



First paleoenvironmental calibrations for modern pollen rain of Tajikistan and Uzbekistan: A case study of pollen - vegetation functional biogeography of Arid Central Asia

Lucas Dugerdil, Odile Peyron, Guillemette Ménot, Dilfuza Egamberdieva, Jakhongir Alimov, Suzanne A.G. Leroy, Eric Garnier, Arkadiusz Nowak, Sébastien Joannin

► To cite this version:

Lucas Dugerdil, Odile Peyron, Guillemette Ménot, Dilfuza Egamberdieva, Jakhongir Alimov, et al.. First paleoenvironmental calibrations for modern pollen rain of Tajikistan and Uzbekistan: A case study of pollen - vegetation functional biogeography of Arid Central Asia. *Global and Planetary Change*, 2025, 252, pp.104857. <10.1016/j.gloplacha.2025.104857>. <hal-05056265>

HAL Id: hal-05056265

<https://hal.science/hal-05056265v1>

Submitted on 5 May 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



Research Article

First paleoenvironmental calibrations for modern pollen rain of Tajikistan and Uzbekistan: A case study of pollen - vegetation functional biogeography of Arid Central Asia

Lucas Dugerdil^{a,b,*}, Odile Peyron^b, Guillemette Ménot^a, Dilmuza Egamberdieva^{c,d}, Jakhongir Alimov^{e,f}, Suzanne A.G. Leroy^{g,h}, Eric Garnierⁱ, Arkadiusz Nowak^{j,k}, Sébastien Joannin^b

^a ENS Lyon, UCBL, CNRS, UMR 5276 LGL-TPE, Lyon, France

^b ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

^c Institute of Fundamental and Applied Research, National Research University TIIAME, Tashkent 100000, Uzbekistan

^d Leibniz Centre for Agricultural Landscape Research (ZALF), 15374 Müncheberg, Germany

^e Faculty of Biology, National University of Uzbekistan, Tashkent 100174, Uzbekistan

^f International Agriculture University, Kibray District, 111200 Tashkent, Uzbekistan

^g Aix Marseille Univ, CNRS, Minist Culture, LAMPEA, UMR 7269, 5 rue du Château de l'Horloge, 13097 Aix-en-Provence, France

^h School of Environmental Sciences, University of Liverpool, Liverpool L69 3GP, UK

ⁱ CEFE Université de Montpellier, CNRS, EPHE, IRD, Montpellier, France

^j Polish Academy of Sciences, Botanical Garden - Center for Biodiversity Conservation, Prawdziwka Str. 2, 02-973 Warszawa, Poland

^k Botanic Garden of the Wrocław University, Sienkiewicza Str. 12, 45-022 Opole, Poland

ARTICLE INFO

Editor: Dr. Sun Jimin

Keywords:

Pollen calibration
Vegetation plot
Tajikistan-Uzbekistan
Arid Central Asia
Plant functional traits
Climate reconstruction
Pollen representativeness

ABSTRACT

Studying modern pollen rain in Middle Asia is crucial for understanding past climate and vegetation changes. This study presents the first dataset of pollen surface samples and vegetation plots from Tajikistan and Uzbekistan, known as the Tajikistan and Uzbekistan Surface Data Base (TUSDB), to enhance our understanding of past climate and vegetation changes in Arid Central Asia (ACA). Multivariate analysis methods, including TWINSpan and CONISS, are used to review the primary vegetation types in Central Asia and assess pollen distribution across these types. Linear relationships, Davis indices, and R-values are applied to evaluate vegetation representativeness based on pollen abundances. Redundancy Analysis is used to compare pollen and climate parameters, and the reliability of TUSDB for local climate reconstructions is tested using transfer functions and machine learning approaches.

The TWINSpan analysis identifies vegetation types consistent with the Uzbek classification, such as desert, steppe, and cryophilous open woodlands. Nine vegetation subtypes are revealed by the pollen samples, with key contributions from trees like *Juniperus* spp. and *Juglans regia*, and non-arboreal plants like Cyperaceae and Poaceae. The study also highlights biases in pollen representation, with some tree taxa overrepresented and certain vegetation-important taxa underrepresented, which can be corrected using pollen R-values.

Abbreviations: ACA, Arid Central Asia; ACADE, Arid Central Asian Database; COSTDB, Cold Steppe Data Base; EAPDB, European Asian Pollen Data Base; SLA, Specific Leaf Area; SSD, Stem Specific Density; TU, Tajikistan-Uzbekistan; TUSDB, Tajikistan-Uzbekistan Surface Data Base.; AP, Arboreal Pollen; NAP, Non-Arboreal Pollen; ShP, Shrub pollen.; AI, Aridity Index; MAAT, Mean Annual Air Temperature; MAP, Mean Annual Precipitation; MPOQ, Mean Precipitation of Coldest Quarter; MPWAQ, Mean Precipitation of Warm Quarter; MTCOQ, Mean Temperature of the Coldest Quarter; MTWAQ, Mean Temperature of the Warmest Quarter; SH, Siberian High.; BRT, Boosted Regression Trees; CA, Correspondence Analysis; CONISS, Constrained hierarchical cluster analysis; CWM, Community-Weighted Mean; FA, Fractional Abundance; MAT, Modern Analogue Technique; NDVI, Normalized Difference Vegetation Index; PCA, Principal Component Analysis; PFT, Plant Functional Type; PolSais, Pollen-Saisie software; PROTEST, PERmutation TEST; RDA, Redundancy Analysis; RF, Random Forest; RMSE, Root Mean Squared Error; RPP, Relative Pollen Productivity; SC, Surface cover; TWINSpan, Two-way Indicator Species Analysis; WAPLS, Weighted Averaging Partial Least Squares.; CLDE, Cold Deciduous Forest; CLMX, Cold Mixed Forest; COCO, Cold Conifer Forest; CODE, Cold Desert; COMX, Cool Mixed Forest; COST, Cold Steppe; HODE, Hot Desert; PION, Pioneer forest; TAIG, Taiga forest; TEDE, Temperate Deciduous Forest; TUND, Tundra; WAMX, Warm Mixed Forest; WAST, Warm Steppe; XERO, Xerophytic Shrubland..

* Corresponding author at: ENS Lyon, UCBL, CNRS, UMR 5276 LGL-TPE, Lyon, France.

E-mail addresses: lucas.dugerdil@ens-lyon.fr (L. Dugerdil), sebastien.joannin@umontpellier.fr (S. Joannin).

<https://doi.org/10.1016/j.gloplacha.2025.104857>

Received 18 December 2024; Received in revised form 25 March 2025; Accepted 22 April 2025

Available online 25 April 2025

0921-8181/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Finally, functional trait aggregation for past pollen sequences shows similarities between modern pollen and vegetation plots. Climate reconstructions validated through transfer functions highlight the complementary role of local and global calibration datasets. These findings confirm the reliability of pollen signals for inferring Holocene climate and vegetation trends, enabling future Uzbek pollen-based reconstructions.

1. Introduction

Arid Central Asia (ACA, covering both Central and Middle Asia; Cowan, 2007) is a region of strong interactions between human societies and their environment over the last millennia (Frachetti, 2008). Especially looking backward, in Tajikistan and Uzbekistan (TU), the question of the rise and collapse of the Oxus valley civilization during the Bronze Age (Luneau, 2019), the socio-economic consequences of the silk roads' dynamics under climate variations during the Middle-Age (d'Alpoim Guedes and Bocinsky, 2018) and more recently the consequences of the Aral Sea desiccation under agricultural management carried out during the Soviet and post-Soviet periods (Micklin, 2007) are all related to climate-vegetation-human societies interactions. Looking forward, despite climatic differences between Central and Middle Asia, ACA offers insights into potential large-scale aridification of Eurasia (Huang et al., 2016; Beck et al., 2018). In ACA, the understanding of grassland ecology has to be deepened since their controlling factor and their diversity are still largely unknown (Wesche et al., 2016) and also because of their critical importance for pastoralism (Gintzburger, 2003; Jamsranjav et al., 2018). During the last decade, several studies were carried out in order to understand the response of ACA vegetation to climate change (Yin et al., 2016; Zhang et al., 2022) and how plants can resist under extreme arid-saline conditions (Shurigin et al., 2022; Egamberdieva et al., 2022). Moreover, the endemic species response to modern climate variations has also been assessed in ACA (Sinaga et al., 2024). It also appears that Quaternary paleoclimate is one of the major drivers of Tajikistan endemism (Radula et al., 2021). That is why long-term climate-vegetation monitoring remains challenging in plant ecology (Seddon et al., 2014). Otherwise, the sedimentological coring conducted in key areas of ACA such as the Caspian Sea (Leroy et al., 2013), in Aral Sea (Aleshinskaya, 1991; Ferronskii et al., 2003; Sorrel et al., 2007), the mountainous lakes from Tian Shan (Spate et al., 2022) and surrounding areas (Leroy and Giral, 2021) can be used to document the complex interactions between climate, vegetation and human populations. The final step is to calibrate past ecological proxies (e.g., pollen) with modern vegetation data. Our study aims to assess the reliability and applicability of these approaches in paleoecology.

To better understand vegetation ecology and paleoenvironmental studies in Tajikistan and Uzbekistan, surface pollen samples and vegetation plots must be compared. First, vegetation type description and classification are the foundation of vegetation ecology, since it helps to reduce the complexity derived from thousands of plant taxa in the environment (Roleček et al., 2009). Numerical algorithms like TWINSPAN can identify indicator species for each vegetation type in phytosociological studies (Nowak et al., 2017, 2024). While this approach has been extensively applied in Tajikistan (Nowak et al., 2017, 2020, 2021, 2022) and Middle Asia (Nowak et al., 2024), it remains under-explored in Uzbekistan.

Vegetation types can be sampled through randomized plots (Kürschner, 2004; Thevs et al., 2008; Nowak et al., 2017; Jamsranjav et al., 2018), vegetation productivity proxies like NDVI (Yin et al., 2016; Zhang et al., 2022) or environmental pollen analysis (Barboni et al., 2004; Ma et al., 2008; Pelánková et al., 2008; Pelánková and Chytrý, 2009; Dehghani et al., 2017; Dugerdil et al., 2021; Robles et al., 2022). After defining and sampling vegetation communities, factors such as soil chemistry (Gintzburger, 2003), salinity (Kürschner, 2004; Dehghani et al., 2017), pH, altitude (Nowak et al., 2017, 2022), hydrology (Thevs et al., 2008), climate (Barboni et al., 2004; Pelánková et al., 2008; Yin et al., 2016; Dehghani et al., 2017; Zhang et al., 2022; Nowak et al.,

2022) and bioclimatic parameters (Jamsranjav et al., 2018) can explain vegetation type distribution. However, these causality or correlation links between vegetation types and the global environment often neglect physiological factors.

This knowledge gap between biotopes and plant ecosystems can be partially fulfilled by studying the physiology of well-known model organisms and their interactions with their milieu: it is the case in Uzbekistan for some halophyte tolerant species (Shurigin et al., 2022; Egamberdieva et al., 2022). However, these studies are time-consuming and cannot catch the whole community features and interactions. In this context, functional ecology, developed in the last decades, is a promising tool to conciliate (1) the observation of global interactions between plant communities and bioclimate parameters and (2) functional, mechanistic and physiological adaptation from individual plants. The individual and measurable features, so-called plant functional traits (Violle et al., 2007), have been standardized and harmonized following common protocols (Pérez-Harguindeguy et al., 2013) and gathered in worldwide databases (Maitner et al., 2018; Kattge et al., 2020; Weigelt et al., 2020; Denelle et al., 2023). These traits are the most physiologically relevant at the species levels: for plant strategies (Grime, 1974), functional typologies (Díaz et al., 2016) and plant adaptation to biotope such as soil features (Borgy et al., 2017) and climate parameters (Prentice et al., 1992; Barboni et al., 2004; Peppe et al., 2011; Nunes et al., 2017) even in arid environment (Frenette-Dussault et al., 2013; Griffin-Nolan et al., 2018). Then, when combining plant traits with functional abundances from vegetation plots, it becomes possible to calculate community-scaled trait values using the Community-Weighted Mean (CWM) traits (Garnier et al., 2004) in order to infer functional responses to climate areas at community scale (Bruehlheide et al., 2018). This relationship was developed for the first time at the ACA scale (Dugerdil et al., 2025), showing the correlation of CWM traits with climate parameters similarly for pollen assemblages and vegetation plots.

This new approach interprets more accurately the interconnections between plant individuals, their traits and their communities with environmental parameters. However, this combined approach still lacks an important dimension: regardless of the scale of observation, from individual to biogeography, each correlation is represented at an almost instantaneous time span. This is one of the most important progress carried in paleoecology (Gauzère et al., 2020; Adeleye et al., 2023): the continuous record of ecological proxy through time with environmental archives (i.e mainly sediment core of lake and peat or loess deposit in drylands; Last et al., 2001) grasps not only the ecosystem properties but also their dynamics through time (Sorrel et al., 2007; Barhoumi et al., 2020; Joannin et al., 2022). Particularly, pollen is usually analysed in sediment to track changes in biodiversity (Birks et al., 2016), local vegetation (Sugita, 2007b; Joannin et al., 2022), regional vegetation (Leroy et al., 2013; Barhoumi et al., 2020), biome distribution (Rudaya et al., 2020), functional diversity (Connor et al., 2018; van der Sande et al., 2021; Gauzère et al., 2020; Adeleye et al., 2023) and human-disturbed vegetation importance (Cromartie et al., 2020; Shumilovskikh et al., 2023). Pollen is also a powerful proxy to reconstruct past climate changes (Chevalier et al., 2020). Moreover, a multi-method approach has shown even more accuracy, especially when selecting the most powerful reconstructions (Dugerdil et al., 2021; d'Oliveira et al., 2023). Although trends on longer time spans are the strength of paleoecology, the latter does suffer from insufficient calibration studies. Modern studies (Pelánková et al., 2008; Ma et al., 2008; van der Sande et al., 2021; Dugerdil et al., 2025) are crucial since biases due to pollen

particularities may mitigate their breakthrough: discrepancies between pollen identifications and plant taxonomy, taphonomy, pollen production and dispersion, among others (Ter Braak et al., 1993; Sugita, 2007a, 2007b; De Klerk and Joosten, 2007).

In this frame, this study attends to merge the strength of the two approaches: (1) paleoecological changes of vegetation and (2) multi-scaled observation in plant sciences. Particularly, we present an interdisciplinary study involving vegetation ecology, paleoenvironmental sciences, functional ecology and climatology in order to precisely infer vegetation biogeography with its temporal dynamics over time (Ma et al., 2008; Pelánková and Chytrý, 2009; van der Sande et al., 2021). Mainly, our approach is based on a new modern dataset of sampling sites in TU. For each site, a vegetation plot and a surface sample for pollen analysis is obtained. Using TWINSpan analysis of vegetation plots, we test the distribution of vegetation types in the TU area and assess

whether they are preserved or biased through pollen rain. Then, we test the applicability of plant functional traits to both vegetation plots and pollen, to measure how reliable the functional pollen signal may be used to predict the functional characteristics of the vegetation. Afterward, we use Tajikistan and Uzbekistan pollen assemblages to calibrate climate reconstruction models (mainly transfer functions and machine learning). The results are compared with those based on datasets of different size (Eurasian, Central Asian and biomes-specific sets). Finally, we test the different approaches to reconstruct past climate changes from the only open-access pollen sequence available in the TU region (i. e. Aral 86 core in Aral Sea; Aleshinskaya, 1991; Ferronskii et al., 2003), to assess the reliability of our approach to reconstruct paleo-climate and environment. In this study, the following issues will be addressed:

1. What are the main ecological controls driving modern biomes and vegetation types distributed along TU climate gradient?

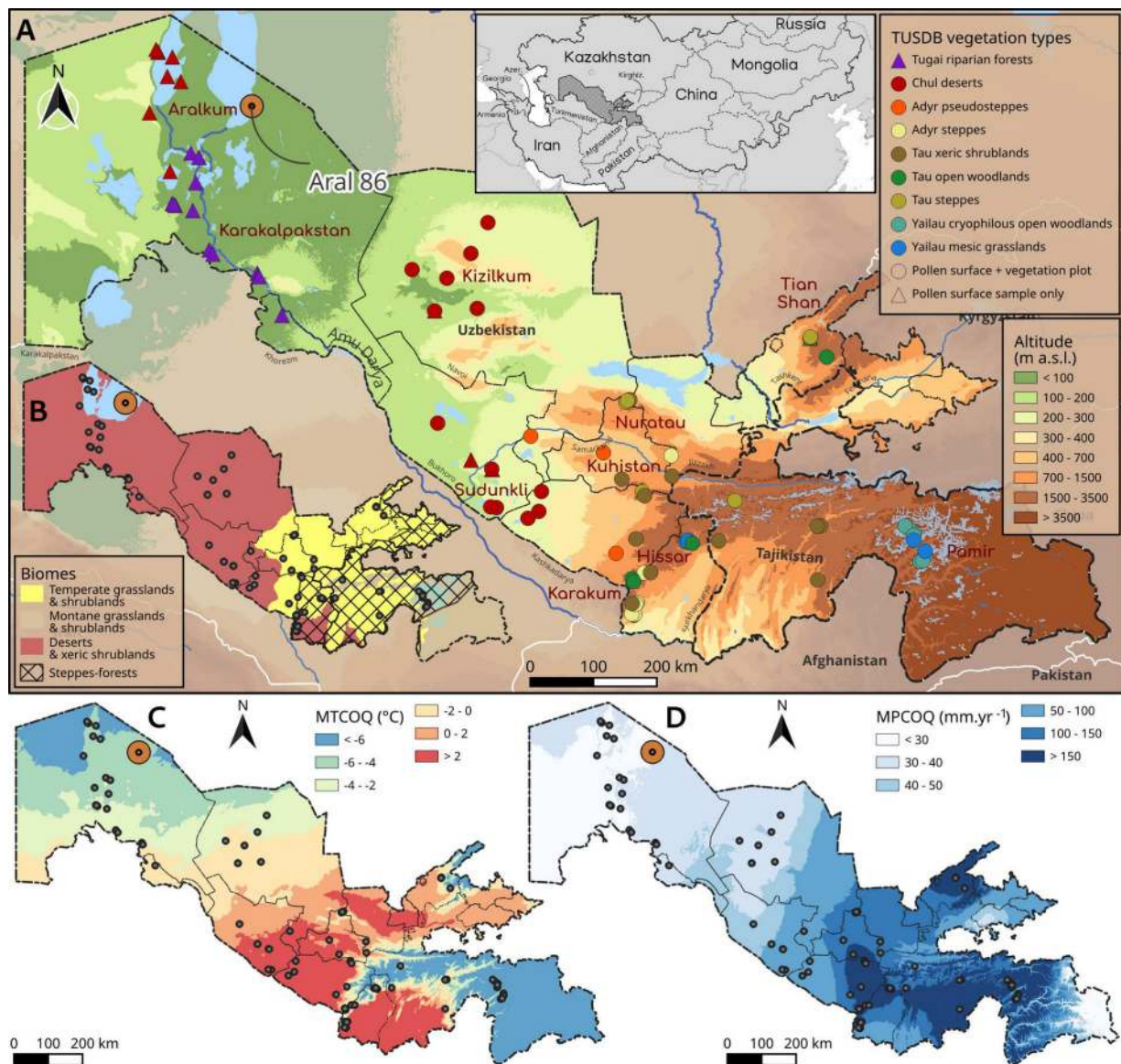


Fig. 1. Tajikistan-Uzbekistan map showing the surfaces samples (TUSDB) on altitude (A) and bioclimate parameters for the study area: the three main biomes (deserts, steppes and mountain mesic grasslands, B); the Mean Temperature of Colder Quarter (MTCOQ, C) and the Mean Precipitation of Colder Quarter (MPCOQ, D). The bioclimatic data come from WorldClim2 (Fick and Hijmans, 2017). The biomes map were extracted from Dinerstein et al. (2017) and the steppes-forests zone (hatched area) from Erdos et al. (2018). All the eight extracted parameters are displayed in Fig. S1. Pink dots show surface samples (soil or moss polsters for pollen sampling associated with a vegetation plot) and yellow dots are some cored wetland or lake. The Aral 86 core (Aleshinskaya, 1991; Ferronskii et al., 2003) used as paleo-application is shown by a large orange dot. The TU area (grey dashed) is also presented in its wider context from the Arid Central Asian area, as proposed in Dugerdil et al. (2025). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. What are the main pollen rain patterns for each of the main TU vegetation types?

3. Is the modern pollen rain representative of these vegetation types, or is it biased?

4. In this case, and after correcting these biases, can we still observe the same plant functional response to climate parameters from two sampling methods: pollen samples and vegetation plots?

5. Finally, can biomes, climate parameters and plant functional diversity be safely reconstructed from Holocene pollen-sequence from TU sedimentological cores?

2. Environmental setting

2.1. Uzbekistan and Tajikistan geography

Tajikistan and Uzbekistan are two countries belonging to the Middle Asian countries (Cowan, 2007) bordered by Kazakhstan in the north, Turkmenistan in the west, Afghanistan in the south, and Kyrgyzstan and China in the east (Fig. 1A; Belolipov et al., 2013), in the core area of ACA (Chen et al., 2015). They are landlocked countries in the central Asian Turan depression, itself encompassed between the Caspian Sea, Ustyurt plateau and the Pamir-Alai range. The altitude is highly contrasted between lowlands (the Turan depression is central Asian largest depression with the lowest point around -10 m a.s.l. at the Sarykamysh Lake, Karakalpakstan, and on average lower than 50 m a.s.l. in the Aralkum and the western Kyzylkum Desert), the midlands composed of plateaus and relic mountains from the Late Carboniferous–Early Permian orogenic period (central Kyzylkum range, Nuratau and Kuratau) and highlands (i.e. Pamir plateau, Alai and Tian Shan ranges with Ismoil Somoni Peak, the highest point culminating at 7495 m a.s.l. in Tajikistan) associated to the Indian subcontinent collision with Asia in the late Cenozoic (Nurtaev et al., 2013).

Loess deposits are important in the TU geography, with a loess deposit belt on the *adyr* piedmont almost everywhere surrounding the Tian Shan and Pamir-Alai mountain complex (Li et al., 2022). These soft sediments are easily watered and transported by water run-off, mainly on the watershed of large rivers (i.e. *rud* in Tajik and *daryo* in Uzbek language) encompassing the areas: the Syrdarya northward, the Amudarya southward and Zarafshan river in between. The strong contrast between the Turan depression and the Pamir-Alai piedmont on one hand, and between irrigated area and drylands on the other hand leads to a gradient from low desert in the north-west, steppe vegetation on the piedmont, forest stands in the mountains lower than 2500 m a.s.l and highland desert in the south-east part of Pamir plateau (Fig. 1B).

2.2. Climate

The special geography of Tajikistan-Uzbekistan leads to a continental climate (Belolipov et al., 2013) with strong temperature seasonality and spatial contrast, mainly due to freezing-cold winter (Fig. 1C; Lioubimtseva et al., 2005) and extremely warm summer (Fig. S1). Mean Annual Precipitation (MAP) occurs almost totally in spring (Fig. 1D). Hyper-arid conditions with $\text{MAP} < 150 \text{ mm.yr}^{-1}$ (Gintzburger, 2003) occur in lowland deserts (Kyzylkum, Sudunkli and Karakum) as well as mid-altitude plateau (Ustyurt Plateau and Fergana valley) and high-elevation rain-shadowed Pamir-Alai range (Sato, 2009) with alpine desert located on the eastern part of the Pamir plateau. However, a temperature gradient divides these drylands following north-south and altitudinal gradients: the Sudunkli (Bukhara and Kashkadarya provinces) and the Karakum (covering mainly the Turkmenistan but only present in Surkhandarya and southern Tajikistan) show the highest temperature (Fig. 1C). From the Köppen-Geiger climate classification, TU is dominated by Arid, desert, cold (BWk) and Arid, steppe, cold (BSk) bioclimate in lowlands; Cold, dry summer (Dsa, Dsb and Dsc) bioclimate in mountainous areas; and Polar, tundra (ET) bioclimate in the alpine belt (Beck et al., 2018). To summarize, the TU area is climatically

separated in cold-wet high mountains (Tian Shan and western Pamir-Alai), cool-arid western plateau (mainly the Ustyurt plateau and the rest of the Karakalpakstan), warm-arid central desert (Kyzylkum and Sudunkli) and warm-semi arid deserts and piedmonts (Karakum, Fergana, Hissar and Turkestan piedmont) as it is shown by the TU bioclimatic space (Principal Component Analysis made on each 1° cell for eight Worldclim2 climate parameters; Fig. 2A).

The climate system in TU area is largely dominated by the Jet Stream influence on wind coming from the west and particularly the winter Westerlies which cross the Alborz or Caucasus slightly southward and

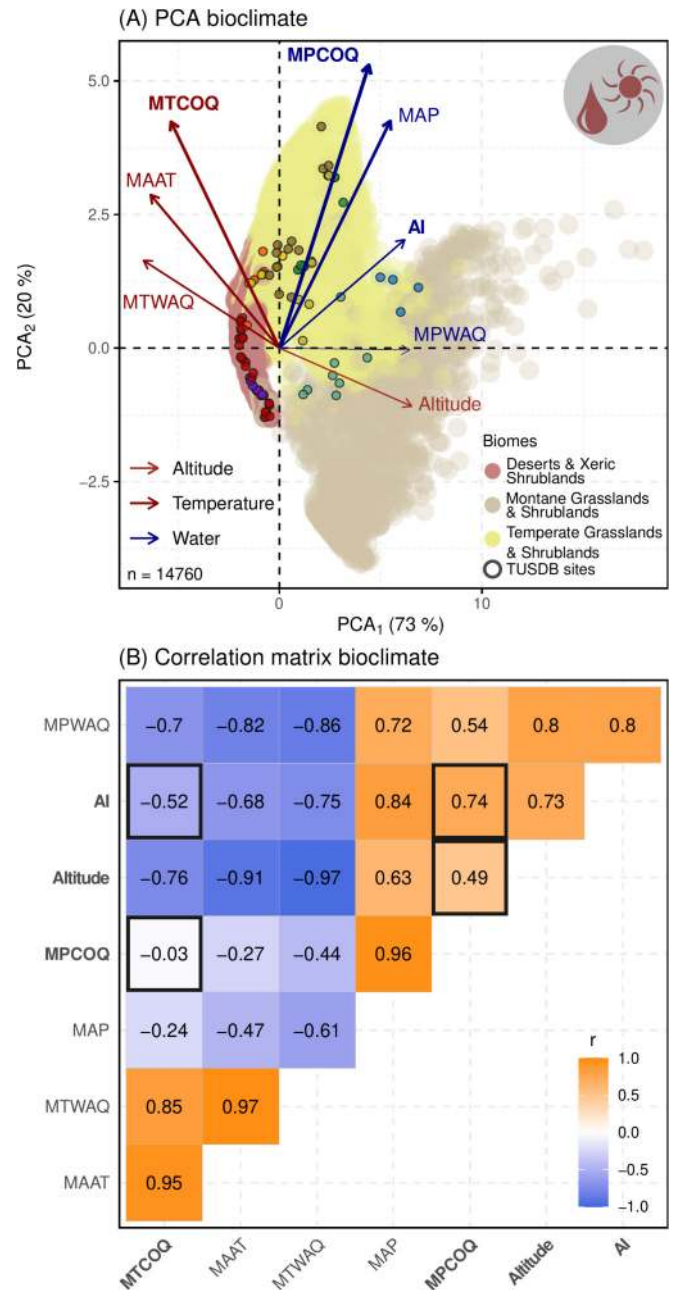


Fig. 2. TU climate parameters co-variation inferred by Principal Component Analysis (A) and matrix of linear correlation (B). For climate parameter deciphering, refer to part 2.6. The statistical analysis was applied to each 1° cell from Worldclim2 climate parameters within the administrative borders of Tajikistan and Uzbekistan. The black dots show the TUSDB sites. Report to Fig. 1 for the vegetation type colour code. Each cell extraction comes from Dinerstein et al. (2017). Bold tiles and labels highlight the climate parameters selected in the study.

bring moisture from the Caspian Sea (Chen et al., 2024). Westerlies are clearly associated with Arabian subtropical anticyclone in TU, these high pressures are distinctly present in Iran and are slightly influential in the Caucasus (Joannin et al., 2014; Robles et al., 2022). Precipitation mainly occurs in the winter period (over than 70 % of the precipitation; Lioubimtseva, 2003) and, with a rain maximum in the early spring (Belolipov et al., 2013). The MAP and the MPOCQ are widely correlated ($r = 0.96$ in Fig. 2B). Siberian Highs (SH) express their influence in summer times (Lioubimtseva et al., 2005) by reducing the rain inflow from the Westerlies in the TU area. The summer Arctic Oscillation drives the SH core south-westward from the Baikal basin to Kazakhstan, freeing summer rain onto the Mongolian Plateau (Dugerdil et al., 2021) and locking the precipitation in Middle Asia (Uzbekistan, Turkmenistan...). In Uzbekistan deserts, the presence of longitudinal sand dunes (Tsoar, 2001) in the core part of the Kyzylkum Desert followed by *barkhan* dunes indicate the importance of northerly winds mainly controlled by the SH (Lioubimtseva, 2003).

Seasonality patterns are very contrasted in ACA and seems to have shifted several times along the Holocene (Tan et al., 2024). Nowadays, in ACA, the precipitation seasonality is mainly controlled by the climate nodes (Chen et al., 2010a, 2024). For instance, in Uzbekistan, the Westerlies vs. SH conflict leads to a spring dominance for annual precipitation (Lioubimtseva et al., 2005). Almost no precipitation occurs in summer, in contrast with the eastern part of ACA (continental China and Mongolia mainly; Li et al., 2022). Around 30 % of the precipitation occur in winter, followed by about 50 % in spring and the rest in autumn (estimation based on the weather station compiled in the CRUTS v.4 database; Harris et al., 2020). This important spring water discharge on the environment preceding a long period of aridity is crucial for plant physiology and is an important explanatory factor of the modern Uzbekistan vegetation biogeography (Gintzburger, 2003). Therefore, the Uzbekistan open woodlands are slightly similar to Mediterranean open woodlands, certainly producing the easternmost analogue (Erdos et al., 2018). The precipitation seasonality pattern is roughly similar for Tajikistan (Radula et al., 2021).

2.3. Regional biomes

From a phytogeographic point of view, TU belong to the Irano-Turanian floristic region (Djamali et al., 2012). Climate conditions mainly control the TU biome distribution: the area is largely dominated by open lands due to aridity (Lioubimtseva, 2003; Gintzburger, 2003). Following Dinerstein et al. (2017), the three main biomes by surface importance are (1) deserts xeric shrublands, (2) temperate grasslands, savannas & shrublands and (3) montane grasslands & shrublands (Fig. 1B). This biome distribution is clearly superposable with the precipitation map (Fig. 1D) and corresponds to the traditional Uzbek descriptions of biomes (Zakirov, 1955). The mountain areas of the Pamir-Alai complex and the western Tian Shan range present wetter conditions ($MAP > 450 \text{ mm.yr}^{-1}$). These conditions allow the establishment of forest-steppes (Erdos et al., 2018). Also, unlike desert and steppes watered mainly in winter and spring; a part of the Pamir and the TU mountain grasslands in general are watered in warmer period (spring mainly but also in summer and fall) as it is shown by the Mean Precipitation of Warm Quarter (MPWAQ) vector on Fig. 2A.

Furthermore, climate forcing on vegetation distribution is mitigated by edaphic conditions and topography. Desert environments have different plant communities between sandy dunes (mainly longitudinal and *barkhans* dunes, not fixed) with psammophytic xeric shrublands; salt-rich depressions around water resurgences (i.e. salt marshes called *solonchak*) or dry clay soils (i.e. seasonally watered depressions mainly dry with desiccation cracks called *taky*) with small halophytic scrubs and shrubs; and deserts on clay or loess soils with *Ferula* spp. / *Artemisia* spp. association with a range of ephemeral herbs (Lioubimtseva, 2003; Toderich et al., 2009). On piedmont, even though the climate parameters are mainly similar with desert, pseudosteppes and steppes are

maintained by better soil conditions with higher organic content and deeper soil horizons supported by stones or rocky outcrops (i.e. *sierozem* soil, Gintzburger, 2003; Belolipov et al., 2013). Among forest-steppe complex, vegetation types are also pondered by altitude gradient. Although the trees need water to be present at lower altitude, humidity and water content are available on the riversides of mountain valleys allowing riparian forests of fruit trees (*Juglans regia*, *Morus alba* and varieties of Rosaceae, *Prunus* spp., *Pyrus* spp., etc.); at higher altitude more rain scattered mosaic woodlands in mountain steppes (*Pistacia vera*, *Crataegus* spp.) and junipers forest-steppes (Erdos et al., 2018).

3. Methods

3.1. Sites

This study is based on the analysis of pollen surface samples and vegetation plots coming from 95 different sampling sites from Tajikistan and Uzbekistan gathered in a new dataset, the Tajik-Uzbek Surface Data Base (TUSDB). Fourteen Tajik and 55 Uzbek sites were collected during two fields trips from summer 2021 and spring 2022, GPS coordinates were recorded by a Garmin eTrex20x with a 3 to 10 m accuracy. In the TUSDB, 12 top-cores (muds or upper lake sediment) without vegetation plots and 57 surface samples (moss polster, soils, peat and dune sands) with vegetation plots are gathered. Since vegetation plots have generally been made on the surrounding of the cored lakes, we have extrapolated the 100 m² vegetation plots for 10 top-cores in statistical analysis, but we removed them for functional traits modelling. In Uzbekistan, three pollen samples from the wide Eurasian Pollen Database compiled by Peyron et al. (2017), modified by Dugerdil et al. (2021) and d'Oliveira et al. (2023) as well as 23 surface samples collected by S. Leroy in 2010 and 2012 in Khorezm and Karakalpakstan have been appended to the TUSDB. No vegetation plots are available for these 26 sites. Each sample is detailed in Tables S1 to S4.

3.2. Vegetation plots

Plant identification. From the field, each plant present on a sampling site has been herborized and photographed. The plant identification was conducted using taxonomic keys from The flora of the USSR (Kirpicznikov, 1969) and the Plants of Central Asia books (Grubov, 1999). The digitized plant herborized were compared with the Moscow Digital Herbarium (<https://plant.depo.msu.ru/>; Seregin, 2024) and the Virtual Flora of Republic of Uzbekistan (<https://floragridmap.uz/>; Floragridmap.uz, 2010). The pictures were compared with illustrated floras (Gintzburger, 2003; Nowak et al., 2020; Sennikov and Tojibaev, 2021; Beshko, 2023) and the Plantarium open online galleries of plants (<https://www.plantarium.ru/>; Plantarium, 2007). All the taxa present in the vegetation plots have been checked with the TNRs package (Boyle et al., 2013).

Vegetation plot. For each site, a vegetation plot surrounding the pollen surface sample was established following a simplified version of the Crackles protocol (Bunting et al., 2013): for each site, at least ten quadrats of one m² in a radius of ten m² surrounding the pollen surface sample were demarcated. Then, in a 100 m² circle, a vegetation plot was examined in each different vegetation type. The taxon abundance was determined using taxon surface cover for two layers: a ground layer (herbs and shrubs below 1.5 m higher) and an arboreal layer above 1.5 m, these two surfaces were then aggregated as a volume using the ratio between the ground layer height (1.5 m) and the arboreal layer (usually around 9 m high). Each taxon was expressed in fractional abundances (FA) following:

$$FA_{veg(i,k)} = \frac{SC_{i,k}}{\sum_{j=1}^{N_{tot,k}} SC_{j,k}} \quad (1)$$

with SC, the surface covered by each i plant species. The $FA_{veg(i,k)}$ was only calculated on the spermatophyte percentage sum in order to compare with the surface pollen samples (bare soil, aquatic plants, lichen, moss and ferns have been removed since they are hardly or not represented in the pollen rain). Then, for each plot, the vegetation value was obtained using a weighted-vegetation model from the ten quadrats (Bunting et al., 2013) using the inverse distance squared (z^{-2}). The rare taxa (i.e. taxa present only in one vegetation plot) were removed for statistical vegetation analysis (Dehghani et al., 2017).

3.3. Vegetation types and subtypes

Based on field observations and using descriptions for Uzbek (Gintzburger, 2003; Belolipov et al., 2013), Tajik (Nowak et al., 2020) and global Middle Asian (Nowak et al., 2024) ecosystems, a vegetation type was assigned to each sampling site. These types were linked to traditional description of Tajik and Uzbek biomes (Zakirov, 1955; Belolipov et al., 2013): (1, *tugai*) desert riparian forest, (2, *chul*) deserts and steppes from the lowlands; (3, *adyr*) steppes and pseudosteppes from the piedmont of all mountain ranges; (4, *tau*) mountain vegetation up to the tree line around 2500 m a.s.l. and (5, *yailau*) high altitude alpine vegetation. Since the Uzbek language transition between Cyrillic to Latin alphabet is still under process, we use in this study the same spelling that the majority of literature, i.e. *tugai* for *to'qay*, *chul* for *cho'l*, *adyr* for *adir*, *tau* for *to'g* and *yailau* for *yaylov* (Zakirov, 1955; Gintzburger, 2003; Belolipov et al., 2013). We also try to provide the modernised phytosociological associations for the TU (Nowak et al., 2024), following the rules of the International Code of Phytosociological

Nomenclature given in its 4th edition (Theurillat et al., 2021) when it is possible. The bioclimate parameters and the soil characteristic (sand, silt, clay, salt- or gypsum-rich) were also taken into account in the vegetation descriptions. We finally labelled our samples in nine vegetation subtypes as follows (full details in Table S5 and illustrations in Fig. 3): (A) TUGAI RIPARIAN FORESTS, Cl. *Populetea euphraticae* Zohary 1962 nom. Inval.; Cl. *Populetea laurifolia-suaveolentis* Hilbig 2000 and Cl. *Tamaricetea arceuthoidis* Akhani et Mucina 2015; (B) CHUL DESERTS, Cl. *Haloxyletea ammodendri* Asri 2003; Cl. *Haloxyletea Salsoletea tomentosae* Akhani 2004; (C) ADYR PSEUDOSTEPPES, Cl. *Stipocapensis-Trachynietea distachyae* S. Brullo in S. Brullo et al. 2001; (D) ADYR STEPPES, Cl. *Artemisia persicae-Stipetea drobovii* A. Nowak, S. Świerszcz, S. Nowak, A. Nobis & M. Nobis in A. Nowak et al., 2024; (E) TAU XERIC SHRUBLANDS, O. *Crataegalia ponticae* A. Nowak, S. Świerszcz, S. Nowak, A. Nobis, E. Klichowska & M. Nobis 2024; (F) TAU OPEN WOODLANDS, Cl. *Junipero-Pinetea sylvestris* Rivas-Mart. 1965; (G) TAU STEPPES, Cl. *Eremogono griffithii-Nepetetea podostachyos* A. Nowak, S. Świerszcz, S. Nowak, A. Nobis, E. Klichowska & M. Nobis 2024 and Cl. *Cleistogenetea squarrosae* Mirkin et al. ex K. Korotkov, O. Morozova, E. Belonovskaja 1991; (H) YAILAU CRYOPHILOUS OPEN WOODLANDS, Cl. *Juniperetea pseudosabinae* Mirkin et al. 1986; (I) YAILAU MESIC GRASSLANDS, Cl. *Festucetea alaico-krylovianae* A. Nowak, S. Świerszcz, S. Nowak, A. Nobis & M. Nobis 2024.

3.4. Pollen surface samples

Analytical protocol. From each site, samples were taken from the upper five cm of soil, moss or mud in a quadrat of one m² in the centre of

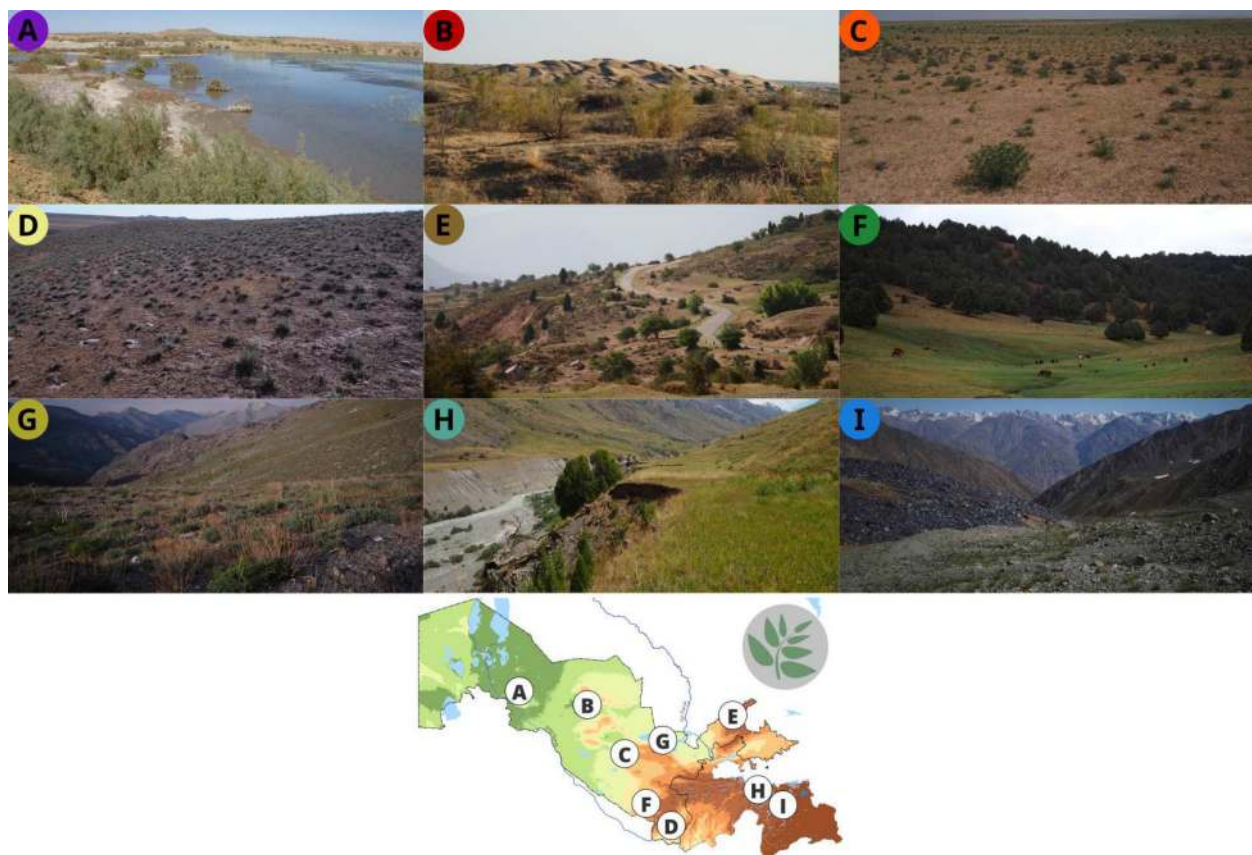


Fig. 3. Main vegetation types of Uzbekistan and Tajikistan in the TUSDB with landscape illustrations from sampling sites. From (A) to (I) following the north-west-south-east bioclimate gradient: (A, MUZSL13) *tugai* riparian forests, (B, MUZT5S01) *chul* deserts, (C, MUZT4M05) *adyr* pseudosteppes, (D, MUZT3M16) *adyr* steppes, (E, MUZT1C02) *tau* xeric shrublands, (F, MUZT3C19) *tau* open woodlands, (G, MUZT6C02) *tau* steppes, (H, MTAT1M02) *yailau* cryophilous open woodlands, (I, MTAT1M05) *yailau* mesic grasslands. Bottom panel, map showing the location of the pictures. The colour codes correspond to the vegetation types colour code from Fig. 1A. Photo credits: (A) S. A. G. Leroy, and all the others panels L. Dugerdil.

the vegetation plot. For top-cores, using a die-cut, one cm³ of the upper part of the core was sampled. Then, the pollen content was extracted from the soil/mud matrix following a complex protocol (Dugerdil et al., 2021) including KOH (10 %), Hydrofluoric acid, heavy liquids (ZnCl₂ + HCL) and acetolysis. The preparations were observed under a BRESSER Science Infinity Microscope with a 40 × magnification lens for the routine count and a 100 × magnification lens for details. For the 23 Karakalpakstan samples, no KOH was used. For mud samples, after deflocculation in Na pyrophosphate, the samples were submitted to cold HCL, HF, HCL and sieving; while acetolysis was used only for moss polsters.

Pollen identifications. The pollen identifications were conducted following common pollen atlases such as the Reille (Reille, 1999) and central Asian pollen determination keys from Mongolia (Murad, 2012), continental China (Lu et al., 2018) and Caucasus (Connor and Kvavadze, 2018). The counts were assisted by the PolSais 4.1 software developed in python 3.0 following the rarefaction curve (which allows to better take into account the total pollen diversity of the sample; Birks and Line, 1992). The average pollen count is 300 grains by samples. However, the total pollen count varies among samples due to their pollen diversity. Pollen count is about 200 for deserts and pseudosteppes, 300 in *tau* steppes or in *yailau* mesic grasslands and 350 in *adyr* steppes, *tau* xeric shrublands & open woodlands, the sites with higher pollen diversity. Pollen absolute concentration was calculated using *Lycopodium* spp. as exotic spore added to each sample (Faegri and Iversen, 1989). The pollen samples are then expressed in fractional abundances (FA_{pol(i,k)}) following the same calculation as for vegetation plots FA_{veg(i,k)} displayed in eq. (1). Pollen types are associated to three major growth types: Arboreal Pollen (AP), Non-Arboreal Pollen (NAP) and Shrub Pollen (ShP).

3.5. Functional traits

Workflow. In order to test a functional approach on both vegetation plots and pollen surface samples, trait values for the TUSDB taxon checklist were extracted from worldwide trait databases. From these, Community-Weighted Mean (CWM) values were calculated for each site (Garnier et al., 2004). We then compared the CWM of vegetation and pollen-based and their relation with climate parameters. The detailed methodological framework applied here was developed by Dugerdil et al. (2025). The methodological workflow is roughly the same, except for the geographical scale (whole ACA against TU territories) and the sampling methods (two different datasets for vegetation plots and pollen surface samples in the Dugerdil et al. (2025) study against a smaller dataset here with both vegetation and pollen on each plot).

Taxon checklist. A taxon checklist is necessary to assign plant species to pollen types. Since no accurate plant taxon checklist exists in the majority of ACA biomes yet and especially in Uzbekistan (Li et al., 2020), we have built a composite plant checklist based on the Tajikistan checklist prepared for the purpose of this paper using the available literature, the checklist of the Chatkal State Biosphere Reserve (Beshko et al., 2024), and we have digitized checklists available for the Tian Shan (Sennikov and Tojibaev, 2021) and Nuratau (Beshko, 2023) mountain ranges as well as the Uzbek provinces of Bukhara (Tojibaev et al., 2021a), Dzhezak (Tojibaev et al., 2019), Kashkadarya (Tojibaev et al., 2021b), and Samarkand (Tojibaev et al., 2018). Then, using the R package GIFT (Denelle et al., 2023), we have extracted the available checklist for our study area used in the GIFT database (Weigelt et al., 2020). The plant occurrences recorded in the Botanical Information and Ecology Network (BIEN) database have also been appended to the ACA checklist using the R package BIEN (Maitner et al., 2018). All the plant record datasets have been merged, and their taxonomy were checked using the TNRS (Boyle et al., 2013) package to correct spelling mistakes and remove synonyms. The varieties, subspecies and non-producing pollen plants were purged, and we finally obtain 7590 species.

Pollen types to plant taxa. A correspondence table between pollen types and plant species has been created using the taxa checklist because of the incompatibility between pollen types and plant species (heterogeneity of the pollen identification level among pollen studies; De Klerk and Joosten, 2007). Subfamily (Asteroideae-type, Cichorioideae-type and Cardueae-type) and sub-genus (*Pinus* Haploxyton-type and *P. Diploxyton*-type) were taken from the National Center for Biotechnology Information's Taxonomy database (NCBI, <https://www.ncbi.nlm.nih.gov>; accessed January 2021). For the pollen types *Aster*-type, *Marricaria*-type and *Senecio*-type, the species from the tribes Asteraceae, Anthemideae and Senecioneae have been used. The Cerealia pollen type was associated to a list of botanical taxa using the USDA Plants Database (Usda, 2024). The assignment of *Ephedra* species to pollen types *E. distachya*-type and *E. fragilis*-type has been made following the scanning electron micrograph of several *Ephedra* species from the Tibetan plateau (Bolinder et al., 2016). Similarly, *Plantago* species have been assigned to *P. coronopus*-type, *P. major*-type and *P. lanceolata*-type using pollen morphological considerations (Gazer and Shalabi, 2016; Mohsenzadeh et al., 2020) and global *Plantago* phylogeny (Shipunov et al., 2020).

Plant trait databases. The plant trait data come mainly from worldwide databases such as TRY (Kattge et al., 2020), BIEN (Maitner et al., 2018) databases and the recent Global Inventory of Floras and Traits (Weigelt et al., 2020) using the GIFT R package (Denelle et al., 2023). The six most commonly covered and used plant traits values (Díaz et al., 2016; Garnier et al., 2004; Bruelheide et al., 2018) were extracted for our TUSDB plant checklist: plant height, seed mass, leaf area, specific leaf area (SLA, i.e. the ratio between the area and the oven-dry mass of the leaf), nitrogen content per leaf mass (N_{leaf}) and stem specific density (SSD, the ratio between stem dry mass and fresh volume of a stem section). The six traits were extracted for each plant in the taxon checklist and an aggregated value were calculated for each pollen types.

Community mean traits. In order to investigate the functional response of the plant community induced by climate or human changes, the mean value of community trait distribution, the CWM, is calculated (Garnier et al., 2004). Trait values for plant species and pollen types are combined with species abundance to calculate CWM values following:

$$CWM_{j,k} = \sum_i^{n_k} p_{i,k} \times t_{i,j} \quad (2)$$

with k each vegetation plot / pollen sample, j each trait, and n the total number of species for the plot k. $p_{i,k}$ is the fractional abundance for species / pollen type i in plot k and $t_{i,j}$ the average value of trait j and species i. Further details can be found in Dugerdil et al. (2025).

3.6. GIS and bioclimate parameters

For each sample, main environmental parameters were extracted using GIS R packages (*rgdal* and *raster*); the altitude from SRTM Digital Elevation Model version 4.1 (Jarvis et al., 2008); six climate parameters from worldclim2.1 (Fick and Hijmans, 2017); the extrapolated Aridity Index (AI) from the CGIAR (Trabucco and Zomer, 2018) and three biomes from Dinerstein et al. (2017). Due to sparsity of weather stations available in Central Asia, extrapolated values from GIS databases were privileged. Bioclimate parameters tested in this study (Figs. 1C-D and S1) are altitude, AI, Mean Annual Air Temperature (MAAT), Mean Annual Precipitation (MAP), Mean Temperature of the Coldest and Warmest Quarter (respectively MTCOQ and MTWAQ, Fig. 1C and D), Mean Precipitation of the Coldest and Warmest Quarter (respectively MPCOQ and MPWAQ). AI is obtained by $AI = 10000 \times MAP/MA - ET_0$ with $MA - ET_0$ the Mean Annual Reference Evapotranspiration (Trabucco and Zomer, 2018). It means that the AI rises up in wet environment when the MAP rises.

3.7. Pollen biomization

In order to subset pollen surface samples by biomes (defined as a large-scale ecosystem occupying large spaces (...) and incorporating a complex of fine-scale biotic communities: [biome] has its characteristic flora and fauna and it is home to characteristic vegetation types and animal communities; Mucina, 2019) and to reconstruct past biomes, we applied the classical biomization method as firstly presented in Prentice et al. (1996). For the TUSDB pollen-based biomization approach and the list of used Plant Functional Types (PFTs) applied in our study, see Peyron et al. (1998) and Tarasov et al. (1998). Mainly the biomization scheme has been carried in R with a process inspired by Guiot and Goeury (1996). We combine pollen types, PFTs, and biomes in order to attribute a dominant biome to each pollen site. First, using a pollen type to PFTs assignation table, we estimated for each surface sample the PFTs score (3):

$$PFT_{j,k} = \sum_i^{n_k} \sqrt{(p_{i,k} - \theta_i)} \quad (3)$$

with k each pollen sample, j each PFT, and n_k the total number of species for the k sample. $p_{i,k}$ is the fractional abundance for pollen type i in plot k and θ_i is a threshold pollen percentage (here 0.5 % following Prentice et al., 1996; Tarasov et al., 1998). Then, using a PFTs-to-biome matrix, the biome affinity scores were calculated for each sample. The selected pollen-based biome is the biome with the highest score as follows (4):

$$Biome_{l,k} = \max \left(\sum_j^{n_{PFT}} PFT_{j,k} \right) \quad (4)$$

with $Biome_{l,k}$ the l biome score for the k surface sample and $PFT_{j,k}$ the associated PFT to the biome. Afterward, we apply the procedure developed by Tarasov et al. (1998) to distinguish cold or warm versions of a biome, particularly steppes or desert. The full assignation tables are available from Prentice et al. (1996) and Tarasov et al. (1998).

3.8. Pollen-inferred past climate reconstructions

Climate reconstruction approaches. In order to reconstruct quantitatively past climate changes from pollen assemblages recorded in environmental archives (lakes, peat...), several methods have been proposed and tested (Chevalier et al., 2020). Among these methods, we selected the three most common approaches: (1) the Weighted Averaging Partial Least Squares (WAPLS) regression, a transfer function based on linear regressions assuming that each pollen type abundance is unimodally related to a climate parameter optimum (Ter Braak et al., 1993), with a climate reconstruction by averaging climate optima of each pollen type weighted by their abundances; (2) the Modern Analogue Technique (MAT), based on the calculation of dissimilarity degree between pollen assemblages in a surface pollen dataset and a fossil pollen sample; with past climate estimates calculated by averaging the values of the climate parameters of the most similar modern samples (i.e. the best analogues; Guiot, 1990; Peyron et al., 2011); and (3) the machine learning-based approach mainly Random Forest (RF, Breiman, 2001; Simon et al., 2023) and Boosted Regression Trees (BRT) following Salonen et al. (2014, 2019). These two methods are based on estimation and combination of many regression trees. Machine learning is based on the separation of the data pool in a training set and a testing set. Each tree is made on a set of pollen sample selected by bootstrapping (Chevalier et al., 2020). The difference between RF and BRT is about the way of choosing the trees: in RF the trees are randomly selected, while in BRT pollen samples under-represented on tree n has a lower probability to be chosen on the tree $n + 1$.

Calibration dataset comparison. These approaches require a modern pollen dataset known as the calibration dataset. Since the choice of the dataset is important for the reliability of the climate reconstruction

(Dugerdil et al., 2021; d'Oliveira et al., 2023), several datasets have been tested here, from global to regional ones:

1. **EAPDB:** European Asian Pollen Data Base (Fig. 4) is the widest dataset ($n_{sites} = 5411$) with samples from the southern Morocco to the Kamchatka peninsula crossing all biomes (Fig. 4). To the first version of this dataset compiled by Peyron et al. (2017), Dugerdil et al. (2021), and d'Oliveira et al. (2023), Dugerdil et al. (2025) have appended numerous sites from the eastern part of ACA (Zheng et al., 2014; Binney et al., 2017; Li, 2018; Sun et al., 2020) in order to spatially equilibrate the database.

2. **ACADB:** Arid Central Asian Data Base ($n_{sites} = 839$ is a selection of the pollen samples of the EAPDB located within the borders of ACA (Fig. 4, dark grey area) following Dugerdil et al. (2025) and Chen et al. (2010a, 2015). Cold steppes and desert, taiga, conifer forest and tundra-like are the main biomes (Fig. 4).

3. **COSTDB:** Cold Steppe Data Base ($n_{sites} = 478$ is another extraction of the EAPDB set, selecting only samples from the cold steppe biome (bigger dots in clear yellow on Fig. 4). The cold steppe sites are mainly located in ACA, Mongolia, Tibet, Kazakhstan up to Siberia and Ukraine including few European sites.

4. **TUSDB:** Tajik-Uzbek Surface Data Base groups the new samples from this study. Here, we have $n_{sites} = 95$ (excepted 3 samples already present in EAPDB, all the samples within the TU borders were new; dark orange area on Fig. 4). Pollen-based biomes are mainly cold steppe, cold desert and tundra-like for mountainous pollen surface samples.

Implementation and statistics. The MAT and WAPLS were conducted with the *rioja* R package (Juggins and Juggins, 2019), the RF with the *randomForest* package and the BRT with the *gbm.step()* function from the *dismo* R package (Hijmans et al., 2017). In our study, MAT, WAPLS, RF and BRT were tested. However, due to RF systematic lower performance, we only show the result for WAPLS, MAT and BRT. Mathematical indices are used to estimate the performance of each method. Since the dataset size is very heterogeneous and the methods are clearly different, it is necessary to use common statistical indices: the Pearson adjusted R^2_{adj} (correlation strength insensitive to the over-parametrization of a model) and the RMSE (given the error interval around the climate parameters value). These indices are easy to calculate for linear regression, but necessitate cross-validation or bootstrapping for transfer functions and machine learning (Chevalier et al., 2020). Moreover, we tested the spatial autocorrelation for each calibration dataset and each method with the calculation of the Moran's I (indicating how much the model is incorrectly enhanced by autocorrelation, Telford and Birks, 2005) using the *Moran.I()* function from *ape* package. Finally, we applied to the fossil pollen sequence the *randomTF()* function from *paleoSIG* package to select the most accurate climate reconstruction (Telford and Birks, 2011). This significance test, based on the comparison between past and modern pollen datasets, (1) estimates the proportion of variance of the past signal explained by climate parameters residuals (using Principal Component Analysis) and (2) compare them with the variance explained by the same transfer function trained on random data.

3.9. Statistical analysis

Pollen representativeness indices. To evaluate the similarity between pollen and vegetation fractional abundances, several methods are proposed, such as the association indices (so-called Davis indices; Kent Davis, 1984; Bunting, 2003; Amami et al., 2010). Indices of association (A_i), under-representation (U_i) and over-representation (O_i) were given for each pollen type i by:

$$A_i = B_0 \times (P_0 + P_1 + B_0)^{-1} \quad (5)$$

$$U_i = P_1 \times (P_1 + B_0)^{-1} \quad (6)$$

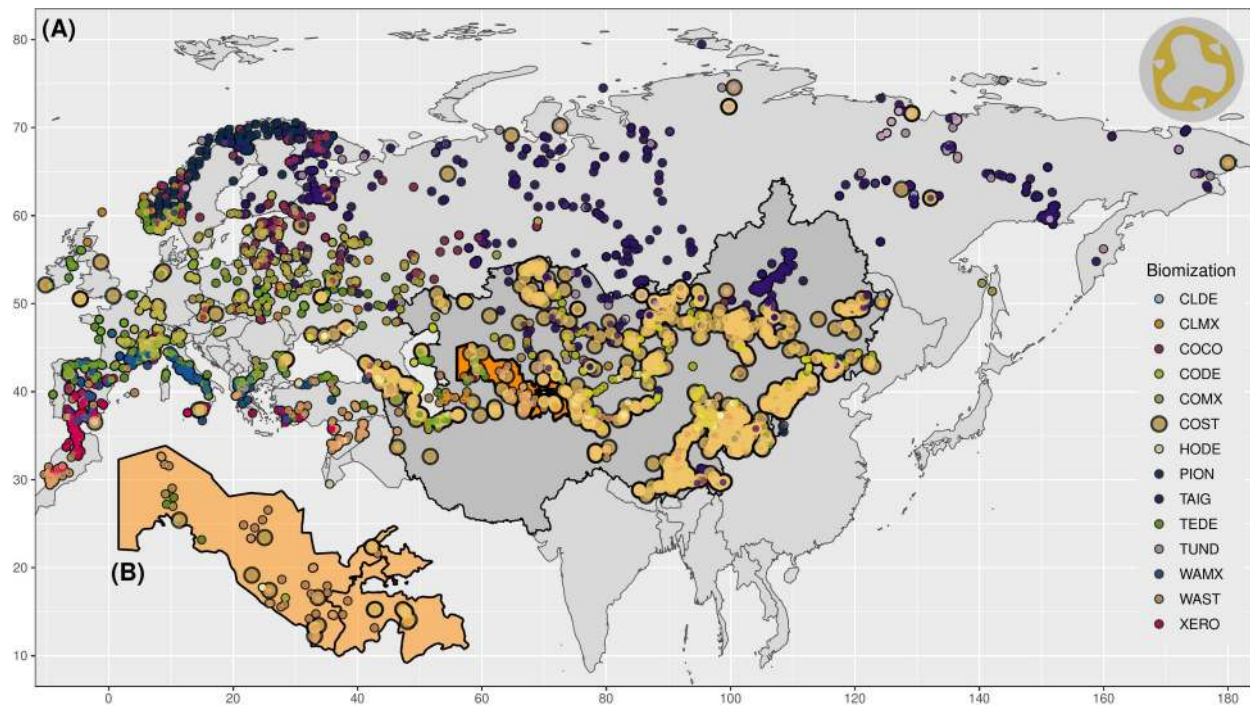


Fig. 4. Eurasian map (A) of the pollen-derived biome values with a zoom of the Tajik-Uzbek samples (B). Light grey enhance EAPDB extent, dark grey the ACADB extent and dark orange the TUSDB. COSTDB samples have large yellow dots. The colour dot refers to the pollen-inferred biome for each site. The biomes are **CLDE**, Cold Deciduous Forest; **CLMX**, Cold Mixed Forest; **COCO**, Cold Conifer Forest; **CODE**, Cold Desert; **COMX**, Cool Mixed Forest; **COST**, Cold Steppe; **HODE**, Hot Desert; **PION** Pioneer forest; **TAIG**, Taiga forest; **TEDE**, Temperate Deciduous Forest; **TUND**, Tundra; **WAMX**, Warm Mixed Forest; **WAST** Warm Steppe; **XERO**, Xerophytic Shrubland following [Prentice et al. \(1996\)](#), [Peyron et al. \(1998\)](#) and [Tarasov et al. \(1998\)](#). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$O_i = P_0 \times (P_0 + B_0)^{-1} \quad (7)$$

with B_0 number of sites in which the taxon i is present in both pollen samples and vegetation plots, P_0 number of the sites of only pollen presence and P_1 the sites in which only vegetation species are present. Only the taxa present both at least in one pollen sample and one vegetation plot were kept for analysis ([Bunting, 2003](#)). Then, [Kent Davis \(1984\)](#) introduced threshold to statistically determine each pollen type representativeness as follows:

- Strongly-associated taxa when A is $>65\%$,
- Associated taxa for $A \in [50;65]\%$,
- Weakly-associated taxa when $A < 50\%$ and both O and U are high,
- Unassociated taxa when $A = 0$ and $O > 50\%$,
- Over-represented taxa if $U = 0$ and $A < 50\%$.

The association indices are based on a binary approach, modelling the pollen-vegetation bias as a *presence/absence* relation. The semi-quantitative approach of pairwise linear relation using Pearson correlation coefficients (r) and p -values have also been calculated. Finally, the R -values, which is a simple and efficient approach to correct the bias between vegetation and pollen ([Parsons and Prentice, 1981](#); [Amami et al., 2010](#)), were also calculated. It is also widely used and comparable with Chinese ([Xia and Ni, 2024](#)) or arid Central Asian ([Wang et al., 2022](#)) studies. Following [Wang et al. \(2022\)](#), we decided to calculate relative R -values (R_{rel}) for each taxon i with:

$$R_i = \frac{(P_i \times V_i^{-1})}{R_{Poaceae}} \quad (8)$$

with P_i the fractional abundance for the pollen i , V_i the fractional abundance of plant taxa and $R_{Poaceae}$ the R -value for *Poaceae*. *Poaceae* have been selected as standard since it is (1) is the most ubiquitous taxon

in both pollen sample and vegetation plots, (2) it has a moderate mean R -value and (3) it is the referential used also in RPP calculation due to its worldwide spread ([Wieczorek and Herzschuh, 2020](#)).

Multivariate analysis. Data cleaning, calculations and statistical analyses were carried in R (version 4.4., [Team, R.C., 2014](#)). Mainly, the vegetation zonation was performed using a Two-way Indicator Species Analysis (TWINSPAN, [Hill and Hill, 1979](#)) using the *Twinspan* R package. This hierarchical and divisive vegetation plot classification algorithm is based on a Canonical Analysis ([Roleček et al., 2009](#)). First a set of cut-level to convert species abundance in pseudo-species were settled as 0.01, 1, 3, 5, 25, 50 and 70 % of abundances. The analysis return clusters for both plant taxa and vegetation plots, with the most important taxa called indicator taxa ([Roleček et al., 2009](#); [Dehghani et al., 2017](#)). The results have been plotted with the *ggplot2* for the heatmap and the *ggdendro* and *dendextend* packages. For pollen and vegetation diagram, the plots were obtained with *ggplot2* and *tidypaleo* ([Dunnington et al., 2022](#)) and the clustering were performed using constrained hierarchical cluster analysis called CONISS ([Grimm, 1987](#)) implemented of the pollen fractional abundances total with the AP/NAP ratio appended. Then, multivariate analyses were performed on vegetation plots (Correspondence Analysis, CA), vegetation plots and surface samples only (Principal Component Analysis, PCA), then combined with bioclimate parameters (Redundancy Analysis, RDA) helped by the *vegan* package ([Dixon, 2003](#)). Using the same package, the two multivariate spaces PCA_{pollen} and PCA_{veg} were formally compared with a Procrustes rotation analyses and these results were tested with PROTEST significance test both implemented with the *vegan* package ([Dixon, 2003](#)).

Application of our interdisciplinary approach to past periods: test on the Aral 86 pollen record. The Holocene time period is poorly documented in Uzbekistan and in Tajikistan; only few analyses on speleothems ([Cheng et al., 2016](#)) and on paleosol-loess sequences ([Welch et al., 2016](#); [Zhang et al., 2020](#)) have been published so far. Due to the rareness of lakes and

wetlands, no sediment core from continental wetlands proxies have been published yet. The only exceptions come from the Aral Sea (Ferronskii et al., 2003; Sorrel et al., 2007; Burr et al., 2019). Actually, the unique open-access pollen sequence available in TU is the Aral 86 sequence in which diatoms and pollen have been counted (Aleshinskaya, 1991). The raw data is available on <https://neotomadb.org/>. Multiple radiocarbon dating have been carried out from this sequence. Among Total Organic Content and carbonates based radiocarbon dating, only nine dates are considered as reliable (Ferronskii et al., 2003). We used a value of 168 ± 53 year to correct the reservoir effect of the carbonate-based ^{14}C ages in the Aral watershed following Burr et al. (2019). Our original approach is thus tested on this sequence.

4. Results

4.1. Vegetation types and plant communities

The TWINSpan analysis of 58 vegetation plots identifies four main vegetation groups in the Tajik-Uzbek biomes (Fig. 5A): deserts (*chul*), steppes (*adyr* and *tau*), *yailau* cryophilous open woodlands, and *yailau* mesic grasslands. This clustering aligns with field observations and altitudinal gradients, representing the main vegetation types and nine subtypes. However, plant abundance variance is insufficient to distinctly separate *adyr* pseudosteppes from *adyr* steppes, or *tau* xeric shrublands from *tau* steppes. Additionally, *tau* open woodlands cluster near the ecotone of *yailau* cryophilous woodlands.

About the plant taxa characterization of our vegetation types, desert sites are well separated from the rest of the plots by the presence of particular indicator taxa such as *Carex physodes*, *Eremopyrum bonaepartis*, *E. orientale*, *Isatis boissieriana* and *Stipagrostis pennata*. These taxa mainly represent the psammophytic vegetation on aeolian sand dunes, while the two first sites (MUZT4S01 and MUZT4S03) are marked by a characteristic halophytic plant community (*Aeluropus litoralis*, *Halocharis hispida* and *Climacoptera lanata*, Fig. S2). The *yailau* ecosystems with no exception (both opened and closed vegetation) are marked by *Carex pachystylis*, *Meniocus linifolius* and *Taraxacum officinale*, while *Alhagi pseudalhagi*, *Bromus tectorum* and *Poa bulbosa* are continuously present from desert to piedmont steppes. For *tau*, no dominant taxa unified the cluster: *Hordeum bulbosum*, *Prunus bucharica* and *P. spinosissima*, *Crataegus pontica*, *Artemisia oliveriana*, *Festuca* spp., *Juglans regia*, *Prunus cerasifera* important in xeric shrublands; *Juniperus seravschanica* and *Origanum vulgare* for open woodlands with *Juniperus semiglobosa* in transition to the *yailau* cryophilous open woodlands. The two highest elevated clusters were also clearly distinct with *Ephedra gerardiana*, *Nepeta podostachys*, *Polygonum coriarium*, *Prangos pabularia*, and *Rosa ecae* as taxa indicator for *yailau* cryophilous open woodlands and *Oxytropis* spp. group for the *yailau* mesic grasslands. To see the full species cortège following the main indicator species for each of the four TWINSpan clusters, please refer to the Figs. S2-S5).

The Correspondence Analysis (CA) shows the clear separation between the four clusters, mainly driven by the presence of *Oxytropis* spp. and *Cousinia laetevirens* as main contributor for *yailau* mesic grasslands on a side (positive CA₁ and CA₂ on Fig. 5B) and *chul* desert samples mainly driven by *Stipagrostis pennata*, *Haloxylon persicum*, *Artemisia diffusa* and *Salsola richteri* on the other one. A third group of *adyr* and *tau* steppes and forest-steppe with negative CA₁ is driven by *Cynodon dactylon* gradually followed by *yailau* cryophilous open woodlands sites (transition to positive CA₁ closer to the *yailau* mesic grasslands with *Stipa tianshanica* and *Betula tianschanica* first; *Polygonum paronychioides* and *Nepeta podostachys* secondly). Among the *adyr/tau* cluster the gradual transition from *chul* deserts to *tau* steppes is following the classical altitudinal succession: deserts, pseudosteppes, *adyr* steppes, *tau* xeric shrublands and open woodlands to finally reach *tau* steppes. Finally, TWINSpan results (Fig. 5A) are really consistent with the CA space projection: within the 15 most contributing taxa to CA, the

majority of them are among TWINSpan indicator species while the others are part of the same cortège as indicator species (i.e. *Cousinia laetevirens* which is not an indicator species, but it is strongly associated with the indicator species *Oxytropis* spp.).

4.2. Pollen distribution

The TUSDB pollen samples are largely dominated by Non-Arboreal Pollen (NAP) such as Amaranthaceae, *Artemisia* spp., Poaceae, Cyperaceae, Fabaceae, Asteroideae-type, *Plantago* spp. and Apiaceae, followed by Arboreal Pollen (AP) *Juniperus* spp., *Betula* spp., *Juglans regia*, *Salix* spp. and Rosaceae (tree)-type (Figs. 6 and S6). The distribution of pollen types follows an altitudinal-bioclimatic transect from the dry-cool conditions in the lowlands of Aral Sea basin in the northwest to the cold-wetter condition in the Pamir high elevation plateau in the southeast, crossing warm-dry deserts in Karakalpakstan *tugai*, Kyzylkum, Sudunkli and Karakum, piedmonts (Nuratau, Zarafshan, Hissar and Tian Shan). Each sample is detailed in Tables S1-S4. Based on CONISS analysis (Fig. 6), we observe nine statistical significant pollen-cluster that globally fit our nine vegetation types which are described in details in Table S6.

4.3. Climatic space

The most contributing climate parameters to explain the bioclimatic particularities of TUSDB have been determined using a PCA performed on the eight major climate parameters for each 1° cells extracted within Tajikistan and Uzbekistan borders (Fig. 2A). The two most contributing factors are the MPOCQ and the MTCOQ (both of them close to their annual average, i.e. MAP and MAAT respectively). Then, the altitude and the AI are also considered as important since they are both high contributors on PCA₁ and because altitude is the most related factor to explain montane grasslands and shrublands site distribution (in clear brown) and AI with temperate grasslands, savannas and shrublands (in yellow, Fig. 2A). The correlation between them has then been assessed, especially due to climate reconstruction necessity: it is important to reconstruct primary and secondary climate parameters (Salonen et al., 2019) while keeping in mind that the two variables have to be as independent as possible. The correlations among the four selected climate parameters have been studied (Fig. 2B). It appears that the winter temperature and precipitation (MTCOQ and MPOCQ) are definitely independent ($r = -0.03$). To summarize, the MPOCQ, MTCOQ, altitude and AI are used in this study to understand the vegetation and pollen distribution but only the MPOCQ and MTCOQ are independent enough to be used for past climate calibration reconstructions.

4.4. Pollen and vegetation environmental gradient

Even though the multivariate statistical analyses are different between CA (Fig. 5B) and PCA (Fig. 7A) site distributions on the multivariate space (the two first axis explaining respectively 8 % and 7 %) are similar: desert and pseudosteppes samples are clustered on the negative PCA₁, the others (steppe, forest and mesic grasslands) on the positive PCA₁ (Fig. 7A). The PCA₂ corresponds to a growth form gradient with woodlands on negative PCA₂ (with mixed samples from *tau* xeric shrublands, open woodlands and *yailau* cryophilous open woodlands). The positive PCA₂ gathers grasslands such as pseudosteppes, steppes, *yailau* mesic grasslands close to PCA₂ = 0, while the *tau* xeric shrublands are spread along the PCA₂ since none of the shrub taxa (i.e. *Pistacia vera*, *Prunus bucharica* or *P. spinosissima*) are clearly significant neither in CA nor in PCA. The same indicator taxa characteristic of vegetation types found in TWINSpan are present as most contributing variables for PCA₁ and PCA₂. Some of these indicator taxa are very well known in the bibliography such as *Carex physodes*, *C. pachystylis*, *Ephedra gerardiana*, *Eremopyrum bonaepartis*, *Meniocus linifolius*, *Nepeta podostachys*, *Origanum vulgare* or *Prangos pabularia* (Gintzburger, 2003; Belolipov et al.,

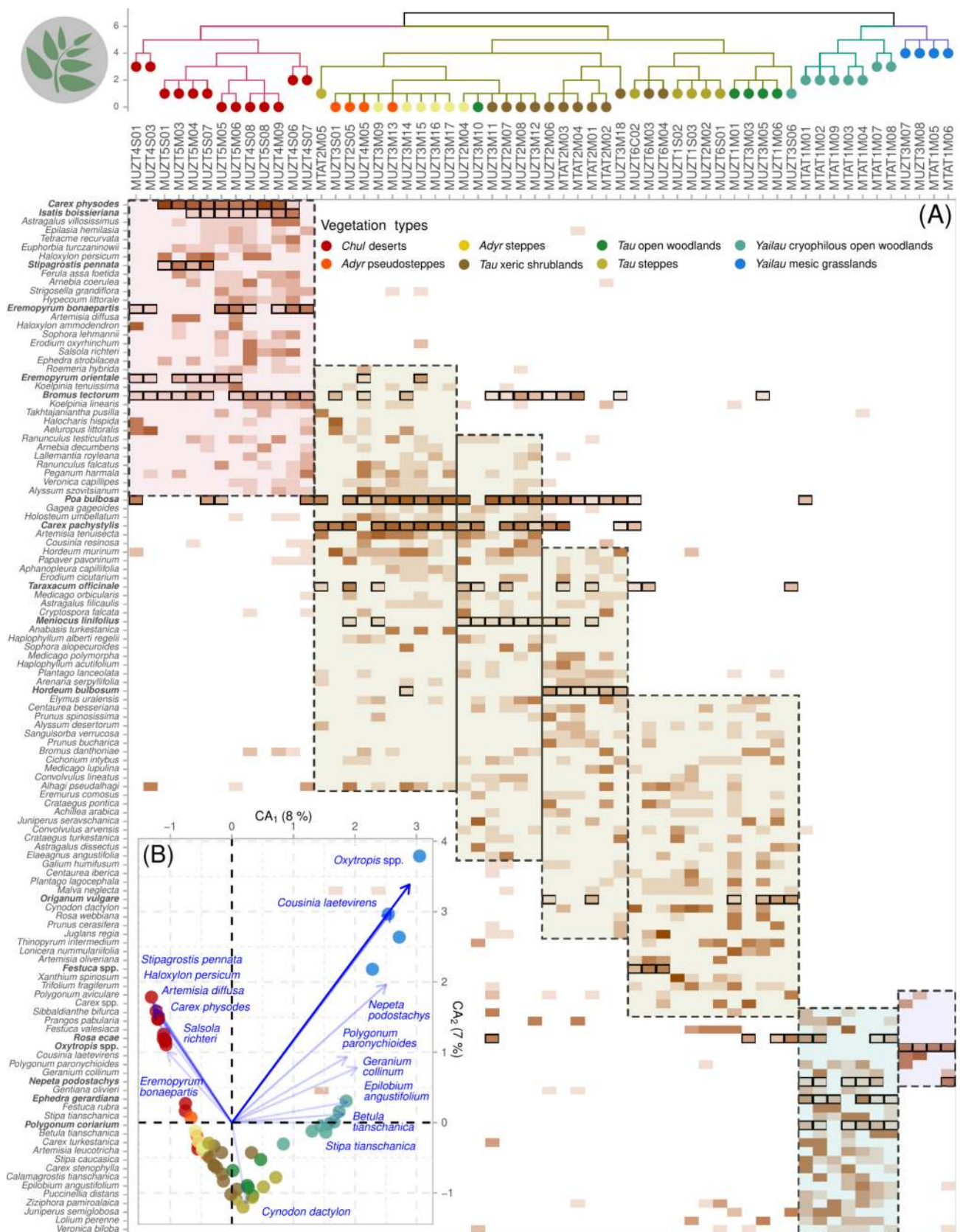


Fig. 5. (A) TWINSpan analysis of the vegetation plots, showing the Tajik-Uzbek major vegetation types. On this plot, only the 100 most relevant taxa are shown. For the TWINSpan analysis conducted on the total set of taxa, please refer to the Figs. S2-S5. Bold labels and tiles show indicator taxa, bullets on the dendrogram leaves highlight the major vegetation types. Colours on dendrogram segments and dashed rectangles highlight the four main TWINSpan clusters. (B) Correspondence Analysis made on 100 taxa, only the 15 most contributing arrows are displayed.

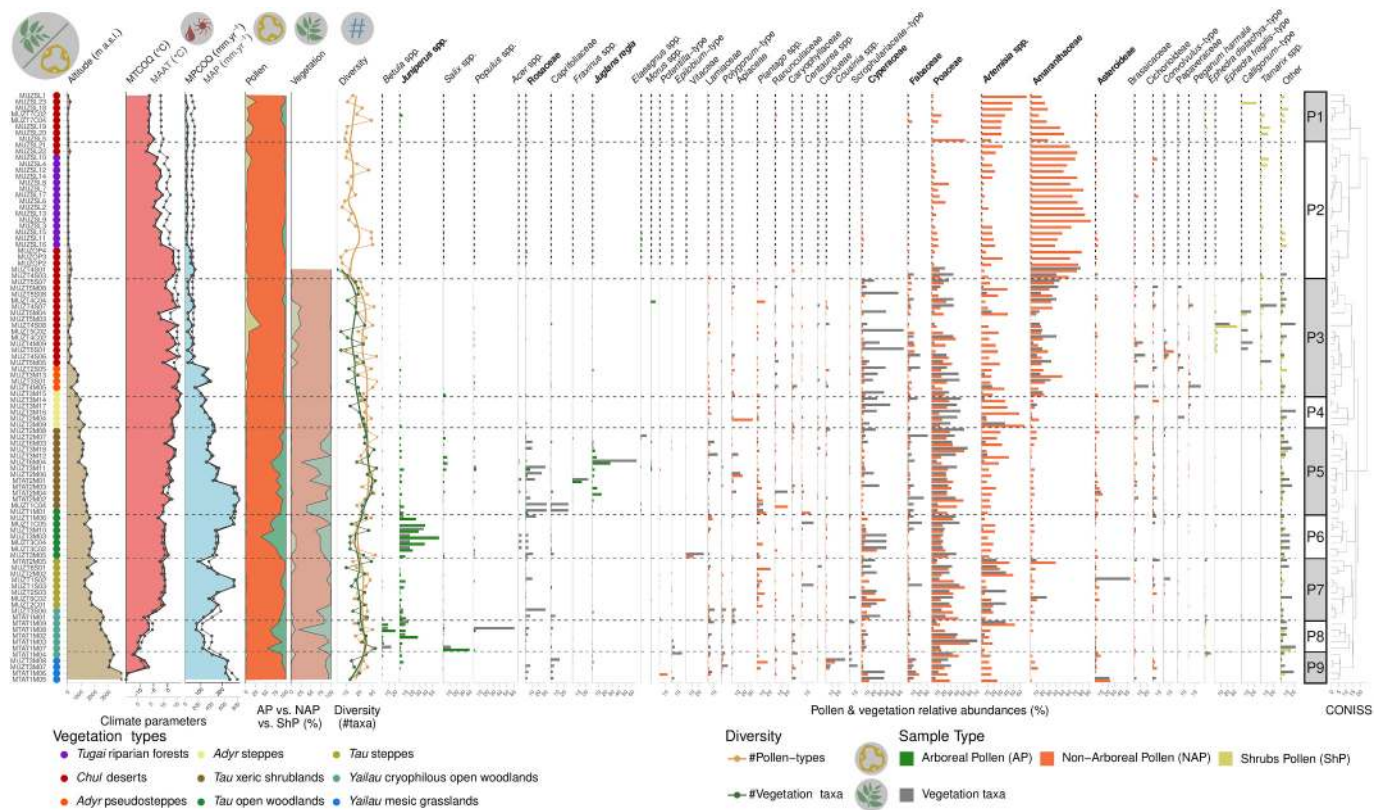


Fig. 6. Synthetic pollen vs. vegetation diagram of the TUSDB surface samples. From left to right, are displayed the coloured dots showing the ecosystems of each sample, the site labels, three bioclimate parameters (MPCOQ, MTCOQ and altitude in coloured area, MAT and MAAT in grey lines; all of them extracted from [Fick and Hijmans, 2017](#)); pollen growth forms (Arboreal, Non-arboreal and Shrubs pollen) and their analogues in vegetation plots; the main pollen types (colours by growth forms, AP green, NAP orange and ShP in yellow) and vegetation taxa (in grey) and the CONISS analysis clustering the pollen samples. The sites are sorted by ecosystems and bioclimatic values (mainly from lowland warm and arid desert in the upper part; to cold-wetter mountainous ecosystems at the bottom). Pollen types in bold are the taxa selected for [Fig. 9](#). The full pollen / vegetation diagram is available as [Fig. S6](#). Both pollen and plant taxa are expressed in fractional abundances (%). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2013; Nowak et al., 2020).

A similar site organization is evidenced on the RDA ($RDA_1 = 6\%$ and $RDA_2 = 5.1\%$ on [Fig. 7B](#)): RDA_1 is a temperature gradient with *yailau* sites on the positive side controlled mainly by altitude, *chul* and xeric shrublands on the negative side controlled mainly by MTCOQ and finally *adyr* and *tau* sites in the middle. The RDA_2 is a precipitation gradient with *chul* desert on the positive side, steppes in the middle and *tau* forests on the negative side. The combination of the two axes identifies and clusters the nine different vegetation types (coloured dots on [Fig. 7A](#) and [B](#)) merging altitudinal gradient (mainly temperature controlled) and the vegetation openness gradient (mainly precipitation controlled). Here, unlike the PCA results, it is possible to differentiate the *tau* and *yailau* open woodlands: *tau* open woodlands is maintained by high MPCOQ (associated to *Juniperus seravschanica*, [Fig. 5A](#)) while *yailau* cryophilous open woodlands benefits from high altitude and cold winter temperature mainly (i.e. *Juniperus semiglobosa* in the global TWINSpan, *Nepeta podostachys* and *Ephedra gerardiana* for indicator species). Similarly, the distinction between *adyr* steppes and *tau* steppes is clearer in RDA than in PCA (*adyr* steppe and pseudosteppes are closer to MTCOQ and *Bromus tectorum* while *tau* steppes are closer to *Origanum vulgare* and altitude).

About the multivariate distribution of pollen surface samples ([Fig. 7C](#)), the PCA_1 explains 23 % of the variance and PCA_2 9 %. PCA_1 is associated to the AP/NAP gradient with forested or mosaic vegetation for PCA_1 positive and mainly open vegetation for PCA_1 negative (i.e. mesic grasslands, deserts, pseudosteppes and steppes). The second axis follows the altitudinal ecological gradient with high altitude (*yailau* and *tau* sites) on PCA_2 negative; *adyr*, *chul* and *tugai* for positive PCA_2 . The

major contributing factors can be clustered in four groups: (1) shrub pollen and *Amaranthaceae* for low elevated and desert sites; (2) AP and *Juniperus* spp. for *tau* shrublands and woodlands; (3) *Poaceae*-*Cyperaceae* association for *tau* and *yailau* steppe and forest and finally (4) the bulk cortège of NAP for steppe/pseudosteppe ecotone. This ensemble may be described as a catch-all band, with the majority of rare pollen types mainly enhanced in the PCA results by the data scaling. Surprisingly, *Artemisia* spp. appears as the last of the 30 most important contributors: it shows that the *Artemisia* spp. signals can poorly describe the TUSDB variance, mainly because the *Artemisia* spp. FA_{pol} is ubiquitous in all vegetation types, and it does not fit with a particular pollen assemblage ([Fig. 6](#)).

The pollen-climate relation is quite strong and mainly driven by the first axis ($RDA_1 = 12\%$ and $R^2 = 0.18$ on [Fig. 7D](#)). Here, the first gradient is the precipitation one (MPCOQ, AI and altitude are really close to the first axis) with forest on the positive RDA_1 , steppes in the middle and desert on negative values. The second axis (*Betula* spp. versus *Asteroidae*-type) is weaker ($RDA_2 = 3\%$) and can be related to the temperature gradient *chul* cold pseudosteppes and *yailau* cryophilous open woodlands that have positive RDA_2 . However, the gradient is not really clear (*yailau* mesic grasslands sites are on the other side). About the pollen types-climate relations, it clearly appears that aridity is the main factor to explain the gradient between on one hand *Amaranthaceae*, ShP (here, it is mainly *Tamarix* spp., *Calligonum*-type, *Convolvulus* spp., both *Ephedra*-types) and on the other *Poaceae*, *Cyperaceae*, AP, (*Juniperus* spp. and *Juglans regia*). *Artemisia* spp. and *Fabaceae* are missing from the major pollen types in the RDA as well.

When comparing the two PCA (pollen and vegetation) using a Pro-

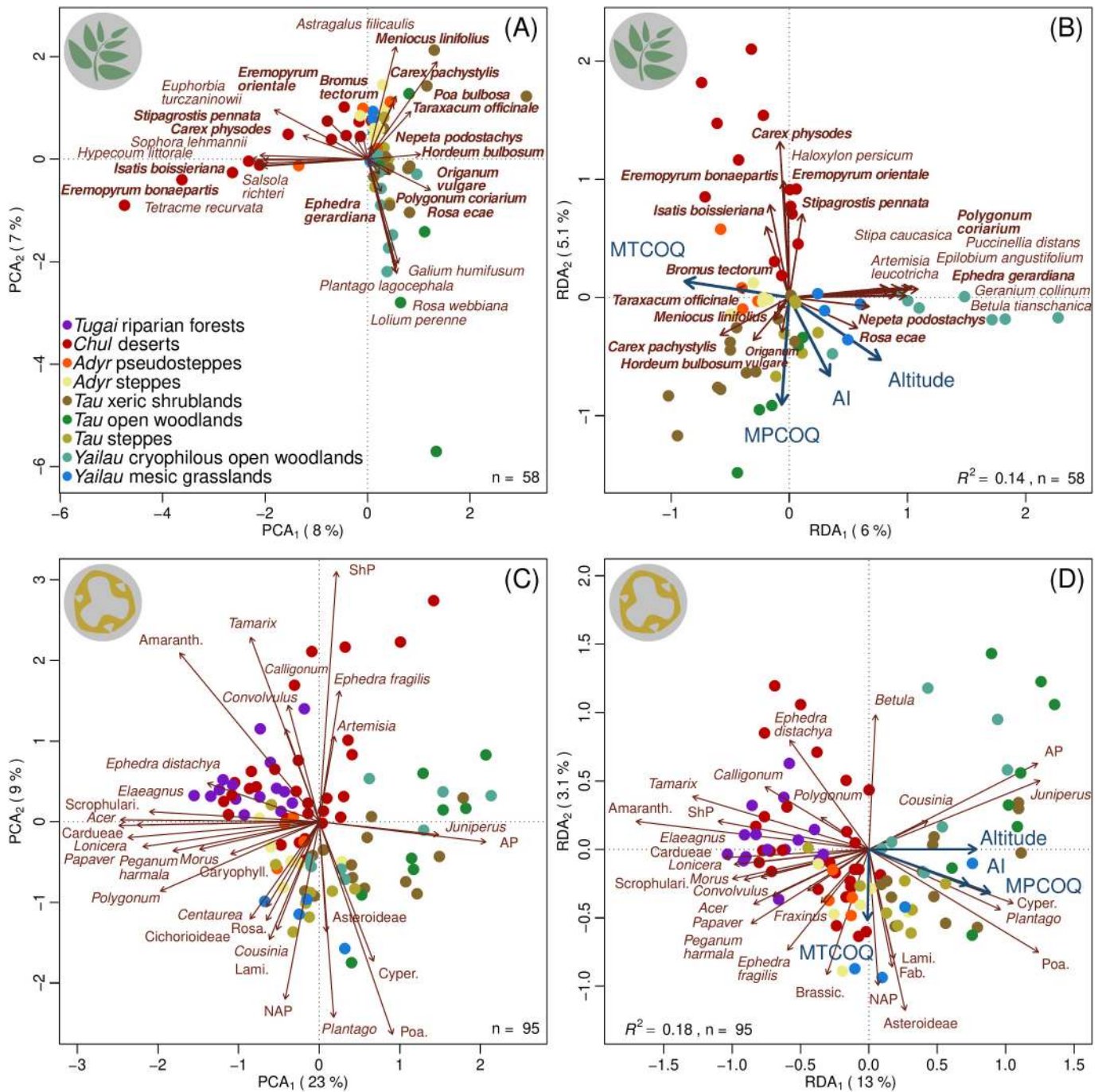


Fig. 7. Multivariate analysis based on vegetation plot, pollen samples and climate parameters: Principal Component Analysis (A) and Redundancy Analysis (B) based on vegetation plots; PCA (C) and RDA (D) based on pollen surface samples. The total variance explained by the two first components are shown on the axes. The dots present the sampling sites and are clustered by vegetation types. To ease clarity, only the 30 most contributing pollen types and vegetation taxa have been displayed as vector. Bold arrows correspond to the plant taxa selected by TWINSpan as ecological indicators (Fig. 5).

crustes rotation analysis (statistic used to compare two PCA spaces; Fig. S7), the two spaces are not perfectly stackable ($m^2_{12} = 140.41$ et PROTEST $r = 0.37$ with p -value < 0.003). However, this comparison is mitigated by the difference in explanatory variables (pollen types and vegetation taxa). When looking at the multivariate space between PCoI₁ and PCoI₂ (Fig. S7), it appears that the main difference comes from the important dispersion of outliers in vegetation plots (extreme values in Fig. 7A) while the pollen PCA (Fig. 7C) is more compact around zero. Mainly, looking at individual Procrustes error (Fig. S7B) about 30 % of the samples are above the mean error, the median is lower than the mean, which is driven by few extreme errors. Fig. S7C shows that the *tau*

xeric shrublands and *tau* open woodlands are the most subject to important variation between pollen and vegetation.

4.5. Pollen representativeness of local vegetation

Comparing vegetation plots and pollen surface samples at the same site gives a good overview of the representativeness biases (Fig. 6): in the TUSDB, Rosaceae (tree)-type, Caprifoliaceae, *Populus* spp., *Elaeagnus* spp., *Epilobium*-type, Scrophulariaceae-type are quite important in the vegetation plots but almost absent from the pollen assemblages. On the opposite, some taxa seem to have the FA_{pol} higher than the FA_{veg} . Davis

indices (Kent Davis, 1984) provide a simple and powerful statistical tool to determine the association between pollen and vegetation taxa, while the relative R-values (Wang et al., 2022) give a first quantitative approximation to correct the pollen production and preservation bias (Li et al., 2005).

Davis indices for TUSDB provide few unassociated taxa (Primulaceae, *Thalictrum* spp. and *Eremurus* spp., Fig. 8A). The unreliability of Primulaceae and *Thalictrum* spp. could be easily explained by the rareness of these taxa both in pollen and vegetation samples. Primulaceae flowering time is early spring, which can explain this rareness. However, *Eremurus* spp. is present in 25 % of the vegetation plots and 16 % of pollen samples. These taxa have no R-value because they are not present in any one location (Fig. 8B). Similarly, over-represented taxa (*Alnus* spp., *Corylus* spp., *Pinus* spp., etc.) are pollen types without any taxa in

vegetation, but *Juglans regia*, *Betula* spp., *Juniperus* spp. and Anacardiaceae are over-represented, even though they are balanced in some sites. Some taxa are strongly associated (Cyperaceae, Fabaceae, Poaceae, Asteroideae-type, Lamiaceae, Brassicaceae and Cichorioideae) while Amaranthaceae, *Artemisia* spp. Caryophyllaceae, Ranunculaceae, Apiaceae and Papaveraceae are associated taxa. These all have $\mu(R_{rel}) > 1$. Following the R-value, *Artemisia* spp. is the most over-represented pollen type followed by Papaveraceae, *Tamarix* spp., Polygonaceae, Boraginaceae, Scrophulariaceae-type and Amaranthaceae. The other pollen types are weakly associated mainly *Populus* spp., *Salix* spp., *Calligonum*-type, *Polygonum*-type, *Hippophae* spp., Rhamnaceae and Caprifoliaceae have also $\mu(R_{rel}) < 1$.

4.6. Spatio-ecological patterns

On Fig. 9, nine pollen types and vegetation taxa have been selected to spatially compare their FA: the dominant pollen types (*Artemisia* spp., Amaranthaceae, Cyperaceae and Poaceae, Fig. 6), the pollen types responding to temperature and precipitation (Asteroideae-type and *Juniperus* spp. respectively, Fig. 7D) and important vegetation taxa (Fabaceae, Rosaceae (tree)-type and *Juglans regia*, Fig. 5). The linear relation between pollen type and vegetation taxa is low for Cyperaceae ($R^2 < 0.01$), Fabaceae ($R^2 = 0.03$) and Asteroideae-type ($R^2 = 0.13$). These are all insignificant relations (p -values > 0.001). *Artemisia* spp. and *Juniperus* spp. have a correlation curve higher than the 1:1 line, showing that they are over-represented, while Rosaceae (tree)-type and *Juglans regia* have correlation curve largely lower than the 1:1 line. The scatter is more centred for Amaranthaceae and Poaceae. The FA distribution for pollen types and vegetation taxa is different among taxa. *Artemisia* spp. have $FA_{pol} > 20\%$ everywhere, while the FA_{veg} is higher in mountains and close to 0 in lowlands, showing the altitudinal dispersion of *Artemisia* spp. pollen grain, certainly due to downdraft transport. The same pattern is observable for Fabaceae and Poaceae. Although Rosaceae (tree)-type have the same FA_{veg} spatial spreading (Tian Shan, Hissar, Kuhistan and Pamir areas), the FA_{pol} is always lower on local area and null on far away sites. For Amaranthaceae, the dispersion comes from desert (high FA_{veg} in Kyzylkum and Sudunkli) to *adyr* steppes (mainly in Kuhistan and Hissar). No pattern is easily observed for Asteroideae-type. The dispersion of *Juglans regia* is also important (much more sites with high FA_{pol} than FA_{veg}), but it is much localized in mountain areas (Nuratau, Kuhistan and Hissar) without spreading in *adyr* or *chul*. This behaviour is shared among other minor trees such as *Fraxinus* spp. and *Salix* spp. (Fig. S8). In vegetation, *Betula* spp. is only observed in the Pamir region at a considerable distance from the location where pollen is found (up to Kyzylkum, Fig. S8).

The FA distribution for pollen types and vegetation taxa is different among taxa. *Artemisia* spp. have $FA_{pol} > 20\%$ everywhere, while the FA_{veg} is higher in mountains and close to 0 in lowlands, showing the altitudinal dispersion of *Artemisia* spp. pollen grain, certainly due to downdraft transport. The same pattern is observable for Fabaceae and Poaceae. Although Rosaceae (tree)-type have the same FA_{veg} spatial spreading (Tian Shan, Hissar, Kuhistan and Pamir areas), the FA_{pol} is always lower on local area and null on far away sites. For Amaranthaceae, the dispersion comes from desert (high FA_{veg} in Kyzylkum and Sudunkli) to *adyr* steppes (mainly in Kuhistan and Hissar). No pattern is easily observed for Asteroideae-type. The dispersion of *Juglans regia* is also important (much more sites with high FA_{pol} than FA_{veg}), but it is much localized in mountain areas (Nuratau, Kuhistan and Hissar) without spreading in *adyr* or *chul*. This behaviour is shared among other minor trees such as *Fraxinus* spp. and *Salix* spp. (Fig. S8). In vegetation, *Betula* spp. is only observed in the Pamir region at a considerable distance from the location where pollen is found (up to Kyzylkum, Fig. S8).

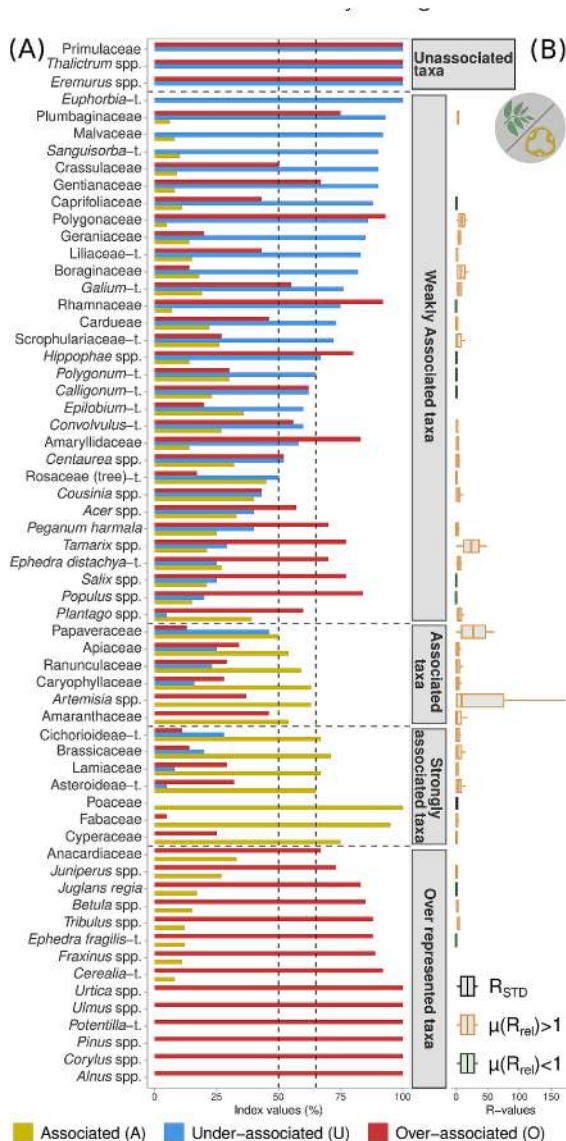


Fig. 8. Graphical representation of the pollen type representativeness factors: (A) the Davis association indices calculated for the taxa present at least once in both pollen samples and vegetation plots; (B) relative R-values (R_{rel} , Wang et al., 2022). Standard R-value is settled to Poaceae. Mean R_{rel} above one shows a higher pollen abundance than vegetation (green boxes); and below one (yellow boxes) is the contrary. Davis threshold are shown in dashed lines and the modelled association category are given in the middle. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

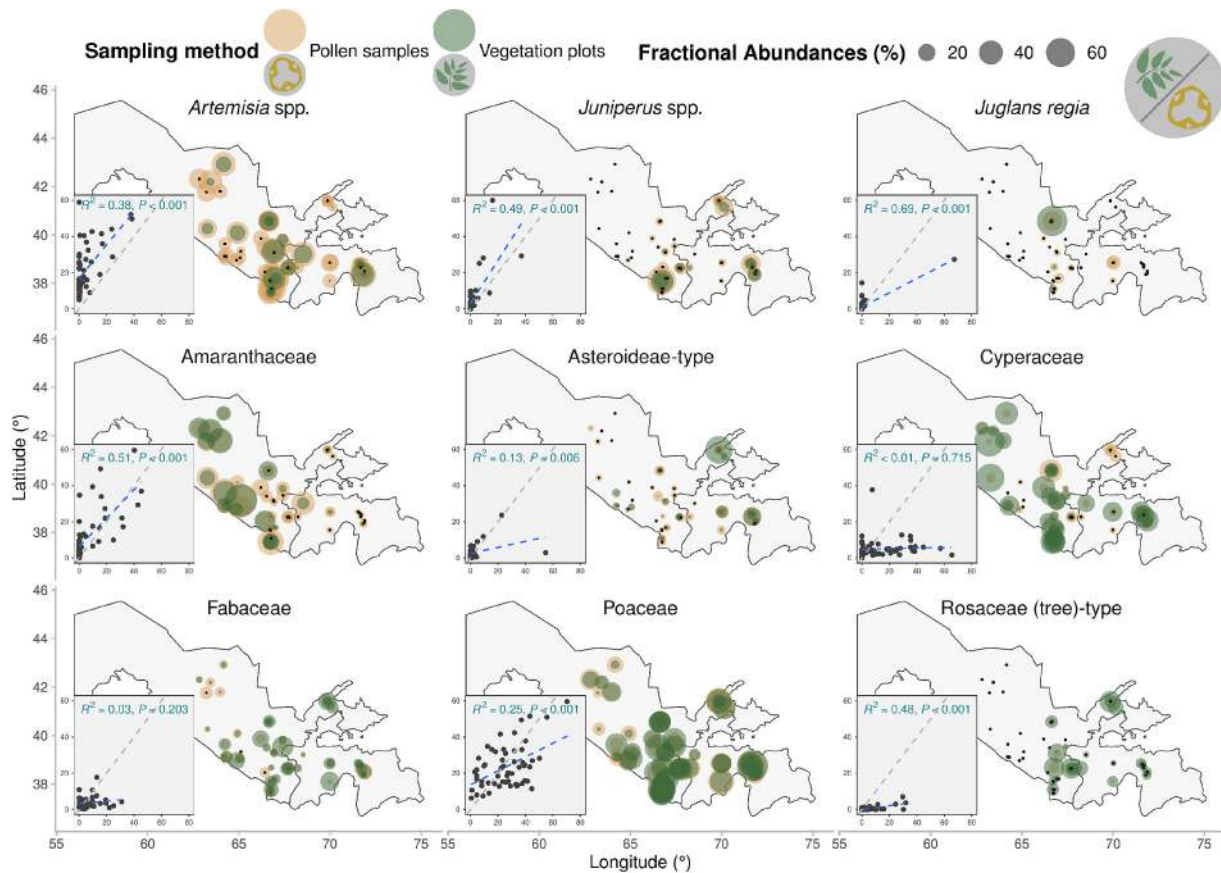


Fig. 9. Tajik-Uzbek map of the pollen (yellow dots) and vegetation (green dots) fractional abundances for nine relevant pollen types. The small black dots display the sites without any pollen and vegetation fractional abundances. The inserts shown the linear relation between plant species from the vegetation plots (x) and the pollen types for the pollen surface samples (y). To see the maps for all the TU pollen types, see Fig. S8. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.7. Plant functional traits

The average CWM values for the six traits were compared for all sites (Fig. 10 A1 to A6). Results show a relatively even distribution of CWM values among communities on the scatter plots, showing that the pollen-based and vegetation-based CWM values do not depend on the type of community at stake. However, for plant height, N_{leaf} and SSD, an open land to forested community gradient is distinguishable. Then, the relationships between pollen and vegetation-based CWM values are always significant, except for N_{leaf} . The relationships are relatively good for specific leaf area ($R^2 = 0.33$) and seed mass ($R^2 = 0.39$) and stronger for stem specific density, leaf area ($R^2 = 0.44$) and plant height ($R^2 = 0.49$).

PCA conducted on pollen-based (Fig. 10B) and vegetation-based CWM values (Fig. 10C) show that the two functional spaces were comparable and consistent with ACA (Dugerdil et al., 2025) and worldwide (Bruehlheide et al., 2018) functional spaces. The first two components explained 72 % and 81 % of the variance for vegetation plots and pollen samples, respectively. In both cases, the CWM trait contribution order is from higher contributors to lower leaf area, SSD and plant height, SLA and finally the lowest contributing factors seed mass and N_{leaf} . In both cases, the PCA_1 reflects a contrast between tall woody plants and small (i.e. herbaceous) plants, while PCA_2 corresponds to a fast-slow gradient of leaf functioning. This second axis also discriminates species on a leaf size gradient. Shrubland and woodland sites have high PCA_1 and PCA_2 (plant woody plant with large leaves), except the *tau* open woodlands which have high plant height and SSD, but low leaf area. Steppes and grasslands show negative values on PCA_1 and PCA_2 (high SLA and leaf

nitrogen content, short plants with low SSD), while desert and pseudosteppes are found for values around zero on PCA_1 and slight negative values on PCA_2 .

However, some differences are visible between pollen and vegetation, as shown by the Procrustes analysis (Fig. 10D). Mainly desert samples have smaller leaf area in pollen samples than in vegetation plots (i.e. desert sites have more negative PCA_2 loading as highlighted by the downstream vectors on Fig. 10D). Moreover, for *Juniperus* spp. dominant vegetation (*tau* open woodlands and *yailau* cryophilous open woodlands) the pollen-based CWM traits have higher PCA_1 than the vegetation-based ones. It means that the recorded plant high and SSD are different among the sampling methods. In any cases, the majority of site projection on the two functional spaces are quite consistent as shown by good Procrustes analysis statistics ($m^2_{12} = 20.3$) and PROTEST significance test results ($r = 0.67$ and $p\text{-value} > 0.001$).

4.8. Pollen climate calibration

The four climate reconstruction methods (MAT, WAPLS, RF and BRT) were tested with four different modern pollen datasets, i.e. from regional to global ones (TUSDB, COSTDB, ACADB and EAPDB) to reconstruct two key climate parameters for the study area (MPCOQ and MTCOQ). The statistical results displayed in Table 1 show that precipitations are better reconstructed than the temperature. About the precipitation calibration effect, the smaller and more local the dataset is, the lower is the RMSE. In consequence, the precipitation RMSE is generally lower for TUSDB, then COSTDB, ACADB and higher for EAPDB. The effect is opposite for the linear correlations (increasing R^2 with the dataset extend). It is particularly visible for the machine

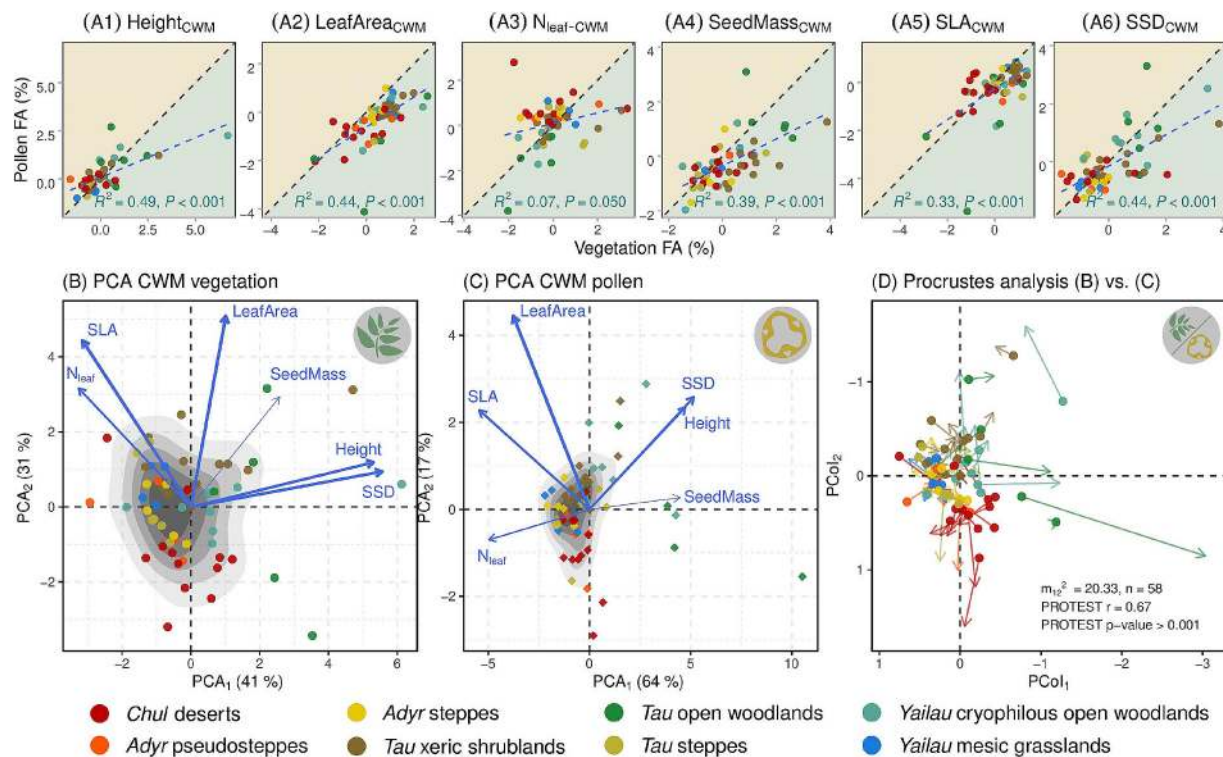


Fig. 10. Comparisons of CWM trait values based on surface pollen samples and vegetation plots. (A) pairwise relationships between vegetation-based CWM (x-axis) and pollen-based CWM values (y-axis) for the six studied traits: stem specific density (SSD, A1), plant height (A2), specific leaf area (SLA, A3), leaf area (A4), N_{leaf} (A5) and seed mass (A6). The multivariate analysis displays principal component analysis (PCA) made on vegetation-based CWM traits (B), pollen-based PCA (C) and the Procrustes analysis comparing the two PCA spaces (D). Here, the dots represent the vegetation scores and the arrows the migration of the pollen score.

learning methods (for BRT and MPCOQ that have $R^2 = 0.83, 0.89, 0.89$ and 0.90 , sorted by increasing spatial extent; for BRT and MTCOQ that have $R^2 = 0.39, 0.90, 0.90$ and 0.92). It appears that the statistical results are always giving the BRT first (for the same calibration dataset and the same climate parameter, BRT has the higher R^2 and the lower RMSE), followed by MAT, WAPLS and RF. The impact of the selected parameters for MAT and WAPLS is also fairly different: for WAPLS the best statistical result (i.e. the k lower number of PLS components to reach the higher R^2 and the k for the lower RMSE) is almost always similar: the same k drives to best R^2 and RMSE (generally around two or three components). The WAPLS is not drastically impacted by the number of components. On the opposite, the k number of analogues to be chosen in the MAT is largely influencing the R^2 and RMSE (Table 1). Finally, the Moran's I show how the climate parameter residuals are spatially auto-correlated for all calibrations. Almost all the Moran's I are statistically relevant (p -values < 0.001) and no datasets are spatially dispersed ($I > 0$). The majority of the calibration are slightly auto-correlated ($I \in [0; 0.3]$). Only the TUSDB (smallest dataset) for MTCOQ are importantly spatially clustered (with $I > 0.4$).

Since the climate reconstruction methods based on machine learning are mostly considered as *black boxes* (Simon et al., 2023), the univariate relationships between pollen types and climate parameters have been explored. Linear relations between climate parameters and key pollen types (Fig. S9) show that (1) *Amaranthaceae* and *Artemisia* spp. are anticorrelated with MPCOQ, MTCOQ, AI and MAP, while *Poaceae* is correlated to them (and even more than in Mongolia; Dugerdil et al., 2021), (2) *Cyperaceae* are correlated with AI and MAP, but anticorrelated with MPCOQ and MTCOQ. One of the important results here is that the trend is fairly similar whatever the size of the calibration dataset, which is not always the case, especially for local climate calibrations (Dugerdil et al., 2021). Only the *Artemisia* spp. relationships with aridity (AI, MPCOQ, MAP) are different between TUSDB (no

correlation) and larger datasets (anticorrelations).

The link between univariate pollen type behaviour and its influence on climate reconstructions is explored. The most contributing taxa within the BRT and RF algorithms and the Pearson correlation r of the main pollen types are displayed in Table 2. *Amaranthaceae*, *Cyperaceae* and *Artemisia* spp. are the most influent pollen types for machine learning. Especially, *Amaranthaceae* is very important in TUSDB with respectively 69, 62 and 18 % of contribution for BRT-based MPCOQ, AI and MTCOQ. For the RF results, we found 14.9, 8.8 and 8.1 %. The *Amaranthaceae* are more contributing to precipitation reconstruction than temperature. Similarly, they are large in TUSDB but also important in ACADB (17, 9 and 7.7 % for AI, MTCOQ and MPCOQ). It becomes less important in COSTDB and even less in EAPDB with 7.2 and 2.3 % for the AI. The weak correlation between climate and *Artemisia* spp. in TUSDB (contrarily with other datasets) affects the *Artemisia* spp. contribution to RF (3.1, 2.8 and 2.6 %) and BRT (6.9, 3.5 and 2.8 %). Surprisingly, even though the linear relation between *Cyperaceae* and climate parameters is not the strongest one ($r \in [-0.20; 0.39]$), its influence in the BRT results with a contribution between 3.6 % for MPCOQ-EAPDB (the most important contribution) to 15 % for MTCOQ-TUSDB. Finally, it appears that the smaller is the dataset, the most the contribution is centred on few taxa. Focused on MPCOQ, with the TUSDB 86.8 % of the BRT contributions are explained by the five key taxa against only 39.9 for COSTDB, 16.1 for ACADB and 8.9 % for EAPDB. However, cross-validation and calibration relative contribution assess the calibration reliability, but do not assess the specific past sequence applicability.

4.9. Aral 86 application test

The Aral 86 core (Aleshinskaya, 1991) has been retrieved before 1991 in the eastern basin of the Aral, a place that dried up in 2007 (aerial view). The age-depth model is displayed in Fig. S10. The sequence covers 8500 yr cal. BP. The pollen record has a very low pollen diversity

Table 1

Cross-validation statistical results for the climate reconstruction training resulting for cross-validation of transfer functions (MAT, WAPLS) or machine learning regressions (RF, BRT) with k, number of parameters (e.g. number of analogues for the MAT and number of PLS components for the WAPLS); R^2 , the linear coefficient of determination; RMSE the calibration error and the Moran's I indicator of spatial autocorrelation with its p-value. Best k shows the number of parameters that infer the best R^2 or RMSE values. The selected k shows the R^2 and RMSE results for the k lowest number of parameter (i.e. the statistical results for the least over-parametrized model).

Climate parameters ^a	Database	N _{sites}	Method	Best k	R^2	Best k RMSE	Selected k	R^2	RMSE	Moran's I ^b
MPCOQ	EAPDB	5411	MAT	4	9	4	4	0.85	58.10	0.09*
			WAPLS	2	2	2	2	0.46	96.70	0.35*
			BRT	–	–	–	–	0.90	56.10	0.11*
			RF	–	–	–	–	0.65	77.40	0.23*
	ACADB	2620	MAT	4	6	4	4	0.83	33.20	0.14*
			WAPLS	2	2	2	2	0.62	45.10	0.25*
			BRT	–	–	–	–	0.89	34.00	0.11*
			RF	–	–	–	–	0.58	46.80	0.25*
	COSTDB	1212	MAT	4	4	4	4	0.73	37.90	0.28*
			WAPLS	5	2	2	2	0.45	53.50	0.35*
			BRT	–	–	–	–	0.83	35.30	0.14*
			RF	–	–	–	–	0.48	46.40	0.3*
	TUSDB ^c	95	MAT	5	10	5	5	0.73	44.20	0.2
			WAPLS	2	2	2	2	0.61	50.90	0.22*
			BRT	–	–	–	–	0.81	44.80	0.29*
			RF	–	–	–	–	0.55	50.70	0.35*
MTCOQ	EAPDB	5411	MAT	4	9	4	4	0.86	4.36	0.14*
			WAPLS	3	3	3	3	0.63	6.29	0.27*
			BRT	–	–	–	–	0.92	3.95	0.14*
			RF	–	–	–	–	0.70	5.58	0.21*
	ACADB	2620	MAT	4	10	4	4	0.75	4.33	0.19*
			WAPLS	2	2	2	2	0.53	5.33	0.32*
			BRT	–	–	–	–	0.86	3.90	0.18*
			RF	–	–	–	–	0.55	5.11	0.25*
	COSTDB	1212	MAT	4	8	4	4	0.61	4.86	0.19*
			WAPLS	2	2	2	2	0.39	5.57	0.29*
			BRT	–	–	–	–	0.80	4.12	0.16*
			RF	–	–	–	–	0.39	5.36	0.23*
	TUSDB	95	MAT	4	6	4	4	0.28	3.99	0.43*
			WAPLS	5	2	2	2	0.13	4.48	0.39*
			BRT	–	–	–	–	0.40	4.13	0.46*
			RF	–	–	–	–	0.23	3.74	0.46*

^a The climate parameters reconstructed are the Mean Temperature of Cold Quarter (MTCOQ) and the Mean Precipitation of Cold Quarter (MPCOQ), basically winter temperature and precipitation. The same table for MAP and AI can be found on Table S7.

^b the * display the Moran's I with p-values <0.001.

^c bold values highlight the statistical results for calibration particularly discussed in the text.

Table 2

Statistical scores for the TUSDB major pollen types (*Artemisia* spp., *Amaranthaceae*, *Cyperaceae*, *Juniperus* spp. and *Poaceae*). The Pearson correlation r and the relative pollen types contribution to the BRT and RF models are shown for each calibration dataset (ACADB, COSTDB, EAPDB and TUSDB) and each selected climate parameters (AI, MTCOQ, MPCOQ). Best values for each column are highlighted in bold.

Climate parameters ^a	Pollen- types	Pearson r				Contrib.				BRT (%) ^b				Contrib. RF (%) ^b			
		ACADB	COSTDB	EAPDB	TUSDB	ACADB	COSTDB	EAPDB	TUSDB	ACADB	COSTDB	EAPDB	TUSDB	ACADB	COSTDB	EAPDB	TUSDB
AI	<i>Artemisia</i> spp.	–0.3	–0.38	–0.39	–0.13	1.7	2.3	1.4	3.5	7.3	8.1	6.1	2.6				
	<i>Amaranthaceae</i>	–0.47	–0.4	–0.39	–0.59	17	7.2	2.7	62	11.4	6.8	6	14.9				
	<i>Cyperaceae</i>	0.15	0.26	–0.06	0.37	6.3	21	8	8.6	2.4	4.7	1.3	7.6				
	<i>Juniperus</i> spp.	0.01	0.01	0.07	0.32	0.14	0.13	0.15	8.4	0.5	1	2.1	9.1				
	<i>Poaceae</i>	0	0.15	0.06	0.5	0.95	1.4	1.2	2.5	1.2	1.8	1.1	5.8				
MTCOQ	<i>Artemisia</i> spp.	–0.09	–0.04	–0.18	0	1.6	3.2	1	6.9	1.2	2.3	1.2	3.1				
	<i>Amaranthaceae</i>	0.12	0.11	0.01	0.01	9	8.6	1.8	18	2.3	1.9	1.1	8.1				
	<i>Cyperaceae</i>	–0.12	–0.1	–0.2	–0.02	8.4	11	4.8	15	1.4	1.7	1.5	5.1				
	<i>Juniperus</i> spp.	0.05	0.07	0.14	–0.05	0.11	0.09	0.12	13	0.5	0.7	2.3	2.5				
	<i>Poaceae</i>	0.09	0.1	0.1	0	1.6	2.1	0.92	6	1	1.8	1.1	4.7				
MPCOQ	<i>Artemisia</i> spp.	–0.19	–0.22	–0.37	–0.12	1.4	2.4	1.5	2.8	2	4.2	4.2	2.8				
	<i>Amaranthaceae</i>	–0.1	–0.03	–0.25	–0.67	7.7	15	2.2	69	1.1	1.3	2	8.8				
	<i>Cyperaceae</i>	–0.13	–0.1	–0.2	0.39	5.8	21	3.6	9.8	1.1	1.2	1	5.5				
	<i>Juniperus</i> spp.	0.13	0.15	0.14	0.42	0.072	0.047	0.28	3.8	2.2	2.7	2.5	12.4				
	<i>Poaceae</i>	0.13	0.29	0.13	0.54	1.1	1.5	1.3	1.4	1.4	3	1.1	10.6				

^a The climate parameters reconstructed are the Aridity Index (AI), the Mean Temperature of Cold Quarter (MTCOQ) and the Mean Precipitation of Cold Quarter (MPCOQ).

^b Taxa contribution for BRT and RF regression machine learning algorithms are relative values for each model and each predicted climate parameter in order to compare them.

with 36 different pollen types, largely dominated by *Artemisia* spp. (> 40 %), *Amaranthaceae* (> 20 %), *Poaceae* (> 10 %), *Cyperaceae*, *Pinus* spp. (undet.) and *Betula* spp. (all $\approx$ 5 %). These key taxa represent an average of 80 % of each pollen sample along the core. NAP range between 80 and 90 % along the core (Fig. S11) and AP are represented by long-spreading arboreal pollen producers (*Pinus* spp. and *Carpinus* spp. are not naturally present in Uzbekistan, and *Betula* spp. and *Alnus* spp. over-represented on Fig. 8). Only the three major taxa (i.e. *Artemisia* spp., *Amaranthaceae* and *Poaceae*) have been plotted in Fig. 11A.

The CONISS analysis clusters the pollen-sequence in five main pollen

units: from 8400 to 4100 yr cal BP, zone largely dominated by *Artemisia* spp. and *Amaranthaceae* (i.e. close to **P3** *chul* warm desert); from 3800 to 2400 yr cal BP with rising *Poaceae*, peak of *Cyperaceae* and AP signal, pollen diversity decreases; from 2100 to 1500 yr cal BP the *Poaceae*, *Cyperaceae* and *Pinus* spp. (undet.) replace the *Artemisia* spp. and *Amaranthaceae* (more similarly to the **P6** or **P7** current pollen signal, such as *tau* open woodlands); and then, from 1500 to 200 yr cal BP, a second phase of *Artemisia* spp. and *Amaranthaceae* similar to **P1** (i.e. *adyr* pseudosteppes) or **P4** (i.e. *adyr* steppes, Fig. 6). Finally, the top zone is made on one sample with an *Artemisia* spp. maximum mainly due to recent human changes. From these five zones, we distinguished three outstanding zones named **Z1** (3800 to 2400 yr cal BP), **Z2** (2100 to 1500 yr cal BP) and **Z3** (200 to -70 yr cal BP).

From Aral 86, the most accurate climate parameters to reconstruct are AI, MTCOQ and MAP (Table 3). The biome reconstruction (Fig. 11B) shows the dominance of two types of steppe (i.e. cold and warm steppes, COST and WAST) followed by cold desert (CODE score $\approx$ 6) and tundra-like (TUND score $\approx$ 4). In **Z3**, the COST/WAST reach their culminating point in parallel with the lowest value for CODE, while in **Z1–2**, the TUND values bump up to the CODE level (TUND score from 4 to 7). About the biome scores, cold and warm steppes are dominant along the core and follow a similar trend (Fig. 11B). Cold deserts present intermediate scores, continuous along the sequence (very similar to the *Amaranthaceae* FA). The last important biome is the tundra-like, especially with important rises during the phases **Z1** and **Z2**. These pollen-based biomes are consistent with the CWM functional traits (Fig. 11C): tundra-like is characterized by high SSD and seed mass (importance of shrubs and scrubs). The plant height remain steady since the biomes remain open (deserts, steppes and tundra-like) along the core. However, the leaf area presents surprising results: it is high during the tundra-like phases (mainly due to *Poaceae* FA).

5. Discussion

5.1. Toward a pollen-based vegetation type reconstruction

5.1.1. Four main vegetation types and a continuous gradient among them

Altitudinal gradient effect and limits. The vegetation plots analysis shows that the traditional Uzbekistan vegetation types clustering based on altitudinal gradient (i.e. *chul*, *adyr*, *yailau* and *tau*; Zakirov, 1955; Gintzburger, 2003; Belolipov et al., 2013) is quite accurate to describe Uzbek vegetation at a coarse division (especially when applied to pollen rain interpretation). At the scale of vegetation types, the altitudinal gradient is not sufficient to explain the wide TU vegetation diversity. Not only azonal vegetation occurs on riversides (Thevs et al., 2008), but also local slope or ground compaction and stability (Nowak et al., 2021) and even the soil thickness (Dulamsuren et al., 2005) disturb the gradient. In the highest elevation area, the level of endemism (Radula et al., 2021; Nowak et al., 2022) also mitigates this simple gradient classification. However, from pollen record, the averaging effect of the pollen rain reduces local vegetation detection (Bunting et al., 2004). Except some particular azonal behaviour in *tugai* community (e.g. Herzsuh et al., 2003) or azonal wet woodlands (e.g. Li et al., 2005), pollen record a coarser vegetation signal.

Vegetation subtypes and phytosociological classes. The distinction in our nine vegetation subtypes (*chul* deserts, *tugai* riparian forests, *adyr* pseudosteppes and steppes, *tau* xeric shrublands, open woodlands and steppes, *yailau* cryophilous open woodlands and mesic grasslands) is less clear, especially since both open and closed vegetation occur in each of the four main vegetation types (i.e. only *chul*, *adyr*, *tau* and *yailau*). Then, it appears that the altitudinal vegetation gradient is less clear and simple than it could be in temperate areas.

For instance, the riverside vegetation is spread in flood plains and along salt marshes from alpine mountain to desert (i.e. the Cl. *Populetea euphraticae* Zohary 1962 from *tugai* in the lowlands; Thevs et al., 2008) to riparian shrublands at higher altitudes (i.e. the Cl. *Salici capusii*-

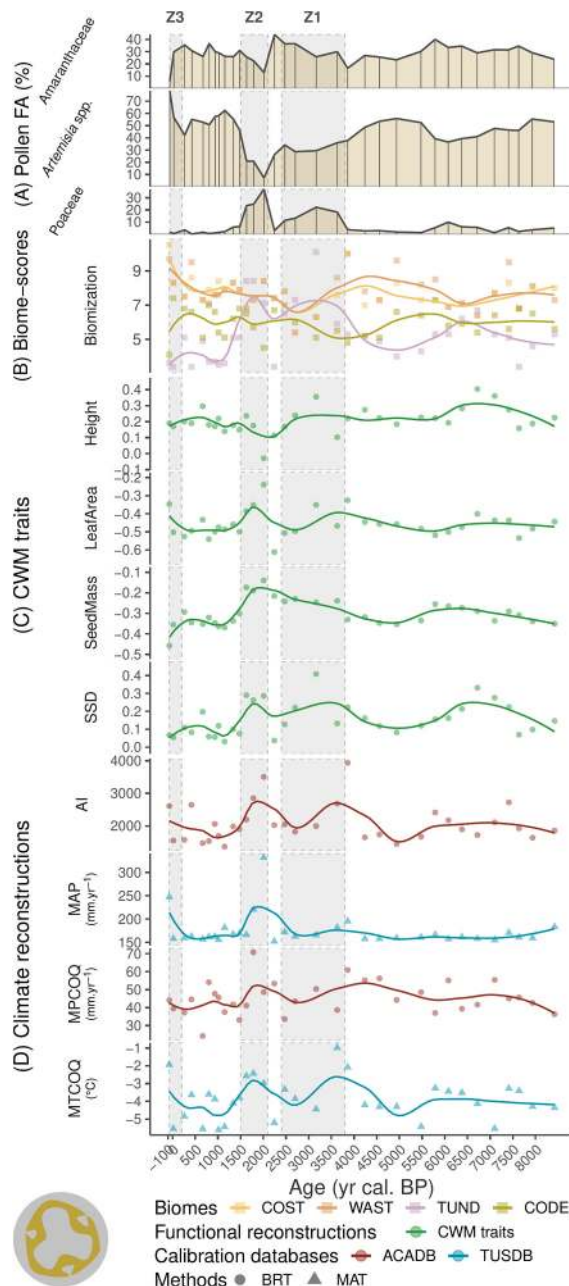


Fig. 11. TUSDB calibration applicability of the Aral 86 core. (A) the three dominant pollen types along the sequence; (B) the pollen-based biomes using PFTs and biomization approach results. Only the four major biome scores are displayed: Cold and Warm Steppe (COST, WAST), Cold Desert (CODE) and Tundra-like (TUND). (C) CWM traits reconstruction (plant height, leaf area, seed mass and SSD) and (D) reconstructed climate parameters based on multi-method and multi-calibration set. The sequence is plotted in age (year cal. BP), refer to Fig. S10 for the age-depth model.

Table 3

Proportion of variance for each Aral 86 climate reconstructions based on pollen transfer functions. These statistical results come from the `randomTF()` significance test developed to estimate the reliability of MAT and WAPLS climate reconstruction for Aral 86 based on the different calibration sets (TUSDB, ACADB and COSTDB) compared to transfer function trained on random calibration sets (for 999 permutations, Telford and Birks, 2011). * displays predictions with p -values >0.1 and bold values highlights the predictions above the 95 % threshold. The complete histograms of the `randomTF()` can be found on Fig. S12.

Database	Reconstruction approach	Climate MTCOQ	parameter MPOQ	AI	MAP	Threshold $p = 0.05^a$	PCA ₁ (var.%) ^b
ACADB	MAT	0.06	0.09	0.11	0.07	0.18	0.52
	WAPLS	0.03	0.06	0.27*	0.21	0.28	0.52
	RF	0.03	0.06	0.15	0.15	0.20	0.52
	BRT	0.02	0.03	0.16	0.08	0.27	0.52
COSTDB	MAT	0.05	0.08	0.08	0.07	0.21	0.52
	WAPLS	0.03	0.11	0.28*	0.22*	0.21	0.52
	RF	0.04	0.05	0.31**	0.1	0.21	0.52
	BRT	0.03	0.02	0.21	0.1	0.27	0.52
TUSDB	MAT	0.03	0.13	0.21	0.18	0.39	0.56
	WAPLS	0.21*	0.21	0.12	0.17	0.26	0.56
	RF	0.07	0.28	0.31	0.27	0.38	0.56
	BRT	0.03	0.08	0.1	0.14	0.41	0.56

^a this threshold corresponds to the 95 % interval values for the 999 permutation of randomly-based transfer functions, i.e. climate reconstructions with a proportion of variance above the proportion of variance of 95 % of the randomly made transfer functions (Chen et al., 2017).

^b this percentage of variance (PCA₁) theoretically corresponds to the maximum of explainable past pollen variance.

Hippophaeetalia rhamnoidis; Nowak et al., 2024) with very different species composition.

Moreover, shrublands are present in (1) *chul* deserts (benefited on the *Haloxylon aphyllum* and *H. persicum*, i.e. black and white *saxaoul* dry-tolerant physiology as well as Fabaceae shrubs, *Calligonum* spp., *Convolvulus* spp. and *Ephedra* spp.; Orlovsky and Birnbaum, 2002), (2) in mosaic within *adyr* steppes or in xeric shrublands using local *tau* wetter conditions (Nowak et al., 2022) and (3) even in the *yailau* belt with mountain riparian shrublands (e.g. with *Salix* spp., *Betula* spp. and *Hippophae rhamnoides*) on within cryophilous open woodlands (e.g. *Juniperus* spp. crawling form or with cushion-shrubs vegetation of *Acantholimon* spp. and *Astracantha* spp.; Nowak et al., 2020).

Then, the indicator taxa are roughly similar between deserts and pseudosteppes (e.g. *Eremopyrum orientale* and *Bromus tectorum*) or between steppes and open woodlands (e.g. *Poa bulbosa* and *Carex pachystylis* in *adyr* and *Origanum vulgare*, *Festuca* spp. and *Rosa ecae* in *tau*; Fig. 5). This effect also reduces the taxa particularities among vegetation subtypes and mitigates the accuracy of threshold settlement between ecotones in TU. To conclude, our study is only based on 58 vegetation plots, which drastically limits its reliability since this sampling cannot cover the whole range of TU vegetation diversity.

Bioclimate-vegetation subtype relationships. However, when analysing the bioclimate parameter influence on the vegetation type distributions, it is then possible to refine the number of possible clusters (i.e. vegetation subtypes). From the four main vegetation types distinguished in the previous part, some ecophysiological particularities separate vegetation subtypes among open woodlands. Especially, on Fig. 7B, it appears that the two TUSDB types of open woodlands (i.e. *yailau* cryophilous and *tau* open woodlands) are clearly different: the first one is suitable for high altitude with very cold winter temperature, while the second one is preferentially found in wetter environment (driven by MAP amount). This is principally due to the different ecophysiological adaptations of *Juniperus seravschanica* and *J. semiglobosa* mainly (Nowak et al., 2020). It is not only juniper characteristics that may explain these differences between *yailau* and *tau* forest-steppes; indeed other taxa within the same PFT such as juniper, i.e. *arctic-alpine dwarf shrub* (aa; Prentice et al., 1996) such as *Betula tianschanica* or *Rhamnus grubovii* are also present in *yailau* open woodlands and can explain this vegetation adaptation to *yailau* extreme climate. Similarly, the *adyr* and *tau* steppes are spread along the bioclimatic gradient on Fig. 7B, highlighting their differences along the altitudinal gradient. Therefore, whenever looking at the bioclimatic condition surrounding the vegetation plots, it clearly becomes possible to refine the number of identifiable vegetation subtypes.

5.1.2. Nine well-marked vegetation subtypes in the environment pollen signal

The vegetation subtypes signal recorded in pollen assemblages. From pollen samples, the plant community signal is slightly different from the one recorded from vegetation plots analysis and facilitates the observation of at least nine vegetation subtypes clusters (mainly from CONISS clustering on Fig. 6). In an indirect way, CONISS results record bioclimatic conditions affecting pollen distribution. CONISS is a constrained hierarchical clustering (unlikely the TWINSpan analysis; Grimm, 1987; Roleček et al., 2009); so the altitudinal gradient controlling the two steppes (*adyr* and *tau* steppes) and the two open woodlands (*yailau* cryophilous and *tau* open woodlands) layers allows distinguishing them. However, the ecologically blind PCA (Fig. 7C) cannot find these ecological boundaries (both on pollen samples and vegetation plots). Therefore, it seems that in the paleo-pollen records, the distinction between these vegetation types cannot be made without the help of rare indicator taxa (Connor and Kvavadze, 2009; Connor et al., 2018). For example, the presence of higher value of Asteroideae-type, *Plantago* spp. and *Juniperus* spp. in *tau* steppe rather than in *adyr* steppe) could be used in the past as a path for vegetation types classification.

Multi-methods classification improvement. The utilization of other proxies can also help to understand local vegetation types such as macro-remains (Birks and Birks, 2000), Non-Pollen Palynomorphs (NPPs) linked to vegetation (Ghosh et al., 2017), organic geochemical proxies (leafwax derived n-alkanes, $\delta^{13}C$, etc.; Aichner et al., 2012) or ancient sedimentary DNA more linked to watershed signal (identification level is not similar between pollen and DNA; Liu et al., 2021). About NPPs, for instance, saprophytic spores such as *Cercophora* spp., *Coniochaeta* cf. *lignaria*, Coniochaetaceae undet. and Xylariaceae undet. Provide information on the presence of dead wood, i.e. the presence of forest in the watershed (Cugny, 2011). Moreover, specific spores of parasitic fungi indicate grasslands or croplands through Poaceae and cereal pathogens (e.g. *Curvularia* spp., *Nigrospora* spp. and *Savoryella* cf. *lignicola*; Gelorini et al., 2011; Wei et al., 2021) or wetlands through pathogens of hydrophytes (Montoya et al., 2010). Finally, the implication of the bioclimatic gradient permitting the distinction between *adyr* and *tau* steppes or between *tau* and *yailau* open woodlands could be used in the past to better interpret pollen signal if and only if a synchronous climate reconstruction is given from the same environmental archives but using an independent climate proxy such as molecular biomarkers (such as brGDGT, Robles et al., 2022; d'Oliveira et al., 2023).

5.1.3. Predicting vegetation types: perspectives

Toward quantitative vegetation type reconstructions. Our vegetation study, comparing modern pollen assemblages, vegetation plots and TU phytosociological knowledge, validates four types and nine vegetation subtypes as possibly observable in pollen signal. Although these observations help the understanding of past pollen rain (cf. the Aral 86 study case, Fig. 11), it however remains a qualitative ecological description. A promising innovation could be to quantitatively attribute a vegetation type score to past pollen samples or, even better, attributing a vegetation type probability estimates (Cairns, 2001). A first approach using machine learning classification converts quantitative proxy values in categorical results. In Cairns (2001), it was tested from vegetation plots only, but it could be attempted to extrapolate to pollen samples. It is increasingly applied to paleo-proxies calibration studies such as geochemical sources (Martínez-Sosa et al., 2023) or biome prediction from pollen samples (Sobol et al., 2019). A second approach does not only return categorical values but also categorical probability estimates. It has recently been used to model open habitat variation through time using a set of different vegetation proxies (Andermann et al., 2022) or biomes from pollen (Sobol et al., 2019).

Modified biomization approaches. More classically, we can think about a biomization-like approach to predict vegetation types (Herzschuh et al., 2004). The biomization method has been developed by (Prentice et al., 1996). The approach is made on PFT assignment to pollen types and biome. Crossing the two matrices (pollen $\times$ PFT and PFT $\times$ biomes) gives biome scores for each sample. The highest biome score defines the site biomes (Peyron et al., 1998; Tarasov et al., 1998). Using a similar table with our vegetation types could be a simple approach to quantitatively estimate TU vegetation types. For instance, the biomization approach has been applied a first time in China by both Ni et al. (2010) and (Chen et al., 2010b). In these studies, they applied the approach to the whole Chinese territory using 76 PFTs to predict a set of 19 biomes (the biomes nomenclature was made on a combination of BIOME4 and Chinese vegetation classification; Ni et al., 2010) to match to the other global biome reconstructions worldwide. Then, in Sun et al. (2020), PFT and biomes classes were totally reshuffled to improve the match between reconstructed vegetation and potential natural vegetation of China. The biomization approach has been widely applied all around the world using quite similar assignment matrix (Peyron et al., 1998; Tarasov et al., 1998; Binney et al., 2017). However, Herzschuh et al. (2004) tried to modify it to reconstruct three Alashan local vegetation types (i.e. oasis, steppe and desert). This approach could be a promising methodological improvement for TU vegetation reconstructions. Especially, the positive results of the reliability of pollen-based CWM traits in TU (Fig. 10) and in the whole ACA (Dugerdil et al., 2025) is a step-forward in the functional trait-based biomization method (Harrison et al., 2010). Although this theoretical background linking plant functional trait to biome distribution has already been applied to future vegetation change models (Boonman et al., 2022), it remains uncertain to explain the present biome distribution (Mucina, 2019) and it has never been done on past pollen records.

5.2. Quantifying pollen and vegetation representativeness

5.2.1. Vegetation plots and pollen surface samples differences

Pollen type representativeness. In TUSDB, the taxa fractional abundances are not always similar between the surface pollen samples and the vegetation plots (Fig. 6) leading to over-representation in the pollen assemblages of specific taxa (e.g. *Pinus* spp., *Ephedra fragilis*-type, *Betula* spp., *Juglans regia*, *Juniperus* spp., etc.) or under-representation of specific taxa (e.g. Rosaceae (tree)-type, *Centaurea* spp., *Cousinia* spp., *Caligonum*-type, *Convolvulus*-type, etc.; Fig. 8).

Origin of these differences among pollen types. These differences could be due to pollen production: taxa producing abundant pollen (e.g. *Pinus* spp.) are more represented in pollen samples than what it is really in the vegetation structure (Wieczorek and Herzschuh, 2020; Xia and Ni,

2024). The pollen dispersion capacity also affects the representativeness: insect-pollinated vs. wind-pollinated (Beer et al., 2007; Bolinder et al., 2016), differential pollen fall speed (Sugita, 2007b) and presence/absence of air-lifting organs on the pollen (such as *Pinus sacchi*; Song et al., 2012). The local microclimate and wind strength or direction could also affect the local pollen deposit (Sugita, 2007b; Maya-Manzano et al., 2017). It could additionally be due to differential pollen preservation whenever settling on the archiving material, e.g. soil, lake sediment or peat organic matter (Li et al., 2005). The exine identifiable structures also have an impact since they allow identifying the producing plant taxa. Furthermore, some biases are independent of the pollen producer: filter effect, such as canopy filter in closed vegetation (Joannin et al., 2012) or macrophytic raft filter (Cromartie et al., 2020) on lake pollen deposition. These filters depend on the vegetation surrounding the coring point, which mainly depends on geography. Finally, the variation among vegetation plots and surface samples could also come from the vegetation plots themselves. Several studies have shown the numerous biases produced among diverse methodological approaches carried to infer the vegetation using vegetation plot analysis: seasonal or annual variation, plot size, method (quadrat, transect...), vegetation particularities (ecotones, shrublands, forests...) could mitigate the vegetation plot accuracy (Godínez-Alvarez et al., 2009; Chytrý et al., 2014).

5.2.2. Spatial biases in TU

Geographical particularities of the biases. These pollen biases considerations have also to be analysed under the light of their spatial discrepancies. The pollen-vegetation abundance bias is not consistent through each TU vegetation type, vegetation subtype and geographical area (Fig. 9). On Fig. 6, it appears that the Fabaceae are a fairly associated taxon in open vegetation, but they are under-represented in semi-closed vegetation (units P5 and P6): this is shown by a bimodal dispersion on the scatter plot Fig. 9 and by a pollen-dominance in Kyzylkum Desert against a vegetation-dominance in Uzbek piedmont. Besides the biogeographical differences in pollen-producing species (e.g. *Ammodendron* spp. in *chul*, *Astragalus* spp. and *Oxytropis* spp. in *tau*), this could be due to the filter effect produced by tree canopy (Sugita, 2007a) or at least a dilution effect in front of important local AP pollen producer (such as *Juniperus* spp.). Also, the Fabaceae identification is difficult in desert samples where the pollen grains are more damaged due to pollen oxidation in dry soil (Lebreton et al., 2010) enhanced by warm/cold and wet/dry alternation in continental climate (Li et al., 2005).

The TU Cyperaceae case. Similarly, for Cyperaceae, this taxon is correctly associated in *yailau* and *tau* ecosystems (units P5 - P9, Fig. 6, Nuratau, Kuhistan and Hissar on Fig. 9) and under-represented in *adyr* and *chul* (no data for P1 and P2 but clear on P3 and P4; pollen almost absent in Sudunkli and Karakum Desert). It means that the *Carex physodes* sand desert ephemeroïd plants are not recorded in the pollen signal, unlike the *Carex pachystylis*. It is mainly due to extreme arid conditions in *chul* and *adyr* which are not favourable for pollen preservation (Horrocks et al., 2000). In a study from China, the same spatial bias affects pollen surface samples, with lower Cyperaceae in Taklamakan Desert sites than in Tian Shan and Qilian Shan alpine mesic grasslands (Huang et al., 2018). For Cyperaceae, the representation bias is more associated to pollen preservation than to pollen production, since it depends on exine thickness, soil oxidation and pH (Li et al., 2005). The same study reveals that Cyperaceae pollen, due to a thin exine, is particularly sensitive to taphonomic changes.

Over-represented pollen types. Finally, over-represented pollen types (through long-distance spreading and high pollen productivity) could be locally well associated, i.e. Amaranthaceae in P3 desert and pseudosteppes ecotone; *Artemisia* spp. in P4 *adyr* steppes and P7 *tau* steppes or *Juniperus* spp. in P6 and P8, similarly to Chinese desert pollen representation, which depends on vegetation subtypes (e.g. shrubby, semi-shrubby or semi-arboreal deserts, Wang et al., 2022). The pollen type representativeness is thus spatially different due to the associated plant

producing the pollen (which depends on biogeography). Comparing this representativeness to surrounding study area is then a key point to better understand our results.

5.2.3. Pollen specificity in Central Asia

Kyrgyz vs. Tajik-Uzbek pollen representativeness. The geographically closest pollen representation study has been carried in Kyrgyzstan (Beer et al., 2007). The study sites (four altitude lakes and moss polsters, $n = 43$) covering an altitudinal range from about 1400 to 2600 m a.s.l., are a good analogue for our *tau* and *yailau* sites. Especially the *tau* xeric shrublands and open woodlands are quite comparable since the dominant taxa in Kyrgyz forest are *Juniperus* spp., *Juglans regia*, *Acer* spp. and *Crataegus* spp. (Beer et al., 2007). They share conclusion about the impact of pollen-dispersion strategy on representativeness scores: wind-dispersed pollen (*Juniperus* spp., *Juglans regia* and in TUSDB *Betula* spp., *Fraxinus* spp., *Alnus* spp. or *Pinus* spp.) can travel on long distance and are likely more over-represented than insect-dispersed (Rosaceae (tree)-type, *Acer* spp.; *Lonicera* spp. and *Pistacia* spp.). Although *Juniperus* spp. is considered by the authors as a *moderately under-represented* [pollen types] in pollen samples, the slope linear coefficient remains higher than in TUSDB (Fig. 9). Moreover, Davis indices classify the *Juniperus* spp. as an *over-represented taxon*. It means that, in more open vegetation type such as open woodlands, *Juniperus* spp. is more down biased than in dense forest (i.e. Kyrgyz forest), but at the same time it has a long-distance dispersion. Our result for *Juglans regia* are very consistent with Beer et al. (2007) but also with the empirically measured anemophilous important pollen production of *J. regia* (Hegi and Conert, 1974): *J. regia* trees can be locally very abundant, and its pollen is commonly spread away to mid-long distances. About NAP, in Beer et al. (2007), the *Artemisia* spp. is largely over-represented (similarly to our study) and create a dilution effect on Poaceae and Cyperaceae grains in pollen samples, which are less abundant in pollen samples than in vegetation plots.

Steppes-forests and open woodlands biogeography. About high altitude studies and good analogues for *yailau* sites, the Altai sites from high elevation steppes and mountainous tundra record similar pollen types (Pelánková and Chytrý, 2009). Especially for shrubs, the *Betula* spp., *Salix* spp., Rosaceae (tree)-type and *Lonicera* spp. are present in vegetation and in pollen samples. For herbs, similar pollen types coexist but not always associated to the same dominant plant species in the Pelánková and Chytrý (2009) study: Apiaceae (mainly *Bupleurum* spp. in Altai, *Ferula* spp. and *Semenovia dasycarpa* here), *Artemisia* spp., Fabaceae (associated to *Onobrychis* spp. in Altai, *Onobrychis* spp., *Oxytropis* spp. and *Astragalus* spp. here), Gentianaceae (*Swertia obtusa* in Altai, *S. lactea* and *Gentiana olivieri* here), Ranunculaceae (*Trollius* spp. and *Aconitum* spp. in Altai, *Adonis turkestanica*), Rosaceae herbs (*Potentilla* spp. in our study, *Potentilla* spp. and *Dryas oxyodonta* in Altai). Cyperaceae is also an abundant and well-associated taxon in Altai but poorly informative since it is ubiquitous from dry mesic grassland grasses (mainly *Carex*) to wetland sedge (*Blasmus* spp. or *Eleocharis* spp.). The same problems are evident in the Tibetan Plateau (Herzschuh et al., 2010) and in Qilian Shan-Tarim basin, (Huang et al., 2018) with Cyperaceae high abundances from desert to alpine mesic grasslands. About trees, the dominant AP from Altai (*Pinus* spp., *Larix sibirica* and *Picea obovata*; Pelánková and Chytrý, 2009) also important in Mongolian Hovsgol nuur area (Ma et al., 2008) or Khentii and Khangai ranges (Dugerdil et al., 2021) are absent in our study. These taxa are representative of the light taiga (*Pinus sylvestris*, *Larix sibirica*; Dulamsuren et al., 2005) and dark taiga (*Pinus sibirica* and *Picea obovata*; Schlütz et al., 2008) steppes-forests and forests. Biogeographically, Tian Shan range is the southern edge for *L. sibirica* and *Picea* spp. extend mainly on north-facing slopes (Erdos et al., 2018). However, the most important AP in Tajikistan and Uzbekistan, i.e. the junipers (*Juniperus* spp. pollen produced by *J. semiglobosa* and *J. seravschanica*) are constrained to the southern steppes-forests area, mainly localized in Anatolia and Iran, with a northern edge located in Tian Shan (Erdos et al., 2018). *Juniperus*

spp. pollen types are present but not dominant in Caucasus Mountains (Connor and Kvavadze, 2009; Robles et al., 2022), absent in central Alborz (Dehghani et al., 2017) and quite important in the Golestan juniper-woodlands (Djamali et al., 2009). To conclude, within ACA, the TU mountains host an important contact area for the biogeography of the most important trees of open woodlands biomes: *Juniperus* spp. southward, with *Picea* spp. and *Larix sibirica* northward.

Cold vs. warm steppes and deserts. The transition between cold steppes (COST clearly distinguished in pollen record in Kazakhstan and Mongolia) and warm steppes (WAST, clearly present in north Africa and Anatolia, Tarasov et al., 1998) should find its location somewhere in ACA. Following Prentice et al. (1992), COST are found for MTWAQ below 22 °C, e.g. Mongolia, Tibet and Pamir, and WAST in the Uzbek lowlands or Tarim basin. At the phytosociological scale, the transition occurs in Tajikistan with the clear separation between the thermophilous secondary perennial in Mediterranean-like climates (including Irano-Turanian) with long dry summer period (i.e. *Cl. Stipo capensis-Trachynietea distachyae* S. Brullo in S. Brullo et al. 2001) and the high altitude desert, steppes and cold semi-deserts of Central Asia (i.e. *Cl. Ajanio-Cleistogenetea songoricae* Mirkin B. in Kashapov R. et al. 1987; Nowak et al., 2024).

Based on pollen samples, for *chul* and *adyr* TU vegetation types, the best analogue studies have been carried across the Mongolian steppes and pseudosteppes up to the Gobi Desert (i.e. COST; Ma et al., 2008; Dugerdil et al., 2021); following an altitudinal gradient transect from mountain to semidesert vegetation in the Georgian (Connor et al., 2004) and Armenian (Robles et al., 2022) Caucasus or in the Iranian Alborz (i.e. WAST; Djamali et al., 2009; Leroy et al., 2013; Dehghani et al., 2017). Such gradient have also been carried out in China among the Tibetan Plateau (COST; Herzschuh et al., 2010) or from the Qilian Shan to the Tarim basin (COST to WAST transition; Huang et al., 2018). In all these cases, the most abundant pollen types are Amaranthaceae, *Artemisia* spp. and Asteroideae-type, despite the dominance of Poaceae in vegetation plots (mainly *Stipa* spp., *Cleistogenes* spp. and *Festuca* spp.; Nowak et al., 2020, 2021). However, Poaceae cannot be neglected in pseudosteppes (sometimes called semi-deserts; Connor et al., 2004; Huang et al., 2018; Robles et al., 2022) and in cold desert (Ma et al., 2008; Dugerdil et al., 2021) pollen samples. Polygonaceae can be found in drylands associated to *Calligonum*-type in cold desert and steppe (Ma et al., 2008), or *Polygonum aviculare*-type in warmer environment recorded in wetland archives (Connor et al., 2004). In this present study, the *Polygonum*-type is mostly associated to the indicator species *Polygonum paronychioides*, *P. polycnemoides* and other *Polygonum* spp. (Nowak et al., 2021).

Dry-cold indicators can be found in *Convolvulus*-type (Ma et al., 2008; Dugerdil et al., 2021), *Ephedra fragilis*-type, *Nitraria* spp., *Tamarix* spp. and Zygophyllaceae (Herzschuh et al., 2010; Huang et al., 2018). While *Centaurea* spp., Lamiaceae and Asteroideae-type are more present in warm steppes and desert (in Iran and in the Caucasus; Djamali et al., 2009; Connor and Kvavadze, 2009; Dehghani et al., 2017; Robles et al., 2022). Also, the AP from wetter surrounding vegetation types spreading from long-distance are not similar since these vegetation types are biogeographically different: we can find *Juniperus* spp., *Quercus*-undet., *Fagus orientalis*, *Carpinus betulus* and *C. orientalis* southward and *Betula* spp., *Pinus* spp. and *Picea* spp. northward. *Salix* spp. and Rosaceae are less abundant but ubiquitous in warm and cold open land. In this context, the TU region appears to be a transition area between cold and warm drylands pollen rain, with indicators coming from the two clusters. However, the very low abundances of long-distance spreading AP (excepted *Juniperus* spp. that is present mainly in WAST but also in the Tibetan and Qilian area, Herzschuh et al., 2010; Huang et al., 2018) and the high abundances of continental Chinese specific pollen indicators (e.g. Fabaceae more than 10 %, or occurrence of *Betula* spp., *Hippophae* spp., *Elaeagnus* spp. or Thymelaeaceae) bring the TU closer to continental Chinese WAST-COST transition area than the WAST Caucasus-Iranian or cold Sibero-Mongolian study cases. Then, based on these conclusions, the pollen-vegetation representativeness could be

quantitatively compared in TU and Chinese mountains and plateaus.

Amaranthaceae are not easy to use to solve this temperature question since they are mainly aridity (Sorrel et al., 2007; Ma et al., 2008; Dugerdil et al., 2021) and salinity (Gintzburger, 2003; Thevs et al., 2008; Toderich et al., 2009) dependent. Consequently, the Amaranthaceae FA_{pol} reach the same range in cold (Ma et al., 2008; Huang et al., 2018) and warm steppes and deserts (Djamali et al., 2009; Dehghani et al., 2017; Robles et al., 2022).

5.2.4. Toward a correction of the past vegetation signal: main results and perspectives

Semi-quantitative vegetation calibration. R-values approach - used as a semi-quantitative vegetation bias correcting factor - has been applied in the whole of China (Xia and Ni, 2024) and more especially in Chinese continental drylands (Li et al., 2005; Wang et al., 2022). Among arid environment pollen types, the most over-represented types in China are *Ephedra* spp., *Artemisia* spp., Amaranthaceae, *Tamarix* spp., Brassicaceae and Asteroideae-type. Results are similar for the TUSDB, especially for the very over-represented ones (*Artemisia* spp., *Tamarix* spp., Amaranthaceae, Asteroideae-type and Brassicaceae). In TUSDB, we also have *Papaver* spp. with high R-value. However, it could be due to miss leading identification (generally not or poorly identified in pollen type; Connor and Kvavadze, 2009; Pelánková and Chytrý, 2009). For Boraginaceae, Plumbaginaceae, Polygonaceae and Scrophulariaceae-type, the R-values are not reliable since these taxa have a low prevalence (high under-associated Davis score in Fig. 8). The two types of *Ephedra* in TUSDB have a low R-value, even smaller than one for *Ephedra fragilis*-type. *Ephedra* spp. over-representation differs from Chinese provinces since it is a middle range pollen spreading type (with very high R-values in desert ecotones, Xia and Ni, 2024) and low R-value further away from deserts. R-values also depend on the desert vegetation subtype, with the highest bias in temperate shrubby desert (Wang et al., 2022). Desert *Ephedra* (i.e. *E. fragilis*-type) are under-represented unlike alpine *Ephedra* (i.e. *E. distachya*-type) highlighting the *Ephedra* spp. pollination particularities. Especially, although no clear link between pollen type and pollination syndrome has been established with *Ephedra*-types, it seems that insect-pollinated *Ephedra* should have differences with wind-pollinated ones (Bolinder et al., 2016). This could also have implication about the R-values differences, since usually wind-pollinated pollen are clearly more over-represented than insect-pollinated ones (Wang et al., 2022).

Quantitative vegetation reconstructions in ACA. Since R-values have been proposed as a correction factor to apply to past pollen sequence to reduce the pollen-vegetation bias in the past (Parsons and Prentice, 1981), few studies attempted this approach in the past. New modern calibration approaches, more complex but more reliable, have been proposed following the R-values such as the extended R-values model owing to a reduction of the pollen-vegetation non-linear issues (Bunting et al., 2004). Among these approaches, the Relative Pollen Productivity (RPP) approach is the most commonly used (Wieczorek and Herzschuh, 2020). Using the maximum likelihood function method, it is possible to have a statistically significant relevant source area of pollen (Bunting et al., 2004) in order to calculate RPP and its standard deviation. Finally, with the help of the RPP, the Landscape Reconstruction Algorithms and REVEALS model may be applied to quantitatively reconstruct regional vegetation (Sugita, 2007a). On the contrary of R-values, using REVEALS, long-trend pollen-based quantitative land-cover reconstruction have been proposed for the Ural region (Barhoumi et al., 2020) and northern ACA (Cao et al., 2018). However, RPP are still complicated to estimate, as their values are importantly fluctuating among study areas and no values are available out of Europe, China and northern Siberia (Wieczorek and Herzschuh, 2020). Moreover, Xia and Ni (2024) have shown that in China, R-values are a quite good and simpler estimation of RPP. In any case, using the Crackles-based vegetation plots from TUSDB, the RPP estimation could be a further development for TU paleo vegetation studies.

5.3. Plant functional ecology in the past

5.3.1. Validation of the CWM traits approach

Pairwise CWM trait relationships. Most relationships between pollen-based and vegetation-based CWM values are significant (Fig. 10A), despite potential biases in pollen representativeness. Carvalho et al. (2019) estimate that these biases are the major obstacle to the use of pollen-based CWM. However, in their study, the utilization of RPP (Wieczorek and Herzschuh, 2020) did not significantly improve the relationships. Moreover, the TU RPP-based correction is not possible yet since RPP are not available in the area yet. In TU, the functional signature of vegetation compares well with that of pollen assemblage, as also found in South America tropical forest (van der Sande et al., 2021) and temperate forest from United Kingdom (Carvalho et al., 2019). Finally, one of the limits of the application of pollen-based CWM emerges from the aggregation effect of the CWM calculation: with CWM, the plant diversity information is lost. Although other metrics accounting for trait variance may be used to capture community functional diversity (cf. Bruehlheide et al., 2018), these are very sensitive to missing values in particular. In line with this, van der Sande et al. (2021) have shown that the information carried by pollen samples is not accurate enough to assess plant community total diversity. Moreover, even the most recent diversity statistical estimation inferred from pollen sample (Birks et al., 2016) cannot overcome this lack of information.

Reliability of the pollen and vegetation based functional spaces. Further, we have shown that in ACA, the aggregation scheme (i.e. which plant taxa are associated to each pollen type in order to aggregate trait values for pollen types) effect is negligible, and that the organization of variables in the phenotypic space of the *global spectrum of plant form and function* (Díaz et al., 2016) derived from pollen surface samples compares well with that derived from vegetation samples (Dugerdil et al., 2025). In this previous study, the community-level functional spaces appeared to be broadly similar among pollen and vegetation plots; this was also consistent with the functional space observed worldwide (Dugerdil et al., 2025). In this previous study, the community-level functional spaces appeared to be broadly similar among pollen and vegetation plots; this was also consistent with the functional space observed worldwide (Bruehlheide et al., 2018).

Toward a pollen-based CWM-environment linkage. All together, this validates the use of CWM calculated on past pollen sequences (further discussed by Dugerdil et al., 2025). However, the response of CWM values of pollen assemblages to environmental parameters, climate in particular, has been investigated so far. Biogeographical factors such as altitude (van der Sande et al., 2021) or latitude (Brussel and Brewer, 2021) have been shown to influence the distribution of pollen-based functional diversity. However, the formal links to climate parameters (e.g. temperature, precipitation and aridity; Peppe et al., 2011; Frenette-Dussault et al., 2013; Nunes et al., 2017) or to human impact (e.g. grazing effect, Jäschke et al., 2020), which are commonly studied in the context of extant vegetation, are still to be established for pollen (Barboni et al., 2004). Overall, our functional approach to pollen samples conducted in TUSD confirms previous calibration tests (van der Sande et al., 2021) and biogeographic validation (Brussel and Brewer, 2021; Dugerdil et al., 2025). More generally, it validates the assumption formulated by Harrison et al. (2010), stating that pollen-based functional ecology holds promises for vegetation and climate reconstructions.

5.3.2. Functional ecology application in the past

Biome and vegetation structure from Aral 86 pollen record. The counter-intuitive finding that CWM_{LeafArea} is higher during the tundra-like phases (i.e. Z1 and Z2 in Fig. 11) than in the steppe phases, shows that the biomization and the functional approach do not match. Indeed, tundra biomes are not characterized by a high average leaf area (Björkman et al., 2018). This biome is indeed composed by a mix of ruderal and stress-tolerant plants (i.e. intermediate SLA but low leaf area; Pierce

et al., 2017). It appears that this high leaf area in tundra-like results from the high proportion of Poaceae and Cyperaceae pollen in the samples (Fig. 11 for Poaceae and Fig. S11 for Cyperaceae). Poaceae and Cyperaceae are important taxa in the grass PFT, which is important in the tundra biome (while only Poaceae counts for steppe score), whose leaf area values are higher than those of Amaranthaceae and *Artemisia* spp.. Pollen-based biome modelling is thus unreliable in this particular case. This is probably due to the low diversity of Aral 86 pollen types combined with the very large Aral watershed, which drains pollen from a vast regional area. This high Poaceae and Cyperaceae phase in the Aral is also associated to organic-rich sediment, which may indicate that these two pollen types could come from very local hydrophytes pollen producers. It seems that, in this context, the modelled leaf area value is more reliable than the *biomized* tundra-like score. Moreover, the plant height presents a discrete decrease from 7000 to 1700 yr cal BP testifying a slight forest retreat within the whole watershed. This information is lost in the biomization process, which is based on a coarser approach. It therefore suggests that biome reconstruction directly from traits is a sound alternative to *biomization* in such complex context (Boonman et al., 2022), as argued by Harrison et al. (2010).

Climate shifting and its functional consequences. The modelled functional traits can also help to choose the best climate calibration to reconstruct past climate parameters (Dugerdil et al., 2025). This is the next step in enhancing the current approach of weighting the transfer function using PFTs (Prentice et al., 1992; Peyron et al., 1998; Tarasov et al., 1998). Many studies have shown the relationships between climate parameters and plant traits (see Garnier et al., 2004), which could theoretically be used to validate transfer functions. For instance, several studies have shown significant relationships between SSD and aridity (Frenette-Dussault et al., 2013; Bruelheide et al., 2018), or leaf area and precipitation (Peppe et al., 2011; Griffin-Nolan et al., 2018). These studies can serve as bases to validate AI reconstructions, with arid periods of Z1 and Z2 characterized by higher average SSD within the Aral watershed. Tundra dominance at this period is unlikely. Moreover, from our local calibration study, it also appears that in the TU region, the MTCOQ favours the semi-forested area (dominance of AP in the higher elevation biomes, Fig. 6)). This relation validates the two reconstructed MTCOQ values in agreement with the plant height data (warmer period with low semi-forested areas in Z1 and Z2, Fig. 11). Using independent multi-proxy climate reconstruction coupled with this functional approach allows determining more precisely whether climate is the major factor controlling factor of the vegetation shifts (Adeleye et al., 2023).

5.4. Pollen-based climate reconstruction

5.4.1. Particular pollen-climate behaviour

Aridity is the first driving factor. In contrast to Europe or North America where temperature is more important (Salonen et al., 2019), in TU, aridity seems to be the main factor controlling pollen distribution (Figs. 6 and 7). It is similar to other climate studies in other areas in ACA from Tibet (Herzschuh, 2007; Herzschuh et al., 2010; Cao et al., 2021), Mongolia (Ma et al., 2008; Dugerdil et al., 2021), Iran (Dehghani et al., 2017) and Caucasus (Connor et al., 2004; Robles et al., 2022). Precipitation is also the most reliable parameter for Chinese transfer functions (Chen et al., 2017). However, here the particularity is the higher importance of MPCOQ than MAP to understand the TU climate system. This could be due to the TUSDB edge position between the Westerly-dominated area and the ACA core area (Tan et al., 2024). The majority of precipitation occurs in winter and spring in the whole of the TU (Belolipov et al., 2013). However, aridity pattern is more marked with MPCOQ than MAP since the two geographic extremes of this study, i.e. the Ustyurt plateau in Karakalpakstan and the Pamir plateau have both relatively more summer precipitation than the central TU zone (this is well-marked on the two precipitation curves on Fig. 6, see also the MPCOQ and MAP maps in Fig. S1B and D).

Artemisia interpretation in the past. A second particularity is about the pollen types / climate parameter interpretation in the past. Unlike a majority of ACA studies (Cao et al., 2021), here, the *Artemisia* spp. pollen type is not climatically relevant. Especially it appears that in TUSDB, *Artemisia* spp. is not correlated to temperature and precipitation ($R^2 \in [0.01; 0.02]$) on the contrary to ACA or of cold-steppes in general ($R^2 \in [0.10; 0.21]$, Fig. S9 and Table 2). First, *Artemisia* spp. is a large group in Uzbekistan (more than 40 species) covering almost all growth forms and present in all vegetation types from drier to wetter biomes (Kapustina et al., 2001). In their study, the distribution of *Artemisia* spp. in Uzbekistan is ubiquitous and shows that the taxon could be present as weed taxon in almost every vegetation types, but it is also a dominant taxon in a large part of a country from *adyr* pseudosteppes (Ustyurt plateau), *chul* warm desert (relic mountains of central Kyzylkum), *adyr* pseudosteppes (Bukhara and Kashkadarya) and of course all *adyr* and *tau* steppes (Nuratau, Kuratau, Hissar or Kuhistan). Moreover, although recent efforts have been carried out to improve the pollen identification level in ACA (Lu et al., 2018), it remains impossible to reach enough precision to distinguish the wide *Artemisia* spp. ecological niche variety (Kapustina et al., 2001).

Other pollen type interpretation in the past. Although *Artemisia* spp. is unreliable to track aridity in TUSDB, Amaranthaceae, Poaceae, Cyperaceae and *Juniperus* spp. are the most associated taxa to MPCOQ. Event though Amaranthaceae is found to be over-represented as pollen in the TUSDB (Fig. 8B) with important dispersion from *chul* to surrounding areas (Fig. 9), this pollen type remains one of the most associated to climate response (Fig. S9) and machine learning contributor (Table 2). Mainly, the ubiquitous Amaranthaceae distribution in TU (Gintzburger, 2003) allows the Amaranthaceae FA_{pom} to accurately respond to local climate amplitude. Moreover, their dominance on *chul* deserts and *adyr* pseudosteppes marks a clear bump in FA_{pom} in hyper-arid and arid environment. Even the different Amaranthaceae communities adapted halophyte, gypsophyte or psammophyte vegetation (Toderich et al., 2009; Dehghani et al., 2017) do not mitigate the link between Amaranthaceae and climate parameters. The Amaranthaceae-Aridity link is well established in TU (Sorrel et al., 2007), in Central Asia (Chen et al., 2010b), as well as the large Northern Hemisphere scale (Salonen et al., 2019).

For both Poaceae and Cyperaceae, the arid ephemeroïd taxa from desert and pseudosteppes (*Carex physodes*, *Eremopyrum bonaepartis*, *E. orientale* or *Bromus tectorum*) are not recorded as much as perennial herbs from wetter environment such as steppe, open woodlands and mesic grasslands (*Carex pachystylis*, *Festuca* spp., *Hordeum bulbosum*, *Stipa caucasica* or *S. tianschanica*). This is similar to Tarim basin - Qilian Shan Cyperaceae pollen grain preservation issue (Huang et al., 2018). Then, the question of the applicability of *Artemisia*/Amaranthaceae ratio as a semi-quantitative aridity index (El-Moslimany, 1990) that has been widely discussed and applied in ACA (Herzschuh, 2007; Chen et al., 2010a; Zhao et al., 2012) in such a situation has to be raised. Zhao et al. (2012) already recalled that this ratio has to be carefully applied in paleo-sequences and calibrated in function of vegetation types for each new study area.

Other pollen types have behaviours slightly different from what we expect, at least compared to European pollen studies. For instance, in ACA *Convolvulus* spp. is associated to drylands (mainly *chul* desert and *adyr* pseudosteppes) which is consistent with Mongolian (Ma et al., 2008; Murad, 2012; Dugerdil et al., 2021) and Chinese dryland pollen rain studies (Ni et al., 2010; Lu et al., 2018). It is also consistent with the ecophysiological studies: the steppes or mesic grasslands *Convolvulus* (i.e. *C. arvensis* and *C. lineatus*) are proportionally less numerous than the *chul* *Convolvulus* (*C. divaricatus*, *C. erinaceus*, *C. fruticosus*, *C. hamadae* and *C. subhirsutus* in Fig. S3). What is more, the latter have a different growth form since they are mainly scrubs with woody shoots; so these are statistically producing more pollen grain than the crawling ones from steppes and mesic grasslands.

Temperature as second driving factor. The independence between primary and secondary climate parameters (Salonen et al., 2019) also allows giving focus to temperature (MTCOQ in our study). *Ephedra fragilis*-type and *Ephedra distachya*-type are in TUSDB one of the best gradients between freezing and mild winters (Fig. 7D). Since these two types are quite rare in pollen assemblages, the temperature interpretation can be reinforced by the presence of Asteroideae-type, Brassicaceae and Lamiaceae which are also associated to milder winter. However, we have to remember that the accuracy of particular taxa to reconstruct climate parameters also depends on their prevalence (i.e. the proportion of sites in which the pollen type is present; Cao et al., 2021). This consideration can partially explain the dominance of major pollen type on BRT and RF climate contribution (Table 1) such as *Artemisia* spp. and Amaranthaceae. The Amaranthaceae response to temperature is not as clear in TU than in the Mongolian Plateaus (FA_{pol} correlated with MAAT; Ma et al., 2008; Dugerdil et al., 2021). However, there, the Gobi Desert mainly explains the Amaranthaceae FA_{pol} distribution. Therefore, the effects of temperature and aridity gradients on Amaranthaceae are not easy to disentangle.

5.4.2. Improving the reliability of pollen-inferred climate reconstructions

Calibration-based statistics. The statistic produced during the cross-validation process (R^2 , RMSE and Moran's I, Table 1) of the different methods and calibration modern pollen datasets are broadly similar to what we found in previous studies (Peyron et al., 2017; Dugerdil et al., 2021; d'Oliveira et al., 2023): (1) the BRT is always the most statistically significant method, far ahead of the RF and the WAPLS. MAT presents intermediate scores; (2) for BRT and MAT, the scores are better for larger datasets (EAPDB > ACADB > COSTDB > TUSDB), while it is the contrary for WAPLS (better scores for local calibration dataset), which is due to the mathematical particularity of the method (Ter Braak et al., 1993; Chevalier et al., 2020); and (3) surprisingly, the *Artemisia* spp. is almost not contributing to the different prediction (Table 2) unlike Amaranthaceae, Cyperaceae, Poaceae and *Juniperus* spp.. To conclude about the calibration process, we can say that the BRT and MAT methods applied on the TUSDB samples give a good trade-off between R^2 , RMSE, Moran's I and dataset size to accurately predict MPCOQ. For MTCOQ, the BRT based on ACADB appears to be the best calibration in TU. However, these general considerations for TU do not consider the sediment archive itself and mainly these statistics only infer the modern-vegetation-climate relationship (Rehfeld et al., 2016).

Past sequence-based statistics. Moreover, the `randomTF()` significance test for the past applicability of transfer functions (Table 3 and Fig. S12) has shown that climatic reconstruction accuracy cannot only depend on the chosen calibration set (Telford and Birks, 2011). Especially, this method allows checking the similarity between the pollen assemblages in the fossil sequence and in the calibration dataset (Chen et al., 2017). Assemblage dissimilarities may come from bioclimatic differences but also from a large panel of biases from pollen identification differences and heterogeneity in pollen taxonomy (De Klerk and Joosten, 2007) or pollen conservation, deposition processes and taphonomy, which are not always similar between surface samples and past pollen archives (especially in large lake Li, 2019). At the same time, the `randomTF()` gives a statistical approach on how to select the appropriate and independent climate parameters to reconstruct, which is a key issue in paleoclimate studies (Salonen et al., 2019). Here also, we have shown that the selected climate parameters independently covering the larger part of climatic variance in TU (i.e. MPCOQ, MTCOQ and AI, Fig. 2) are not necessarily the most accurate climate parameters to reconstruct from the Aral 86 sequence (AI, MTCOQ and MAP, Table 3). Finally, particular attention have to be given if human impact is reported to disturb the normal pollen-climate relationships, in order to avoid unreliable climate reconstructions (Peyron et al., 1998; Cromartie et al., 2020; Jäschke et al., 2020; Shumilovskikh et al., 2023).

6. Conclusions

Based on a new interdisciplinary approach combining both pollen surface samples and plots from Tajikistan and Uzbekistan, our study has shown some limitations which are: (1) the difficulty to superimpose plant-based vegetation classification to pollen-based one, partly due to (2) the spatially differential biases between pollen and vegetation. Nevertheless, some significant results indicate the validity of our pollen-vegetation approach:

1. Pollen assemblages are a valuable tool to reconstruct reliable and accurate vegetation type changing, functional diversity, biomes and key climate parameters in Tajikistan and Uzbekistan.
2. Based on key pollen types, i.e. *Juniperus* spp., *Juglans regia*, Cyperaceae, Poaceae and Amaranthaceae, the major Uzbek vegetation types (*chul*, *adyr*, *tau* and *yailau*) as well as nine vegetation subtypes can be tracked in the past.
3. Davis indices and R-values of these pollen types are a first easy-to-use correcting factor to estimate the real vegetation abundances of over-represented pollen types (*Betula* spp., *J. regia* and *Pinus* spp.) and under-represented pollen types (*Calligonum*-type, *Centaurea* spp., *Convolvulus*-type, *Cousinia* spp. and Rosaceae), locally changing in Tajikistan and Uzbekistan.
4. CWM traits based on pollen abundances show similar trends to vegetation-based one, i.e. the univariate trait relationships as well as the global plant functional spaces are respected both at the species scales and at the plant community scale. Applied to the pollen assemblages of the Aral 86 sequence, it also has shown its strength to validate or invalidate the pollen-based biome scores produce by the *biomization* approach.
5. Climate reconstruction methods should always be tested on several modern calibration datasets in order to produce the most appropriate climate estimates: here we have shown that the MAT and BRT produce the best statistical scores, and that the TUSDB is more reliable to inferred past MPCOQ while larger databases, such as the ACADB performed better on the MTCOQ. Nevertheless, the aforementioned results must be weighted in light of their applicability to past sequences, as determined through past sequence applicability testing.

Author contributions: CRediT

LD, GM, OP and SJ conceived the ideas and designed the followed methodological framework. SJ acquired the financial support for this project. DE, JA, SJ and LD organized the two field trips and LD, SJ and JA collected the surface samples. SL collect sample from a third field trip. LD and AN worked on the plant taxa identification. LD carried the laboratory work, the pollen counting and realized the R and LaTeX-scripts. LD compile previous pollen data from OP and SL as well as the Aral 86 core from <https://neotomadb.org/>. LD analysed the data with the help of SJ, OP and EG and prepared the visualization with SJ, OP, EG and SL. LD, OP and SJ led the writing of the original draft. All authors contributed critically to the manuscript and gave final approval for publication.

CRediT authorship contribution statement

Lucas Dugerdil: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Odile Peyron:** Writing – review & editing, Validation, Methodology, Conceptualization. **Guillemette Ménot:** Writing – review & editing, Validation, Conceptualization. **Dilfuza Egamberdieva:** Writing – review & editing, Resources, Project administration. **Jakhongir Alimov:** Writing – review & editing, Resources, Project administration. **Suzanne A.G. Leroy:** Writing – review & editing, Resources, Data curation. **Eric Garnier:** Writing – review & editing, Validation, Methodology. **Arkadiusz Nowak:** Writing – review & editing, Resources, Data curation. **Sébastien Joannin:** Writing –

review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Code availability

The R scripts used to clean, analysed the TUSDB data and plot all the figures presented in this study (except Fig. 1 realized in QGIS) is totally open and available on the following GitHub repository: https://github.com/LucasDugerdil/TUSD_R_script/.

Funding sources

This research was funded, in whole or in part, by [ANR, Grant ANR-22-CE27-0018 and Grant ANR-20-FRAL-0006]. A CC-BY public copy-right license has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission, in accordance with the grant's open access conditions. Most of the surface samples in N. Uzbekistan were taken owing to the EU-FP7-IRSES-2009-CLIMSEAS project 247,512 and in the frame of fieldwork organized by P. Zavalov (Shirshov Institute of Oceanology, Moscow).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the TRY initiative on plant traits (<http://www.try-db.org>); all the datasets from TRY used in this study are listed Dugerdil et al. (2025). We are grateful to *La Tendresse*, *Le BIB* and Zoé Hernandez for their precious supports in France; Nozimbek Namozovich and Sokhib Abdusamatov in Uzbekistan and Gulmirzozoda Masnavi Gulmirzo in Tajikistan. We are grateful to Vivien Mai Yung Sen for his help in high mountain sampling in Hissar and Pamir. Thanks to Elodie Brisset (IMBE) and Johanna Lhuillier (Archéorient) for their support during the Uzbekistan field trips. This is an ISEM contribution XXXX.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gloplacha.2025.104857>.

Data availability

Sample location and features, vegetation plots and pollen surface sample relative abundances are available on pangaea dataset XXXX.

References

- Adeleye, M.A., Haberle, S.G., Gallagher, R., Andrew, S.C., Herbert, A., 2023. Changing plant functional diversity over the last 12,000 years provides perspectives for tracking future changes in vegetation communities. *Nat. Ecol. Evol.* <https://doi.org/10.1038/s41559-022-01943-4>.
- Aichner, B., Herzsich, U., Wilkes, H., Schulz, H.M., Wang, Y., Others, 2012. Ecological development of lake donggi cona, north-eastern Tibetan Plateau, since the late glacial on basis of organic geochemical proxies and non-pollen palynomorphs. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 313, 140–149. <https://doi.org/10.1016/j.palaeo.2011.10.015>.
- Aleshinskaya, Z., 1991. Diatomic and palynologic characteristic of bottom sediments. *Hist. Lake Sevan Issyk-Kul Balkhash Zaisan Aral* 243–245.
- Amami, B., Muller, S.D., Rhazi, L., Grillas, P., Rhazi, M., Bouahim, S., 2010. Modern pollen-vegetation relationships within a small mediterranean temporary pool (western Morocco). *Rev. Palaeobot. Palynol.* 162, 213–225. <https://doi.org/10.1016/j.revpalbo.2010.06.012>.
- Andermann, T., Strömberg, C.A., Antonelli, A., Silvestro, D., 2022. The origin and evolution of open habitats in North America inferred by bayesian deep learning models. *Nat. Commun.* 13, 4833. <https://doi.org/10.1038/s41467-022-32300-5>.
- Barboni, D., Harrison, S.P., Bartlein, P.J., Jalut, G., New, M., Prentice, I.C., Sanchez-Goni, M.F., Spessa, A., Davis, B., Stevenson, A.C., 2004. Relationships between plant traits and climate in the Mediterranean region: a pollen data analysis. *J. Veg. Sci.* 15, 635–646. <https://doi.org/10.1111/j.1654-1103.2004.tb02305.x>.
- Barhoumi, C., Ali, A.A., Peyron, O., Dugerdil, L., Borisova, O., Golubeva, Y., Subetto, D., Kryshen, A., Drobyshev, I., Ryzhkova, N., Joannin, S., 2020. Did long-term fire control the coniferous boreal forest composition of the northern ural region (Komi republic, Russia)? *J. Biogeogr.* <https://doi.org/10.1111/jbi.13922> n/a.
- Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* 5, 180214. <https://doi.org/10.1038/sdata.2018.214>.
- Beer, R., Heiri, O., Tinner, W., 2007. Vegetation history, fire history and lake development recorded for 6300 years by pollen, charcoal, loss on ignition and chironomids at a small lake in southern Kyrgyzstan (alay range, central asia). *Holocene* 17, 977–985. <https://doi.org/10.1177/0959683607082413>.
- Belolipov, I.V., Zaurov, D.E., Eisenman, S.W., 2013. The geography, climate and vegetation of Uzbekistan. In: Eisenman, S.W., Zaurov, D.E., Struwe, L. (Eds.), *Medicinal Plants of Central Asia: Uzbekistan and Kyrgyzstan*. Springer New York, New York, NY, pp. 5–7. https://doi.org/10.1007/978-1-4614-3912-7_2.
- Beshko, N., 2023. *Flora of Central Asia: The Nuratau Mountains*. Korea National Arboretum, Pocheon.
- Beshko, N., Gaziev, A., Kodyrov, U., Kosimov, Z., Nazarov, B., 2024. Checklist of the flora of the chatkal state biosphere reserve (uzbekistan). <https://doi.org/10.15468/q2d6gd>.
- Binney, H., Edwards, M., Macias-Fauria, M., Lozhkin, A., Anderson, P., Kaplan, J.O., Andreev, A., Bezrukova, E., Blyakharchuk, T., Jankovska, V., Khazina, I., Krivonogov, S., Kremenetski, K., Nield, J., Novenko, E., Ryabogina, N., Solovieva, N., Willis, K., Zernitskaya, V., 2017. Vegetation of eurasia from the last glacial maximum to present: Key biogeographic patterns. *Quat. Sci. Rev.* 157, 80–97. <https://doi.org/10.1016/j.quascirev.2016.11.022>.
- Birks, H.H., Birks, H.J.B., 2000. Future uses of pollen analysis must include plant macrofossils. *J. Biogeogr.* 27, 31–35. <https://doi.org/10.1046/j.1365-2699.2000.00375.x>.
- Birks, H.J.B., Line, J.M., 1992. The use of rarefaction analysis for estimating palynological richness from quaternary pollen-analytical data. *Holocene* 2, 1–10. <https://doi.org/10.1177/095968369200200101>.
- Birks, H.J.B., Felde, V.A., Bjune, A.E., Grytnes, J.A., Seppä, H., Giesecke, T., 2016. Does pollen-assemblage richness reflect floristic richness? A review of recent developments and future challenges. *Rev. Palaeobot. Palynol.* 228, 1–25. <https://doi.org/10.1016/j.revpalbo.2015.12.011>.
- Bjorkman, A.D., Myers-Smith, I.H., Elmendorf, S.C., Normand, S., Rüger, N., Beck, P.S., Blach-Overgaard, A., Blok, D., Cornelissen, J.H.C., Forbes, B.C., 2018. Plant functional trait change across a warming tundra biome. *Nature* 562, 57–62. <https://doi.org/10.1038/s41586-018-0563-7>.
- Bolinder, K., Norbäck Ivarsson, L., Humphreys, A.M., Ickert-Bond, S.M., Han, F., Hoorn, C., Rydin, C., 2016. Pollen morphology of *ephedra* (gnetales) and its evolutionary implications. *Grana* 55, 24–51. <https://doi.org/10.1080/00173134.2015.1066424>.
- Boonman, C.C., Huijbregts, M.A., Benítez-López, A., Schipper, A.M., Thuiller, W., Santini, L., 2022. Trait-based projections of climate change effects on global biome distributions. *Divers. Distrib.* 28, 25–37. <https://doi.org/10.1111/ddi.13431>.
- Borgy, B., Violle, C., Choler, P., Denelle, P., Munoz, F., Kattge, J., Lavorel, S., Loranger, J., Amiaud, B., Bahn, M., 2017. Plant community structure and nitrogen inputs modulate the climate signal on leaf traits. *Glob. Ecol. Biogeogr.* 26, 1138–1152. <https://doi.org/10.1111/geb.12623>.
- Boyle, B., Hopkins, N., Lu, Z., Raygoza Garay, J.A., Mozzherin, D., Rees, T., Matasci, N., Narro, M.L., Piel, W.H., McKay, S.J., Lowry, S., Freeland, C., Peet, R.K., Enquist, B.J., 2013. The taxonomic name resolution service: An online tool for automated standardization of plant names. *BMC Bioinf.* 14, 16. <https://doi.org/10.1186/1471-2105-14-16>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/a:1010933404324>.
- Brulheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S.M., Botta-Dukát, Z., Chytrý, M., Field, R., Jansen, F., Kattge, J., Pillar, V.D., Schrod, F., Mahecha, M.D., Peet, R.K., Sandel, B., Van Bodegom, P., Altman, J., Alvarez-Dávila, E., Arfin Khan, M.A.S., Attorre, F., Aubin, I., Baraloto, C., Barroso, J.G., Bauters, M., Bergmeier, E., Biurrun, I., Bjorkman, A.D., Blonder, B., Čarni, A., Cayuela, L., Černý, T., Cornelissen, J.H.C., Craven, D., Dainese, M., Derroire, G., De Sanctis, M., Díaz, S., Doležal, J., Farfan-Rios, W., Feldpausch, T.R., Fenton, N.J., Garnier, E., Guerin, G.R., Gutiérrez, A.G., Haider, S., Hattab, T., Henry, G., Hérault, B., Higuchi, P., Hölzel, N., Homeier, J., Jentsch, A., Jürgens, N., Kačák, Z., Karger, D.N., Kessler, M., Kleyer, M., Knollová, I., Korolyuk, A.Y., Kühn, I., Laughlin, D.C., Lens, F., Loos, J., Louault, F., Lyubenova, M.I., Malhi, Y., Marcenó, C., Mencuccini, M., Müller, J.V., Munzinger, J., Myers-Smith, I.H., Neill, D. A., Niinemets, Ü., Orwin, K.H., Ozinga, W.A., Penuelas, J., Pérez-Haase, A., Petřík, P., Phillips, O.L., Pärtel, M., Reich, P.B., Römermann, C., Rodrigues, A.V., Sabatini, F.M., Sardans, J., Schmidt, M., Seidler, G., Silva Espejo, J.E., Silveira, M., Smyth, A., Sporbert, M., Svenning, J.C., Tang, Z., Thomas, R., Tsiripidis, I., Vassilev, K., Violle, C., Virtanen, R., Weiher, E., Welk, E., Wesche, K., Winter, M., Wirth, C., Jandt, U., 2018. Global trait-environment relationships of plant communities. *Nat. Ecol. Evol.* 2, 1906–1917. <https://doi.org/10.1038/s41559-018-0699-8>.

- Brussel, T., Brewer, S.C., 2021. Functional paleoecology and the pollen-plant functional trait linkage. *Front. Ecol. Evol.* 8, 564609. <https://doi.org/10.3389/fevo.2020.564609>.
- Bunting, M.J., 2003. Pollen-vegetation relationships in non-arboreal moorland taxa. *Rev. Palaeobot. Palynol.* 125, 285–298. [https://doi.org/10.1016/S0034-6667\(03\)00005-8](https://doi.org/10.1016/S0034-6667(03)00005-8).
- Bunting, M.J., Gaillard, M.J., Sugita, S., Middleton, R., Broström, A., 2004. Vegetation structure and pollen source area. *Holocene* 14, 651–660. <https://doi.org/10.1191/0959683604hl744rp>.
- Bunting, M.J., Farrell, M., Broström, A., Hjelte, K., Mazier, F., Middleton, R., Nielsen, A., Rushton, E., Shaw, H., Twiddle, C., 2013. Palynological perspectives on vegetation survey: a critical step for model-based reconstruction of quaternary land cover. *Quat. Sci. Rev.* 82, 41–55. <https://doi.org/10.1016/j.quascirev.2013.10.006>.
- Burr, G., Kuzmin, Y., Krivonogov, S., Gusskov, S., Cruz, R., 2019. A history of the modern Aral Sea (central asia) since the late pleistocene. *Quat. Sci. Rev.* 206, 141–149. <https://doi.org/10.1016/j.quascirev.2019.01.006>.
- Cairns, D.M., 2001. A comparison of methods for predicting vegetation type. *Plant Ecol.* 156, 3–18. <https://doi.org/10.1023/A:1011975321668>.
- Cao, X., Tian, F., Li, F., Gaillard, M.J., Rudaya, N., Herzschuh, U., 2018. Pollen-based quantitative land-cover reconstruction for northern Asia during the last 40ka. *Clim. Past Discuss.* 1–53. <https://doi.org/10.5194/cp-2018-111>.
- Cao, X., Tian, F., Li, K., Ni, J., Yu, X., Liu, L., Wang, N., 2021. Lake surface sediment pollen dataset for the alpine meadow vegetation type from the eastern Tibetan Plateau and its potential in past climate reconstructions. *Earth Syst. Sci. Data* 13, 3525–3537. <https://doi.org/10.5194/essd-13-3525-2021>.
- Carvalho, F., Brown, K.A., Waller, M.P., Bunting, M.J., Boom, A., Leng, M.J., 2019. A method for reconstructing temporal changes in vegetation functional trait composition using holocene pollen assemblages. *PLoS One* 14, e0216698. <https://doi.org/10.1371/journal.pone.0216698>.
- Chen, F.H., Chen, J.H., Holmes, J., Boomer, I., Austin, P., Gates, J.B., Wang, N.L., Brooks, S.J., Zhang, J.W., 2010a. Moisture changes over the last millennium in arid central asia: a review, synthesis and comparison with monsoon region. *Quat. Sci. Rev.* 29, 1055–1068. <https://doi.org/10.1016/j.quascirev.2010.01.005>.
- Chen, Y., Ni, J., Herzschuh, U., 2010b. Quantifying modern biomes based on surface pollen data in China. *Glob. Planet. Chang.* 74, 114–131. <https://doi.org/10.1016/j.gloplacha.2010.09.002>.
- Chen, J., Chen, F., Feng, S., Huang, W., Liu, J., Zhou, A., 2015. Hydroclimatic changes in China and surroundings during the medieval climate anomaly and little ice age: Spatial patterns and possible mechanisms. *Quat. Sci. Rev.* 107, 98–111. <https://doi.org/10.1016/j.quascirev.2014.10.012>.
- Chen, J., Lv, F., Huang, X., Birks, H.J.B., Telford, R.J., Zhang, S., Xu, Q., Zhao, Y., Wang, H., Zhou, A., Huang, W., Liu, J., Wei, G., 2017. A novel procedure for pollen-based quantitative paleoclimate reconstructions and its application in China. *Sci. China Earth Sci.* 60, 2059–2066. <https://doi.org/10.1007/s11430-017-9095-1>.
- Chen, S., Chen, J., Ding, G., Ma, S., Ji, P., Zhou, A., Wu, D., Khormali, F., Hou, J., Chen, F., 2024. Dipole pattern of holocene hydroclimate variations across the asian drylands: critical evidence from West Asia. *J. Geophys. Res. Atmos.* 129, e2023JD039413. <https://doi.org/10.1029/2023JD039413>.
- Cheng, H., Spötl, C., Breitenbach, S.F., Sinha, A., Wassenburg, J.A., Jochum, K.P., Scholz, D., Li, X., Yi, L., Peng, Y., Others, 2016. Climate variations of central asia on orbital to millennial timescales. *Sci. Rep.* 6, 36975. <https://doi.org/10.1038/srep36975>.
- Chevalier, M., Davis, B.A., Heiri, O., Seppä, H., Chase, B.M., Gajewski, K., Lacourse, T., Telford, R.J., Finsinger, W., Guiot, J., Kühl, M., Mäezumi, S.Y., Tipton, J.R., Carter, V.A., Brussel, T., Phelps, L.N., Dawson, A., Zanon, M., Vallé, F., Nolan, C., Mauri, A., De Vernal, A., Izumi, K., Holmström, L., Marsicek, J., Goring, S., Sommer, P.S., Chaput, M., Kupriyanov, D., 2020. Pollen-based climate reconstruction techniques for late quaternary studies. *Earth Sci. Rev.* 210, 103384. <https://doi.org/10.1016/j.earscirev.2020.103384>.
- Chytrý, M., Tichý, L., Hennekens, S.M., Schaminée, J.H., 2014. Assessing vegetation change using vegetation-plot databases: a risky business. *Appl. Veg. Sci.* 17, 32–41. <https://doi.org/10.1111/avsc.12050>.
- Connor, S.E., Kvavadze, E.V., 2009. Modelling late quaternary changes in plant distribution, vegetation and climate using pollen data from Georgia, caucasus. *J. Biogeogr.* 36, 529–545. <https://doi.org/10.1111/j.1365-2699.2008.02019.x>.
- Connor, S.E., Kvavadze, E.V., 2018. Pollen microphotographs from Georgia, caucasus. *Connor, S.E., Thomas, I., Kvavadze, E.V., Arabuli, G.J., Avakov, G.S., Sagona, A., 2004. A survey of modern pollen and vegetation along an altitudinal transect in southern Georgia, caucasus region. Rev. Palaeobot. Palynol.* 129, 229–250. <https://doi.org/10.1016/j.revpalbo.2004.02.003>.
- Connor, S.E., Colombaroli, D., Confortini, F., Gobet, E., Ilyashuk, B.P., Ilyashuk, E.A., Van Leeuwen, J.F.N., Lamentowicz, M., Van Der Knaap, W.O., Malysheva, E., Marchetto, A., Margalitadze, N., Mazei, Y., Mitchell, E.A.D., Payne, R.J., Ammann, B., 2018. Long-term population dynamics: Theory and reality in a peatland ecosystem. *J. Ecol.* 106, 333–346. <https://doi.org/10.1111/1365-2745.12865>.
- Cowan, P.J., 2007. Geographic usage of the terms middle asia and central asia. *J. Arid Environ.* 69, 359–363. <https://doi.org/10.1016/j.jaridenv.2006.09.013>.
- Cromartie, A., Blanchet, C., Barhoumi, C., Messenger, E., Peyron, O., Ollivier, V., Sabatier, P., Etienne, D., Karakhanyan, A., Khatchadourian, L., Smith, A.T., Badalyan, R., Perello, B., Lindsay, I., Joannin, S., 2020. The vegetation, climate, and fire history of a mountain steppe: a holocene reconstruction from the south caucasus, shenkani, Armenia. *Quat. Sci. Rev.* 246, 106485. <https://doi.org/10.1016/j.quascirev.2020.106485>.
- Cugny, C., 2011. The Use of Non-Pollen Palynomorphs for Reconstructing the History of Pastoral Activities in the Pyrenees : From Modern Datasets to Reconstruction of the Past. Theses. Université Toulouse le Mirail - Toulouse II.
- d'Alpoim Guedes, J., Bocinsky, R.K., 2018. Climate change stimulated agricultural innovation and exchange across asia. *Sci. Adv.* 4, eaar4491. <https://doi.org/10.1126/sciadv.aar4491>.
- De Klerk, P., Joosten, H., 2007. The difference between pollen types and plant taxa: a plea for clarity and scientific freedom. *Quat. Sci. J.* 56, 162–171. <https://doi.org/10.3285/eg.56.3.02>.
- Dehghani, M., Djamali, M., Gandouin, E., Akhiani, H., 2017. A pollen rain-vegetation study along a 3600 m mountain-desert transect in the irano-turanian region; implications for the reliability of some pollen ratios as moisture indicators. *Rev. Palaeobot. Palynol.* 247, 133–148. <https://doi.org/10.1016/j.revpalbo.2017.08.004>.
- Denelle, P., Weigelt, P., Kreft, H., 2023. GIFT—an R package to access the global inventory of floras and traits. *Methods Ecol. Evol.* 14, 2738–2748. <https://doi.org/10.1111/2041-210X.14213>.
- Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., Garnier, E., Bönsch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Joseph Wright, S., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D., Gorné, L.D., 2016. The global spectrum of plant form and function. *Nature* 529, 167–171. <https://doi.org/10.1038/nature16489>.
- Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E.C., Jones, B., Barber, C.V., Hayes, R., Kormos, C., Martin, V., Crist, E., Sechrest, W., Price, L., Baillie, J.E.M., Weeden, D., Suckling, K., Davis, C., Sizer, N., Moore, R., Thau, D., Birch, T., Potapov, P., Turubanova, S., Tyukavina, A., De Souza, N., Pinter, L., Brito, J.C., Llewellyn, O.A., Miller, A.G., Patzelt, A., Ghazanfar, S.A., Timberlake, J., Klöser, H., Shennan-Farpon, Y., Kindt, R., Lillesso, J.P.B., Van Breugel, P., Graudal, L., Voge, M., Al-Shammari, K.F., Saleem, M., 2017. An ecoregion-based approach to protecting half the terrestrial realm. *Bioscience* 67, 534–545. <https://doi.org/10.1093/biosci/bix014>.
- Dixon, P., 2003. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927–930 (doi: 10/dz9txb, arXiv:3236992).
- Djamali, M., de Beaulieu, J.L., Campagne, P., Andrieu-Ponel, V., Ponel, P., Leroy, S., Akhiani, H., 2009. Modern pollen rain-vegetation relationships along a forest-steppe transect in the golestan national park, NE Iran. *Rev. Palaeobot. Palynol.* 153, 272–281. <https://doi.org/10.1016/j.revpalbo.2008.08.005>.
- Djamali, M., Brewer, S., Breckle, S.W., Jackson, S.T., 2012. Climatic determinism in phytogeographic regionalization: a test from the irano-turanian region, SW and central asia. *Flora* 207, 237–249. <https://doi.org/10.1016/j.flora.2012.01.009>.
- d'Oliveira, L., Dugerdil, L., Ménot, G., Evain, A., Muller, S.D., Ansanay-Alex, S., Azuara, J., Bonnet, C., Bremond, L., Shah, M., Peyron, O., 2023. Reconstructing 15 000 years of southern France temperatures from coupled pollen and molecular (branched glycerol dialkyl glycerol tetraether) markers (canroute, massif central). *Clim. Past* 19, 2127–2156. <https://doi.org/10.5194/cp-19-2127-2023>.
- Dugerdil, L., Joannin, S., Peyron, O., Jouffroy-Bapicot, I., Vannière, B., Boldgiv, B., Unkelbach, J., Behling, H., Ménot, G., 2021. Climate reconstructions based on GDGT and pollen surface datasets from Mongolia and baikal area: Calibrations and applicability to extremely cold-dry environments over the late holocene. *Clim. Past* 17, 1199–1226. <https://doi.org/10.5194/cp-17-1199-2021>.
- Dugerdil, L., Peyron, O., Violle, C., Joannin, S., Ménot, G., Denelle, P., Bruehlheide, H., Chytrý, M., Field, R., Hatim, M.Z., Gholizadeh, H., Dolezal, J., Pillar, V.D., Shaltout, K.H., Schrödt, F., Garnier, E., 2025. Functional Signatures of Surface Pollen and Vegetation are Broadly similar: Good News for Past Reconstructions of Vegetation. *J. Biogeogr.* jbi.15100. <https://doi.org/10.1111/jbi.15100>.
- Dulamsuren, C., Hauck, M., Mühlenberg, M., 2005. Vegetation at the taiga forest—steppe borderline in the western khentey mountains, northern Mongolia. In: *Annales Botanici Fennici*. JSTOR, pp. 411–426.
- Dunnington, D.W., Libera, N., Kurek, J., Spooner, I.S., Gagnon, G.A., 2022. Tidypaleo: Visualizing paleoenvironmental archives using ggplot2. *J. Stat. Softw.* 101, 1–20. <https://doi.org/10.18637/jss.v101.i07>.
- Egamberdieva, D., Alimov, J., Shurigin, V., Alaylar, B., Wirth, S., Bellingrath-Kimura, S. D., 2022. Diversity and plant growth-promoting ability of endophytic, halotolerant bacteria associated with tetragonia tetragonioides (pall.) kuntze. *Plants* 11, 49. <https://doi.org/10.3390/plants11010049>.
- El-Moslimany, A.P., 1990. Ecological significance of common nonarborescent pollen: examples from drylands of the middle east. *Rev. Palaeobot. Palynol.* 64, 343–350. [https://doi.org/10.1016/0034-6667\(90\)90150-h](https://doi.org/10.1016/0034-6667(90)90150-h).
- Erdos, L., Ambarli, D., Anenkhonov, O.A., Bátori, Z., Cserhalmi, D., Kiss, M., Kröel-Dulay, G., Liu, H., Magnes, M., Molnár, Z., 2018. The edge of two worlds: a new review and synthesis on eurasian forest-steppes. *Appl. Veg. Sci.* 21, 345–362 doi: 10/ggspn6.
- Faegri, K., Iversen, J., 1989. *Textbook of Pollen Analysis*.
- Feroniskii, V.I., Polyakov, V.A., Brezgunov, V.S., Vlasova, L.S., Karpychev, Y.A., Bobkov, A.F., Romaniovskii, V.V., Johnson, T., Ricketts, D., Rasmussen, K., 2003. Variations in the hydrological regime of kara-bogaz-gol gulf, lake issyk-kul, and the Aral Sea assessed based on data of bottom sediment studies. *Water Res.* 30, 252–259. <https://doi.org/10.1023/A:1023826011601>.
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315 doi: 10/gb2jng.
- Floragridmap.uz, 2010. Virtual flora of republic of Uzbekistan. <https://floragridmap.uz/>.

- Frachetti, M.D., 2008. *Pastoralist Landscapes and Social Interaction in Bronze Age Eurasia*. University of California press, Berkeley.
- Frenette-Dussault, C., Shipley, B., Meziane, D., Hingrat, Y., 2013. Trait-based climate change predictions of plant community structure in arid steppes. *J. Ecol.* 101, 484–492. <https://doi.org/10.1111/1365-2745.12040>.
- Garnier, E., Cortez, J., Billes, G., Navas, M.L., Roumet, C., Debussche, M., Laurent, G., Blanchard, A., Aubry, D., Bellmann, A., Neill, C., Toussaint, J.P., 2004. Plant functional markers capture ecosystem properties during secondary succession. *Ecology* 85, 2630–2637. <https://doi.org/10.1890/03-0799>.
- Gaüzère, P., Iversen, L.L., Seddon, A.W.R., Violle, C., Blondel, B., 2020. Equilibrium in plant functional trait responses to warming is stronger under higher climate variability during the holocene. *Glob. Ecol. Biogeogr.* 29, 2052–2066. <https://doi.org/10.1111/geb.13176>.
- Gazer, M.H., Shalabi, L.F., 2016. The role of pollen morphology in the identification and classification of plantago (plantaginaceae). *Egypt. J. Exp. Biol.* 12, 125–132. <https://doi.org/10.5455/egyeb.20160628055424>.
- Gelorini, V., Verbeke, A., van Geel, B., Cocquyt, C., Verschuren, D., 2011. Modern non-pollen palynomorphs from east african lake sediments. *Rev. Palaeobot. Palynol.* 164, 143–173. <https://doi.org/10.1016/j.revpalbo.2010.12.002>.
- Ghosh, R., Paruya, D.K., Acharya, K., Ghorai, N., Bera, S., 2017. How reliable are non-pollen palynomorphs in tracing vegetation changes and grazing activities? Study from the Darjeeling himalaya, India. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 475, 23–40. <https://doi.org/10.1016/j.palaeo.2017.03.006>.
- Gintzburger, G., 2003. *Rangelands of the Arid and Semi-Arid Zones in Uzbekistan* (Editions Quae).
- Godínez-Alvarez, H., Herrick, J.E., Mattocks, M., Toledo, D., Van Zee, J., 2009. Comparison of three vegetation monitoring methods: their relative utility for ecological assessment and monitoring. *Ecol. Indic.* 9, 1001–1008. <https://doi.org/10.1016/j.ecolind.2008.11.011>.
- Griffin-Nolan, R.J., Bushey, J.A., Carroll, C.J., Challis, A., Chieppa, J., Garbowski, M., Hoffman, A.M., Post, A.K., Slette, I.J., Spitzer, D., 2018. Trait selection and community weighting are key to understanding ecosystem responses to changing precipitation regimes. *Funct. Ecol.* 32, 1746–1756. <https://doi.org/10.1111/1365-2435.13135>.
- Grime, J.P., 1974. Vegetation classification by reference to strategies. *Nature* 250, 26–31. <https://doi.org/10.1038/250026a0>.
- Grimm, E.C., 1987. CONISS: a FORTRAN 77 program for stratigraphically constrained cluster analysis by the method of incremental sum of squares. *Comput. Geosci.* 13, 13–35. [https://doi.org/10.1016/0098-3004\(87\)90022-7](https://doi.org/10.1016/0098-3004(87)90022-7).
- Grubov, V.I., 1999. *Plants of Central Asia: Plant Collections from China and Mongolia. Volume 1: Introduction, Ferns, Bibliography*. Science Publishers, Inc.
- Guiot, J., 1990. Methodology of the last climatic cycle reconstruction in France from pollen data. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 80, 49–69. [https://doi.org/10.1016/0031-0182\(90\)90033-4](https://doi.org/10.1016/0031-0182(90)90033-4).
- Guiot, J., Goeury, C., 1996. PPPBASE, a software for statistical analysis of paleoecological and paleoclimatological data. *Dendrochronologia* 14, 300.
- Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* 7, 1–18. <https://doi.org/10.1038/s41597-020-0453-3>.
- Harrison, S.P., Prentice, I.C., Barboni, D., Kohfeld, K.E., Ni, J., Sutra, J.P., 2010. Ecophysiological and bioclimatic foundations for a global plant functional classification. *J. Veg. Sci.* 21, 300–317. <https://doi.org/10.1111/j.1654-1103.2009.01144.x>.
- Hegi, G., Conert, H., 1974. *Illustrierte Flora von Mitteleuropa; Pteridophyta, Spermatophyta, Band 3 Angiospermae, Dicotyledones 1; Teil 1. Parey*.
- Herzschuh, U., 2007. Reliability of pollen ratios for environmental reconstructions on the Tibetan Plateau. *J. Biogeogr.* 34, 1265–1273. <https://doi.org/10.1111/j.1365-2699.2006.01680.x>.
- Herzschuh, U., Kürschner, H., Ma, Y., 2003. The surface pollen and relative pollen production of the desert vegetation of the Alashan Plateau, western Inner Mongolia. *Chin. Sci. Bull.* 48, 1488–1493. <https://doi.org/10.1360/02wd0256>.
- Herzschuh, U., Tarasov, P., Wünnemann, B., Hartmann, K., 2004. Holocene vegetation and climate of the alashan plateau, NW China, reconstructed from pollen data. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 211, 1–17. <https://doi.org/10.1016/j.palaeo.2004.04.001>.
- 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. <https://doi.org/10.1111/j.1365-2699.2009.02245.x>.
- Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., Hijmans, M.R.J., 2017. Package 'dismo'. *Circles* 9, 1–68.
- Hill, M.O., Hill, M.O., 1979. TWINSPAN: A FORTRAN Program of Arranging Multivariate Data in an Ordered Two Way Table by Classification of Individual and Attributes.
- Horrocks, M., Deng, Y., Ogden, J., Sutton, D., 2000. A reconstruction of the history of a holocene sand dune on great barrier island, northern New Zealand, using pollen and phytolith analyses. *J. Biogeogr.* 27, 1269–1277. <https://doi.org/10.1046/j.1365-2699.2000.00508.x>.
- Huang, J., Yu, H., Guan, X., Wang, G., Guo, R., 2016. Accelerated dryland expansion under climate change. *Nat. Clim. Chang.* 6, 166–171. <https://doi.org/10.1038/nclimate2837>.
- Huang, X., Chen, X., Du, X., 2018. Modern pollen assemblages from human-influenced vegetation in northwestern China and their relationship with vegetation and climate. *Veg. Hist. Archaeobotany* 27, 767–780. <https://doi.org/10.1007/s00334-018-0672-0>.
- Jamsranjav, C., Reid, R.S., Fernández-Giménez, M.E., Tsevee, A., Yadamsuren, B., Heiner, M., 2018. Applying a dryland degradation framework for rangelands: the case of Mongolia. *Ecol. Appl.* 28, 622–642. <https://doi.org/10.1002/eap.1684>.
- Jarvis, A., Guevara, E., Reuter, H., Nelson, A., 2008. Hole-filled SRTM for the globe : Version 4 : Data grid.
- Jäschke, Y., Heberling, G., Wesche, K., 2020. Environmental controls override grazing effects on plant functional traits in tibetan rangelands. *Funct. Ecol.* 34, 747–760. <https://doi.org/10.1111/1365-2435.13492>.
- Joannin, S., Brugiapaglia, E., de Beaulieu, J.L., Bernardo, L., Magny, M., Peyron, O., Goring, S., Vannière, B., 2012. Pollen-based reconstruction of holocene vegetation and climate in southern Italy: the case of lago trifoglietti. *Clim. Past* 8, 1973–1996. doi: 10/f4hw4z.
- Joannin, S., Ali, A.A., Ollivier, V., Roiron, P., Peyron, O., Chevaux, S., Nahapetyan, S., Tozalakyan, P., Karakhanyan, A., Chataigner, C., 2014. Vegetation, fire and climate history of the lesser caucasus: a new holocene record from zarishat fen (Armenia). *J. Quat. Sci.* 29, 70–82. doi: 10/f5m7vt.
- Joannin, S., Capit, A., Ollivier, V., Bellier, O., Brossier, B., Mourier, B., Tozalakian, P., Colombié, C., Yevadian, M., Karakhanyan, A., Gasparyan, B., Malinsky-Buller, A., Chataigner, C., Perello, B., 2022. First pollen record from the late holocene forest environment in the lesser caucasus. *Rev. Palaeobot. Palynol.* 304, 104713. <https://doi.org/10.1016/j.revpalbo.2022.104713>.
- Juggins, S., Juggins, M.S., 2019. Package 'rioja' (R package).
- Kapustina, L.A., Torrell, M., Vallès, J., 2001. *Artemisia communities in arid zones of Uzbekistan (central asia)*. In: USDA for. Serv. Proc. RMRS-P-21, pp. 104–112.
- Kattge, J., Bönsch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Tautenhahn, S., Werner, G.D., Aakala, T., Abedi, M., Others, 2020. TRY plant trait database—enhanced coverage and open access. *Glob. Chang. Biol.* 26, 119–188. <https://doi.org/10.1111/gcb.14904>.
- Kent Davis, O., 1984. Pollen frequencies reflect vegetation patterns in a great basin (U.S. a.) mountain range. *Rev. Palaeobot. Palynol.* 40, 295–315. [https://doi.org/10.1016/0034-6667\(84\)90013-7](https://doi.org/10.1016/0034-6667(84)90013-7).
- Kirpichnikov, M.E., 1969. The flora of the USSR. *Taxon* 18, 685–708. <https://doi.org/10.2307/1218928>.
- Kürschner, H., 2004. Phytosociological studies in the alashan gobi - a contribution to the flora and vegetation of Inner mongolia (NW China). *Phytocoenologia* 34, 169–224. <https://doi.org/10.1127/0340-269X/2004/0034-0169>.
- Last, W.M., Smol, J.P., Birks, H.J.B., 2001. *Terrestrial, Algal, and Siliceous Indicators*. Kluwer Academic, Dordrecht; Boston.
- Lebreton, V., Messager, E., Marquer, L., Renault-Miskovsky, J., 2010. A neotaphonomic experiment in pollen oxidation and its implications for archaeopalynology. *Rev. Palaeobot. Palynol.* 162, 29–38. <https://doi.org/10.1016/j.revpalbo.2010.05.002>.
- Leroy, S.A., Giral, S.R., 2021. Humid and cold periods in the last 5600 years in arid central asia revealed by palynology of *picea schrenkiana* from issyk-kul. *Holocene* 31, 380–391. <https://doi.org/10.1177/0959683620972776>.
- Leroy, S.A., Kakkroodi, A.A., Kroonenberg, S., Lahijani, H.K., Alimohammadian, H., Nigarov, A., 2013. Holocene vegetation history and sea level changes in the SE corner of the Caspian Sea: Relevance to SW asia climate. *Quat. Sci. Rev.* 70, 28–47. <https://doi.org/10.1016/j.quascirev.2013.03.004>.
- Li, Q., 2018. Spatial variability and long-term change in pollen diversity in nam Co catchment (central tibetan plateau): Implications for alpine vegetation restoration from a paleoecological perspective. *Sci. China Earth Sci.* 61, 270–284. <https://doi.org/10.1007/s11430-017-9133-0>.
- Li, Q., 2019. Distribution and vegetation representation of pollen assemblages from surface sediments of nam Co, a large alpine lake in the central tibetan plateau. *Veg. Hist. Archaeobotany* 28, 365–377. <https://doi.org/10.1007/s00334-018-0699-2>.
- Li, Y.C., Xu, Q.H., Yang, X.L., Chen, H., Lu, X.M., 2005. Pollen-vegetation relationship and pollen preservation on the northeastern Qinghai-tibetan plateau. *Grana* 44, 160–171. <https://doi.org/10.1080/00173130500230608>.
- Li, W., Tojibaev, K.S., Hisoriev, H., Shomurov, K.F., Luo, M., Feng, Y., Ma, X., 2020. Mapping asia plant's current status of floristic information for central asian flora. *Global Ecol. Conserv.* 24, e01220. <https://doi.org/10.1016/j.gecco.2020.e01220>.
- Li, Y., Song, Y., Kaskaoutis, D.G., Zhang, X., Chen, X., Shukurov, N., Orozbaev, R., 2022. Atmospheric dust dynamics over central asia: a perspective view from loess deposits. *Gondwana Res.* 109, 150–165. <https://doi.org/10.1016/j.jgr.2022.04.019>.
- Lioubimtseva, E., 2003. 12 arid environments. In: *The Physical Geography of Northern Eurasia*, 3. Oxford University Press on Demand, p. 267.
- Lioubimtseva, E., Cole, R., Adams, J., Kapustin, G., 2005. Impacts of climate and land-cover changes in arid lands of central asia. *J. Arid Environ.* 62, 285–308. <https://doi.org/10.1016/j.jaridenv.2004.11.005>.
- Liu, S., Li, K., Jia, W., Stoof-Leichsenring, K.R., Liu, X., Cao, X., Herzschuh, U., 2021. Vegetation reconstruction from siberia and the Tibetan Plateau using modern analogue technique—comparing sedimentary (ancient) DNA and pollen data. *Front. Ecol. Evol.* 9, 668611. <https://doi.org/10.3389/fevo.2021.668611>.
- Lu, K.Q., Xie, G., Li, M., Li, J.F., Trivedi, A., Ferguson, D.K., Yao, Y.F., Wang, Y.F., 2018. Pollen spectrum, a cornerstone for tracing the evolution of the eastern central asian desert. *Quat. Sci. Rev.* 186, 111–122. <https://doi.org/10.1016/j.quascirev.2018.02.009>.
- Luneau, É., 2019. Climate change and the rise and fall of the oxus civilization in southern central asia. In: Yang, L.E., Bork, H.R., Fang, X., Mischke, S. (Eds.), *Socio-Environmental Dynamics along the Historical Silk Road*. Springer International Publishing, Cham, pp. 275–299. https://doi.org/10.1007/978-3-030-00728-7_14.
- Ma, Y., Liu, K.B., Feng, Z., Sang, Y., Wang, W., Sun, A., 2008. A survey of modern pollen and vegetation along a south–north transect in Mongolia. *J. Biogeogr.* 35, 1512–1532. <https://doi.org/10.1111/j.1365-2699.2007.01871.x>.
- Maitner, B.S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S.M., Guaderrama, D., Hinchliff, C.E., Jørgensen, P.M., Kraft, N.J., McGill, B., Merow, C.,

- Moruela-Holme, N., Peet, R.K., Sandel, B., Schildhauer, M., Smith, S.A., Svenning, J. C., Thiers, B., Violle, C., Wiser, S., Enquist, B.J., 2018. The BIEN R package: a tool to access the botanical information and ecology network (BIEN) database. *Methods Ecol. Evol.* 9, 373–379. <https://doi.org/10.1111/2041-210X.12861>.
- Martínez-Sosa, P., Tierney, J.E., Pérez-Angel, L.C., Stefanescu, I.C., Guo, J., Kirkels, F., Sepúlveda, J., Peterse, F., Shuman, B.N., Reyes, A.V., 2023. Development and application of the branched and isoprenoid GDGT machine learning classification algorithm (BIGMac) for paleoenvironmental reconstruction. *Paleoceanogr. Paleoclimatol.* 38, e2023PA004611. <https://doi.org/10.1029/2023PA004611>.
- Maya-Manzano, J.M., Sady, M., Tormo-Molina, R., Fernández-Rodríguez, S., Oteros, J., Silva-Palacios, I., Gonzalo-Garijo, A., 2017. Relationships between airborne pollen grains, wind direction and land cover using GIS and circular statistics. *Sci. Total Environ.* 584–585, 603–613. <https://doi.org/10.1016/j.scitotenv.2017.01.085>.
- Micklin, P., 2007. The Aral Sea disaster. *Annu. Rev. Earth Planet. Sci.* 35, 47–72 doi: 10/dw337.
- Mohsenzadeh, S., Sheidai, M., Ghahremaninejad, F., Koohdar, F., 2020. A palynological study of the genus *plantago* (plantaginaceae). *Grana* 59, 454–465. <https://doi.org/10.1080/00173134.2020.1736147>.
- Montoya, E., Rull, V., van Geel, B., 2010. Non-pollen palynomorphs from surface sediments along an altitudinal transect of the Venezuelan Andes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 297, 169–183. <https://doi.org/10.1016/j.palaeo.2010.07.026>.
- Mucina, L., 2019. Biome: Evolution of a crucial ecological and biogeographical concept. *New Phytol.* 222, 97–114. <https://doi.org/10.1111/nph.15609>.
- Murad, W., 2012. Late Quaternary Vegetation History and Climate Change in the Gobi Desert, South Mongolia. Ph.D. thesis. Göttingen Centre for Biodiversity and Ecology.
- Ni, J., Yu, G., Harrison, S.P., Prentice, I.C., 2010. Palaeovegetation in China during the late quaternary: Biome reconstructions based on a global scheme of plant functional types. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 289, 44–61. <https://doi.org/10.1016/j.palaeo.2010.02.008>.
- Nowak, A., Nobis, M., Nowak, S., Gębala, M., Nobis, A., 2017. Phytosociology and ecology of deciduous forests in Tajikistan (middle asia). *Phytocoenologia* 47, 67–94. <https://doi.org/10.1127/phyto/2017/0084>.
- Nowak, A., Nowakowska, S., Nobis, A., Wróbel, A., Świercz, S., Klichowska, E., Dembic, I., Kusza, G., Nobis, M., 2020. Illustrated Flora of Tajikistan and Adjacent Areas. PAN, Polish Academy of Sciences, Warsaw Cracow Opole.
- Nowak, A., Świercz, S., Nowak, S., Nobis, M., 2021. Vegetation diversity of scree and taluses of the pamir and south-western tian Shan in middle asia. *Folia Geobot.* 56, 43–67. <https://doi.org/10.1007/s12224-021-09392-w>.
- Nowak, A., Świercz, S., Nowak, S., Nobis, A., Klichowska, E., Nobis, M., 2022. Syntaxonomy, diversity and distribution of the vegetation of rock crevices, clefts and ledges in the colline and montane belts of central pamir-alai and western tian Shan mts (middle asia). *Plant Biosyst. – Int. J. Deal. all Asp. Plant Biol.* 156, 808–823. <https://doi.org/10.1080/11263504.2021.1922531>.
- Nowak, A., Świercz, S., Nowak, S., Nobis, A., Klichowska, E., Nobis, M., 2024. Conspectus of the vegetation types of Tajikistan and adjacent areas with special focus on phytosociological classes. *Acta Soc. Bot. Pol.* 93, 1–16. <https://doi.org/10.5586/asbp/191675>.
- Nunes, A., Köbel, M., Pinho, P., Matos, P., de Bello, F., Correia, O., Branquinho, C., 2017. Which plant traits respond to aridity? A critical step to assess functional diversity in Mediterranean drylands. *Agric. For. Meteorol.* 239, 176–184. <https://doi.org/10.1016/j.agrformet.2017.03.007>.
- Nurtaev, B., Kharin, V., McCann, T., Valdivia-Manchego, M., 2013. The north nuratau fault zone, Uzbekistan—structure and evolution of a palaeozoic suture zone. *J. Geodyn.* 64, 1–14. <https://doi.org/10.1016/j.jog.2012.11.002>.
- Orlovsky, N., Birnbaum, E., 2002. The role of *haloxylon* species for combating desertification in central asia. *Plant Biosyst. – Int. J. Deal. all Asp. Plant Biol.* 136, 233–240. <https://doi.org/10.1080/11263500212331351139>.
- Parsons, R.W., Prentice, I.C., 1981. Statistical approaches to R-values and the pollen–vegetation relationship. *Rev. Palaeobot. Palynol.* 32, 127–152. [https://doi.org/10.1016/0034-6667\(81\)90001-4](https://doi.org/10.1016/0034-6667(81)90001-4).
- Pelánková, B., Chytrý, M., 2009. Surface pollen–vegetation relationships in the forest–steppe, taiga and tundra landscapes of the russian Altai mountains. *Rev. Palaeobot. Palynol.* 157, 253–265. <https://doi.org/10.1016/j.revpalbo.2009.05.005>.
- Pelánková, B., Kunes, P., Chytrý, M., Jankovská, V., Ermakov, N., Svobodová-Svitavská, H., 2008. The relationships of modern pollen spectra to vegetation and climate along a steppe–forest–tundra transition in southern siberia, explored by decision trees. *Holocene* 18, 1259–1271. <https://doi.org/10.1177/0959683608096600>.
- Peppe, D.J., Royer, D.L., Cariglino, B., Oliver, S.Y., Newman, S., Leight, E., Enikolopov, G., Fernandez-Burgos, M., Herrera, F., Adams, J.M., Correa, E., Currano, E.D., Erickson, J.M., Hinojosa, L.F., Hoganson, J.W., Iglesias, A., Jaramillo, C.A., Johnson, K.R., Jordan, G.J., Kraft, N.J.B., Lovelock, E.C., Lusk, C.H., Niinemets, Ü., Penuelas, J., Rapson, G., Wing, S.L., Wright, I.J., 2011. Sensitivity of leaf size and shape to climate: Global patterns and paleoclimatic applications. *New Phytol.* 190, 724–739. <https://doi.org/10.1111/j.1469-8137.2010.03615.x>.
- Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M.S., Cornwell, W.K., Craine, J.M., Gurvich, D.E., Urcelay, C., Veneklaas, E.J., Reich, P.B., Poorter, L., Wright, I.J., Ray, P., Enrico, L., Pausas, J.G., de Vos, A.C., Buchmann, N., Funes, G., Quétier, F., Hodgson, J.G., Thompson, K., Morgan, H.D., ter Steege, H., van der Heijden, M.G.A., Sack, L., Blonder, B., Poschlod, P., Vaieretti, M.V., Conti, G., Staver, A.C., Aquino, S., Cornelissen, J.H.C., Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M.S., Cornwell, W.K., Craine, J.M., Gurvich, D.E., Urcelay, C., Veneklaas, E.J., Reich, P.B., Poorter, L., Wright, I.J., Ray, P., Enrico, L., Pausas, J.G., de Vos, A.C., Buchmann, N., Funes, G., Quétier, F., Hodgson, J.G., Thompson, K., Morgan, H.D., ter Steege, H., van der Heijden, M.G.A., Sack, L., Blonder, B., Poschlod, P., Vaieretti, M.V., Conti, G., Staver, A.C., Aquino, S., Cornelissen, J.H.C., 2013. New handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* 61, 167–234. <https://doi.org/10.1071/BT12225>.
- 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. *Quat. Res.* 49, 183–196. <https://doi.org/10.1006/qres.1997.1961>.
- Peyron, O., Goring, S., Dormoy, I., Kotthoff, U., Pross, J., de Beaulieu, J.L., Drescher-Schneider, R., Vannière, B., Magny, M., 2011. Holocene seasonality changes in the central mediterranean region reconstructed from the pollen sequences of lake accessa (Italy) and tenaghi philippou (Greece). *Holocene* 21, 131–146. <https://doi.org/10.1177/0959683610384162>.
- Peyron, O., Combourieu-Nebout, N., Brayshaw, D., Goring, S., Andrieu-Ponel, V., Desprat, S., Fletcher, W., Gambin, B., Ioakim, C., Joannin, S., Kotthoff, U., Kouli, K., Montade, V., Pross, J., Sadori, L., Magny, M., 2017. Precipitation changes in the Mediterranean basin during the holocene from terrestrial and marine pollen records: a model-data comparison. *Clim. Past* 13, 249–265. <https://doi.org/10.5194/cp-13-249-2017>.
- Pierce, S., Negreiros, D., Cerabolini, B.E.L., Kattge, J., Díaz, S., Kleyer, M., Shipley, B., Wright, S.J., Soudzilovskaia, N.A., Onipchenko, V.G., van Bodegom, P.M., Frenette-Dussault, C., Weiher, E., Pinho, B.X., Cornelissen, J.H.C., Grime, J.P., Thompson, K., Hunt, R., Wilson, P.J., Buffa, G., Nyakunga, O.C., Reich, P.B., Caccianiga, M., Mangili, F., Ceriani, R.M., Luzzaro, A., Brusa, G., Siefert, A., Barbosa, N.P.U., Chapin III, F.S., Cornwell, W.K., Fang, J., Fernandes, G.W., Garnier, E., Le Stradic, S., Penuelas, J., Melo, F.P.L., Slaviero, A., Tabarelli, M., Tampucci, D., 2017. A global method for calculating plant CSR ecological strategies applied across biomes worldwide. *Funct. Ecol.* 31, 444–457. <https://doi.org/10.1111/1365-2435.12722>.
- Plantarium, 2007. Plants and lichens of Russia and neighboring countries: Open online galleries and plant identification guide. <https://www.plantarium.ru/>.
- Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A., Solomon, A.M., 1992. Special paper: a global biome model based on plant physiology and dominance, soil properties and climate. *J. Biogeogr.* 19, 117. <https://doi.org/10.2307/2845499> (arXiv:2845499).
- Prentice, C., Guiot, J., Huntley, B., Jolly, D., Cheddadi, R., 1996. Reconstructing biomes from palaeoecological data: a general method and its application to european pollen data at 0 and 6 ka. *Clim. Dyn.* 12, 185–194. <https://doi.org/10.1007/BF00211617>.
- Radula, M., Świercz, S., Nobis, M., Nowak, S., Nobis, A., Nowak, A., 2021. Palaeoclimate has a major effect on the diversity of endemic species in the hotspot of mountain biodiversity in Tajikistan. *Sci. Rep.* 11, 18684. <https://doi.org/10.1038/s41598-021-98027-3>.
- Rehfeld, K., Trachsel, M., Telford, R.J., Laepple, T., 2016. Assessing performance and seasonal bias of pollen-based climate reconstructions in a perfect model world. *Clim. Past* 12, 2255–2270. <https://doi.org/10.5194/cp-12-2255-2016>.
- Reille, M., 1999. Pollen et Spores d'europe et d'afrigue Du NordFeniXX. Laboratoire de Botanique historique et Palynologie, Marseille, p. 543.
- Robles, M., Peyron, O., Brugiapaglia, E., Ménot, G., Dugerdil, L., Ollivier, V., Ansanay-Alex, S., Develle, A.L., Tozalakyan, P., Meliksetian, K., 2022. Impact of climate changes on vegetation and human societies during the holocene in the south caucasus (vanevan, Armenia): a multiproxy approach including pollen, NPPs and brGDGTs. *Quat. Sci. Rev.* 277, 107297. <https://doi.org/10.1016/j.quascirev.2021.107297>.
- Roleček, J., Tichý, L., Zelený, D., Chytrý, M., 2009. Modified TWINSpan classification in which the hierarchy respects cluster heterogeneity. *J. Veg. Sci.* 20, 596–602. <https://doi.org/10.1111/j.1654-1103.2009.01062.x>.
- Rudaya, N., Sergey, K., Michal, S., Xianyang, C., Snezhana, Z., 2020. Postglacial history of the steppe Altai: climate, fire and plant diversity. *Quat. Sci. Rev.* 249, 106616. <https://doi.org/10.1016/j.quascirev.2020.106616>.
- Salonen, J.S., Luoto, M., Alenius, T., Heikkilä, M., Seppä, H., Telford, R.J., Birks, H.J.B., 2014. Reconstructing palaeoclimatic variables from fossil pollen using boosted regression trees: Comparison and synthesis with other quantitative reconstruction methods. *Quat. Sci. Rev.* 88, 69–81. <https://doi.org/10.1016/j.quascirev.2014.01.011>.
- Salonen, J.S., Korpela, M., Williams, J.W., Luoto, M., 2019. Machine-learning based reconstructions of primary and secondary climate variables from north american and european fossil pollen data. *Sci. Rep.* 9, 15805. <https://doi.org/10.1038/s41598-019-52293-4>.
- Sato, T., 2009. Influences of subtropical jet and Tibetan Plateau on precipitation pattern in Asia: Insights from regional climate modeling. *Quat. Int.* 194, 148–158 doi: 10/c56mvd.
- Schlütz, F., Dulamsuren, C., Wieckowska, M., Mühlenberg, M., Hauck, M., 2008. Late holocene vegetation history suggests natural origin of steppes in the northern mongolian mountain taiga. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 261, 203–217. <https://doi.org/10.1016/j.palaeo.2007.12.012>.
- Seddon, A.W.R., Mackay, A.W., Baker, A.G., Birks, H.J.B., Breman, E., Buck, C.E., Ellis, E. C., Froyd, C.A., Gill, J.L., Gillson, L., Johnson, E.A., Jones, V.J., Juggins, S., Macias-Fauria, M., Mills, K., Morris, J.L., Nogués-Bravo, D., Punyasena, S.W., Roland, T.P., Tanentzap, A.J., Willis, K.J., Aberhan, M., van Asperen, E.N., Austin, W.E.N., Battarbee, R.W., Bhagwat, S., Belanger, C.L., Bennett, K.D., Birks, H.H., Ramsey, C. B., Brooks, S.J., de Bruyn, M., Butler, P.G., Chambers, F.M., Clarke, S.J., Davies, A.L., Dearing, J.A., Ezard, T.H.G., Feurdean, A., Flower, R.J., Gell, P., Hausmann, S., Hogan, E.J., Hopkins, M.J., Jeffers, E.S., Korhola, A.A., Marchant, R., Kiefer, T., Lamentowicz, M., Larocque-Tobler, I., López-Merino, L., Liow, L.H., McGowan, S., Miller, J.H., Montoya, E., Morton, O., Nogués, S., Onoufriou, C., Boush, L.P., Rodríguez-Sánchez, F., Rose, N.L., Sayer, C.D., Shaw, H.E., Payne, R., Simpson, G., Sohar, K., Whitehouse, N.J., Williams, J.W., Witkowski, A., 2014. Looking forward

- through the past: Identification of 50 priority research questions in palaeoecology. *J. Ecol.* 102, 256–267. <https://doi.org/10.1111/1365-2745.12195>.
- Sennikov, A.N., Tojibaev, K.S., 2021. Checklist of Vascular Plants of the Tian-Shan Mountain System. Korea National Arboretum, Pocheon.
- Seregin, A.P., 2024. Moscow Digital Herbarium: Electronic Resource. <https://plant.depo.msu.ru/>.
- Shipunov, A., Fernández, A.J.L., Hassemmer, G., Alp, S., Lee, H.J., Pay, K., 2020. How to map a plantain: Phylogeny of the diverse plantagineae (lamiales). *bioRxiv* 2020–2027. <https://doi.org/10.1101/2020.07.31.230813>.
- Shumilovskikh, L., Djmalali, M., de Beaulieu, J.L., Ponel, P., Nokandeh, J., Omrani Rekavandi, H., Sauer, E., 2023. Woodland use in treeless landscapes? A compilation of botanical data from NE Iran during the sasanian empire. *Quat. Int.* 659, 34–44. <https://doi.org/10.1016/j.quaint.2023.02.013>.
- Shurigin, V., Alimov, J., Davranov, K., Gulyamova, T., Egamberdieva, D., 2022. The diversity of bacterial endophytes from iris pseudacorus L. and their plant beneficial traits. *Curr. Res. Microb. Sci.* 3, 100133. <https://doi.org/10.1016/j.crmicr.2022.100133>.
- Simon, S.M., Glaum, P., Valdovinos, F.S., 2023. Interpreting random forest analysis of ecological models to move from prediction to explanation. *Sci. Rep.* 13, 3881. <https://doi.org/10.1038/s41598-023-30313-8>.
- Sinaga, P., Klichowska, E., Nowak, A., Nobis, M., 2024. Hybridization and introgression events in cooccurring populations of closely related grasses (poaceae: Stipa) in high mountain steppes of central asia. *PLoS One* 19, e0298760. <https://doi.org/10.1371/journal.pone.0298760>.
- 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. <https://doi.org/10.1016/j.quascirev.2019.03.027>.
- Song, U., Park, J., Song, M., 2012. Pollen morphology of *pinus* (pinaceae) in Northeast China. *For. Sci. Technol.* 8, 179–186. <https://doi.org/10.1080/21580103.2012.704973>.
- Sorrel, P., Popescu, S.M., Klotz, S., Suc, J.P., Oberhänsli, H., 2007. Climate variability in the Aral Sea basin (central asia) during the late holocene based on vegetation changes. *Quat. Res.* 67, 357–370. <https://doi.org/10.1016/j.yqres.2006.11.006>.
- Spate, M., Leipe, C., Motuzaite Matuzeviciute, G., 2022. Reviewing the palaeoenvironmental record to better understand long-term human-environment interaction in inner asia during the late holocene. *Front. Ecol. Evol.* 10, 939374. <https://doi.org/10.3389/fevo.2022.939374>.
- Sugita, S., 2007a. Theory of quantitative reconstruction of vegetation I: Pollen from large sites REVEALS regional vegetation composition. *Holocene* 17, 229–241. <https://doi.org/10.1177/0959683607075837>.
- Sugita, S., 2007b. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. *Holocene* 17, 243–257. <https://doi.org/10.1177/0959683607075838>.
- Sun, Y., Xu, Q., Zhang, S., Li, Y., Li, M., Li, Y., Wang, T., Zhang, X., Wang, Y., Huang, R., Hao, J., Zheng, Z., Zhou, Z., 2020. A novel procedure for quantitative regional paleoclimatic reconstruction using surface pollen assemblages. *Quat. Sci. Rev.* 240, 106385. <https://doi.org/10.1016/j.quascirev.2020.106385>.
- Tan, L., Cheng, H., Li, D., Orozbaev, R., Li, Y., Xu, H., Edwards, R.L., Song, Y., Ma, L., Lin, F., Sinha, A., An, Z., 2024. Hydroclimatic changes on multiple timescales since 7800 y BP in the winter precipitation-dominated central asia. *Proc. Natl. Acad. Sci.* 121, e2321645121. <https://doi.org/10.1073/pnas.2321645121>.
- Tarasov, P.E., Cheddadi, R., Guiot, J., Bottema, S., Peyron, O., Belmonte, J., Ruiz-Sanchez, V., Saadi, F., Brewer, S., 1998. A method to determine warm and cool steppe biomes from pollen data; application to the Mediterranean and Kazakhstan regions. *J. Quat. Sci.* 13, 335–344. [https://doi.org/10.1002/\(SICI\)1099-1417\(199807/08\)13:4<335::AID-JQS375>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1099-1417(199807/08)13:4<335::AID-JQS375>3.0.CO;2-A).
- Team, R.C., 2014. R: a language and environment for statistical computing. *MSOR Connections* 1.
- Telford, R.J., Birks, H.J.B., 2005. The secret assumption of transfer functions: Problems with spatial autocorrelation in evaluating model performance. *Quat. Sci. Rev.* 24, 2173–2179. <https://doi.org/10.1016/j.quascirev.2005.05.001>.
- Telford, R.J., Birks, H.J.B., 2011. A novel method for assessing the statistical significance of quantitative reconstructions inferred from biotic assemblages. *Quat. Sci. Rev.* 30, 1272–1278. <https://doi.org/10.1016/j.quascirev.2011.03.002>.
- Ter Braak, C.J.F., Juggins, S., Birks, H.J.B., Van der Voet, H., 1993. Weighted averaging partial least squares regression (WA-PLS): Definition and comparison with other methods for species-environment calibration. In: *Multivariate Environmental Statistics*, 6. Elsevier, pp. 525–560.
- Theurillat, J.P., Willner, W., Fernández-González, F., Bültmann, H., Čarni, A., Gigante, D., Mucina, L., Weber, H., 2021. International code of phytosociological nomenclature. 4th edition. *Appl. Veg. Sci.* 24, e12491. <https://doi.org/10.1111/avsc.12491>.
- Thevs, N., Zerbe, S., Peper, J., Succow, M., 2008. Vegetation and vegetation dynamics in the tarim river floodplain of continental-arid Xinjiang, NW China. *Phytocoenologia* 38, 65. <https://doi.org/10.1127/0340-269X/2008/0038-0065>.
- Toderich, K., Shuyskaya, E., Shoaib, I., Gismatullina, L., Rajabov, T., Bekchanov, B., Aralova, D., 2009. Phytogenic resources of halophytes of central asia and their role for rehabilitation of sandy desert degraded rangelands. *Land Degrad. Dev.* 20, 386–396. <https://doi.org/10.1002/ldr.936>.
- Tojibaev, K., Beshko, N., Kodirov, U., R., B., Mirzalieva, D., 2018. Inventory of the Flora of Uzbekistan. Samarkand Region.
- Tojibaev, K., Beshko, N., Shomurodov, H., Kodirov, U., Turginov, O., Others, 2019. Inventory of the Flora of Uzbekistan. Kashkadarya Region.
- Tojibaev, K., Beshko, N., Esankulov, A., Avazbek, B., Azimova, D., 2021a. Inventory of the Flora of Uzbekistan: Dzhihazak Region.
- Tojibaev, K., Beshko, N., Esanov, H., 2021b. Inventory of the Flora of Uzbekistan: Bukhara Region.
- Trabucco, A., Zomer, R., 2018. Global aridity index and potential evapo-transpiration (ET0) climate database v2. In: CGIAR consortium for spatial information (CGIAR-CSI).
- Tsoar, H., 2001. Types of aeolian sand dunes and their formation. In: Beig, R., Ehlers, J., Frisch, U., Hepp, K., Hillebrandt, W., Imboden, D., Jaffe, R.L., Kippenhahn, R., Lipowsky, R., Löhneysen, H.V., Ojima, I., Weidenmüller, H.A., Wess, J., Zittartz, J., Balmforth, N.J., Provenzale, A. (Eds.), *Geomorphological Fluid Mechanics*, 582. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 403–429. https://doi.org/10.1007/3-540-45670-8_17.
- Usda, N., 2024. The PLANTS Database. <http://plants.usda.gov>, 02/01/2024.
- van der Sande, M.T., Bush, M.B., Urrego, D.H., Silman, M., Farfan-Rios, W., Garcia Cabrera, K., Shenkin, A., Malhi, Y., McMichael, C.H., Gosling, W., 2021. Modern pollen rain predicts shifts in plant trait composition but not plant diversity along the Andes–amazon elevational gradient. *J. Veg. Sci.* 32, e12925. <https://doi.org/10.1111/jvs.12925>.
- Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E., 2007. Let the concept of trait be functional! *Oikos* 116, 882–892. <https://doi.org/10.1111/j.0030-1299.2007.15559.x>.
- Wang, Q., Li, J.F., Lu, K.Q., Xie, G., Qin, F., Ferguson, D.K., Wang, G.H., Yao, Y.F., Wang, Y.F., 2022. Pollen R-values in arid central asia for quantitative palaeo-vegetation reconstruction. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 596, 110993. <https://doi.org/10.1016/j.palaeo.2022.110993>.
- Wei, H., Duan, R., Xu, Q., Yang, S., Fan, Q., Hou, G., Du, Y., Qin, Z., Gao, J., 2021. Fungal spore indicators of vegetation and highland pastoralism in modern topsoil and dung, eastern Tibetan Plateau. *CATENA* 202, 105231. <https://doi.org/10.1016/j.catena.2021.105231>.
- Weigelt, P., König, C., Kreft, H., 2020. GIFT – a global inventory of floras and traits for macroecology and biogeography. *J. Biogeogr.* 47, 16–43. <https://doi.org/10.1111/jbi.13623>.
- Welc, F., Toychiew, K., Suska-Malawska, M., Marks, L., Mętrak, M., 2016. Paleoclimatological and geochronological significance of the holocene loess–soil successions of the tien Shan foothills of Uzbekistan. *Stud. Quat.* 33, 57–68. <https://doi.org/10.1515/squa-2016-0006>.
- Wesche, K., Ambarli, D., Kamp, J., Török, P., Treiber, J., Dengler, J., 2016. The palaeartic steppe biome: a new synthesis. *Biodivers. Conserv.* 25, 2197–2231. <https://doi.org/10.1007/s10531-016-1214-7>.
- Wieczorek, M., Herzsuh, U., 2020. Compilation of relative pollen productivity (RPP) estimates and taxonomically harmonised RPP datasets for single continents and northern hemisphere extratropics. *Earth Syst. Sci. Data* 12, 3515–3528. <https://doi.org/10.5194/essd-12-3515-2020>.
- Xia, J., Ni, J., 2024. Regional features of pollen r-values in China. *Sci. China Earth Sci.* 67, 295–308. <https://doi.org/10.1007/s11430-022-1191-8>.
- Yin, G., Hu, Z., Chen, X., Tiyp, T., 2016. Vegetation dynamics and its response to climate change in central asia. *J. Arid. Land* 8, 375–388. <https://doi.org/10.1007/s40333-016-0043-6>.
- Zakirov, K.Z., 1955. Flora and vegetation of the zarafshan river basin. *Uzb. SSR Acad. Sci. Press Tashkent (Russ.)*.
- Zhang, J., Zhou, X., Long, H., 2020. Late quaternary loess accumulation at the rudak section in Uzbekistan, central asia: Chronology and palaeoclimate implications. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 109695. <https://doi.org/10.1016/j.palaeo.2020.109695>.
- Zhang, Q., Sun, C., Chen, Y., Chen, W., Xiang, Y., Others, 2022. Recent oasis dynamics and ecological security in the tarim river basin, central asia. *Sustainability* 14, 3372. <https://doi.org/10.3390/su14063372>.
- Zhao, Y., Liu, H., Li, F., Huang, X., Sun, J., Zhao, W., Herzsuh, U., Tang, Y., 2012. Application and limitations of the artemisia/chenopodiaceae pollen ratio in arid and semi-arid China. *Holocene* 22, 1385–1392. <https://doi.org/10.1177/0959683612449762>.
- Zheng, Z., Wei, J., Huang, K., Xu, Q., Lu, H., Tarasov, P., Luo, C., Beaudouin, C., Deng, Y., Pan, A., Zheng, Y., Luo, Y., Nakagawa, T., Li, C., Yang, S., Peng, H., Cheddadi, R., 2014. East asian pollen database: Modern pollen distribution and its quantitative relationship with vegetation and climate. *J. Biogeogr.* 41, 1819–1832. <https://doi.org/10.1111/jbi.12361>.