SCIENCE CHINA
Earth Sciences

•RESEARCH PAPER•

March 2021 Vol.64 No.3: 425–439
https://doi.org/10.1007/s11430-020-9693-y

Modern pollen assemblages of the surface lake sediments from the steppe and desert zones of the Tibetan Plateau

Feng QIN*

Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

Received May 19, 2020; revised September 21, 2020; accepted October 21, 2020; published online January 21, 2021

Abstract Modern pollen analysis is the basis for revealing the palaeovegetation and palaeoclimate changes from fossil pollen spectra. Many studies pertaining to the modern pollen assemblages on the Tibetan Plateau have been conducted, but little attention has been paid to pollen assemblages of surface lake sediments. In this study, modern pollen assemblages of surface lake sediments from 34 lakes in the steppe and desert zones of the Tibetan Plateau are investigated and results indicate that the two vegetation zones are dominated by non-arboreal pollen taxa and show distinctive characteristics. The pollen assemblages from the desert zone contain substantially high relative abundance of Chenopodiaceae while those from the steppe zone are dominated by Cyperaceae. Pollen ratios show great potential in terms of separating different vegetation zones and to indicate climate changes on the Tibetan Plateau. The *Artemisia*/Chenopodiaceae ratio and arboreal/non-arboreal pollen ratio could be used as proxies for winter precipitation. *Artemisia*/Cyperaceae ratio and the sum of relative abundance of xerophilous elements increase with enhanced warming and aridity. When considering the vegetation coverage around the lakes, hierarchical cluster analysis suggests that the studied sites can be divided into four clusters: meadow, steppe, desert-steppe, and desert. The pollen-based vegetation classification models are established using a random forest algorithm. The random forest model can effectively separate the modern pollen assemblages of the steppe zone from those of the desert zone on the Tibetan Plateau. The model for distinguishing the four vegetation clusters shows a weaker but still valid classifying power. It is expected that the random forest model can provide a powerful tool to reconstruct the palaeovegetation succession on the Tibetan Plateau when more pollen data from surface lake sediments are included.

Keywords Modern pollen assemblage, Tibetan Plateau, Surface lake sediment, Pollen-climate relationship, Pollen-based vegetation classification model

Citation: Qin F. 2021. Modern pollen assemblages of the surface lake sediments from the steppe and desert zones of the Tibetan Plateau. Science China Earth Sciences, 64(3): 425–439, https://doi.org/10.1007/s11430-020-9693-y

1. Introduction

Pollen analysis is one of the most important aspects of Quaternary botany, and has achieved significant contribution to modern ecology and biogeography (Birks, 2019). Recovering the past vegetation and climate changes from fossil pollen spectra must rely on the understanding of the relationship between modern pollen and contemporary vegetation and climate. Modern pollen datasets are the basis for developing numerical models of pollen-climate (Birks, 1998; Jackson and Williams, 2004; Birks et al., 2010) and pollen-vegetation (Sugita, 1994; Sugita, 2007a; Sugita, 2007b; Gaillard et al., 2008).

To gain insight into the relationship between pollen and vegetation/climate, many studies on modern pollen assemblages have been performed in different regions of China (e.g., Liu et al., 1999; Luo et al., 2010; Zhang et al., 2010; Zheng et al., 2014; Xu et al., 2016). The modern pollen

* Corresponding author (email: qinfeng@igsnrr.ac.cn)

assemblages of the Tibetan Plateau have received considerable attention due to the important role of the Tibetan Plateau in modifying the climate of Asia (An et al., 2001; Molnar et al., 2010; An et al., 2015).

Modern pollen analyses have been conducted at different spatial scales on the Tibetan Plateau. Some studies involved large amounts of samples from extensive areas of the Tibetan Plateau. Yu et al. (2001) illustrated the spatial patterns in relative abundance of pollen taxa, and explored their relationships with different vegetation zones. Shen et al. (2006) analysed the pollen-climate relationship based on modern pollen assemblages from the eastern Tibetan Plateau and developed the pollen-climate transfer functions. Herzschuh and her co-authors conducted a series of studies on the modern pollen assemblages of the surface lake sediments in 113 lakes from the central and eastern Tibetan Plateau and: (1) verified the reliability of pollen ratios for reflecting climate changes (Herzschuh, 2007); (2) evaluated the significance of pollen taxa for indicating modern vegetation types and climate variations (Herzschuh and Birks, 2010); and (3) developed transfer functions for reconstructing past climate changes on the Tibetan Plateau (Herzschuh et al., 2010). Lu et al. (2011) compiled a large number of modern pollen assemblages over the whole Tibetan Plateau and investigated the relationship between pollen and climate. Consequently, based on the modern pollen dataset, they established transfer functions for predicting environmental changes by using different algorithms and compared the performance of these transfer functions. Li et al. (2020) explored the relationships between distribution of tree pollen and the pathway of the Asian Summer Monsoon based on the modern pollen assemblages from the eastern Tibetan Plateau.

Except for these synthesised works, modern pollen assemblages are also studied at small spatial scales in different parts of the Tibetan Plateau, such as Kunlun Mountain in the northwest (Weng et al., 1993; Cour et al., 1999), the Qaidam Basin and adjacent areas (Wei et al., 2011; Zhao and Herzschuh, 2009) and the Lake Qinghai basin (Shang et al., 2009) in the northeast, a longitudinal transect across the central part (Zhang et al., 2015), lakes in the central and western part (Ma et al., 2017a; Ma et al., 2017b), two altitudinal transects in the eastern part (Zhang et al., 2017), the Lhasa Valley in the central-southern region (Zhang et al., 2018), and the Yadong area in the southern part (Zhang et al., 2020).

Notably, the modern pollen assemblages of the above studies are mainly from surface soil, and only a few are from surface lake sediments, moss polster and airborne pollen traps. Some studies (Wilmshurst and McGlone, 2005; Zhao et al., 2009; Lisitsyna et al., 2012; Qin et al., 2015) have demonstrated that different pollen signals can be observed in different types of sediment. Considering that most fossil pollen sequences are taken from the lake cores, modern pollen assemblages of the surface lake sediments are the most appropriate reference to interpret the fossil pollen data. However, only a few studies are focused on the modern pollen assemblages from surface lake sediments on the Tibetan Plateau, including Herzschuh and her co-authors' investigations in the central and eastern region (Herzschuh et al., 2010; Herzschuh, 2007; Herzschuh and Birks, 2010) and reports of Ma et al. (2017a, 2017b) in the central and western region. More studies are required to investigate the modern pollen signal in surface lake sediments on the Tibetan Plateau.

In this study, the pollen assemblages from surface lake sediments of 34 lakes in the steppe and desert zones on the Tibetan Plateau are presented, and numerical techniques have been used to: reveal the characters of modern pollen assemblages from the steppe and desert zones on the Tibetan Plateau; investigate the relationships between pollen ratios (*Artemisia*/Chenopodiaceae, *Artemisia*/Cyperaceae, and arboreal/non-arboreal) and climate factors; and evaluate the pollen-vegetation relationships and establish the pollen-based vegetation classification model through machine learning.

2. Materials and methods

2.1 Study region

The Tibetan Plateau covers an area of approximately 2.54 million km^2 (Zhang et al., 2014) with an average elevation of >4000 m a.s.l. Along the thermal and moisture gradient from southeast to northwest on the Tibetan Plateau, the zonation of vegetation generally shows a change of forest-meadow-steppe-desert (Zhang, 1978; Zhang, 2007). The studied lakes are located in the northern, central, and western Tibetan Plateau (Figure 1), scattered in the steppe zone and desert zone.

The steppe zone of the Tibetan Plateau is mainly dominated by plant from genera *Stipa* (*S. purpurea*, *S. bungeana*, *S. subsessiliflora* var. basiplumosa, *S. roborowskyii*, and *S. capillacea*), *Artemisia* (*A. wellbyi*, *A. younghusbandii*, and *A. stracheyi*) and *Carex* (*C. moorcroftii* and *C. montis-everestii*). Other important components include grasses such as *Littledalea racemose*, *Orinus thoroldii*, *Pennisetum flaccidum*, and *Aristida adscensionis*, and shrubs such as *Caragana versicolor*, *Sophora moorcroftiana*, *Sabina pingii* var. *wilsonii*, and *Potentilla fruticosa* (Zhang, 2007).

The dominant species of the desert zone of the Tibetan Plateau is mainly the plants from Chenopodiaceae family, such as *Ceratoides* (*C. compacta*, *C. latens*), *Salsola* (*S. abrotanoides*), *Haloxylon* (*H. ammodendron*), and *Kalidium* (*K. foliatum* and *K. cuspidatum*). Other xerophilous plants like *Ephedra* (*E. przewalskii*), *Zygophyllum xanthoxylon*, *Nitraria* (*N. roborowskii*, *N. sibirica*), *Ajania tibetica* and *Artemisia* (*A. rhodantha*, *A. arenaria*) can also be the

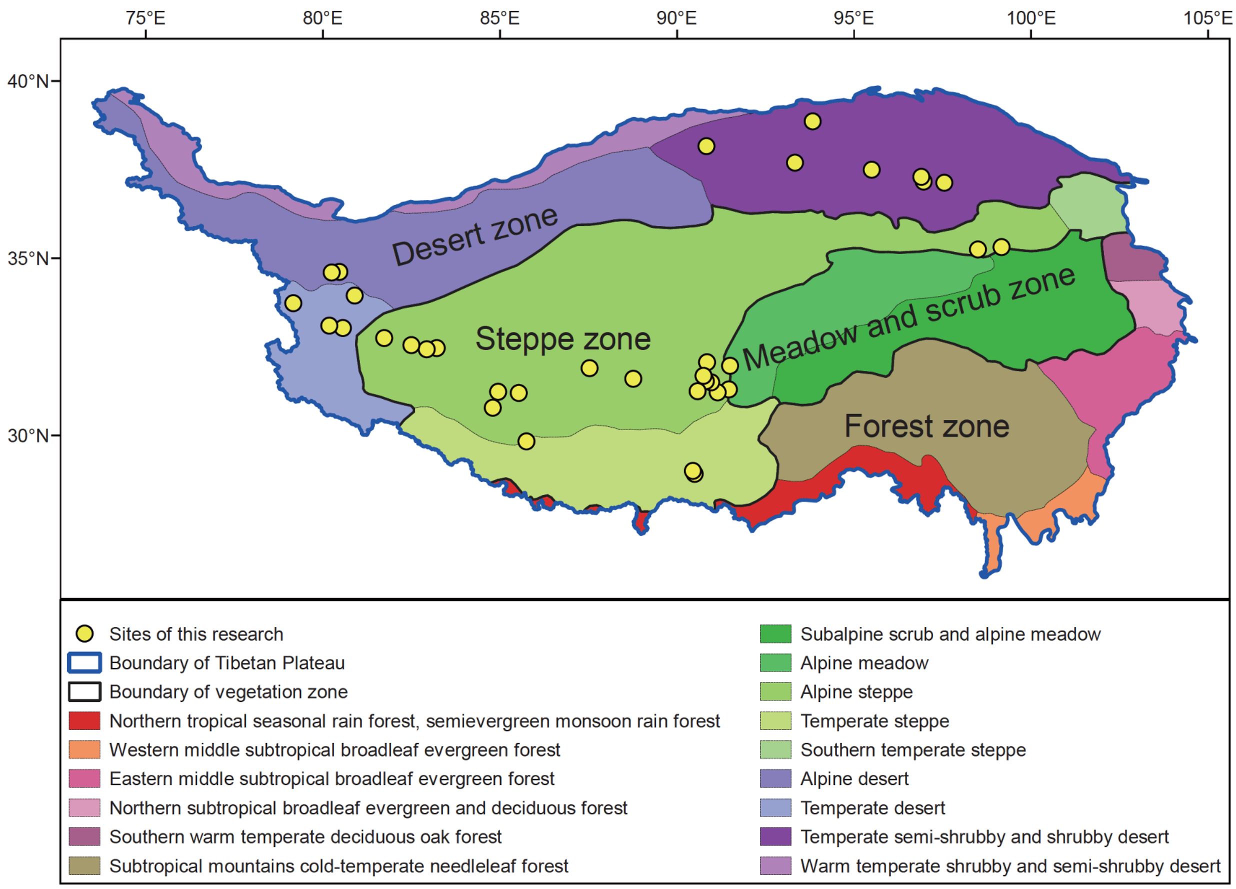

Figure 1 Map showing the locations of studied lakes and the vegetation zonation of the Tibetan Plateau. Boundary of Tibetan Plateau is from Zhang et al. (2014). Vegetation zonation is from The Editorial Committee of Vegetation Map of China, Chinese Academy of Sciences (2007).

dominant species. Some grasses like *Stipa purpurea* and *S. glareosa* and sedges like *Carex moorcroftii* together with the above-mentioned typical desert plants can form the desert-steppe vegetation (Zhang, 2007).

2.2 Samples and pollen analysis

Surface lake sediments (the uppermost 1–2 cm of lake sediment) were collected for pollen analysis at the centre of 34 Tibetan Plateau lakes (Figure 1, Table 1), of which 22 and 12 are located in the steppe zone and desert zone, respectively. Notably, four lakes from the steppe zone (Lake Kuhai, Lake Donggi Cona, Lake Cuona, and Lake Nairiping Co) are situated close to the boundary of the steppe and meadow zones.

Pollen samples were treated following a modified acetolysis procedure (Moore et al., 1991), including chemical treatments with 10% HCl, 10% NaOH, and 40% HF, followed by acetolysis treatment, sieving with a (10 μm) mesh in an ultrasonic bath, and finally mounting in glycerol. Before chemical treatment, a known number of exotic *Lycopodium* spores (12542 grains/sample) were added to each sample to enable estimation of pollen concentrations. The pollen grains were observed and counted using an optical microscope at 400× magnification. Pollen identifications were based on published palynological literature (Xi and Ning, 1994; Wang et al., 1995; Tang et al., 2016) and reference collections from the study area. At least 100 terrestrial pollen grains were counted for each sample with an average amount of 415 grains. The relative abundance (expressed as percentage) of each pollen taxon in a sample was calculated against the sum of all terrestrial pollen. Tilia 1.7.16 was used to construct the pollen diagram (Grimm, 2011). The taxonomic treatment of angiosperm families referred to the APG IV (The Angiosperm Phylogeny Group, 2016) with one exception for retaining Chenopodiaceae to facilitate the comparison with previous studies.

2.3 Climate and vegetation data

To explore the relationships between pollen assemblages and their corresponding climate, six climate factors for each site were extracted from interpolating the meteorological data, namely, the mean annual temperature (T_{ann}), mean July temperature (T_{Jul}), mean January temperature (T_{Jan}), mean annual precipitation (P_{ann}), mean July precipitation (P_{Jul}) and

Table 1 Locations of the sampling sites on the Tibetan Plateau[a)]

No.	Lakes	Latitude (°N)	Longitude (°E)	Elevation (m a. s. l.)	Area (km^2)*	Vegetation zone
1	Lake Kuhai	35.32	99.17	4132	44.4	steppe
2	Lake Donggi Cona	35.25	98.5	4086	232.2	steppe
3	Lake Cuona	31.97	91.5	4590	182.4	steppe
4	Lake Nairiping Co	31.3	91.47	4530	69.6	steppe
5	Lake Beng Co	31.22	91.15	4671	141.3	steppe
6	Lake Peng Co	31.5	90.97	4534	135.7	steppe
7	Lake Zigetang Co	32.07	90.85	4573	191.4	steppe
8	Lake Jiang Co	31.53	90.82	4603	36.1	steppe
9	Lake Daru Co	31.68	90.75	4688	54.2	steppe
10	Lake Bamu Co	31.25	90.58	4565	190.9	steppe
11	Lake Chen Co	28.92	90.5	4437	39.1	steppe
12	Lake Kongmu Co	29.48	90.45	4450	40.4	steppe
13	Lake Cuoe	31.6	88.77	4568	269	steppe
14	Lake Daze Co	31.9	87.53	4470	244.7	steppe
15	Lake Dajia Co	29.83	85.75	5151	114.5	steppe
16	Lake Qige Co	31.2	85.53	4667	20.3	steppe
17	Lake Dawa Co	31.23	84.95	4628	114.4	steppe
18	Lake Youbu Co	30.78	84.8	4645	64.3	steppe
19	Lake Darebu Co	32.47	83.22	4438	21	steppe
20	Lake Bieruoze Co	32.43	82.93	4407	33.2	steppe
21	Lake Zhacangchaka	32.55	82.5	4354	128.25	steppe
22	Lake Aweng Co	32.75	81.73	4430	58.6	steppe
23	Lake Gahai	37.13	97.55	2852	32	desert
24	Lake Tuosu	37.17	96.97	2808	165.9	desert
25	Lake Hurleg	37.3	96.9	2817	56.7	desert
26	Lake Xiaochaidan	37.5	95.5	3177	71.5	desert
27	Lake Sugan	38.87	93.83	2796	106	desert
28	Lake Xitaijinaier	37.7	93.33	2688	126	desert
29	Lake Gasikule	38.17	90.83	2857	123.8	desert
30	Lake Jiezechaka	33.95	80.9	4530	107.6	desert
31	Lake Rebang Co	33.03	80.57	4326	31.6	desert
32	Lake Longmu Co	34.62	80.47	5010	97	desert
33	Lake Sumxi Co	34.6	80.25	5057	24.6	desert
34	Lake Zongxiong Co	33.1	80.18	4351	7.2	desert

a) * Refer to Wang and Dou (1998).

mean January precipitation (P_{Jan}). The climate data of each lake were interpolated using a plate spline regression based on the meteorological records of thirty years (1981–2010) (http://data.cma.cn).

Vegetation data were derived from the 1:1,000,000 vegetation map of China (The Editorial Committee of Vegetation map of China, Chinese Academy of Sciences, 2007). The land cover grid data on the vegetation map were extracted by the buffers surrounding the lake shore with 1-km steps up to a distance of 50 km. A total of 90 land cover types were recorded (details see Supplementary Tables S1 and S2, http://link.springer.com), and they were grouped into 9 vegetation groups (conifer forest, shrubland, desert, desert-steppe, steppe, meadow, swamp, alpine vegetation, and cultural vegetation) and one non-vegetated group (including water bodies and bare land). The coverages of land cover were merged into coverages of the vegetation group, which were used as vegetation coverages in further analysis.

The plants growing nearby the sampling site contribute more pollen to the pollen assemblage than those growing distantly. Therefore, distance weighting was performed to correct the differences in vegetation contributions at different distances to the pollen assemblage (Gaillard et al., 2008). The vegetation coverage was weighted using the inverse distance weighting method (Prentice and Webb, 1986), which divided the vegetation coverages by their distances to the lake shore. Finally, the coverage percentages were calculated and used in the numerical analyses.

2.4 Numerical analyses

Ordination techniques were used to examine the relationships among pollen assemblages, as well as that between pollen assemblages and climate factors. A preliminary detrended correspondence analysis (DCA) showed the length of standard deviation unit on the first axis was 1.58. It indicated that a linear-model based ordination method would be suitable for analysing the pollen dataset; therefore, principal components analysis (PCA) was adopted. Only those pollen taxa with relative abundance higher than 2% in at least one sample were included in ordinations, and the pollen data were standardised by a square root transformation before running DCA and PCA.

Generally, the correlation between pollen and plant abundances improves with increasing survey extent of vegetation around the sediment site up to a given area beyond which the correlation does not continue to improve (Sugita, 2007b). This area was defined as the relevant source area of pollen (RSAP) by Sugita (1994), and the pollen assemblage can reflect the variation in vegetation within the RSAP (Sugita et al., 1999). The extent of RSAP can be estimated by using the extended *R*-value model and maximum likelihood method based on the abundances of modern pollen and plants (Parsons and Prentice, 1981; Sugita, 1994; Bunting et al., 2013). Unfortunately, detailed plant abundance information for the vegetation map was not available. Therefore, I adopted a compromised strategy to evaluate the correlation between pollen assemblage dataset and the vegetation coverage dataset using the Mantel test (Mantel, 1967; Legendre and Fortin, 1989), which is a procedure to analyse the relationship between two dissimilarity matrices. Vegetation data within different distances were tested with the pollen data to determine the appropriate distance at which the pollen assemblages and the vegetation coverages showed the highest correlation. Both raw vegetation coverages and distance-weighted coverage data were tested with pollen data. The percentages of pollen and vegetation coverages were normalised, then the "Bray-Curtis" dissimilarity matrices were calculated. The Mantel test was performed on the dissimilarity matrices of the two types of datasets, and the Mantel statistic was calculated based on Pearson's product-moment correlation.

Hierarchical cluster analysis was used to classify the sites based on their vegetation coverages. The coverage data were normalised and the "Bray-Curtis" dissimilarity were calculated. Clusters were constructed using the UPGMA (unweighted pair group method with arithmetic mean) algorithm.

Random forest (Breiman, 2001), a machine learning algorithm, was adopted to establish the pollen-vegetation model and evaluate the possibility of distinguishing vegetation types using pollen data. Two vegetation classification schemes were used as the labels for the sampling sites. One scheme simply divided the studied sites into two categories (desert and steppe) based on the vegetation zone in which they were located (vegetation zone scheme). The other scheme adopted a more detailed classification for the vegetation according to the hierarchical cluster analysis of vegetation coverage around the sampling sites (vegetation cluster scheme).

All numerical analyses were performed in R version 3.5.1 (R Core Team, 2018). Package vegen 2.5-2 (Oksanen et al., 2018) was involved in ordination analysis, the Mantel test, and hierarchical cluster analysis. The random forest algorithm was implemented with randomForest 4.6-14 (Liaw and Wiener, 2002).

3. Results

3.1 Pollen assemblages

A total of seventy-two terrestrial pollen taxa were found in the studied surface lake samples. Pollen grains derived from non-arboreal plants (non-arboreal pollen) dominated the pollen assemblages with an average relative abundance of 91.36% (ranged from 77.88% to 100%), among which *Artemisia*, Cyperaceae, Chenopodiaceae, and Poaceae were the most important. Arboreal pollen contributed a small proportion to the pollen assemblages (ranged from 0 to 19.82%, averaged 6.93%), among which *Pinus* and *Quercus* were the important elements. Pollen assemblages from different vegetational zones showed distinct characteristics (Figure 2).

Steppe zone: The pollen assemblages from the steppe zone were dominated by *Artemisia* (1.47–62.95%) and Cyperaceae (2.27–69.85%). Poaceae also represented a significant proportion that ranged from 2.27% to 18.81%. Chenopodiaceae (0.46–16.36%), *Aster*-type (0–4.59%), Ranunculaceae (0.45–6.36%), *Thalictrum* (0–3.38%), and Rosaceae (0–1.92%) were commonly found. *Anthemis*-type (0–19.09%), Crassulaceae (0–14.22%), and *Myricaria* (0–7.19%) usually accounted for negligible amounts of pollen, however, these represented high relative abundance in certain samples. Arboreal pollen was usually found, including *Pinus* (0.18–13.64%), *Quercus* (0–5.93%), *Alnus* (0–3.69%), *Betula* (0–3.69%), and *Picea* (0–2.73%).

Desert zone: Chenopodiaceae (15.71–56.19%) and *Artemisia* (7.71–44%) were dominant elements in the pollen assemblages from the desert zone. Cyperaceae generally had much lower relative abundances in the samples from the desert zone (1.61–31.41%) than those from steppe zone, but

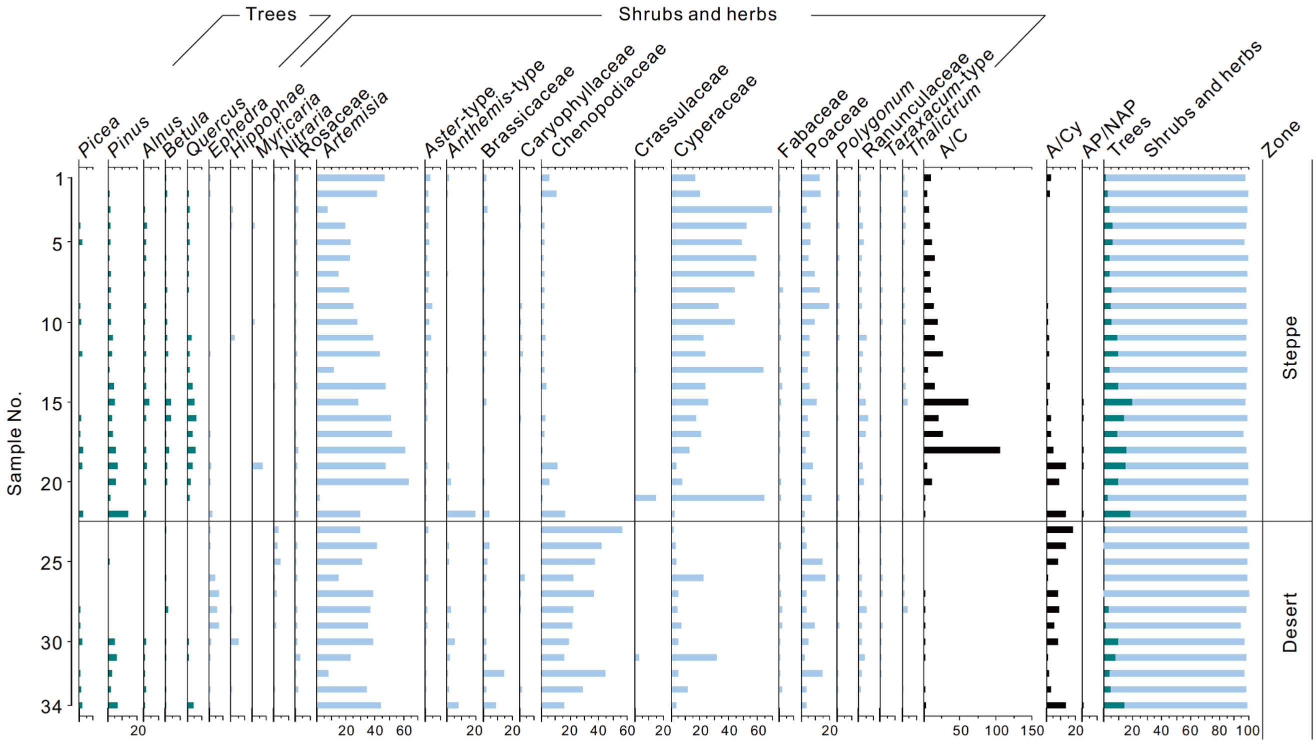

Figure 2 Pollen spectra of the steppe zone and desert zone on the Tibetan Plateau.

presented high relative abundances in two samples (22.22% in Lake Xiaochaidan and 31.41% in Lake Rebang Co.). Relative pollen abundance of Poaceae ranged from 2.09% to 16.19%. Brassicaceae (0–14.00%) and *Anthemis*-type (0–7.33%) could contribute important proportions in several samples. Xerophilous elements such as *Ephedra* (0–6.61%) and *Nitraria* (0–4.14%) were commonly found. Arboreal pollen (0–14.67%) covered lower proportions in the pollen assemblages of the desert zone than those of the steppe zone. Among the arboreal pollen, *Pinus* (0–6.67%) and *Quercus* (0–4%) were the most important.

3.2 PCA results

PCA suggested that the first PCA axis (axis-1 for short) represented 47.79% of the variations in the pollen data, and second PCA axis (axis-2 for short) explained 18.28% (Figure 3). Samples from desert zone lakes (desert samples for short) distributed on the left part of axis-1. Samples from the steppe zone lakes (steppe samples for short) were generally situated to the right of those from desert zone, although some of them could distribute on the negative side of axis-1. On axis-2, those desert samples mostly distributed in the lower part with two exceptions, while the steppe samples were found on both upper and lower parts.

Among the pollen taxa, Chenopodiaceae, Cyperaceae, and *Artemisia* showed the greatest variances. Poaceae and Cyperaceae, which were dominant in the steppe zone of Tibetan Plateau, were in the bottom right of the PCA diagram. The xerophilous taxa like Chenopodiaceae, *Ephedra*, and *Nitraria* located in the bottom left of the diagram. All arboreal taxa were located on the positive side of axis-2, including *Quercus*, *Pinus*, *Betula*, *Picea*, and *Alnus*.

For the climatic factors, P_{ann} and P_{Jul} showed positive correlation, while T_{ann} and T_{Jul} showed negative correlation with axis-1. Desert samples correspond to high T_{ann} and T_{Jul}, while steppe samples correlated with high P_{ann} and P_{Jul}.

3.3 Mantel test results

When contrasting pollen assemblages against the distance-weighted vegetation data, the Mantel test indicated that the correlation coefficient between pollen assemblage and vegetation coverage exhibited a rising trend with increasing buffer extent until reaching the peak at 9 km (Figure 4). Then, the correlation coefficient declined gradually. For the Mantel test including raw vegetation data, the correlation coefficient showed a rapid increase in the first 3 km, followed by a sharp decrease (Figure 4). The pollen assemblages showed stronger correlation with the distance-weighted vegetation coverages than the raw vegetation coverages. Consequently, the distance-weighted vegetation data were adopted for further analyses rather than the raw data. The vegetation coverages within a 9-km buffer around the lakes were thought to be the most appropriate to study the pollen-vegetation relationship.

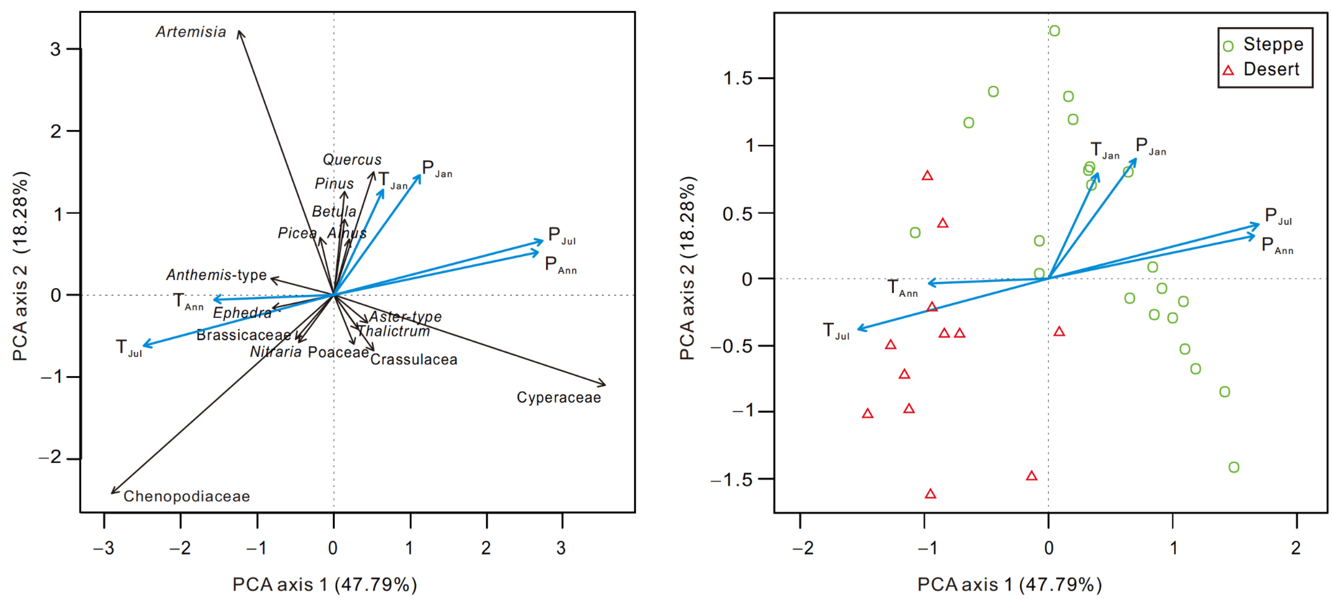

Figure 3 PCA plots of the pollen assemblage from the surface lake sediments on the Tibetan Plateau.

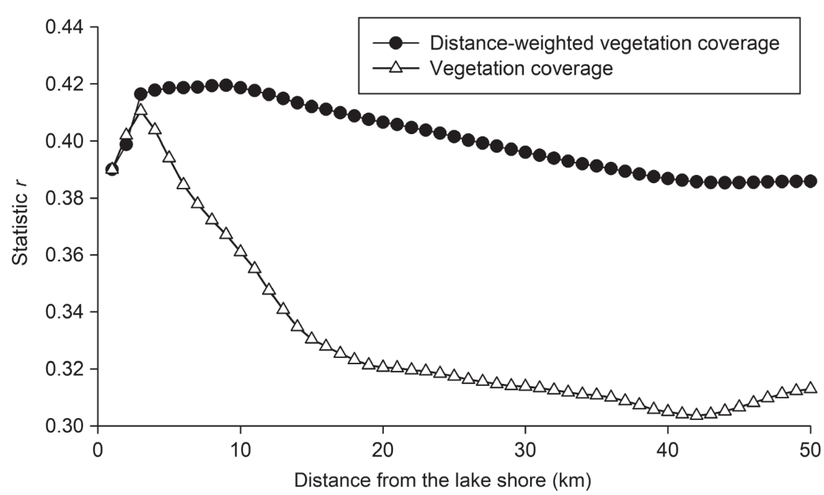

Figure 4 Relationship between pollen assemblage and vegetation revealed by the Mantel test.

3.4 Hierarchical cluster analysis results

Hierarchical cluster analysis on the vegetation coverage data within 9-km buffers around the lakes showed four main clusters (Figure 5). The four clusters could reflect the differences in vegetation around the studied lakes (Supplementary Figure S1). Cluster 1 contained eleven sites surrounded primarily by desert. The second cluster consisted of only two sites characterised by meadow. Cluster 3 included sixteen sites, and all of them were surrounded by significant amounts of steppe. The last cluster involved five sites, which were surrounded mainly by desert-steppe. Therefore, the vegetation of the studied lakes was assigned following the results of cluster analysis as desert, meadow, steppe and desert-steppe. This classification was adopted in the random forest algorithm.

3.5 Random forest results

When the vegetation of the sampling sites was simply divided into steppe or desert (vegetation zone scheme), the number of trees grown in the model and number of the pollen taxa randomly sampled as predict variables at each split were set to 100 and 20, respectively. The model showed a robust prediction power with a very small out-of-bag (OOB) estimate of error rate (2.94%). Nearly all samples were correctly classified, except one steppe sample, which was misidentified as desert (Table 2). The importance of pollen taxa to accurately distinguish the vegetation zones were shown in Figure 6. Chenopodiaceae, with a substantially higher mean

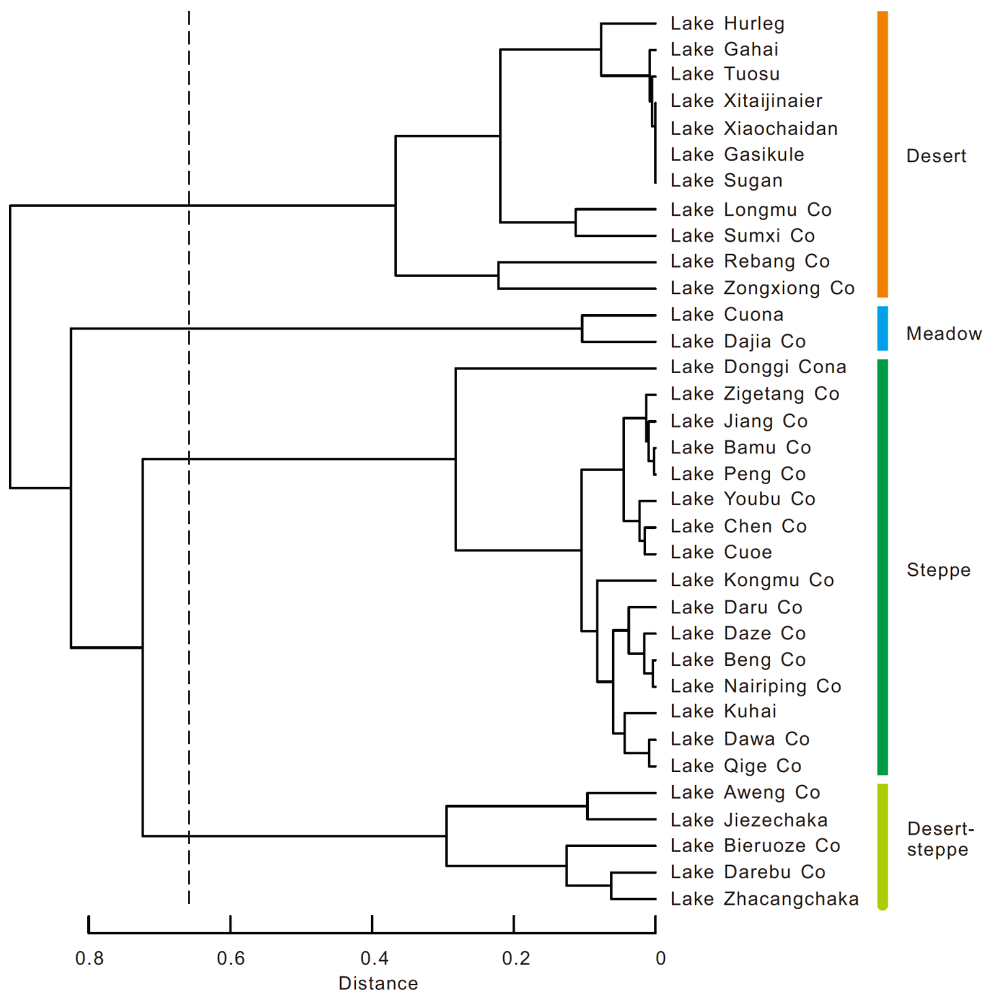

Figure 5 Dendrogram of hierarchical cluster analysis on the vegetation data within 9-km buffers.

Table 2 Confusion matrix showing the performance of the random forest model for classifying the steppe and desert zones of the Tibetan Plateau using modern pollen assemblages

Observed versus predicted vegetation zone (Model OOB error 2.94%)		Predicted vegetation zone		OOB error
		steppe	desert	
Observed vegetation zone	steppe	21	1	0.045
	desert	0	12	0

decrease in accuracy (19.39) than other pollen taxa, played the most important role in classifying the pollen assemblages.

The model for classifying the four vegetation clusters was developed by growing 400 trees with the number of the pollen taxa randomly sampled as predict variables at each split set as five. The performance of the vegetation cluster model, with an OOB estimate of error rate of 23.53%, was not as good as that of vegetation zone model. All of the steppe samples were correctly identified, while all meadow samples were misidentified as steppe (Table 3). Two desert samples out of eleven were misidentified as desert-steppe. Desert-steppe samples fell into three categories, desert (2 samples), desert-steppe (1 sample), and steppe (2 samples). Chenopodiaceae was still the most important element with the highest mean decrease in accuracy of 11.96 (Figure 6). *Ephedra* (8.99) and Cyperaceae (8.51) contributed significantly to the accuracy of the model. *Anthemis*-type (6.97), *Alnus* (6.14), and *Pinus* (5.31) also played important roles.

4. Discussion

4.1 Modern pollen assemblages from the steppe and desert zones of the Tibetan Plateau

Non-arboreal taxa dominated the pollen assemblages of the

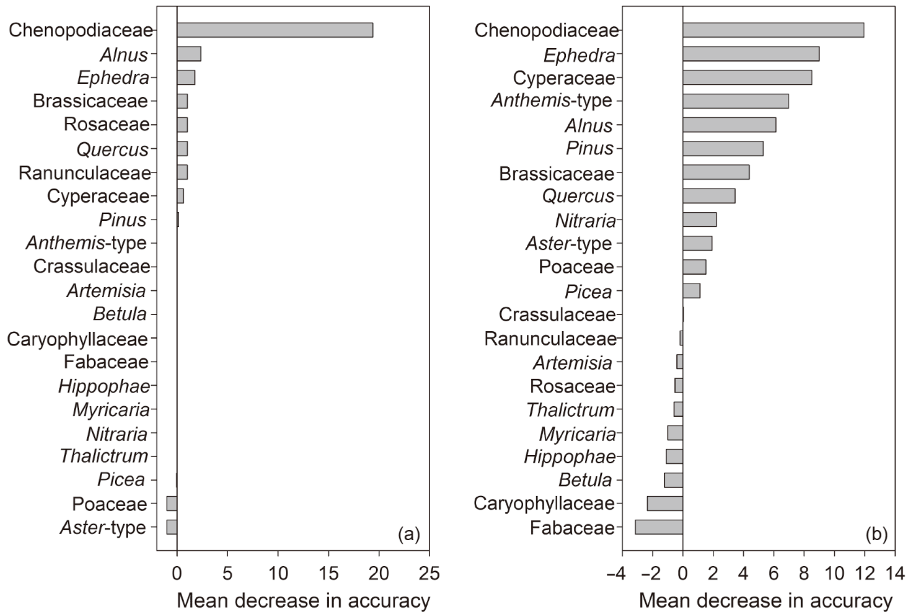

Figure 6 Importance of pollen taxa in pollen-based classification models for (a) vegetation zones and (b) vegetation clusters measured by the mean decrease in accuracy.

Table 3 Confusion matrix showing the performance of the random forest model for classifying vegetation clusters of the Tibetan Plateau using modern pollen assemblages

Observed versus predicted vegetation zone (Model OOB error 23.53%)		Predicted vegetation cluster				OOB error
		meadow	steppe	desert-steppe	desert	
Observed vegetation cluster	meadow	0	2	0	0	1
	steppe	0	16	0	0	0
	desert-steppe	0	2	1	2	0.80
	desert	0	0	2	9	0.18

surface lake sediments from both the steppe zone and the desert zone of the Tibetan Plateau, which represented the vegetation components of open landscape. The pollen assemblages from the two different vegetation zones share some common characteristics (Figure 2). *Artemisia* maintains remarkable proportions in modern pollen assemblages of both vegetation zones, and Poaceae pollen always contributes a certain amount. Some other pollen taxa contribute similar proportions to modern pollen assemblages of the two different vegetation zones, including Rosaceae, Ranunculaceae, *Thalictrum*, and *Polygonum*.

Obvious differences exist between the modern pollen assemblages of the two vegetation zones. Substantially higher relative abundances of Chenopodiaceae characterised the pollen assemblages of desert zone of the Tibetan Plateau, and the xerophilous elements, like *Ephedra* and *Nitraria*, exhibit higher relative abundances in desert samples. These taxa, which are drought-tolerant plants, prevail in the Tibetan Plateau desert. In steppe samples, Cyperaceae becomes the dominant element. The pollen of broadleaved trees, such as *Quercus*, *Betula*, and *Alnus*, presented with higher relative abundances than in desert samples.

The differences of modern pollen assemblages of the surface lake sediments between steppe zone and desert zone on the Tibetan Plateau are also seen in the published studies (sampling sites see Supplementary Figure S2). All investigations of modern pollen assemblages of the surface lake sediments indicate the dominance of Chenopodiaceae pollen with high relative abundance of *Artemisia* pollen in the desert vegetation on the Tibetan Plateau (this study; Herzschuh, 2007; Ma et al., 2017a, 2017b). In contrast, Chenopodiaceae pollen never exceeds 20% in the modern pollen assemblages of the surface lake sediments from the steppe vegetation. Instead, Cyperaceae, *Artemisia*, and Poaceae become the dominant pollen taxa, although some differences in their relative abundances can be observed in different parts of the steppe zone. The steppe sites of Herzschuh (2007) are located in the east of the steppe zone (Supplementary Figure S2), and her results show that Cyperaceae pollen plays a more important role than Poaceae and *Artemisia*. However, modern pollen assemblages from the west of the steppe zone (this study; Ma et al., 2017a) are dominated by *Artemisia* pollen, while Cyperaceae and Poaceae are not as important as those in the east part. This phenomenon can be attributed to the increase of aridity along a southeast-northwest gradient on the Tibetan Plateau (Wu et al., 2005), which may result in the flourish of Cyperaceae plant in the east of the steppe zone and the thriving of *Artemisia* in the west.

Modern pollen assemblages of the surface lake sediments from the steppe zone on the Tibetan Plateau are distinct from those of the temperate steppe region in northern China, although they have some similar characteristics. For instance, *Artemisia* pollen is also the dominant element, and Poaceae always occurs in a certain amount in the surface lake sediments of the temperate steppe region (Zhao et al., 2009; Han et al., 2017). However, obvious differences can be found in the relative abundances of Cyperaceae and Chenopodiaceae in the two steppe regions. Cyperaceae dominates the modern pollen assemblages from the steppe zone of the Tibetan Plateau, but is rarely an important element in the modern pollen assemblages of the temperate steppe region in northern China. The large amount of Cyperaceae pollen in the lake sediments of the steppe zone on the Tibetan Plateau should be attributed to the Cyperaceae plants as important components of the alpine steppe community (mainly *Carex*), or as dominant species in the alpine meadow community (mainly *Kobresia*) scattered around the lake. On the contrary, Chenopodiaceae accounts for a small proportion in the pollen assemblages of the steppe zone of the Tibetan Plateau, but can occupy significant relative abundance in the samples from the temperate steppe region (Zhao et al., 2009; Han et al., 2017). Studies on modern pollen assemblages of the temperate steppe region have suggested that high relative abundance of Chenopodiaceae pollen is correlated with intense human disturbance (Liu et al., 2006; Zhang et al., 2010). Therefore, one possible reason for lower Chenopodiaceae pollen in surface lake sediment from the steppe zone of the Tibetan Plateau is that the Tibetan Plateau sustained fewer human activities than the temperate steppe region.

Unlike the relationship of modern pollen assemblages between the steppe zone of the Tibetan Plateau and the temperate steppe region, modern pollen assemblages of the surface lake sediments from the desert zone of the Tibetan Plateau are similar to those from the warm temperate and temperate desert regions in northwestern China. Chenopodiaceae and *Artemisia* are also the dominant taxa in pollen assemblage of the warm temperate and temperate desert regions, while xerophilous taxa such as *Ephedra* and *Nitraria* are frequently found (Zhao et al., 2009; Qin et al., 2015). This implies that the dominant taxa and characteristic elements of the pollen assemblages of the surface lake sediments from deserts in China are generally the same despite their geographic differences.

In summary, the modern pollen assemblages of the surface lake sediments from the steppe zone are distinct from those of the desert zone on the Tibetan Plateau.

4.2 Pollen ratios and their indications to vegetation and climate

PCA results (Figure 3) show that P_{ann}, P_{Jul}, and T_{Jul} are the most important climate factors controlling the variations in pollen assemblages. PCA axis-1 explains nearly half of the pollen dataset variations, while the three climate factors show the highest correlation with PCA axis-1. The strong influence of P_{ann} on modern pollen assemblages of the Tibetan Plateau was also seen in previous studies like Shen et al. (2006), Herzschuh et al. (2010), and Herzschuh and Birks (2010). Samples from the steppe zone of the Tibetan Plateau correlated to higher P_{ann} and P_{Jul} and lower T_{Jul}, that is, wetter and cooler conditions. By contrast, samples from the desert zone correlated to higher T_{Jul}, and lower P_{ann} and P_{Jul}, that is, dryer and warmer environments. Therefore, the pollen assemblages can reflect the environmental differences between the two vegetation zones of the Tibetan Plateau.

Pollen ratios are useful indices to indicate vegetation and/or climate changes, among which ratios of *Artemisia*/Chenopodiaceae (A/C), arboreal/non-arboreal pollen (AP/NAP), and *Artemisia*/Cyperaceae (A/Cy) are the most frequently used ones with respect to the Tibetan Plateau.

A/C is often used as a humidity index in arid and semi-arid regions (El-Moslimany, 1990; Zhao et al., 2012). The correlation analysis (Table 4) suggests that A/C shows strong positive correlation with mean January precipitation (P_{Jan}, r^2=0.50). The possible explanation for the close relationship between A/C and winter precipitation is that the higher winter precipitation leads to greater winter snow depth in the study region, which in turn may bring about higher growing season soil moisture and nutrient availability for the vegetation (Peng et al., 2010). Meanwhile, the results show that A/C is weakly correlated with mean annual precipitation (P_{ann}, r^2=0.08) and mean July precipitation (P_{Jul}, r^2=0.07). Herzschuh (2007) indicated that A/C had a significant correlation with P_{ann} (r^2=0.25) in pollen assemblages of surface lake sediments from the eastern and central Tibetan Plateau. Ma et al. (2017b) also observed a high correlation between A/C and P_{ann} (r^2=0.56) in the central and western Tibetan Plateau. Ma et al. (2017a) found no obvious correlation between A/C and P_{ann} (r^2=0.21) in the southwestern Tibetan Plateau. Thus, A/C may be used to reflect winter precipitation change on the Tibetan Plateau, however, it should be used with caution to indicate mean annual precipitation.

Furthermore, A/C in samples from the steppe zone of the Tibetan Plateau (average 17.97) is much higher than that from the desert zone (1.24), although some exceptions are observed (Figure 7). This phenomenon is consistent with previous observations. Herzschuh (2007) suggested that A/C

Table 4 Coefficients of determination (r^2) between pollen ratios and climate factors[a)]

	P_{Jan}	P_{Jul}	P_{ann}	T_{Jan}	T_{Jul}	T_{ann}
A/C	0.50***	0.07	0.08	0.19***	0.03	0.005
	+	+	+	+	−	+
A/Cy	0.08	0.36***	0.37***	0.0001	0.43***	0.23**
	−	−	−	−	+	+
AP/NAP	0.26**	0.001	0.0002	0.09	0.02	0.004
	+	+	+	+	−	−
Sum_{xero}	0.26**	0.45***	0.41***	0.001	0.63***	0.40***
	−	−	−	−	+	+

a) * p<0.05, ** p<0.01, *** p<0.001, + positive correlation, − negative correlation.

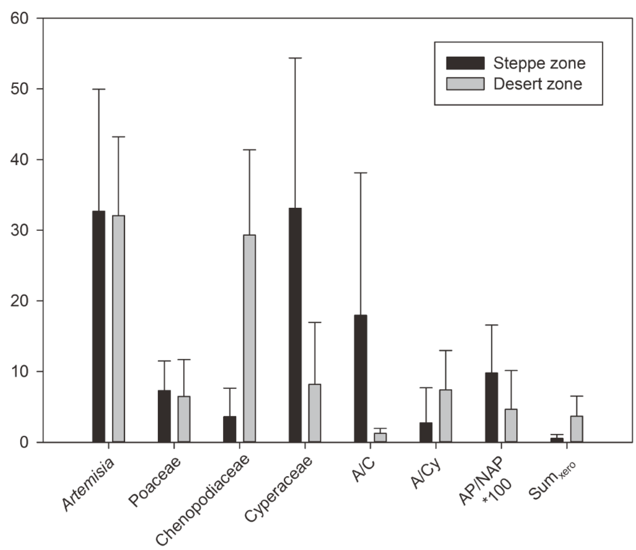

Figure 7 Values of dominant pollen taxa and pollen ratios for the steppe and desert zones on the Tibetan Plateau. Error bar=standard deviation.

values increased from desert to steppe in the eastern and central Tibetan Plateau. Ma et al. (2017b) indicated that an A/C value of 1.2 could be the threshold to separate desert and steppe vegetation in the western and central Tibetan Plateau. Therefore, A/C has great potential as a useful index to distinguish steppe from desert on the Tibetan Plateau.

The arboreal/non-arboreal pollen ratio (AP/NAP) was initially used to indicate the openness of the landscape around the sampling site. In this study, all samples were from the open landscape. The nearest forest community was recorded at least 20 km from the lake (Lake Hurleg) on the vegetation map. Arboreal pollen (AP) should be transported far from the source plants to the studied lakes. Previous studies have discussed the upward transportation of arboreal pollen on the Tibetan Plateau (Cour et al., 1999; Herzschuh and Birks, 2010; Zhang et al., 2017). Some authors also find the potential of arboreal pollen to indicate vegetation and climate changes on the Tibetan Plateau. Herzschuh (2007) supposed that sum of arboreal pollen alone could indicate the regional precipitation changes. Lu et al. (2008) supposed a threshold percentage for *Picea* and *Abies* pollen to recognise the tree line. Li et al. (2020) illustrated that the relative abundance of AP was highly coupled with the path and intensity of the Asian Summer Monsoon.

In this study, relative abundances of arboreal pollen range from 0 to 19.82% with an average of 6.93%. Accordingly, AP/NAP only presented very low values (0–0.25). Steppe samples yielded a higher average AP/NAP than desert samples (Figure 7), which may partly explained by the differences of their distance to the forest in the southeastern part of the Tibetan Plateau (Figure 1). However, some desert samples from the western Tibetan Plateau (lakes Jiezechaka, Rebang Co, Longmu Co, Sumxi Co and Zongxiong Co) can yield AP/NAP ratios as high as the steppe samples. Cour et al. (1999) found that AP/NAP could be high in areas with low plant coverage, because of the very low pollen input from local plants. This should also be true in the desert samples of this study. Moreover, the correlation analysis only yielded a relatively strong positive correlation (r^2=0.26) for AP/NAP and P_{Jan}, which suggests that AP/NAP may be used as an indicator for winter precipitation.

The A/Cy was introduced by Herzschuh et al. (2006) to be used as a summer temperature index in the central and eastern Tibetan Plateau based on the pollen percentage distribution of the two taxa. Later, research by Herzschuh (2007) on the pollen assemblages of surface lake sediments in the central and eastern Tibetan Plateau revealed that A/Cy was <1 in most alpine meadow samples and >1 in temperate steppe and desert samples. She also found that A/Cy was significantly correlated with T_{Jul} and weakly correlated with P_{ann}. Zhao and Herzschuh (2009) confirmed that A/Cy was much lower in samples from meadow than those from steppe, steppe-desert, and desert in the Qaidam Basin, northeastern Tibetan Plateau. A/Cy in their pollen assemblages showed a positive correlation with T_{Jul} and a negative correlation with P_{ann}. Ma et al. (2017a) found an increasing trend of A/Cy from alpine meadow and steppe ecotone, to sub-alpine shrub steppe, to alpine steppe, then to mountain and alpine desert in the central and western Tibetan Plateau. Their results also indicated a significant correlation between A/Cy and P_{ann}.

Results of this study generally show much lower A/Cy in samples from the steppe zone than those from the desert zone (Figure 7), which implies it can be used to distinguish steppe from desert on the Tibetan Plateau. The correlation analysis (Table 4) suggests that A/Cy is negatively correlated with P_{ann} (r^2=0.37) and P_{Jul} (r^2=0.36), but positively correlated with T_{ann} (r^2=0.23) and T_{Jul} (r^2=0.43). Therefore, A/Cy seems to increase with drying and warming conditions.

Pollen from xerophilous plants, including *Ephedra*, *Nitraria*, *Tamarix* and Zygophyllaceae, characterised the modern pollen assemblages of the desert zone, although their relative abundance never exceeded 7% in a single sample (Figure 2). The sum of the relative abundances of these xerophilous taxa (Sum_{xero} for short hereafter) is much higher in samples from the desert zone (average 3.69%) than those from the steppe zone (average 0.55%) (Figure 7). Thus, Sum_{xero} is supposed to be a useful index to separate desert from steppe on the Tibetan Plateau. In addition, Sum_{xero} shows strong positive correlation with T_{Jul} (r^2=0.63) and T_{ann} (r^2=0.40), while it is negatively correlated with the P_{ann} (r^2=0.41), P_{Jul} (r^2=0.45), and P_{Jan} (r^2=0.26) (Table 4). Higher values of Sum_{xero} correspond to more arid and warmer condition.

In conclusion, A/C, A/Cy, and Sum_{xero} can be used to distinguish steppe and desert on the Tibetan Plateau. A/C and AP/NAP could be used as winter precipitation index. A/Cy and Sum_{xero} increases with enhanced aridity and warming.

4.3 Pollen-vegetation relationship

4.3.1 Relationship between modern pollen assemblage and vegetation

The vegetation type for a pollen sample is normally recognised based on the field survey at the sampling site. However, the vegetation type for a sample collected from lakes is hard to determine following this strategy, as it is arduous to survey all vegetation around the lake. Therefore, customarily, the vegetation type for a lake sample is assigned according to the vegetation zone/region in which the lake is located. This assignment method loses detailed vegetation information, because different types of vegetation can be simultaneously distributed around a lake.

The studied lakes can be surrounded by at most six vegetation formations within a 1-km buffer around the lake shore on the 1:1,000,000 vegetation map. Considering the detailed vegetation information can improve understanding of the pollen-vegetation relationships. For instance, the substantially higher A/C ratio of Lake Zongxiong Co and Lake Jiezechaka than other desert lakes could be attributed to the considerable amount of desert-steppe and steppe around them (see Supplementary Figure S1), although both lakes are located in the desert zone. In contrast, Lake Aweng Co is located in the steppe zone, but its modern pollen assemblage contained a high percentage of Chenopodiaceae, similar to those in the desert zone lakes. This should derive from the significant amount of desert-steppe surrounding the lake.

Another problem of vegetation type assignment for pollen samples is that the coverages of different vegetation can change with increasing extent of vegetation survey. The types of vegetation formation increased with increasing buffer area surrounding the lake, and the dominant formation (the formation covered the largest area) also changed. For example, six kinds of formations occurred within a 1-km buffer of Lake Tuosu, while the number of formation types increased to eleven within a 10-km buffer. Consequently, the most important vegetation switched from *Nitraria sibirica* desert to *Haloxylon ammodendron* desert.

In this study, the most appropriate extent for studying the pollen-vegetation relationship was determined using the Mantel test. The results indicated that the vegetation coverage within a 9-km buffer around the lakes showed the strongest correlation with modern pollen assemblages on the Tibetan Plateau. This extent is not the same as the relevant source area of pollen (RSAP, Sugita, 1994), which concerns the relationship between abundance of pollen and plant. Nevertheless, it could still be a substitute to RSAP when the procedure for estimating RSAP is not applicable. Therefore, the vegetation types of the studied lakes were assigned based on the vegetation coverages within a 9-km buffer around them.

The cluster analysis suggested that the studied sites were merged into four clusters based on the vegetation coverages within a 9-km buffer. The sampling sites within one cluster were all dominated by the same vegetation type, although they could be located in different vegetation zones (Supplementary Figure S1). Therefore, each cluster was thought to represent one of the four vegetation types: desert, desert-steppe, steppe and meadow.

The pollen assemblages of the four clusters show distinct characteristics (Figure 8). The meadow cluster has the highest average relative abundance of Cyperaceae and the lowest average relative abundance of Chenopodiaceae. In contrast, the desert cluster yields the lowest Cyperaceae and highest Chenopodiaceae. Additionally, the desert-steppe cluster contains a higher Chenopodiaceae percentage than the steppe cluster, but lower Cyperaceae percentage than it. A/C exhibits a decreasing trend when the vegetation type changes from meadow, steppe, desert-steppe, to desert, which corresponds to the humidity variation in the four vegetation types. A/Cy shows higher value in the desert-steppe and desert clusters than in the meadow and steppe clusters. The desert cluster consists of higher Sum_{xero} than the other

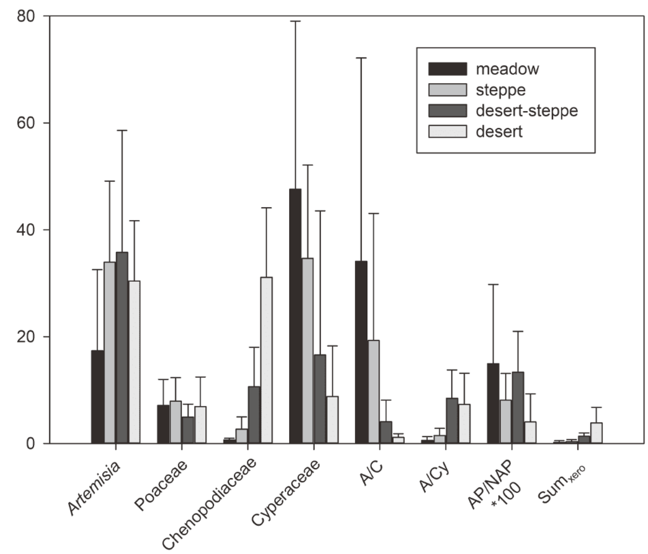

Figure 8 Values of dominant pollen taxa and pollen ratios for the four vegetation clusters. Error bar=standard deviation.

three clusters.

4.3.2 Pollen-based vegetation classification model

Sobol and Finkelstein (2018) developed pollen-based vegetation classification models for Africa and Arabia using different algorithms and tested their performance of classifying vegetation types. They found that the random forest model was the most accurate and precise among different models. Then, using the random forest algorithm, Sobol et al. (2019) established pollen-based vegetation classification models in southern Africa, and applied the models to fossil pollen spectra of Wonderkrater for reconstructing its palaeovegetation succession. Here, I used the random forest algorithm to develop models for classifying the alpine vegetation on the Tibetan Plateau through pollen data.

At the vegetation zone level, the random forest model exhibited a very high accuracy with a very low OOB estimate of error rate (Table 2). The only misidentification (Lake Aweng Co) may be attributed to the significant contribution of the desert community distributed around the lake (Supplementary Figure S1), although it is located in the steppe zone.

The performance of the random forest model for classifying the four vegetation clusters was not as good as that for the vegetation zones (Table 3). Samples of the steppe cluster are all correctly identified, but the other three clusters all included misidentified samples. Two misidentified samples in the desert cluster fall into desert-steppe cluster. The reason may be that desert and desert-steppe share some common taxa (Zhang, 2007). For instance, *Stipa glareosa* desert-steppe can include a large number of xerophilous species such as *Ceratoides latens* and *Ajania fruticulosa*. Meanwhile, these xerophilous species are important elements in some desert communities. For the same reason, the desert-steppe can present features of both desert and steppe. Consequently, samples of desert-steppe encounter the most confused classification that classified the samples into three categories. Samples of the meadow cluster were all wrongly assigned to steppe. The possible reasons may be that: (1) Cyperaceae, the dominant plant family in meadows, showed no distinguishing relative abundance in pollen assemblages of the meadow cluster compared to the steppe cluster (Figure 2); (2) many lakes in the steppe cluster are surrounded by significant amounts of meadow vegetation (Supplementary Figure S1), which introduced obvious pollen signal of the meadow to the steppe samples; and (3) the meadow cluster contains only two samples, which may not offer enough representativeness of this cluster to the random forest algorithm.

It should be noted that the sample size for developing the vegetation classification model is small, especially when the samples are divided into four groups. A better performance can be expected when more samples are included in the training set in the future.

According to the results of the random forest, Chenopodiceae shows extraordinary importance in separating the two vegetation zones on the Tibetan Plateau, while the other pollen taxa contribute much less to the accuracy of the classification model (Figure 6). This confirms that the relative abundance of Chenopodiaceae pollen is crucial for distinguishing steppe and desert on the Tibetan Plateau. Chenopodiaceae are also important in classifying the four vegetation clusters. However, the importance of the other pollen taxa increases when categories for classifying increase from two to four. Cyperaceae and *Ephedra* have substantial influence on the accuracy of the classification model for the four vegetation clusters. Meanwhile, *Pinus*, *Alnus*, and *Anthemis*-type also show great importance although their relative abundances in the modern pollen assemblages are often low.

In summary, the vegetation data within a 9-km buffer around the lake is suitable for exploring the relationship between pollen and vegetation in this study. Pollen-based vegetation classification models established using the random forest algorithm are shown to be promising and have potential to be used to reconstruct the past vegetation based on fossil pollen spectra on the Tibetan Plateau.

5. Conclusion

The modern pollen assemblages of surface lake sediments can be used to distinguish the steppe and desert zones of the Tibetan Plateau, although some common characteristics can also be found between modern pollen assemblages of the two vegetation zones. The *Artemisia*/Chenopodiaceae ratio and arboreal/non-arboreal pollen ratio could be used to indicate winter precipitation change. High value of *Artemisia*/Cyperaceae ratio and the sum of relative abundance of xerophilous elements can reflect warm and arid conditions. The pollen-based vegetation classification models are established by using a random forest algorithm, which can effectively separate the steppe zone from the desert zone on the Tibetan Plateau based on modern pollen assemblages. When the vegetation types of the studied sites are divided into four clusters according to the hierarchical cluster analysis, the random forest model shows a weaker but still valid classifying power. The random forest algorithm has significant potential to develop the pollen-based vegetation classification model for reconstructing the past vegetation succession based on fossil pollen spectra on the Tibetan Plateau.

Acknowledgements *The author wishes to thank Prof. Yan Zhao (Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China) for her suggestions regarding the manuscript, Prof. Juzhi Hou (Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China) for his generous provision of the samples, Dr. M. Jane Bunting (University of Hull, Hull, UK) for her helpful discussion, Dr. Jian Yang (Institute of Botany, Chinese Academy of Sciences, Beijing, China) for his suggestions on the numerical analysis, Dr. Chen Liang (Heibei GEO University, Shijiazhuang, China) for her help with interpolating the climate data, and two anonymous reviewers for their constructive comments. This work was supported by the National Natural Science Foundation of China (Grant Nos. 41671202 & 41690113), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA20070101) and the National Key Research and Development Program of China (Grant No. 2016YFA0600501).*

References

An Z S, Kutzbach J E, Prell W L, Porter S C. 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. Nature, 411: 62–66

An Z S, Wu G X, Li J P, Sun Y B, Liu Y, Zhou W J, Cai Y J, Duan A M, Li L, Mao J Y, Cheng H, Shi Z G, Tan L C, Yan H, Ao H, Chang H, Feng J. 2015. Global monsoon dynamics and climate change. Annu Rev Earth Planet Sci, 43: 29–77

Birks H J B. 2019. Contributions of Quaternary botany to modern ecology and biogeography. Plant Ecol Diversity, 12: 189–385

Birks H J B. 1998. Numerical tools in palaeolimnology—Progress, potentialities, and problems. J Paleolimn, 20: 307–332

Birks H J B, Heiri O, Seppä H, Bjune A E. 2010. Strengths and weaknesses of quantitative climate reconstructions based on late-Quaternary biological proxies. Open Ecol J, 279–280: 52–52

Breiman L. 2001. Random forests. Machine Learning, 45: 5–32

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. Quat Sci Rev, 82: 41–55

Cour P, Zheng Z, Duzer D, Calleja M, Yao Z. 1999. Vegetational and climatic significance of modern pollen rain in northwestern Tibet. Rev Palaeobot Palynol, 104: 183–204

El-Moslimany A P. 1990. Ecological significance of common nonarboreal pollen: Examples from drylands of the Middle East. Rev Palaeobot Palynol, 64: 343–350

Gaillard M J, Sugita S, Bunting M J, Middleton R, Broström A, Caseldine C, Giesecke T, Hellman S E V, Hicks S, Hjelle K, Langdon C, Nielsen A B, Poska A, von Stedingk H, Veski S. 2008. The use of modelling and simulation approach in reconstructing past landscapes from fossil pollen data: A review and results from the POLLANDCAL network. Veget Hist Archaeobot, 17: 419–443

Grimm E C. 2011. Tilia 1.7.16. Springfield: Illinois State Museum

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. Holocene, 27: 1567–1577

Herzschuh U. 2007. Reliability of pollen ratios for environmental reconstructions on the Tibetan Plateau. J Biogeogr, 34: 1265–1273

Herzschuh U, Birks H J B. 2010. Evaluating the indicator value of Tibetan pollen taxa for modern vegetation and climate. Rev Palaeobot Palynol, 160: 197–208

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. J Biogeogr, 37: 752–766

Herzschuh U, Winter K, Wunnemann B, Li S. 2006. A general cooling trend on the central Tibetan Plateau throughout the Holocene recorded by the Lake Zigetang pollen spectra. Quat Int, 154-155: 113–121

Jackson S T, Williams J W. 2004. Modern analogs in Quaternary paleoecology: Here today, gone yesterday, gone tomorrow? Annu Rev Earth Planet Sci, 32: 495–537

Legendre P, Fortin M J. 1989. Spatial pattern and ecological analysis. Vegetatio, 80: 107–138

Li J F, Xie G, Yang J, Ferguson D K, Liu X D, Liu H, Wang Y F. 2020. Asian Summer Monsoon changes the pollen flow on the Tibetan Plateau. Earth-Sci Rev, 202: 103114

Liaw A, Wiener M. 2002. Classification and regression by randomForest. R News, 2: 18–22

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. Veget Hist Archaeobot, 21: 187–199

Liu H Y, Cui H T, Pott R, Speier M. 1999. The surface pollen of the woodland-steppe ecotone in southeastern Inner Mongolia, China. Rev Palaeobot Palynol, 105: 237–250

Liu H Y, Yi W, Tian Y H, Zhu J L, Wang H Y. 2006. Climatic and anthropogenic control of surface pollen assemblages in East Asian steppes. Rev Palaeobot Palynol, 138: 281–289

Lu H Y, Wu N Q, Liu K B, Zhu L P, Yang X D, Yao T D, Wang L, Li Q, Liu X Q, Shen C M, Li X Q, Tong G B, Jiang H. 2011. Modern pollen distributions in Qinghai-Tibetan Plateau and the development of transfer functions for reconstructing Holocene environmental changes. Quat Sci Rev, 30: 947–966

Lu H, Wu N Q, Yang X D, Shen C M, Zhu L P, Wang L, Li Q, Xu D K, Tong G B, Sun X J. 2008. Spatial pattern of *Abies* and *Picea* surface pollen distribution along the elevation gradient in the Qinghai-Tibetan Plateau and Xinjiang, China. Boreas, 37: 254–262

Luo C X, Zheng Z, Tarasov P, Nakagawa T, Pan A D, Xu Q H, Lu H Y, Huang K Y. 2010. A potential of pollen-based climate reconstruction using a modern pollen-climate dataset from arid northern and western China. Rev Palaeobot Palynol, 160: 111–125

Ma Q F, Zhu L P, Lü X M, Wang Y, Guo Y, Wang J B, Ju J T, Peng P, Tang L Y. 2017a. Modern pollen assemblages from surface lake sediments and their environmental implications on the southwestern Tibetan Plateau. Boreas, 46: 242–253

Ma Q F, Zhu L P, Wang J B, Ju J T, Lü X M, Wang Y, Guo Y, Yang R M, Kasper T, Haberzettl T, Tang L Y. 2017b. *Artemisia*/Chenopodiaceae ratio from surface lake sediments on the central and western Tibetan Plateau and its application. Palaeogeogr Palaeoclimatol Palaeoecol, 479: 138–145

Mantel N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res, 27: 209–220

Molnar P, Boos W R, Battisti D S. 2010. Orographic controls on climate and paleoclimate of Asia: Thermal and mechanical roles for the Tibetan

Plateau. Annu Rev Earth Planet Sci, 38: 77–102
Moore P D, Webb J A, Collinson M E. 1991. Pollen Analysis. 2nd ed. Oxford: Blackwell
Oksanen J, Blanchet F G, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P R, O'Hara R B, Simpson G L, Solymos P, Stevens H H, Szoecs E, Wagner H. 2018. Vegan: Community Ecology Package. R Package Version 2.5-2
Parsons R W, Prentice I C. 1981. Statistical approaches to *R*-values and the pollen-vegetation relationship. Rev Palaeobot Palynol, 32: 127–152
Peng S S, Piao S L, Ciais P, Fang J Y, Wang X H. 2010. Change in winter snow depth and its impacts on vegetation in China. Glob Change Biol, 16: 3004–3013
Prentice I C, Webb T III. 1986. Pollen percentages, tree abundances and the Fagerlind effect. J Quat Sci, 1: 35–43
Qin F, Zhao Y, Li Q, Cai M T. 2015. Modern pollen assemblages from surface lake sediments in northwestern China and their importance as indicators of vegetation and climate. Sci China Earth Sci, 58: 1643–1655
R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing
Shang X, Li X Q, An Z S, Ji M, Zhang H B. 2009. Modern pollen rain in the Lake Qinghai basin, China. Sci China Ser D-Earth Sci, 52: 1510–1519
Shen C M, Liu K B, Tang L Y, Overpeck J T. 2006. Quantitative relationships between modern pollen rain and climate in the Tibetan Plateau. Rev Palaeobot Palynol, 140: 61–77
Sobol M K, Finkelstein S A. 2018. Predictive pollen-based biome modeling using machine learning. PLoS ONE, 13: e0202214
Sobol M K, Scott L, Finkelstein S A. 2019. Reconstructing past biomes states using machine learning and modern pollen assemblages: A case study from Southern Africa. Quat Sci Rev, 212: 1–17
Sugita S. 2007a. Theory of quantitative reconstruction of vegetation I: Pollen from large sites REVEALS regional vegetation composition. Holocene, 17: 229–241
Sugita S. 2007b. Theory of quantitative reconstruction of vegetation II: All you need is LOVE. Holocene, 17: 243–257
Sugita S. 1994. Pollen representation of vegetation in Quaternary sediments: Theory and method in patchy vegetation. J Ecol, 82: 881–897
Sugita S, Gaillard M J, Broström A. 1999. Landscape openness and pollen records: A simulation approach. Holocene, 9: 409–421
Tang L Y, Mao L M, Shu J W, Li C H, Shen C M, Zhou Z Z. 2016. A Illustrated Handbook of Quaternary Pollen and Spores in China (in Chinese). Beijing: Science Press
The Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc, 181: 1–20
The editorial committee of vegetation map of China, Chinese Academy of Sciences. 2007. Vegetation Map of the People's Republic of China (1 : 1000000) (in Chinese). Beijing: Geology Press
Wang F X, Chien N F, Zhang Y L, Yang H Q. 1995. Pollen Flora of China (in Chinese). 2nd ed. Beijing: Science Press
Wang S M, Dou H S. 1998. The Lakes of China (in Chinese). Beijing: Science Press
Wei H C, Ma H Z, Zheng Z, Pan A D, Huang K Y. 2011. Modern pollen assemblages of surface samples and their relationships to vegetation and climate in the northeastern Qinghai-Tibetan Plateau, China. Rev Palaeobot Palynol, 163: 237–246
Weng C Y, Sun X J, Chen Y S. 1993. Numerical characteristics of pollen assemblages of surface samples from the west Kunlun Mountains (in Chinese). Acta Bot Sin, 35: 69–79
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. Rev Palaeobot Palynol, 136: 1–15
Wu S H, Yin Y H, Zheng D, Yang Q Y. 2005. Climate changes in the Tibetan Plateau during the last three decades (in Chinese). Acta Geogr Sin, 60: 3–11
Xi Y Z, Ning J C. 1994. Pollen morphology of plants from Chinese arid and semiarid areas (in Chinese). Yushania, 11: 119–191
Xu Q H, Zhang S R, Gaillard M J, Li M Y, Cao X Y, Tian F, Li F R. 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. Quat Sci Rev, 149: 151–166
Yu G, Tang L Y, Yang X D, Ke X K, Harrison S P. 2001. Modern pollen samples from alpine vegetation on the Tibetan Plateau. Glob Ecol Biogeogr, 10: 503–519
Zhang R, Tian F, Xu Q, Zhou X, Liu X, Cao X. 2020. Representation of modern pollen assemblage to vertical variations of vegetation and climate in the Yadong area, eastern Himalaya. Quat Int, 536: 45–51
Zhang X S. 2007. Vegetation Map of China and Its Geographic Pattern-Illustration of the Vegetation Map of the People's Republic of China (1:1000000) (in Chinese). Beijing: Geology Press
Zhang X S. 1978. The plateau zonality of vegetation in Xizang (in Chinese). Acta Bot Sin, 20: 140–149
Zhang Y, Kong Z, Wang G, Ni J. 2010. Anthropogenic and climatic impacts on surface pollen assemblages along a precipitation gradient in north-eastern China. Glob Ecol Biogeogr, 19: 621–631
Zhang Y, Kong Z, Yang Z, Wang L, Duan X. 2017. Surface pollen distribution from alpine vegetation in Eastern Tibet, China. Sci Rep, 7: 586
Zhang Y J, Duo L, Pang Y Z, Felde V A, Birks H H, Birks H J B. 2018. Modern pollen assemblages and their relationships to vegetation and climate in the Lhasa Valley, Tibetan Plateau, China. Quat Int, 467: 210–221
Zhang Y L, Li B Y, Zheng D. 2014. Datasets of the boundary and area of the Tibetan Plateau (in Chinese). Acta Geogr Sin, 69: 65–68
Zhang Z Y, Harrison S P, Mosbrugger V, Ferguson D K, Paudayal K N, Trivedi A, Li C S. 2015. Evaluation of the realism of climate reconstruction using the coexistence approach with modern pollen samples from the Qinghai-Tibetan Plateau. Rev Palaeobot Palynol, 219: 172–182
Zhao Y, Herzschuh U. 2009. Modern pollen representation of source vegetation in the Qaidam Basin and surrounding mountains, north-eastern Tibetan Plateau. Veget Hist Archaeobot, 18: 245–260
Zhao Y, Liu H Y, Li F R, Huang X Z, Sun J H, Zhao W W, Herzschuh U, Tang Y. 2012. Application and limitations of the *Artemisia*/Chenopodiaceae pollen ratio in arid and semi-arid China. Holocene, 22: 1385–1392
Zhao Y, Xu Q H, Huang X Z, Guo X L, Tao S C. 2009. 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. Rev Palaeobot Palynol, 156: 519–524
Zheng Z, Wei J H, Huang K Y, Xu Q H, Lu H Y, Tarasov P, Luo C X, Beaudouin C, Deng Y, Pan,A D, Zheng Y W, Luo Y L, Nakagawa, T, Li C H, Yang S X, Peng H H, Cheddadi R. 2014. East Asian pollen database: Modern pollen distribution and its quantitative relationship with vegetation and climate. J Biogeogr, 41: 1819–1832

(Responsible editor: Xiaoqiang LI)