**[DRAFT 2]**

**Database of the Humans on the Plant Earth – Long-term impacts on biosphere dynamics (HOPE)**

**What is HOPE?**

HOPE is a European Research Council Advanced Grant-funded research project, which aims to answer two fundamental questions in biosphere dynamics in the context of ongoing anthropogenic environmental crises. With a focus on ecological dynamics, the major research questions of the project, formulated as hypotheses, are: (1) ecological processes changed through time due to human activity, and (2) inter-relationships between these processes changed with time. HOPE is an innovative, interdisciplinary, and nearly global project (includes five continents: Asia, Europe, North America, Latin America, and Oceania), and uses information from various sources, such as palaeoecology, archaeology and palaeoclimatology to test the above-mentioned hypotheses. Fossil pollen archives are the major sources of information in quantifying the properties of the terrestrial plant ecosystems during the Holocene (past 8000 years), in which these past ecosystem properties are analysed in the form of ecological properties of fossil pollen assemblages, so-called pollen-assemblage properties (PAPs) (Bhatta et al. 2023). Applying consistent state-of the-art numerical methods and rigorous statistical modelling techniques, HOPE aims to quantify 25 PAPs such as taxonomic richness, and evenness, compositional changes, compositional turnover, and rates of change, at a scale of an individual pollen record, ecological or climate zone, and continent. Spatio-temporal trends of these properties are then analysed with respect to proxies from different sources such as climate, human population density, and presence of human indicator taxa in a pollen record (human events).

Data available in online palaeoecological databases such as Neotoma Palaeoecology Database (Williams et al. 2018) and Pangaea (PANGAEA® - Data Publisher for Earth & Environmental Science https://www.pangaea.de/) are the key sources in testing the above-mentioned hypotheses. These databases not only manage the information on fossil pollen assemblages in the form of fossil pollen records, but also make available vital information on the age-chronology and meta-information regarding these records. However, HOPE also uses fossil pollen data from individual data contributors and publications, especially from the regions with large data-gaps.

**Why HOPE?**

Historical ecological legacies, such as past climatic and anthropogenic land-use changes, can play a significant role in influencing present-day ecosystem composition and dynamics (e.g., Jackson & Hobbs 2009; Jackson & Blois 2015; Birks 2019; Divíšek et al. 2020; Fordham et al. 2020). Studies so far suggest that there have been marked changes in natural ecosystems throughout the globe during the Holocene, especially with the beginning of human activities during the early- to mid-Holocene (e.g., Fuller et al. 2011; Fyfe et al. 2015; Pennisi 2015; Nolan et al. 2018; Kuneš et al. 2019; Stephens et al. 2019; Ellis et al. 2021). Moreover, a broad-scale study by Lyons et al. (2016) indicates that basic species-assemblage patterns that remained relatively consistent for many millions of years, changed significantly in the mid-Holocene because of anthropogenic impacts. In this context, cross-scale assessment of the ecosystem properties and their interrelations with respect to climatic and human land-use regimes is of paramount importance to get much needed scientific information for effective ecosystem management.

Tests of the above mentioned and similar hypotheses in geoscience including palaeoecology and palaeoclimatology are based on a fundamental principle of ‘uniformitarianism’, which assumes that *observations and measurements of contemporary Earth processes and products can be used to explain the formation of similar products by similar processes operating in the past through the action of ‘natural laws’* (Hutton 1795; Lyell 1830-33; Knight & Harrison 2014). Uniformitarian approach (the present is the key to the past), especially ‘methodological uniformitarianism’ (a concept that *present-day observations on contemporary processes can be used to infer past or future conditions*) (cf. Gould 1965; Baker 2014), is the basic logic and conceptual methodology by which the past events and underlying processes can be reconstructed and is thus the basic research approach and philosophy of Earth scientists. However, recent research (Dietl 2016; Lyons et al. 2016) proposes that human activities in the last 6000 years have modified the ecological properties of ecosystems to an extent that using past ‘natural experiments’ to predict the future changes is flawed and the use of uniformitarianism is outdated. Therefore, the hypotheses to be tested by HOPE are important for Earth System science, especially in cross verifying the associated proposal that methodological uniformitarianism should be discarded (cf. Dietl 2016). HOPE’s aim and potential impact thus extends well beyond the frontiers of human influence on ecosystem properties and mechanisms towards the core of Earth science philosophy and its conceptual framework.

**Methods**

HOPE is entirely based on three types of data: fossil pollen data, palaeoclimate data, and archaeological data of human density. Of these, fossil pollen data constitutes a major component of the HOPE-database, which is used to quantify PAPs and to detect periods of human influence (human-events) on natural ecosystems during the Holocene. Data of palaeoclimate, archaeological data on human density, and human-events are used as explanatory variables in the test of hypotheses.

***Data compilation***

*Fossil pollen data*

We compiled 2998 raw fossil pollen records from the Neotoma Palaeoecology Database (accessed on 28.02.2021), covering six continents (Asia, Europe, North America, Latin America, Africa, and Oceania) (Figure 1). However, we also compiled 202 raw fossil pollen records from the Pangaea database (accessed on 13.01.2020), publications (6 records), and private data contributors (451 records) for the regions with data-gaps (Figure 1). Data on age chronology (chronology table) and meta information (site name, country, continent, location, altitude, depositional environment, etc.) for each fossil record are also compiled together with the fossil pollen data. Most of the fossil pollen datasets are raw counts. However, a few pollen percentage datasets from some regions with data-gaps are also compiled in the database (Figure 1). Raw datasets, except those obtained from the private data contributors are available at xxx.

*Palaeoclimate data*

We extracted the Holocene-wide (covering past 12000 years) global climatic data from the CHELSA-TraCE21k database (Karger et al. 2023). CHELSA-TraCE21k takes account of orographic variation and provides monthly climate values for temperature and precipitation at 30 arcsecond (approximately 1 kilometre) spatial resolution (Karger et al. 2023) in 100-year time periods for the last 21000 years. From these data, we predicted climatic variables for each sample of the processed (see below) pollen records using natural splines. Climate data of HOPE are located at xxx.

*Archaeological data of summed probability densities (SPDs)*

We quantified SPDs for the selected pollen records using the global dataset of archaeological artefacts-based radiocarbon dates (archaeo-radiocarbon) from Bird et al. (2022). Archaeo-radiocarbon dates with valid geographical location, and ‘LocAccuracy’ > 0 were selected for estimation of SPDs. For each pollen record, archaeo-radiocarbon dates were classified by the geographical distance to the pollen record. A representative SPD estimation was selected for each pollen record as an indicator of human presence or anthropogenic activity for that record. For further details of SPD estimation, see Felde et al. (xxx). Archaeo-radiocarbon and SPD data for HOPE are available at xxx.

*Data of present climatic classification*

To study the patterns in PAPs at different spatial scales, we classify all the continents into present climatic zones. For this, we adopted new version of present (1980–2016) Köppen-Geiger climate classification at 1-km resolution (Beck et al. 2018). Geo-tiff files of global maps of the Köppen-Geiger climate classification (published in Beck et al. 2018) that were used for climatic zonation in our project area are available at xxx.

**Data screening**

All HOPE data compiled from various sources was processed consistently applying a set of data robust processing criteria which were set prior to data processing. Important criteria for data processing include data type and variable element (pollen versus other), depositional environment, ecological groups, number of samples in each pollen record, number of taxa in each record, number of chronology control points, type of chronology control point, gap of chronology control points, age-limit of the youngest and oldest samples in each record, and pollen counts in each sample of a pollen record. A workflow was devised to describe these criteria and to provide a procedure for their application in a stepwise manner during data processing (Flantua et al. 2023). The workflow was then implemented using ‘RFossilpol version 1.0.0’ package in R (Flantua et al. 2023). ‘RFossilpol’ processes and standardises fossil pollen data applying the above-mentioned criteria in a sequential manner and ensures reproducibility of data processing outputs. Estimation of age (calibrated years before present or cal yr BP) of each sample of the pollen records and standardisation of the nomenclature and taxonomy of pollen taxa in the records are the most important steps during data processing, which are described below.

*Age-depth modelling*

We estimated age of each sample in each pollen record by the Bchron (Haslett & Parnell 2008) age-depth modelling approach, in which we calibrated the dates of the selected chronology-control point types, and then predicted calibrated age for each sample-depth in all records. We performed age-depth modelling using the “bchron” package in R (Haslett & Parnell 2008). Details of the age-depth modelling for the fossil pollen data in HOPE are given in Flantua et al. (2023), and the age-depth model for each record are available at xxx.

*Taxonomic harmonisation*

Fossil pollen taxa are pollen morphotypes identified by palynologists based on the morphology of pollen found in the sediment samples. Nomenclature and taxonomy of pollen taxa differ among sequences and among pollen analysts (Birks et al. xxx). Therefore, standardisation of taxonomy and nomenclature of pollen taxa in the fossil pollen records compiled from large geographical territories is an important step in data processing, called ‘taxonomic harmonisation’. For taxonomic harmonisation, we followed a hierarchical nomenclature system of the pollen morpho-types, in which closely related morpho-types are merged into a higher accepted pollen taxon or morpho-type (Birks et al. xxx). We harmonised the pollen morpho-types to the taxonomically highest-precision level of pollen morpho-types identified by palynologists. Based on the present flora and hence the pollen taxonomy, and the principles and traditions of identifying, naming, and classifying pollen morphotypes by different pollen analysts, we delineated phytogeographic regions (harmonisation regions) globally, and harmonised pollen taxa for each harmonisation region separately. Details of the taxonomic harmonisation of fossil pollen data in HOPE is given in Birks et al. (xxx).

*Pollen diagrams*

We produced pollen diagrams of the processed pollen records, which were analysed by experts for overall data quality and for detection of human influence on vegetation based on pollen taxa. Pollen diagrams of the fossil pollen records are available at *xxx.*

**Processed data**

There are 1377 fully processed fossil pollen records in HOPE database, covering 6 regions namely, Asia, Europe, North America, Latin America, Africa, and Oceania (Figure 2). However, Africa includes only 20 records, which are not harmonised taxonomically, and may not be representative of the entire region. Therefore, we exclude data from Africa for further analyses.

Each processed dataset consists of three main components: pollen counts/percentages, age-chronology, and metadata, and altogether, there are 38 data parameters (Table 1). Important data parameters are summarised in Tables 2-5 and Figures 3-8.

Number of harmonised taxa, number of samples, and pollen counts per record is rather heterogeneous across records within each region (Figures 3, 4, 5). Average number of harmonised taxa per region varies from 41 in Asia and North America to 64 in Europe (Table 3) and average number of samples in pollen records ranges from 45 in Oceania to 84 in Europe (Table 3). Average number of pollen counts varies from 339 in Oceania to 885 in Europe (Table 3). Pollen counts in some records from Asia and Oceania are consistent (100) across multiple records because these records are pollen percentages, in which pollen count in each sample is 100 (Figure 5). A few pollen records from Latin America, North America and Europe have exceptionally high pollen counts. Similarly, there is high variation in number of chronology control points among chronology tables from each region (Figure 6), where average number of chronology control points varies from 8 each in Latin America, North America, and Oceania to 13 in Asia (Table 3). Calibrated age of pollen records in each region ranges from 0 cal yr BP to 12000 cal yr BP and is relatively consistent across pollen records in all the regions, except in Oceania (Figure 7). However, there is big variation in the calibrated age uncertainty of the samples of pollen records in each region (Figure 8, table 3), where average calibrated age uncertainty per region varies from 829 years in Latin America to 990 years in North America (Table 3). Depositional environment types for each pollen record are listed in Table 4. There are 34, 33, 39, 47, and 26 depositional environments of pollen records, respectively from Asia, Europe, North America, Latin America, and Oceania (Table 4). Similarly, there are 20, 14, 12, 11, and 8 chronology control point types in the age-chronology tables, respectively from Asia, Europe, North America, Latin America, and Oceania (Table 5).

**An example of use of HOPE-data**

Fossil pollen data in HOPE is mainly used for quantification of PAPs for each pollen record, which thereby are sued as response variables in the test of main HOPE-hypotheses (see above and Felde et al. xxx). Additionally, PAPs for a particular region are analysed at different spatial and temporal scales to elucidate spatio-temporal dynamics of pollen-vegetation in the region during the Holocene (e.g., Bhatta et al. 2023). Here we present an example of spatio-temporal patterns of compositional turnover and rate of pollen-compositional change (RoC) using fossil pollen data from Europe (Figures 9, 10, 11).

We estimated compositional turnover in each processed pollen record by performing a detrended canonical correspondence analysis (DCCA) (ter Braak 1986; ter Braak & Šmilauer 2012) of pollen compositional data (with Hill’s detrending, non-linear rescaling, and no down-weighting of rare taxa (Hill & Gauch 1980)), where we used time (sample age) as the sole explanatory variable. We estimated the total gradient length by calculating range of the sample scores (*CaseR*) of the first DCCA axis, whereas the individual ‘*CaseR*’ scores for a pollen record represent sample-to-sample compositional turnover based on all pollen taxa within the record (Birks 2007; Felde et al. 2020). We performed DCCA analysis using CANOCO version 4.5 (ter Braak & Šmilauer 2002) via “fit\_ordination” function in R-Ecopol package in R version 4.1.1 (R Core Team 2022) using a second-order polynomial predictor of age.

Using a procedure by Mottl et al. (2021) we estimated the RoC in each pollen record using the authors’ “R-Ratepol” package (Mottl et al. 2021). We used an age bin-size of 500 years and 5 window shifts and estimated RoC using the Chi-squared metric (Prentice 1980) between consecutive time intervals. We transformed pollen counts to proportions and applied the “Age-weighted average” smoothing method with maximal age range for 500 years for each pollen type. For RoC estimation, we selected the random sample in each bin as a representative of that bin and standardised each “Working Unit” to 150 individuals (otherwise, to the smallest number of pollen grains in record) using random resampling without replacement. We set the number of iterations to 1,000. A detailed procedure for RoC estimation, and reasons for estimating RoC over 500-year bins are described in Mottl et al. (2021).

Spatial variation in pollen compositional turnover, as revealed by total gradient length, is non-linear in Europe (Figure 9). Along latitudinal gradient, compositional turnover increases from c. 40o N to c. 50o N and then declines up to c. 60o N. There is a gradual increase in total compositional turnover towards northern extreme of the latitudinal gradient above c. 60o N (Figure 9b). Similarly, total compositional turnover increases gradually from c. -10o E to c. 5o E longitude, and then decreases linearly eastwards along the latitudinal gradient (Figure 9c).

Temporally, there is a marked increase in compositional turnover (DCCA axis 1 *CaseR* scores) during the Last Glacial Maximum (LGM)-Holocene transition (c. 12000-10000 cal yr BP) (Figure 10), and then compositional turnover decreases linearly during the Holocene. This pattern of temporal variation in compositional turnover is consistent for all climate zones. However, in the ‘Arid’ zone, increase in turnover is observed until middle phase of the early-Holocene (c. 12000-8000 cal yr BP). In the ‘Cold’ climate zone, there is a gradual increase in compositional turnover from the late-Holocene (c. 1000 cal yr BP) to present (Figure 10).

Overall, we observe a gradual increase in RoC from the LGM-Holocene transition to the recent time, with an acceleration in RoC during the late-Holocene (Figure 11). This pattern is consistent across the climate zones except in ‘Arid’ zone, in which the pattern is rather irregular (Figure 11).

**References**

Baker, V.R. 2014. Uniformitarianism, earth system science, and geology. *Anthropocene* 5: 76–79.

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. *Scientific Data* 5: 1–12.

Bhatta, K.P., Mottl, O., Felde, V.A., Flantua, S.G.A., Birks, H.H., Cao, X., Chen, F., Grytnes, J.-A., Seddon, A.W.R., & Birks, H.J.B. 2023. Exploring spatio-temporal patterns of palynological changes in Asia during the Holocene. *Frontiers in Ecology and Evolution* 11: DOI 10.3389/fevo.2023.1115784.

Bird, D., Miranda, L., Linden, M. Vander, Robinson, E., Bocinsky, R.K., Nicholson, C., Capriles, J.M., Finley, J.B., Gayo, E.M., Gil, A., Guedes, J. d’Alpoim, Hoggarth, J.A., Kay, A., Loftus, E., Lombardo, U., Mackie, M., Palmisano, A., Solheim, S., Kelly, R.L., & Freeman, J. 2022. p3k14c, a synthetic global database of archaeological radiocarbon dates. *Scientific Data* 9:.

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

Birks, H.J.B. 2007. Estimating the amount of compositional change in late-Quaternary pollen-stratigraphical data. *Vegetation History and Archaeobotany* 16: 197–202.

ter Braak, C.J.F. 1986. Canonical Correspondence Analysis : A New Eigenvector Technique for Multivariate Direct Gradient Analysis. *Ecology* 67: 1167–1179.

ter Braak, C.J.F., & Šmilauer, P. 2002. Canoco for Windows Version 4.5.

ter Braak, C.J.F., & Šmilauer, P. 2012. Canoco Reference Manual and User’s Guide: software for ordination (version 5.0). Microcomputer Power, Ithaca, USA.

Dietl, G.P. 2016. Different worlds. *Nature* 529: 29–30.

Divíšek, J., Hájek, M., Jamrichová, E., Petr, L., Večeřa, M., Tichý, L., Willner, W., & Horsák, M. 2020. Holocene matters: Landscape history accounts for current species richness of vascular plants in forests and grasslands of eastern Central Europe. *Journal of Biogeography* 47: 721–735.

Ellis, E.C., Gauthier, N., Klein Goldewijk, K., Bliege Bird, R., Boivin, N., Díaz, S., Fuller, D.Q., Gill, J.L., Kaplan, J.O., Kingston, N., Locke, H., McMichael, C.N.H., Ranco, D., Rick, T.C., Shaw, M.R., Stephens, L., Svenning, J.-C., & Watson, J.E.M. 2021. People have shaped most of terrestrial nature for at least 12,000 years. *Proceedings of the National Academy of Sciences* 118: e2023483118.

Felde, V.A., Flantua, S.G.A., Jenks, C.R., Benito, B.M., de Beaulieu, J.-L., Kuneš, P., Magri, D., Nalepka, D., Risebrobakken, B., ter Braak, C.J.F., Allen, J.R.M., Granoszewski, W., Helmens, K.F., Huntley, B., Kondratienė, O., Kalniņa, L., Kupryjanowicz, M., Malkiewicz, M., Milner, A.M., Nita, M., Noryśkiewicz, B., Pidek, I.A., Reille, M., Salonen, J.S., Šeirienė, V., Winter, H., Tzedakis, P.C., & Birks, H.J.B. 2020. Compositional turnover and variation in Eemian pollen sequences in Europe. *Vegetation History and Archaeobotany* 29: 101–109.

Flantua, S.G.A., Mottl, O., Felde, V.A., Bhatta, K.P., Birks, H.H., Grytnes, J.-A., Seddon, A.W.R., & Birks, H.J.B. 2023. A guide to the processing and standardization of global palaeoecological data for large-scale syntheses using fossil pollen. 1–18.

Fordham, D.A., Jackson, S.T., Brown, S.C., Huntley, B., Brook, B.W., Dahl-Jensen, D., Gilbert, M.T.P., Otto-Bliesner, B.L., Svensson, A., Theodoridis, S., Wilmshurst, J.M., Buettel, J.C., Canteri, E., McDowell, M., Orlando, L., Pilowsky, J., Rahbek, C., Nogues-Bravo, D., Thomas Gilbert, M.P., Otto-Bliesner, B.L., Svensson, A., Theodoridis, S., Wilmshurst, J.M., Buettel, J.C., Canteri, E., McDowell, M., Orlando, L., Pilowsky, J., Rahbek, C., & Nogues-Bravo, D. 2020. Using paleo-archives to safeguard biodiversity under climate change. *Science* 369: eabc5654.

Fuller, D.Q., van Etten, J., Manning, K., Castillo, C., Kingwell-Banham, E., Weisskopf, A., Qin, L., Sato, Y.-I., & Hijmans, R.J. 2011. The contribution of rice agriculture and livestock pastoralism to prehistoric methane levels:An archaeological assessment. *The Holocene* 21: 743–759.

Fyfe, R.M., Woodbridge, J., & Roberts, N. 2015. From forest to farmland: Pollen-inferred land cover change across Europe using the pseudobiomization approach. *Global Change Biology* 21: 1197–1212.

Gould, S.J. 1965. Is uniformitarianism necessary? *American Journal of Science* 263: 223–228.

Haslett, J., & Parnell, A. 2008. A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society. Series C: Applied Statistics* 57: 399–418.

Hill, M.O., & Gauch, H.G. 1980. Detrended Correspondence Analysis, an improved ordination technique. *Vegetatio* 42: 47–58.

Hutton, J. 1795. *Theory of the Earth; with proofs and illustrations*. Creesh, Edinburgh.

Jackson, S.T., & Blois, J.L. 2015. Community ecology in a changing environment: Perspectives from the Quaternary. *Proceedings of the National Academy of Sciences of the United States of America* 112: 4915–4921.

Jackson, S.T., & Hobbs, R.J. 2009. Ecological restoration in the light of ecological history. *Science* 325: 567–569.

Karger, D.N., Nobis, M.P., Normand, S., Graham, C.H., & Niklaus, E. 2023. CHELSA-TraCE21k v1.0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the Past* 19: 439–456.

Knight, J., & Harrison, S. 2014. Limitations of uniformitarianism in the Anthropocene. *Anthropocene* 5: 71–75.

Kuneš, P., Abraham, V., & Herben, T. 2019. Changing disturbance-diversity relationships in temperate ecosystems over the past 12000 years. *Journal of Ecology* 107: 1678–1688.

Lyell, C. 1830. *Principles of Geology*. Murray, London.

Lyons, S.K., Amatangelo, K.L., Behrensmeyer, A.K., Bercovici, A., Blois, J.L., Davis, M., Dimichele, W.A., Du, A., Eronen, J.T., Tyler Faith, J., Graves, G.R., Jud, N., Labandeira, C., Looy, C. V., McGill, B., Miller, J.H., Patterson, D., Pineda-Munoz, S., Potts, R., Riddle, B., Terry, R., Tóth, A., Ulrich, W., Villaseñor, A., Wing, S., Anderson, H., Anderson, J., Waller, D., & Gotelli, N.J. 2016. Holocene shifts in the assembly of plant and animal communities implicate human impacts. *Nature* 529: 80–83.

Mottl, O., Grytnes, J.A., Seddon, A.W.R., Steinbauer, M.J., Bhatta, K.P., Felde, V.A., Flantua, S.G.A., & Birks, H.J.B. 2021. Rate-of-change analysis in paleoecology revisited: A new approach. *Review of Palaeobotany and Palynology* 293:.

Nolan, C., Overpeck, J.T., Allen, J.R.M., Anderson, P.M., Betancourt, J.L., Binney, H.A., Brewer, S., Bush, M.B., Chase, B.M., Cheddadi, R., Djamali, M., Dodson, J., Edwards, M.E., Gosling, W.D., Haberle, S., Hotchkiss, S.C., Huntley, B., Ivory, S.J., Kershaw, A.P., Kim, S.-H., Latorre, C., Leydet, M., Lézine, A.-M., Liu, K.-B., Liu, Y., Lozhkin, A. V, McGlone, M.S., Marchant, R.A., Momohara, A., Moreno, P.I., Müller, S., Otto-Bliesner, B.L., Shen, C., Stevenson, J., Takahara, H., Tarasov, P.E., Tipton, J., Vincens, A., Weng, C., Xu, Q., Zheng, Z., & Jackson, S.T. 2018. Past and future global transformation of terrestrial ecosystems under climate change. *Science* 361: 920–923.

PANGAEA® - Data Publisher for Earth & Environmental Science. PANGAEA® - Data Publisher for Earth & Environmental Science doi:10.1594/PANGAEA.

Pennisi, E. 2015. Human impacts on ecosystems began thousands of years ago: Early humans broke up existing plant and animal networks, perhaps boosting extinction risks. *Science* 350: 1452.

Prentice, I.C. 1980. Multidimensional scaling as a research tool in Quaternary palynology: A review of theory and methods. *Review of Palaeobotany and Palynology* 31: 71–104.

R Core Team. 2022. R: A language and environment for statistical computing.

Stephens, L., Fuller, D., Boivin, N., Rick, T., Gauthier, N., Kay, A., Marwick, B., Geralda, C., Armstrong, D., Barton, C.M., Denham, T., Douglass, K., Driver, J., Janz, L., Roberts, P., Rogers, J.D., Thakar, H., Altaweel, M., Johnson, A.L., Sampietro Vattuone, M.M., Aldenderfer, M., Archila, S., Artioli, G., Bale, M.T., Beach, T., Borrell, F., Braje, T., Buckland, P.I., Jiménez Cano, N.G., Capriles, J.M., Diez Castillo, A., Çilingiroğlu, Ç., Negus Cleary, M., Conolly, J., Coutros, P.R., Covey, R.A., Cremaschi, M., Crowther, A., Der, L., di Lernia, S., Doershuk, J.F., Doolittle, W.E., Edwards, K.J., Erlandson, J.M., Evans, D., Fairbairn, A., Faulkner, P., Feinman, G., Fernandes, R., Fitzpatrick, S.M., Fyfe, R., Garcea, E., Goldstein, S., Goodman, R.C., Dalpoim Guedes, J., Herrmann, J., Hiscock, P., Hommel, P., Horsburgh, K.A., Hritz, C., Ives, J.W., Junno, A., Kahn, J.G., Kaufman, B., Kearns, C., Kidder, T.R., Lanoë, F., Lawrence, D., Lee, G.-A., Levin, M.J., Lindskoug, H.B., López-Sáez, J.A., Macrae, S., Marchant, R., Marston, J.M., McClure, S., McCoy, M.D., Miller, A.V., Morrison, M., Motuzaite Matuzeviciute, G., Müller, J., Nayak, A., Noerwidi, S., Peres, T.M., Peterson, C.E., Proctor, L., Randall, A.R., Renette, S., Robbins Schug, G., Ryzewski, K., Saini, R., Scheinsohn, V., Schmidt, P., Sebillaud, P., Seitsonen, O., Simpson, I.A., Sołtysiak, A., Speakman, R.J., Spengler, R.N., Steffen, M.L., Storozum, M.J., Strickland, K.M., Thompson, J., Thurston, T.L., Ulm, S., Ustunkaya, M.C., Welker, M.H., West, C., Williams, P.R., Wright, D.K., Wright, N., Zahir, M., Zerboni, A., Beaudoin, E., Munevar Garcia, S., Powell, J., Thornton, A., Kaplan, J.O., Gaillard, M.-J., Klein Goldewijk, K., & Ellis, E. 2019. Archaeological assessment reveals Earth’s early transformation through land use. *Science* 365: 897–902.

Williams, J.W., Grimm, E.G., Blois, J., Charles, D.F., Davis, E., Goring, S.J., Graham, R., Smith, A.J., Anderson, M., Arroyo-Cabrales, J., Ashworth, A.C., Betancourt, J.L., Bills, B.W., Booth, R.K., Buckland, P., Curry, B., Giesecke, T., Hausmann, S., Jackson, S.T., Latorre, C., Nichols, J., Purdum, T., Roth, R.E., Stryker, M., & Takahara, H. 2018. The Neotoma Paleoecology Database: A multi-proxy, international community-curated data resource. *Quaternary Research* 89: 156–177.

**A map of the world with different colored squares

Description automatically generated**

**Figure 1.** Distribution of raw pollen records across climate zones (upper panel); number of raw fossil pollen records based on data source and type of pollen data.

**A close-up of a map

Description automatically generated**

**Figure 2.** Distribution of processed pollen records across climate zones (upper panel); number of fully processed fossil pollen records based on data source and type of pollen data.

**A group of colorful lines

Description automatically generated with medium confidence**

**Figure 3.** Number of harmonised taxa per pollen record in each climate zone.

A group of graphs with different colored lines

Description automatically generated with medium confidence

**Figure 4.** Number of samples per pollen record in each climate zone.

A group of graphs with numbers

Description automatically generated with medium confidence

**Figure 5.** Meanpollen count per pollen record in each climate zone

A group of graphs with different colored lines

Description automatically generated

**Figure 6.** Number of chronology control points per pollen record

A group of colorful lines

Description automatically generated with medium confidence

**Figure 7.** Calibrated age-range (cal yr BP) per pollen record.

**A group of colorful lines

Description automatically generated with medium confidence**

**Figure 8**. Mean of calibrated age uncertainty (yrs) per pollen record.

A diagram of different weather conditions

Description automatically generated with medium confidence

**Figure 9.** Spatial pattern of total compositional turnover (gradient length of detrended canonical correspondence analysis (DCCA)) revealed by fossil pollen assemblages in Europe.

A graph of different colored lines

Description automatically generated

**Figure 10.** Temporal pattern of compositional turnover (*CaseR* scores of the first axis of detrended canonical correspondence analysis (DCCA) revealed by fossil pollen assemblages in Europe. Trends for each climate zone are fitted by four (one for each climate-zone) hierarchical generalised additive models (hGAMs), accounting for the random effect of location. The red line (loess smoother) represents overall temporal pattern of compositional turnover in Europe.

A graph of different colored lines

Description automatically generated

**Figure 11.** Temporal pattern of rate of pollen-compositional change (RoC) in Europe. Trends for each climate zone are fitted by four (one for each climate-zone) hierarchical generalised additive models (hGAMs), accounting for the random effect of location. The red line (loess smoother) represents overall temporal pattern of RoC in Europe.

**Table 1.** Parameters of HOPE fossil pollen data

|  |  |
| --- | --- |
| Data parameters | Data parameters |
| dataset\_id | age\_type |
| handle | curve\_name |
| siteid | postbomb\_curve\_name |
| sitename | raw\_counts |
| long | pollen\_percentage |
| lat | young\_age |
| altitude | old\_age |
| depositionalenvironment | end\_of\_interest\_period |
| region | source\_of\_data |
| country | data\_publicity |
| harmonisation\_region | doi |
| levels | counts\_harmonised |
| n\_sample\_counts | publications |
| age\_min | publication |
| age\_max | source\_of\_data\_simple |
| age\_uncertainty | counts\_to\_plot |
| chron\_control\_format | ecozone\_koppen\_5 |
| n\_chron\_control | ecozone\_koppen\_30 |
| chron\_control\_limits | ecozone\_koppen\_15 |

**Table 2.** Summary of processed pollen data per pollen record.

‘.csv’ file with: dataset\_id, long, lat, depositionalenvironment, region, ecozone\_koppen\_5, pollen\_percentage (TRUE/FALSE), n\_taxa\_harmonised\_per\_sequence, n\_sample\_filtered\_per\_sequence, n\_chron\_control\_filtered\_per\_sequence, max\_pol\_counts\_per\_sequence, min\_pol\_counts\_per\_sequence, mean\_pol\_counts\_per\_sequence, max\_chronology\_age\_per\_sequence, min\_chronology\_age\_per\_sequence, mean\_chronology\_age\_per\_sequence, max\_chronology\_age\_error\_per\_sequence, min\_chronology\_age\_error\_per\_sequence, mean\_chronology\_age\_error\_per\_sequence, max\_calibrated\_age\_per\_sequence, min\_calibrated\_age\_per\_sequence, mean\_calibrated\_age\_per\_sequence, calibrated\_age\_range\_per\_sequence, max\_calibrated\_age\_uncertainty\_per\_sequence, min\_calibrated\_age\_uncertainty\_per\_sequence, mean\_calibrated\_age\_uncertainty\_per\_sequence

**Table 3.** Summary of processed pollen data per region.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| region | Europe | North America | Latin America | Oceania | Asia |
| max\_n\_taxa\_harmonised\_per\_continent | 130 | 113 | 229 | 123 | 139 |
| min\_n\_taxa\_harmonised\_per\_continent | 16 | 13 | 9 | 16 | 7 |
| mean\_n\_taxa\_harmonised\_per\_continent | 64 | 41 | 61 | 63 | 41 |
| max\_n\_sample\_filtered\_per\_continent | 858 | 551 | 306 | 397 | 612 |
| min\_n\_sample\_filtered\_per\_continent | 13 | 6 | 5 | 7 | 5 |
| mean\_n\_sample\_filtered\_per\_continent | 84 | 47 | 52 | 45 | 53 |
| max\_n\_chron\_control\_filtered\_per\_continent | 405 | 42 | 50 | 67 | 356 |
| min\_n\_chron\_control\_filtered\_per\_continent | 3 | 3 | 3 | 3 | 3 |
| mean\_n\_chron\_control\_filtered\_per\_continent | 11 | 8 | 8 | 8 | 13 |
| max\_pol\_counts\_per\_continent | 16390.00 | 12540.00 | 33470.00 | 2800.00 | 7556.00 |
| min\_pol\_counts\_per\_continent | 26.00 | 26.00 | 26.00 | 4.00 | 25.00 |
| mean\_pol\_counts\_per\_continent | 885.00 | 514.00 | 690.00 | 339.00 | 446.00 |
| max\_chronology\_age\_per\_continent | 45800.00 | 46000.00 | 44100.00 | 45900.00 | 44230.00 |
| min\_chronology\_age\_per\_continent | -67.00 | -74.00 | -525.00 | -65.00 | -2000.00 |
| mean\_chronology\_age\_per\_continent | 5161.05 | 5893.77 | 5620.09 | 6711.86 | 6430.29 |
| max\_chronology\_age\_error\_per\_continent | 2744.00 | 3000.00 | 2700.00 | 2400.00 | 2410.00 |
| min\_chronology\_age\_error\_per\_continent | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| mean\_chronology\_age\_error\_per\_continent | 126.16 | 141.43 | 93.69 | 143.56 | 122.55 |
| max\_calibrated\_age\_per\_continent | 12000.00 | 11999.00 | 12000.00 | 11994.50 | 12000.00 |
| min\_calibrated\_age\_per\_continent | -72.00 | -72.00 | -72.00 | -72.00 | -72.00 |
| mean\_calibrated\_age\_per\_continent | 5360.69 | 5501.25 | 4488.30 | 4575.53 | 5163.71 |
| max\_calibrated\_age\_uncertainty\_per\_continent | 13080.83 | 9678.42 | 7463.69 | 4730.51 | 17416.75 |
| min\_calibrated\_age\_uncertainty\_per\_continent | 0.50 | 0.00 | 0.00 | 0.00 | 1.50 |
| mean\_calibrated\_age\_uncertainty\_per\_continent | 795.61 | 990.28 | 829.49 | 921.16 | 879.33 |
| mean\_calibrated\_age\_range\_per\_continent | 10543.10 | 11058.70 | 9924.00 | 10108.41 | 10258.85 |

**Table 4.** Deposition environment of the processed pollen records

|  |  |
| --- | --- |
| region | depositional environment |
| Asia | Bog |
| Asia | Bog sediment |
| Asia | Coastal |
| Asia | Drained Lake |
| Asia | Eolian deposit |
| Asia | Fen |
| Asia | Glacial Origin Lake |
| Asia | Lacustrine |
| Asia | Lake |
| Asia | Lake sediment |
| Asia | Marsh |
| Asia | Mire |
| Asia | Moraine Dammed Lake |
| Asia | Natural Lake |
| Asia | Palustrine |
| Asia | Peat bog |
| Asia | Peat sediment |
| Asia | Peatland |
| Asia | Playa |
| Asia | Raised Bog |
| Asia | River basin |
| Asia | Swamp |
| Asia | Tectonic Origin Lake |
| Asia | Terrestrial |
| Asia | Thermokarst Lake |
| Asia | Wetland |
| Asia | coastal sediment |
| Asia | lake |
| Asia | peat |
| Europe | Blanket Bog |
| Europe | Bog |
| Europe | Cirque Lake |
| Europe | Coastal lagoon |
| Europe | Drained Lake |
| Europe | Explosion Crater Lake |
| Europe | Fen |
| Europe | Floodplain Mire |
| Europe | Glacial Meltwater Channel Lake |
| Europe | Glacial Origin Lake |
| Europe | Glacial Scour Lake |
| Europe | Kettle Lake |
| Europe | Lacustrine |
| Europe | Landslide Origin Lake |
| Europe | Marsh |
| Europe | Mire |
| Europe | Moraine Dammed Lake |
| Europe | Natural Lake |
| Europe | Natural Lake (Origin Unknown) |
| Europe | Open-water Transition Mire |
| Europe | Palustrine |
| Europe | Periglacial Origin Lake |
| Europe | Raised Bog |
| Europe | Soligenous Mire |
| Europe | Swamp |
| Europe | Tectonic Origin Lake |
| Europe | Valley Glacier Lake |
| Europe | Valley Mire |
| Europe | Volcanic Origin Lake |
| Latin America | Blanket Bog |
| Latin America | Bog |
| Latin America | Cienega |
| Latin America | Coastal swamp |
| Latin America | Ephemeral Lake |
| Latin America | Explosion Crater Lake |
| Latin America | Fen |
| Latin America | Floodplain Mire |
| Latin America | Glacial Origin Lake |
| Latin America | Glacial Scour Lake |
| Latin America | Lacustrine |
| Latin America | Lake |
| Latin America | Lake Dammed by Alluvial Fans |
| Latin America | Lake Marginal Fen |
| Latin America | Mangrove peat |
| Latin America | Marsh |
| Latin America | Mire |
| Latin America | Moraine Dammed Lake |
| Latin America | Natural Lake |
| Latin America | Natural Lake (Origin Unknown) |
| Latin America | Ombrotrophic lake |
| Latin America | Oxbow Lake |
| Latin America | Palm swamp |
| Latin America | Peat |
| Latin America | Peat bog |
| Latin America | Peatbog |
| Latin America | Playa |
| Latin America | Raised Bog |
| Latin America | Sedge wetland |
| Latin America | Solution Origin Lake |
| Latin America | Swamp |
| Latin America | Tectonic Origin Lake |
| Latin America | Volcanic Origin Lake |
| Latin America | Volcanic basin |
| Latin America | Wetland |
| Latin America | lacustrine |
| Latin America | peatbog |
| Latin America | peatland |
| Latin America | shallow lake |
| Latin America | wetland |
| North America | Bog |
| North America | Caldera Lake |
| North America | Cirque Lake |
| North America | Drained Lake |
| North America | Drift Filled Valley Lake |
| North America | Dune Dammed Lake |
| North America | Explosion Crater Lake |
| North America | Fen |
| North America | Floating Mire |
| North America | Glacial Meltwater Channel Lake |
| North America | Glacial Origin Lake |
| North America | Glacial Outburst Flood Lake |
| North America | Glacial Scour Lake |
| North America | Ice Thrust Lake |
| North America | Interdunal Lake |
| North America | Kettle Lake |
| North America | Lacustrine |
| North America | Lake Marginal Fen |
| North America | Landslide Origin Lake |
| North America | Marsh |
| North America | Mire |
| North America | Moraine Dammed Lake |
| North America | Natural Lake |
| North America | Natural Lake (Origin Unknown) |
| North America | Palustrine |
| North America | Periglacial Origin Lake |
| North America | Playa |
| North America | Poor Fen |
| North America | Raised Bog |
| North America | Soligenous Mire |
| North America | Solution Origin Lake |
| North America | Swamp |
| North America | Tectonic Origin Lake |
| North America | Thermokarst Lake |
| North America | Volcanic Origin Lake |
| North America | Wind Origin Lake |
| Oceania | Caldera Lake |
| Oceania | Cirque Lake |
| Oceania | Glacial Origin Lake |
| Oceania | Mire |
| Oceania | Natural Lake |
| Oceania | Peat sediment |
| Oceania | Solution Origin Lake |
| Oceania | Valley Mire |
| Oceania | Wind Origin Lake |
| Oceania | coastal |
| Oceania | lacustrine |
| Oceania | lacustrine, drained lake |
| Oceania | lacustrine, natural open-water |
| Oceania | lacustrine, playa |
| Oceania | lacustrine, volcanic lake |
| Oceania | terrestrial |
| Oceania | terrestrial, mire (i.e. peatland, >30cm peat) |
| Oceania | terrestrial, mire, bog |
| Oceania | terrestrial, mire, bog, blanket bog |
| Oceania | terrestrial, mire, fen |
| Oceania | terrestrial, mire,swamp (forested wetland or peatland) |

**Table 5.** Chronology control point types in processed data

|  |  |
| --- | --- |
| region | chroncontroltype |
| Asia | AMS 14C |
| Asia | AMS14C |
| Asia | Calendar\_age\_BP |
| Asia | Core top |
| Asia | Core\_top |
| Asia | Guess |
| Asia | Lead-210 |
| Asia | OSL |
| Asia | Radiocarbon |
| Asia | Radiocarbon, average of two or more dates |
| Asia | Radiocarbon, calibrated |
| Asia | Tephra |
| Asia | c14 |
| Asia | cal\_age |
| Asia | calibrated |
| Asia | core\_top |
| Asia | radiocarbon |
| Asia | varve\_age |
| Europe | Annual laminations (varves) |
| Europe | Annual laminations (varves)/Sedimentation rate |
| Europe | Core top |
| Europe | Core top, estimated |
| Europe | Guess |
| Europe | Historical documentation |
| Europe | Lead-210 |
| Europe | Radiocarbon |
| Europe | Radiocarbon, average of two or more dates |
| Europe | Radiocarbon, calibrated |
| Europe | Radiocarbon, reservoir correction |
| Europe | Section top |
| Europe | Tephra |
| Latin America | Collection data |
| Latin America | Collection date |
| Latin America | Core top |
| Latin America | Core top, estimated |
| Latin America | Guess |
| Latin America | Radiocabon |
| Latin America | Radiocarbon |
| Latin America | Radiocarbon, calibrated |
| Latin America | Radiocarbon, reservoir correction |
| Latin America | Tephra |
| Latin America | tephra/Radiocarbon |
| North America | Annual laminations (varves) |
| North America | Core top |
| North America | Core top, estimated |
| North America | Guess |
| North America | Historical fire |
| North America | Lead-210 |
| North America | Radiocarbon |
| North America | Radiocarbon, average of two or more dates |
| North America | Radiocarbon, calibrated |
| North America | Radiocarbon, infinite |
| North America | Radiocarbon, reservoir correction |
| North America | Tephra |
| Oceania | AMS |
| Oceania | C14 date (corrected) |
| Oceania | C14 date (uncorrected) |
| Oceania | Core top |
| Oceania | Radiocarbon |
| Oceania | Radiocarbon, calibrated |
| Oceania | Tephra |
| Oceania | c14 |