and University of South Bohemia, Ceské Budejovice.

Jackson, S.T. & Williams, J.W. (2004) Modern analogs in Quaternary paleoecology: here today, gone yesterday, gone tomorrow? *Annual Review of Earth and Planetary Sciences*, **32**, 495–537.

Juggins, S. (2003) *C² user guide. Software for ecological and palaeoecological data analysis and visualisation.* University of Newcastle, Newcastle upon Tyne.

Larocque, I., Hall, R.I. & Grahn, E. (2001) Chironomids as indicators of climate change: a 100-lake training set from a subarctic region of northern Sweden (Lapland). *Journal of Paleolimnology*, **26**, 307–322.

Lotter, A.F., Birks, H.J.B., Hofmann, W. & Marchetto, A. (1997) Modern diatom, cladocera, chironomid and chrysophyte cyst assemblages as quantitative indicators for the reconstructions of past environmental conditions in the Alps I. Climate. *Journal of Paleolimnology*, **18**, 395–420.

Lotter, A.F., Birks, H.J.B., Eicher, U., Hofmann, W., Schwander, J. & Wick, L. (2000) Younger Dryas and Allerød summer temperatures at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **159**, 349–361.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. & Wanner, H. (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, **303**, 1499–1503.

Luterbacher, J., Xoplaki, E., Casty, C., Wanner, H., Pauling, A., Küttel, M., Rutihauser, T., Brönnimann, S., Fischer, E., Fleitmann, D., Gonzalez-Rouco, F.J., Garcia-Herrera, R., Barriendos, M., Rodrigo, F., Gonzalez-Hidalgo, J.C., Saz, M.A., Gimeno, L., Ribera, P., Brunet, M., Paeth, H., Rimbu, N., Felis, T., Jacobeit, J., Dünkeloh, A., Zorita, E., Guiot, J., Türkes, M., Alcoforado, M.J., Trigo, R., Wheeler, D., Tett, S., Mann, M.E., Touchan, R., Shindell, D.T., Silenzi, S., Montagna, P., Camuffo, D., Mariotti, A., Nanni, T., Brunetti, M., Maugeri, M., Zerefos, C., De Zolt, S. & Lionello, P. (2006). Mediterranean climate variability over the last centuries: a review. *The Mediterranean climate: an overview of the main characteristics and issues* (ed.

by P. Lionello, P. Malanotte-Rizzoli and R. Boscolo), pp. 27–148. Elsevier, Amsterdam.

Moore, P.D., Webb, J.A. & Collinson, M.E. (1991) *Pollen analysis*, 2nd edn. Blackwell Science, Oxford.

Oksanen, J. & Minchin, P.R. (2002) Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling*, **157**, 119–129.

Petrarca, S., Cogliani, E., Mancini, M. & Spinelli, F. (1999) *Profilo climatico d'Italia*. ENEA, Rome.

Peyron, O., Guiot, J., Cheddadi, R., Tarasov, P., Reille, M., de Beaulieu, J.-L., Bottema, S. & Andrieu, V. (1998) Climatic reconstruction in Europe for 18,000 yr B.P. from pollen data. *Quaternary Research*, **49**, 183–196.

Peyron, O., Begeot, C., Brewer, S., Heiri, O., Magny, M., Millet, L., Ruffaldi, P., Van Campo, E. & Yu, G. (2005) Late-Glacial climatic changes in Eastern France (Lake Lautrey) from pollen, lake-levels, and chironomids. *Quaternary Research*, **64**, 197–211.

Pignatti, S. (1982) *Flora d'Italia*, 1st edn. Edagricole, Bologna.

Pignatti, S. (1997) *Ecologia del Paesaggio*. UTET, Turin.

Pigott, C.D. & Pigott, S. (1993) Water as a determinant of the distribution of trees at the boundary of the Mediterranean zone. *Journal of Ecology*, **81**, 557–566.

Piovesan, G. & Schirone, B. (2000) Winter North Atlantic oscillation effects on the tree rings of the Italian beech (*Fagus sylvatica* L.). *International Journal of Biometeorology*, **44**, 121–127.

Pons, A. & Quézel, P. (1998) A propos de la mise en place du climat méditerranéen. *Comptes Rendu de l'Academie des Sciences de Paris, Sciences de la Terre et des Planètes*, **327**, 755–760.

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. & Helmisaari, H. (1991) Silvics of north European trees: compilation, comparisons and implications for forest succession modelling. *Forest Ecology and Management*, **42**, 79–93.

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**, 507–519.

Punt, W., Blackmore, S. & Clarke, G.C.S. (1976–1996) *Northwest European pollen flora*. Volumes 1 (1976), 2 (1980), 3 (1981), 4 (1984), 5 (1988), 6 (1991), 7 (1996). Elsevier, Amsterdam.

Reille, M. (1992) *Pollen et spores d'Europe et d'Afrique du nord*. Laboratoire de Botanique Historique et Palynologie, Marseille.

Renberg, I. (1991) The HON-Kajak sediment corer. *Journal of Paleolimnology*, **6**, 167–170.

Rosen, P., Hall, R.I., Korsman, T. & Renberg, I. (2000) Diatom transfer-functions for quantifying past air temperature, pH and total organic carbon concentration from lakes in northern Sweden. *Journal of Paleolimnology*, **24**, 109–123.

Sadori, L., Giraudi, C., Petitti, P. & Ramrath, A. (2004) Human impact at Lago di Mezzano (central Italy) during the Bronze Age: a multidisciplinary approach. *Quaternary International*, **113**, 5–17.

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**, 251–267.

Sugita, S. (1994) Pollen representation of vegetation in Quaternary sediments — theory and method in patchy vegetation. *Journal of Ecology*, **82**, 881–897.

Telford, R.J. (2006) Limitations of dinoflagellate cyst transfer functions. *Quaternary Science Reviews*, **25**, 1375–1382.

Telford, R.J. & Birks, H.J.B. (2005) The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance. *Quaternary Science Reviews*, **24**, 2173–2179.

ter Braak, C.J.F. & Prentice, I.C. (1988) A theory of gradient analysis. *Advances in Ecological Research*, **18**, 271–317.

ter Braak, C.J.F. & Šmilauer, P. (2003) *CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5)*. Centre for Biometry Wageningen, Ithaca, NY.

Tinner, W., Hubschmid, P., Wehrli, M., Ammann, B. & Conedera, M. (1999) Long-term forest fire ecology and dynamics in southern Switzerland. *Journal of Ecology*, **87**, 273–289.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.-M., Hoegh-Guldberg, O. & Bairlein, F. (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–395.

BIOSKETCHES

Walter Finsinger is a post-doctoral researcher with interests in palaeoecology, vegetation dynamics, fire history, climate change, human impact on terrestrial ecosystems and the development of proxies for quantitative reconstructions.

Oliver Heiri is a post-doctoral researcher at Palaeoecology, Utrecht University, The Netherlands. His research interests include palaeolimnology, early human influence on lake ecosystems, Quaternary entomology and quantitative environmental reconstruction based on the fossil record.

Verushka Valsecchi has completed her PhD thesis at the Institute of Plant Sciences, University of Bern. Her interests are human impact and climate change in southern Europe and the Near East.

Willy Tinner is a lecturer in Geobotany and Palaeoecology at the University of Bern, Switzerland. His interests focus on the long-term interactions of climate, land-use, vegetation composition and the fire regime.

André F. Lotter is Professor of Palaeoecology at Utrecht University, The Netherlands. His research interests focus on the biotic response of terrestrial and aquatic systems to natural and anthropogenic disturbances during the late Quaternary and on quantitative environmental reconstruction.

Editor: Martin Sykes

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 Locations, names, abbreviations, and elevations of the 79 sites used in the study.

Appendix S2 Summary of the environmental variables of the 79-lake set.

Appendix S3 Species-specific response models of pollen taxa with more than 10 occurrences.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1466-8238.2007.00313.x

(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.