Review of Palaeobotany and Palynology 328 (2024) 105156

Contents lists available at ScienceDirect

Review of Palaeobotany and Palynology

journal homepage: www.elsevier.com/locate/revpalbo

A new approach to counting fossil and modern pollen grains: The orderly count

Tutku Tuncalı Yaman [a], Bikem Ekberzade [b], Hülya Caner [c,*], Ruya Y. Dagdeviren [d], Sena Inkaya [e], Nurgul K. Kılıc [d], Suzanne A.G. Leroy [f,g], Fabienne Marret [g], Cetin Senkul [h], Jessie Woodbridge [i], Meral Avci [j]

[a] Marmara University, Faculty of Business Administration, Department of Management Information Systems, Istanbul, Turkey
[b] Istanbul Technical University, Eurasia Institute of Earth Sciences, Climate and Marine Sciences Department, Istanbul, Turkey
[c] Istanbul University, Institute of Marine Sciences and Management, Department of Marine Geology and Geophysics, Istanbul, Turkey
[d] Istanbul University-Cerrahpasa, Faculty of Forestry, Department of Forest Botany, Istanbul, Turkey
[e] Istanbul University, Social Sciences Institute, Geography Programme, Istanbul, Turkey
[f] Aix Marseille Univ, CNRS, Minist. Culture & Com., LAMPEA, 13094 Aix-en-Provence, France
[g] University of Liverpool, School of Environmental Sciences, Department of Geography and Planning, Liverpool, UK
[h] Süleyman Demirel University, Department of Geography, Isparta, Turkey
[i] University of Plymouth, School of Geography, Earth and Environmental Sciences, Plymouth, UK
[j] Istanbul University, Department of Geography, Istanbul, Turkey

ARTICLE INFO

Keywords:
Pollen analysis
Pollen count
Orderly count
Minimum count

ABSTRACT

Palynology, the study of pollen and spores, plays a crucial role in various scientific disciplines, including earth sciences (paleovegetation and paleoclimatology), botany, allergy, archaeology, forensic sciencs and cosmetics. This study delves into the critical question in fossil pollen analysis studies: the minimum count of pollen grains required for accurate estimation of vegetation composition. Various statistical methods have been proposed over the years to address this question. Our research introduces an alternative technique, the orderly count, tailored to the nature of palynological analysis. We apply this method to diverse sediment catchments, including peat bogs, marine and lake sediments, from different geographical locations. Additionally, we revisit the reliability coefficients and propose adjustments for more accurate results. Our findings suggest that relying on statistical methods without considering the specific characteristics of palynological data may lead to low reliability. We advocate for the integration of dissimilarity criteria and the orderly count in sample size assessments for enhanced accuracy in palynological analyses. Our study emphasizes the importance of choosing appropriate methodologies aligned with the unique aspects of palynology to ensure robust and reliable results.

1. Introduction

Palynology is the science of understanding and reconstructing past vegetation by means of proxy data, which in this case it is the pollen grains that are produced by seed plants and that become preserved in sediments from different environments, e.g. lakes, marine and peatlands (Erdtman, 1952, 1954, 1957; Faegri and Iversen, 1989; Moore et al., 1991). Besides various areas of application, the reconstruction of Quaternary vegetation change is a frequent aim of palynologists. Pollen samples collected from lakes, peat bogs, coastal and marine sediments suggest potential ecological interplays between individual plant types, advance and retreat of forests, the structure and floristic composition of vegetation formations for a specific place and through time, which in turn may indicate past changes in climate, human land use and biodiversity (Chevalier et al., 2020; Woodbridge et al., 2021).

The use of quantitative methods in palynology dates back to the 1930s. With roots in geology, at its inception, the science was used as an aide by geologists to determine changes in land cover, reconstruct past climate, and as a dating tool for bog stratigraphy (Birks, 2019). From the 1950s onwards, the proxy nature of pollen grains helped scientists understand spatio-temporal palaeovegetation (past vegetation) dynamics along with anthropogenic effects on land cover (Faegri and Iversen,

* Corresponding author.
E-mail address: hcaner@istanbul.edu.tr (H. Caner).

https://doi.org/10.1016/j.revpalbo.2024.105156
Received 13 June 2024; Received in revised form 1 July 2024; Accepted 3 July 2024
Available online 6 July 2024

Fig. 1. Locations of the cores which were included in the analyses for this study and their respective codes used in plots: **1.** Karataş Lake, Turkey (37°23′08"N – 29°58′18″E), KRTS(1/2/3); **2.** Akgöl Lake, Turkey (41°02′49"N – 30°33′55″E), AKGL (0/40/100); **3.** Tuzla Lake, Turkey (38°50′00"N – 35°20′00″E), TZLL (1/2/3); **4.** Aktaş Lake, Turkey (41°12′54"N – 43°13′01″6E), AKTS (1/42/81); **5.** NE Black Sea (44°13′76.67" - 38°38'05"), NEBS (21/45/133); **6.** Coldhabour Moss, UK (53°26′2.07"N – 1°53′7.74"W), CLDM (1.01 1/1.1 36/1.21108); **7.** Lake Constance, Alp Mountains (47°42′35"N – 9°03′41″E), CNST (4/362/592); **8.** Issyk Lake, Kyrgyzstan (a. 42°20′09.2"N – 76°44′30.8″E, b. 42°25′49.8"N – 76°47′42.7″E, c. 42°17′18.2"N – 77°26′41.3″E), ISSK (14/16/50); **9.** Region of Issyk-Kul, Kyrgyzstan (42°44′23.5"N – 78°13′17.8″E) RIKA; **10.** Region of Issyk-Kul, Kyrgyzstan (42°22′40.6"N – 77°54′59.1″E), RIKB; **11.** Region of Issyk-Kul, Kyrgyzstan (42°18′46.4"N – 76°28′08.2″E), RIKC; **12.** Caspian Sea (45°00′55.93"N – 48°28′17.96″E), CSPS (7,5/22,8).

1964). Faegri and Iversen's book highlighted the ecological importance of palaeovegetation reconstruction through detailed taxonomic identification of individual pollen grains while recognizing the statistical importance of the number of pollen grains counted – referred to as "the pollen count" (Birks and Berglund, 2018).

1.1. *In search of the minimum count*

The typical palynological analysis in a laboratory setting is an arduous and time demanding task and inherits a considerable amount of bias: a trained eye (i.e. experienced in the taxonomy of the pollen in the study area) may identify a larger number of different species whereas an untrained eye may miss or misidentify several taxa. Thus, a minimum pollen count is important in establishing an optimum standard for the study, and contributes in large part to the accuracy of reconstruction by means of proxy (i.e. pollen as a proxy for vegetation). A large number of studies to date attempted to address the question of the minimum count for accurate palynological analyses (Erdtman, 1952, 1954; Faegri and Iversen, 1964, 1989; Moore et al., 1991; Djamali and Cilleros, 2020; Pardoe et al., 2021). The first statistical study on the number of pollen grains that should be counted in a sample was carried out by Fred Alexander Barkley, where he compared equal numbers of pollen grains from different depths in peat samples in central Pennsylvania, using both the *Pearson Correlation Coefficient* with the *Spearman-Brown Prophecy* formula, and concluded that a count of 175 to 200 would be sufficient to attain a 90% accuracy in reconstructing vegetation composition (Barkley, 1934). In the 1970s, Anne P. Bonny published a procedure for obtaining an estimate of fossil pollen per unit sediment volume using a method involving the addition of an estimated number of exotic or marker pollen grains to a measured volume of sediment. With the added markers, Bonny tried to calculate an approximate confidence interval for a pollen count by determining the potential error margin in a count (Bonny, 1972).

Mosimann, Birks and Gordon, proposed different statistical methods to be introduced to palynological analysis including, but not limited to, *confidence intervals, binomial* and *multinomial distribution, difference measurement equations* (Mosimann, 1965; Gordon and Birks, 1974). Applications of multivariate analysis methods (i.e. *cluster analysis* and *principal component analysis*) to pollen data were initiated in the 1970s (Mosimann, 1965; Dale and Walker, 1970; Adam, 1974; Walker and Wilson, 1978; Webb III and Brayson, 1972; Gordon and Birks, 1974). In 2006, Weng made use of *rarefaction analysis* to compare diversity between different-sized pollen samples, an extension of the prior statistical approaches to determine ecological variety by means of diversity indices. While explaining his approach, he highlighted the necessity to minimize the effects of external factors (i.e., the different pollen production rates of plants, dispersal ability of pollen grains and additional environmental factors) for a well-designed methodology (Weng et al., 2006). Most recently, Morteza Djamali and Kevin Cilleros revisited *the Pearson Correlation Coefficient* and *Spearman-Brown* methodology to determine the minimum count level for a reliable pollen analysis (Djamali and Cilleros, 2020).

1.2. *The importance of the minimum count*

Much as the statistical methods applied to pollen count data, the minimum number of pollen grains to be counted matters greatly. Counting a large number of pollen grains in each sample takes a considerable amount of time, but more importantly in certain samples the amount of pollen grains and their diversity at different layers of an

individual core may be different, reflecting temporal changes in vegetation. Therefore, a minimum number to be counted becomes critical for sustaining continuity in cross comparison.

In light of this, Sears suggested that counting up to 100 to 200 pollen grains in each sample would be sufficient for an accurate analysis (Sears, 1930). Erdtman used a minimum count of 150 in his studies (Erdtman, 1931). Lewis and Cocke determined the minimum number to be 800, while Bowman recommended a minimum count of 1000–1800 pollen grains for accuracy (Lewis and Cocke, 1929; Bowman, 1931). In 1946, Iversen published a method where he referred to percentages of taxa in a sample instead of absolute counts, highlighting that for further statistical analysis, it was necessary to count approximately 500 pollen grains per sample (Iversen, 1946; Manten, 1966). Djamali and Cilleros, suggested that for samples from Lake Urmia (NW Iran) that were rich in pollen diversity, a minimum pollen count of 174 and 83 is sufficient to obtain 0.95 and 0.90 reliability coefficients respectively. The authors further proposed, based on a statistical analysis methodology that made use of *Spearman-Brown* with *Pearson Correlation Coefficient*, that counting 67 pollen grains was sufficient for a reliable result, while the number could differ for individual studies (Djamali and Cilleros, 2020).

In this study, we attempted to underline the occasional misuse of certain statistical methods in palynological analyses, while suggesting an additional technique applied to the pollen count, that takes its name from the nature of the analysis itself: *the orderly count*. To show that our method can be applied to palaeo research involving different types of sediment catchments, we conducted our analyses on different core samples from peat bogs, marine and lake sediments and from a number of different geographies (the Anatolian Peninsula, Black Sea, Caspian Sea, Alpine lakes, and mires from northern England). Through our analyses, we also offer an alternative answer to the age-old question: when the reconstruction of floristic composition and structure of vegetation based on dominant pollen types is the main objective, what should be the minimum number of pollen grains counted?

2. Methodology

After a sample is taken from a sediment core for pollen analysis, numerous chemical processes are applied before the sample is mounted on the slide, with the grains ready to be identified and counted under a light microscope. All inferences made by means of the final count data and inferential statistical methods that are applied during the analysis are based on the sampling theory assumption that the sample should be randomly selected from the population (Alf and Lohr, 2007). However, in the case of pollen grains, due to their nature, the assumption of randomness may not always apply. Pollen grains are small in size (10–100 μm) and since specific gravities greater than unity when wet follow *Stokes' Law of Resistance*, the morphologically identical grains tend to attract each other (Brush and Brush, 1994). Thus, the assumption that the sample taken from the core will contain all pollen species randomly – for a given time period, for a given area – may not always be accurate.

Thus, to increase the number of samples to run our tests on, and to see whether or not any depth within the core, or the source of the sample (lake, bog or marine) would have a differentiating affect in our analyses, where possible, count data were obtained from different depths in a sample core. Fig. 1 shows the location of the study sites, their respective coordinates and reference names with a number designating core depth as they appear in the plots.

2.1. Study area and sample selection

The pollen samples considered in our sets of analyses are from sample cores collected for different studies from diverse ecosystems: lakes in the Anatolian Peninsula, Eastern Mediterranean Basin; a peat bog from the Peak District National Park, UK; northeast Black Sea; the Caspian Sea; the shores of Lake Issyk-Kul, Krygyzstan; and Gnadensee in Lake Constance, German Alps. Lake sediments and peat bogs depleted of oxygen are favored among palynologists for spatio-temporal analysis of changes in vegetation composition, as they provide better preservation of the organic deposition – including pollen grains – and thus a more detailed temporal record of past vegetation (Moore et al., 1991). Marine cores offer the possibility to study long-term vegetation changes at a regional scale, usually in a multidisciplinary approach (biogeochemistry, sedimentology, etc.). They can also be the only archives of palaeoenvironmental conditions available where there are no appropriate sites to study (see Hooghiemstra et al., 2006). In general, preservation of pollen grains is good in marine sediments, as they settle rapidly in the water column. On the other hand, diversity is often lower than on the continent due to factors such as pollen production and transport (Holmes, 1994).

By applying our method to samples taken from different geographic regions and accumulation areas with varying dynamics, we tested the similarities and differences of samples obtained from diverse environments. The Anatolian Peninsula is remarkably rich and diverse in terms of vegetation. It encompasses various climatic regions from east to west and from north to south, each characterized by distinct types of vegetation. In the north, plant species typical of the Euro-Siberian region can be found (Fig. 1, sample 4; Karlıoğlu Kılıç et al., 2018), while Central (Fig. 1, sample 3; Şenkul et al., 2018), Eastern, and Southeastern Anatolia are characterized by species from the Irano-Turanian region. In the south, Mediterranean vegetation predominates (Fig. 1, sample 1; Şenkul and Kalıpçı, 2019). The Marmara region is a cross-over climate zone with vegetation showing characteristics of the Mediterranean and the Black Sea regions (Fig. 1, sample 2; Karlıoğlu Kılıç et al., 2020). A marshland area in England, characterized by the prevalence of aquatic plants, constitutes another of our study sites (Fig. 1, sample 6). In another of our sample sites, the area around Lake Issyk-Kul, reeds, rushes, and various aquatic plants are commonly found (Fig. 1, samples 8–11). In another of our sample sites, the Caspian Sea, halophytic plants are present (Fig. 1, sample 12). In the German Alps near Lake Constance, Gnadensee is another of our sample sites characterized by alpine and subalpine vegetation (Fig. 1, sample 7). The northern Black Sea sample (Fig. 1, sample 5) is a marine core, and includes information on both the transformation of the region's now Colchic forests, and dinoflagellate cyst assemblages (Marret et al., 2019).

2.2. Statistical measures and ecological indices used in this study

While the use of quantitative methods and statistics to determine an optimum pollen count dates back to the 1930s, different approaches are recommended by different researchers for the same problem with occasional inconsistencies between the theory behind and the application of their proposed technique: i.e., the debate in determining the distribution of pollen samples (Mosimann, 1965; Birks and Gordon, 1984; Overpeck et al., 1985; Maher, 1972; Birks and Birks, 1980; Keen et al., 2014). Furthermore, although the fundamental theory of the frequently used *Pearson Product–Moment Correlation Coefficient* relies on the assumption of the provision of a normal distribution within the observed data, the coefficient is often used to determine an association between pollen samples whose distribution is not known (Barkley, 1934; Djamali and Cilleros, 2020) or is likely not normally distributed. We summarize the statistical theory behind different *Measures of Association* that are applied to assess sample reliability which we used in our comparative analyses in detail in Appendix A.

In this study, we use the terms reliability and validity with reference to their definitions in measurement theory. In measurement theory, reliability refers to the consistency of a measurement across repeated assessments. Reliable measures have minimal random errors. One of the methods for evaluating reliability is the Split-Halves test. Validity, on the other hand, pertains to the accuracy of a measurement in capturing what it is intended to measure (Nelson Jr., 1980).

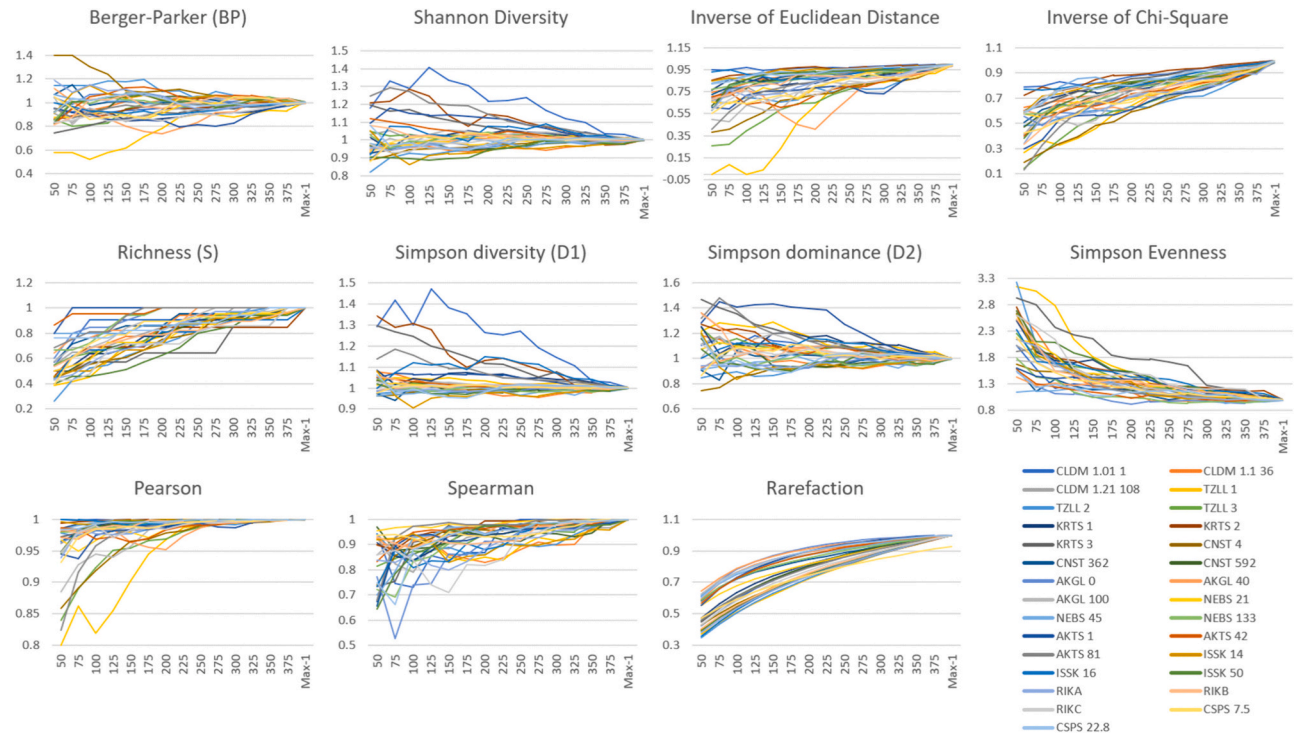

Fig. 2. Changes in the indices shown on the y-axis, with respect to sample (random count) size. Sample names are followed by a number designating sample level in the core, their locations detailed in Fig. 1. In the case of inverse of *Euclidean Distance*, since proportional frequency is used, we adjusted our calculations of the count data accordingly.

Corelation matrix	Pearson's R	Spearman Correlation	Chi-Square	Inverse of Chi-Square	Entropy	Richness (S)	Simpson diversity (D1)	Simpson dominance (D2)	Berger-Parker (BP)	Simpson Evenness	Rarefaction
Pearson's R	1.000										
Spearman Correlation	0.097	1.000									
Chi-Square	-0.704	-0.621	1.000								
Inverse of Chi-Square	0.704	0.621	-1.000	1.000							
Entropy	-0.695	-0.014	0.429	-0.429	1.000						
Richness (S)	-0.412	0.198	0.100	-0.100	0.735	1.000					
Simpson diversity (D1)	-0.660	-0.038	0.440	-0.440	0.972	0.630	1.000				
Simpson dominance (D2)	-0.824	0.079	0.415	-0.415	0.908	0.691	0.853	1.000			
Berger-Parker (BP)	0.733	-0.031	-0.430	0.430	-0.958	-0.658	-0.971	-0.925	1.000		
Simpson Evenness	-0.736	-0.108	0.484	-0.484	0.603	0.010	0.650	0.701	-0.697	1.000	
Rarefaction	-0.403	0.196	0.099	-0.099	0.730	0.999	0.625	0.681	-0.650	0.001	1.000

Fig. 3. The Correlation Matrix of indices calculated with the maximum random count data of 400, from 24 sample sets with *Pearson Correlation Coefficient.*

2.3. *The orderly count*

Weng et al. (2006) stated that while evaluating palynological richness where the number of pollen counts is limited only to a few hundred in the majority of samples, all taxa for the study region may not be identified during the counting process. However, when the samples are counted "sequentially", the number of taxa identified in any given sample increases with larger sample size, assuming the sample is randomly selected The *orderly count* approach builds up on this idea. It requires the analyst to identify and record pollen grains sequentially (as counting progresses) without tallying by taxa, whereas in the classical approach the sample is randomly tallied as count progresses (from now on called the random count).A non-statistical approach is illustrated in Pardoe et al. (2021), where additional taxa are recorded line by line, allowing visualization of the increase in diversity obtained by additional counts. Here in the orderly count, each taxon is recorded one after the other as counting progresses.

Therefore, in the data preparation process of the reliability coefficient calculation, instead of dividing the total count data into two equal parts randomly, we first gave a sequence number to the pollen taxa during the count of the entire sample – i.e. if the first pollen type spotted in the sample belonged to Quercus L. that became number 1, the second being another Quercus L. pollen (or any other taxon) that became number 2, etc. For the random list representing the classical methodology, the *orderly count* list was randomized using the Excel function = *RAND()* which assigned a random number between 0 and 1 to the records on the count list and then relisted the count data in ascending order. We split the total lists into two halves and performed our calculations on these two sequential samples: while calculating the relationship measures to determine the reliability of a sample size of 400, the first 200 pollen grains from numbers 1–200 were added to the first group and the second 200 to the second group. To test the effects of an orderly vs a random count we used *Spearman-Brown* along with a similarity measure (inverse of *Chi-square Distance*) in addition to its native *Pearson Correlation*, and we tested our results with the *Split-Half* approach to further verify sample adequacy and reliability. Our results from different reliability measures (which are detailed in Appendix A) that we applied to samples which were split randomly vs orderly are explained in detail in the Results and Discussion section below. The minimum (for the purposes of this study also the maximum) count number of 400 per sample was a mutual decision by the authors of this study: an even number with which statistical analyses could conveniently be run on the

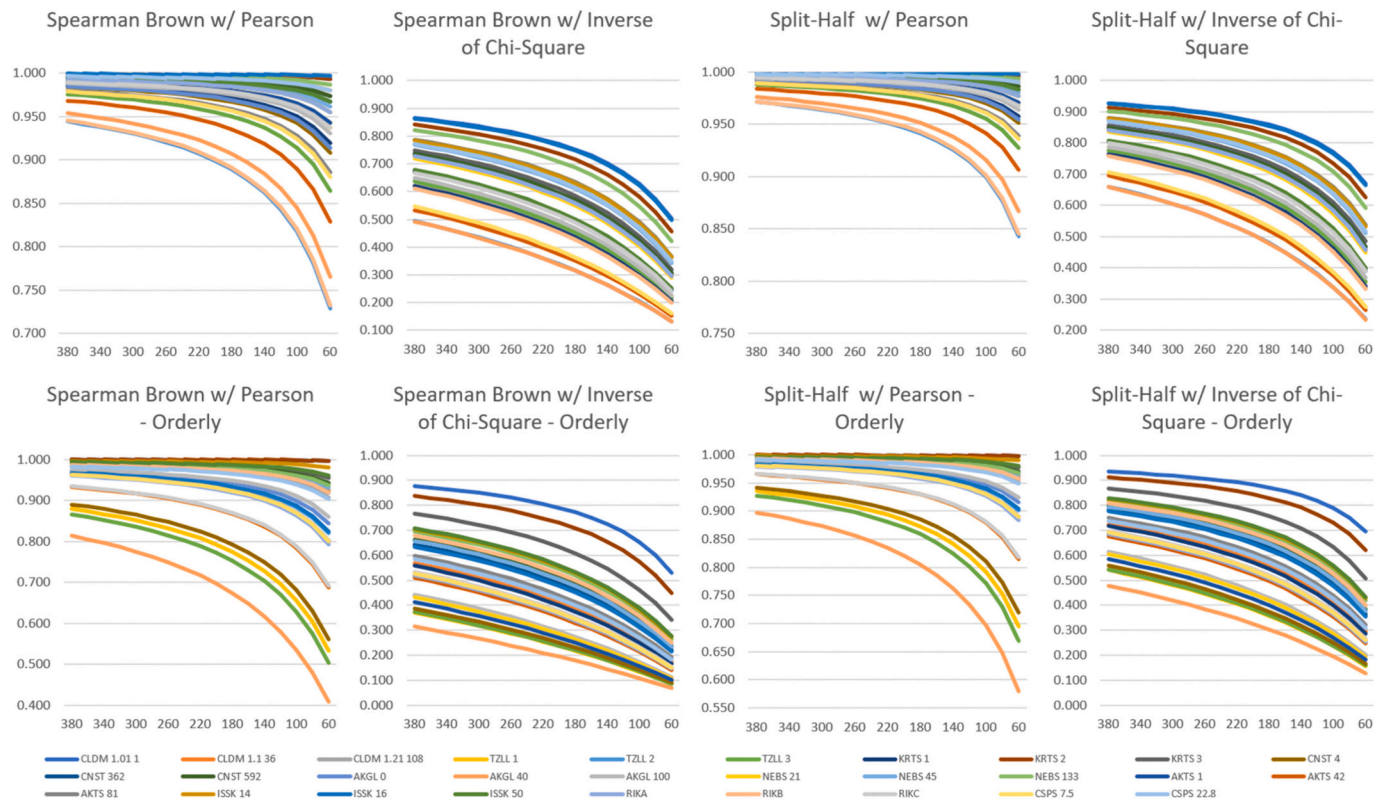

Fig. 4. *Spearman-Brown Prophecy* and Split-Half calculated with two different association coefficients, *Pearson Correlation* and the inverse of *Chi-square Distance* respectively. The y-axis shows reliability whereas the x axis shows the sample size. Lake names are followed by a number designating sample level in the core from 1 to 3: 3 being the deepest sample. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

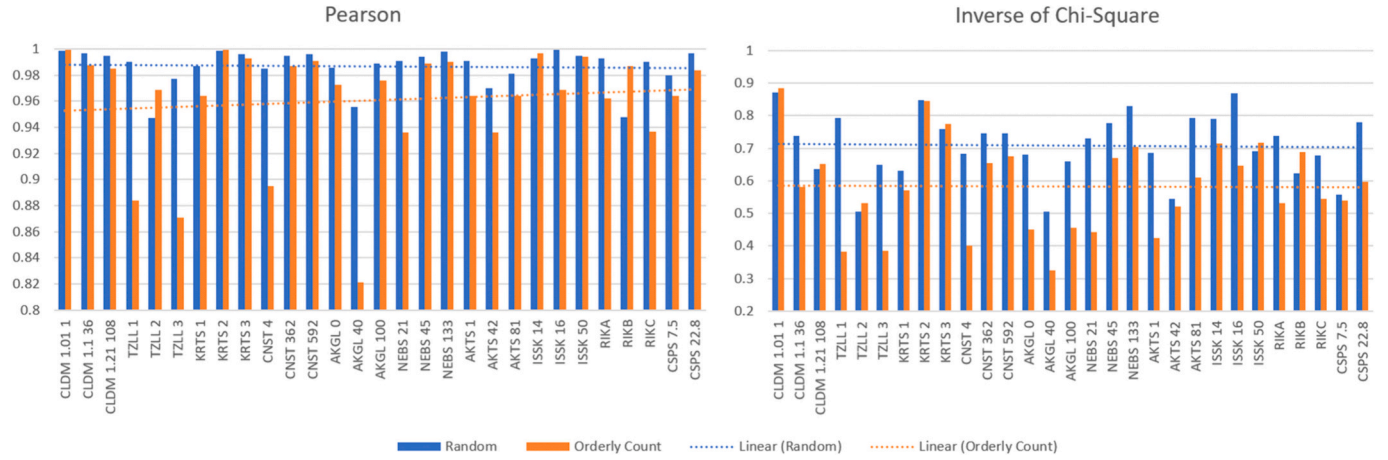

Fig. 5. Sample reliability coefficients (y-axis denotes index values) with Pearson Correlation and inverse of Chi-square values for the two halves of optimum count data (400 for this study) for individual samples.

split halves; high enough to give sufficient samples for comparative analyses; and a level within the range of a large number of palynological studies (Erdtman, 1952; Faegri and Iversen, 1964; Bonny, 1972; Birks and Birks, 1980; Prentice, 1982a, 1982b; Moore et al., 1991; Reille, 1992; Djamali and Cilleros, 2020). Once a count of 400 was reached per sample, the count data list was reworked, for the *orderly count* erasing the final counts by increments of 25 (leaving sets of 375, 350, and so on) until there was a count of 50 left. In the random count list, increments of 25 were erased starting from the top of this randomized list, resulting in an equal number of samples as in the orderly count. In the case of the orderly count, erasing increments of 25 from the end of the list reflects the real-life practice of using smaller numbers as a minimum count. For the randomized list, it makes no difference from which end we shorten our list (Fig. 2).

Based on the frequency distribution of these 15 different sample sizes and 24 different pollen count data sets, as well as the sample with the optimum count of 400 for this study, *Pearson's Correlation, Spearman's Correlation, Euclidean* and *Chi-square distances*, *Richness (S), Shannon's Entropy (H'), Simpson diversity (D1), Simpson dominance (D2), Berger-Parker (BP)* indices, as well as *(Simpson) Evenness* and *Rarefaction Coefficients* were calculated. All results gathered from all samples were indexed respectively. This allowed us to see how the indices changed along with sample size, mainly when different sample sizes were compared with the base index from the optimum count data of 400

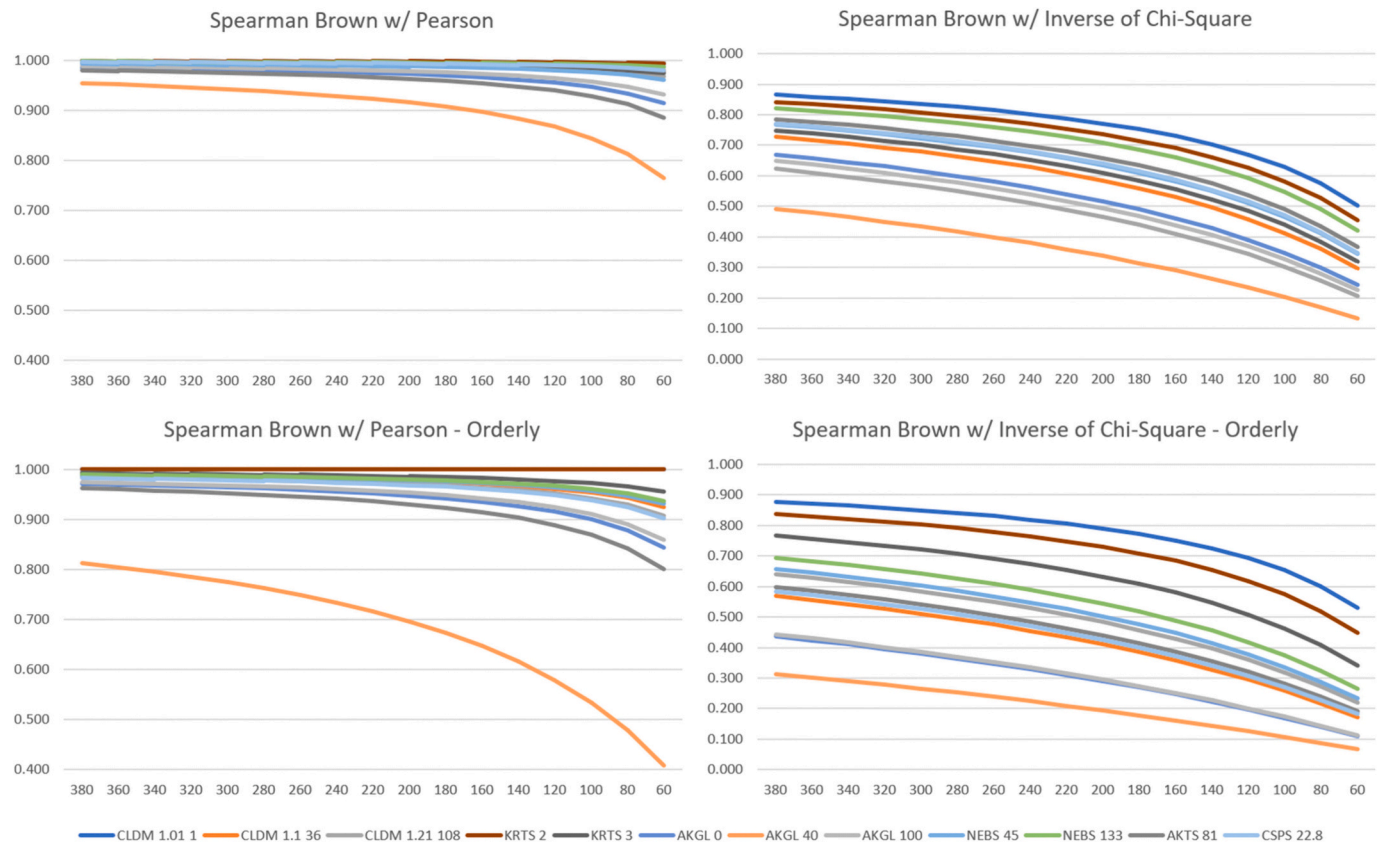

Fig. 6. *Spearman-Brown Prophecy* calculated for the samples with low richness index with two different association coefficients, *Pearson Correlation* and the inverse of *Chi Square Distance* respectively. The y-axis shows reliability whereas the x axis shows the sample size. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Fig. 3).

3. Results and discussion

In palynology, the pollen sample is a representation of the vegetation composition in a given region, over a certain time period, and a representative sample is one that reflects the population's characteristics. Characteristics of the population can be described through its natural diversity, and representativeness can be evaluated by comparing different measures of the diversity of the samples against the population. Thus, its accurate interpretation is of outmost importance, as it not only gives us an idea of past vegetation distributions, but also, by means of proxy, helps us interpret past changes in vegetation diversity, land-use change, and climate (Pardoe et al., 2021; Cleal et al., 2021).

While making interpretations, the researcher needs to have all the available data, and thus the minimum number for a pollen count becomes extremely important: it needs to be within a range where all key vegetation types are observed. Since regions may differ in richness of taxa, accepting a single number for the minimum count becomes risky, and flexibility needs to be observed as different count levels may be necessary for regions with higher floristic diversity vs those where the samples do not exhibit similar levels of diversity.

Fig. 2 shows the results from the random count data, in which, despite our samples representing different vegetation compositions from different geographies with different fossil pollen reservoirs and different accumulation rates, convergence started to occur in sample sizes of 300 and above, except for *Pearson's Correlation.* When we revisited the use of the reliability coefficient (r_p) in the *Spearman-Brown* approach with the *Pearson Correlation*, it became apparent that using the *Prophecy Coefficient of Reliability* with an appropriate similarity measure increased accuracy through a more sensitive detection of reliability than the *Pearson Correlation.* This upheld our prior arguments that (1) *Pearson Correlation* is not a statistically appropriate method for this type of data, and (2) specifically for the field of palynology, to assess the reliability of the size of the sample, an appropriate *Measure of Association* should replace *Pearson Correlation* if *Spearman-Brown's Prophecy* is the method of choice.

Our main criticism in using *Spearman-Brown's Prophecy* with *Pearson Correlation* formula (Appendix A: Eq. 11 – an adaptation of the *Pearson Product Moment Correlation* (r_{ij}) in Eq. 1) to count data – first by Barkley, and later by his successors – is that, as aforementioned, *Pearson Correlation* provides a correlation coefficient for data types measured at interval or ratio levels and upholds an assumption that the variables have a normal distribution (Allen and Yen, 2002). In palynology, the count data are derived from a limited sample, one which has been sampled from a sediment core. In theory, the range of a count variable lies between 0 and infinity, therefore one should be able to display these count data as a one-way frequency distribution, similarly to any kind of nominal variable. However, in practice, palynological data are obtained from a limited sample, the range of which varies between samples, and the variability of which is dependent on the sample size. Even though, in theory, it is commonly accepted that the distribution of count data types is multinomial, we cannot and should not make that assumption in the case of a limited pollen sample from a sediment core. Also, per theory, *Poisson Distribution* serves as a robust alternative to model count data (Tang et al., 2012; Addison-Smith et al., 2020).

Thus, we recommend that in the operationalization of the *Spearman-Brown* approach, first an appropriate association measure needs to be selected between two pollen count datasets, obtained by dividing the original pollen sample into two halves. This leaves us with two one-way frequency distributions of taxa where x represents the first half and y represents the second half. If we suppose x has s, and y has r different

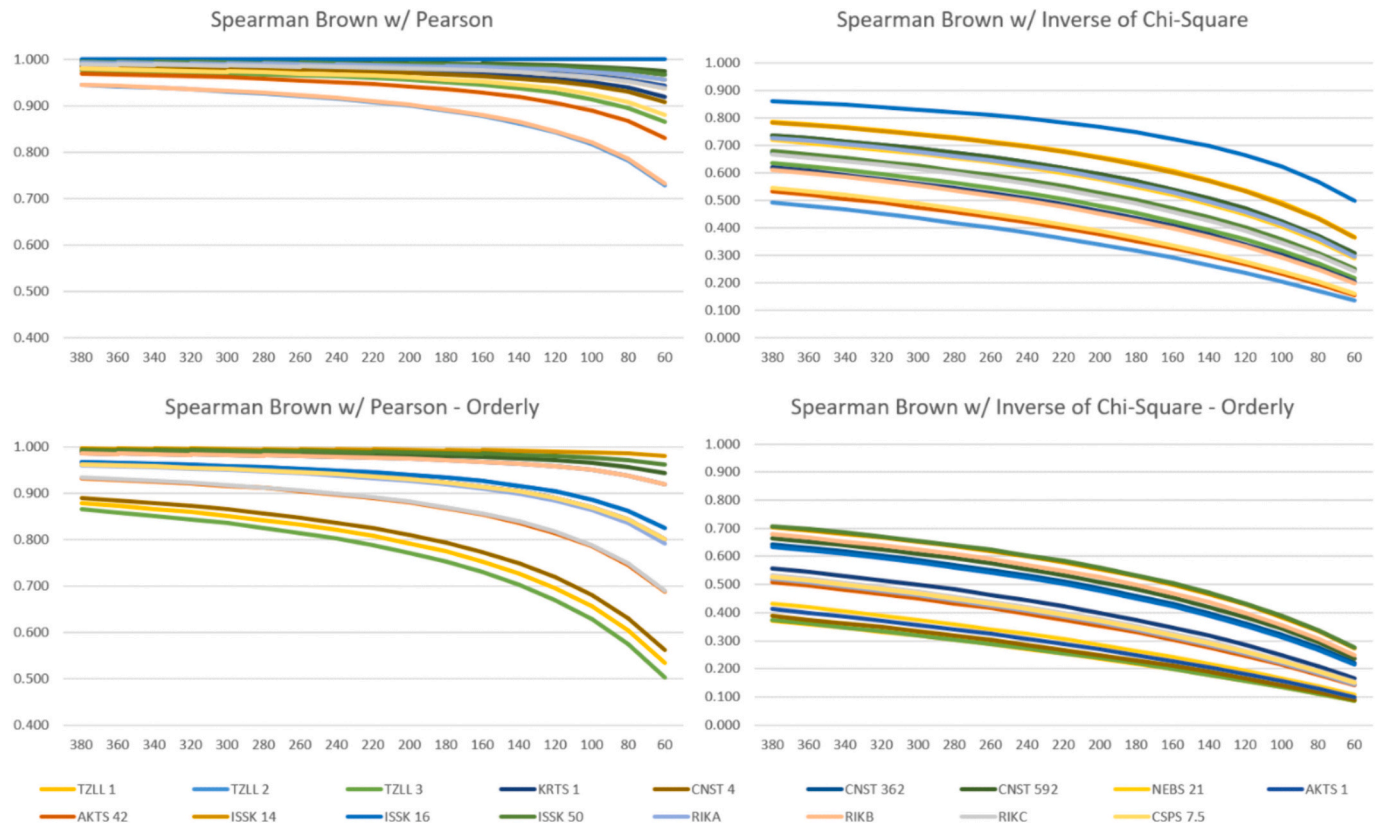

Fig. 7. *Spearman-Brown Prophecy* calculated for only the samples with high richness index with two different association coefficients, *Pearson Correlation* and the inverse of *Chi-square Distance* respectively. The y-axis shows reliability whereas the x axis shows the sample size. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

levels (also note that neither the total number of taxa nor the identified pollen types in the two halves need to be identical), the outcomes of pairs can be represented in an $s \times r$ contingency table. Now, the *Pearson's Chi-Square Statistic* will be ideal for testing the general association of row and column variables (Tang et al., 2012).

Fig. 3 shows the correlation matrix of indices calculated by *Pearson Correlation Coefficient* from the 29 samples with a maximum count data of 400. This matrix validates our use of the *Inverse of Chi-square Distance* (a similarity measure which does not rely on a specific distribution) as an appropriate *Measure of Association* for this type of count data if *Spearman-Brown* is the chosen method, and shows that the theoretical basis for each measure has been provided for.

When *Spearman-Brown* and *Split-Half* methodologies (as explained in Appendix A) were applied with both *Pearson Correlation* and our proposed inverse of *Chi-square Distance* as a measure of association to both the random count and the orderly count lists, a noticeable difference in the level of reliability was observed, even at our maximum sample size of the 400-grain count (Fig. 4). When a sample of 400 is considered and *Spearman-Brown Prophecy* is calculated with the inverse of *Chi-square Distance* with the random count list, a cut of 20 in count number indicated a decrease in the confidence level within a range of 15–55%. When *Pearson Correlation* was applied, this decrease was by 5% at most. As a means of verification when *Split-Half* was calculated using these two different association coefficients, a similar pattern emerged. Alternately, when the same methodologies were compared on the orderly count data, a significant decrease in reliability coefficients from both techniques were observed. This was indeed a carry-over effect of the change, depending on whether the *Pearson Correlation* or the inverse of *Chi-square Distance* was used on the *orderly count* data, as we explain further below.

The Spearman-Brown method inherently assumes random sampling, which is the basis of parametric methods in statistics. However, while sampling for palynological analysis, this randomness hardly applies due to a series of constraints from the morphological qualities of the pollen grains to the nature of the science itself – which requires that the grains in a sample are counted in an orderly manner, although this order is not reflected in the final count record. Thus, we initially suggested as a correction to formulate *Spearman-Brown* with a *Measure of Association* (inverse of *Chi-square Distance*), as opposed to its native *Pearson Correlation*; and that integrating the *orderly count* into future methodologies may avert this statistical method's weakness when evaluating pollen count data for reliability. Orderly count is in fact in tune with the nature of the pollen count: in the classical method the taxa are listed as they are observed and then strikes are added next to them as more of them are seen in the sample, whereas in the orderly count, the taxa are listed by their assigned numbers per their order of observation, the numbers extending from 1 to the chosen optimum count level (400 in our case).

When our lists of count data (both orderly count and random count) that were split into two halves to further assess the sample's reliability with a statistical method such as *Spearman-Brown* using *Pearson Correlation*, the coefficients changed in a much more conservative manner from each other and were higher when the data had been randomized; as opposed to the significantly lower index values observed in the orderly count (Fig. 5). This was also in support of our suggestion that random sampling cannot be assumed for palynological samples. Our orderly count data additionally showed that when the samples are split into half, the two half-samples are significantly dissimilar when inverse of *Chi-square* is applied instead of *Pearson Correlation*. This result also highlighted the importance of suggesting and/or selecting the correct range for a minimum count for maximum representativeness of a sample: not necessarily an absolute number, but a range which needs to be appropriately high enough for accurate construction of pollen assemblage

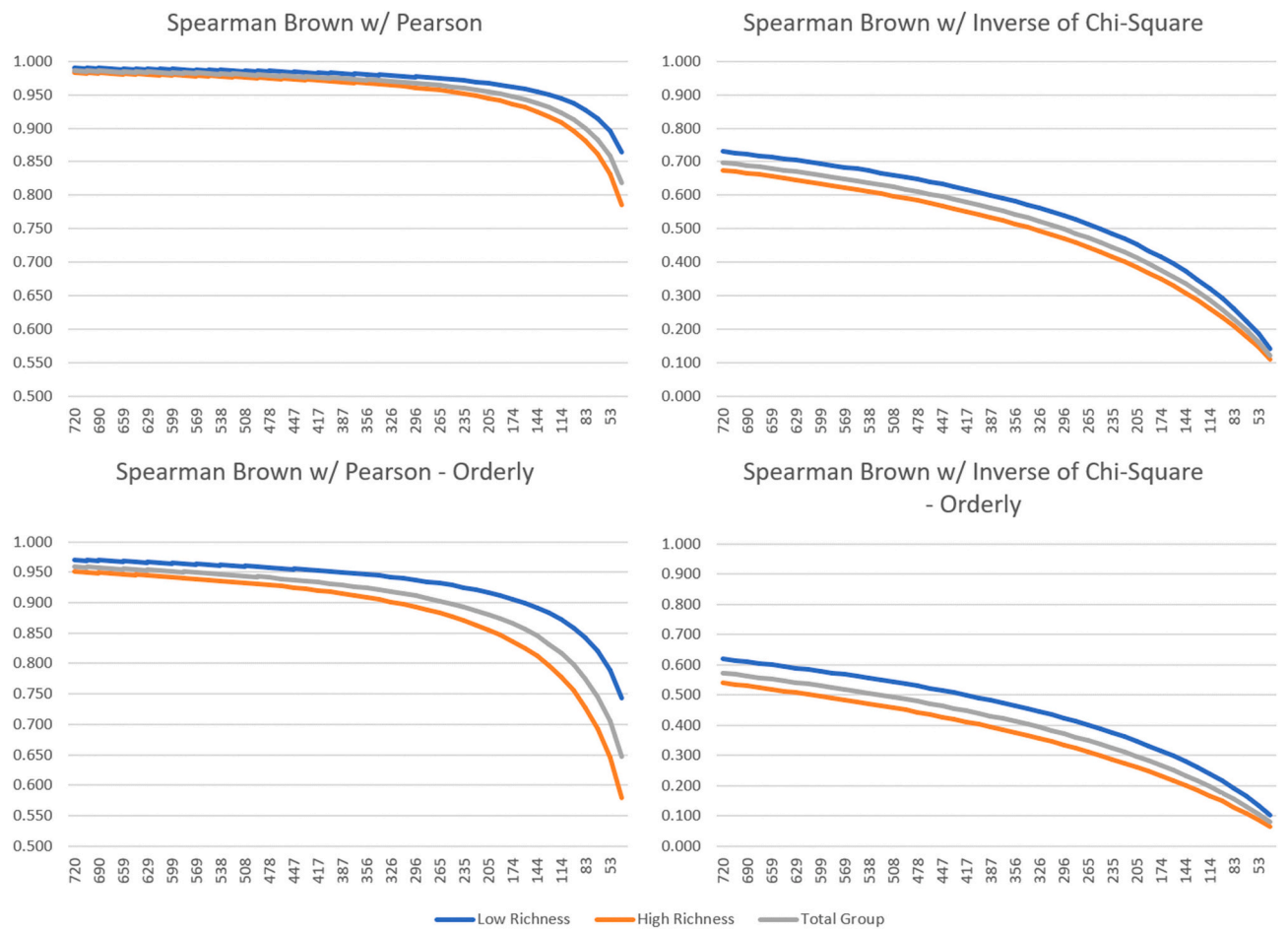

Fig. 8. *Spearman-Brown Prophecy* calculated for all sample groups, including the two sub-samples taking the average of the reliability coefficients calculated with two different association coefficients, *Pearson Correlation* and the inverse of *Chi-square Distance* respectively. The y-axis shows reliability whereas the x axis shows the sample size. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

composition.

To see if we could confidently suggest a minimum count range, we further analyzed our simulation results, looking for changes in reliability levels if the samples exhibited different levels of diversity. As is stated in several studies (Magurran, 1988; Weng et al., 2006), sample size has a direct effect on the number of taxa identified in a sample. However, in palynology, the size of the sample is by and large limited (Moore et al., 1991). *Richness* (explained in Appendix A) as an indicator of diversity helps the researcher easily identify whether they are looking at a sample representing a high diversity in vegetation composition, or vice-versa. To test if our proposed methodology can be applied to samples with different levels of diversity, we first calculated the *Richness* (*S*) index values for the samples, and then divided the entire sample base into two halves using the median *S* value (21 in our case) as a threshold. We then compared both the random and the orderly count groups from these two sample pools of high vs low *S* using *Spearman-Brown* with *Pearson Correlation* and alternately with the inverse of *Chi-square Distance*. Comparison of the results from these groups rendered results in agreement with the above-mentioned findings that samples exhibiting a high diversity of taxa require a higher sample size for accurate detection of taxa. This, from a palynological perspective can be interpreted as: a higher (minimum) count level in turn renders more accurate reconstructions of paleovegetation. Fig. 6 shows *Spearman-Brown Prophecy* results for the group of 12 samples with low richness index, and Fig. 7 shows the second group of 17 samples exhibiting a high richness index.

Fig. 8 shows the average of the reliability coefficients for all sample sub-groups (high *S*, low *S*) and total samples from all core-depths. In this final plot, it becomes clearer that when the reliability coefficients are calculated using *Pearson Correlation Coefficient* a reliability of around 80% can quickly be reached at relatively small count numbers, and in all groups. The reliability level drops to around 20% for a similar range of minimum count for the random count data when the reliability coefficient is calculated using the inverse of *Chi-square Distance,* and drops further below 20% when the same technique is used with the *orderly count*. Higher reliability levels – over 70% – are reached only at count numbers above 700 for the random count data, and at much higher minimum count numbers by using *orderly count* (i.e., for a reliability coefficient of 70%, more than 1000 pollen grains need to be counted). The reliability for the count numbers higher than 400 were calculated by inversing the Spearman-Brown formula (as shown in Appendix A); increasing reliability (instead of the count) and testing for higher count levels to see where convergence would occur.

4. Conclusion

Using statistical methods with little to no adjustments for distributions, and/or data types that do not necessarily meet the criteria behind the approach – i.e., parametric methods with the inherent assumption that the data should follow a certain distribution, vs non-parametric methods where the distribution is not a concern – is a common pitfall in quantitative scientific analysis that makes use of statistics. In this study, we show how such automated applications of statistical approaches may render results with low reliability, whereas adjustments reflecting the nature of the work may significantly improve results'

accuracy – i.e. ascertaining sample representativeness.

Thus, in addition to integrating a dissimilarity criterion into the assessment of reliability of sample sizes if *Spearman-Brown Prophecy* is the chosen method for comparison in the palynological analysis, we recommend that the potential effect of the *orderly count* should be considered. Our results also support this approach, which holds true for fossil pollen grains from different sediment catchment types and throughout our geographically disparate study points. The nature of the science provides that after a certain number of pollen grains are counted, additional counts of the sample does not provide significant additional information about the sample's representativeness of the population, an assumption also supported by our results: this threshold acts as the optimum (minimum) count.

When we are trying to ascertain the reliability, and with that the representativeness of the sample for the whole population, if we use the *orderly count* as our chosen technique we will drive at more accurate results, albeit at higher minimum count numbers. This also holds true in our analyses, where samples counted in an orderly manner render smaller reliability ratios than those that are counted randomly (classical method).

Additionally, if we are attempting to assess the representativeness of a sample in palynological interpretation, then we should keep in mind that our sample in question might not be random. This is partially due to the nature of the pollen counting activity and the tendency of pollen grains from the same taxa to clump together (Brush and Brush, 1994). All this strengthens our argument that a palynological sample is not randomly selected, which in turn makes the application of statistical methods which assume random sampling inappropriate in its analysis. However, if such a method is used, then as an amelioration, it can be adjusted to reflect the nature of the study, in our case using a dissimilarity measure along with our own methodology: *the orderly count*.

Author Contributions

T. Tuncalı Yaman developed the orderly count technique, ran the statistical analysis, prepared the plots, wrote the methodology; B. Ekberzade wrote the article, and edited the methodology; H. Caner edited and wrote the references; M. Avci prepared Fig. 1, the vegetation map, and edited the references; TTY, BE, HC and MA, within their fields of expertise contributed to the overall design of the article, participated in all revisions; RYD and SI conducted an initial literature review; HC, RYD, NKK, and CS participated in the initial discussions for the article, provided pollen data from previous studies and recounted their data according to the methodology of the orderly count; JW, SAGL, and FM provided and recounted the pollen data from previous studies using the orderly count, revised the manuscript critically and improved the final version for publication. BE, FM, HC, JW, MA, SAGL made intellectual contributions before submission. All authors have read and approved the final version of the manuscript.

Declaration of competing interest

The authors have no relevant financial or non-financial interests to disclose.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank Nüzhet Dalfes and Sedat Avcı for their kind input and critical feedback. In this study, some of the reference preparations used in the sequential counts were obtained from our project titled "Pollen Atlas of Important Woody Plants of Istanbul." Therefore, the authors would like to thank the Scientific Research Projects Coordination Unit of Istanbul University for supporting the aforementioned project. This study was funded by Scientific Research Projects Coordination Unit of Istanbul University. Project number: FBA-2020-36100. Bikem Ekberzade was funded under ITU-BAP grant number MDK-2023-44546, and TUBITAK Grant 123O248. The financial support received from TUBITAK does not mean that the content of the publication is approved in a scientific sense by TUBITAK.

Appendix A. Appendix

A.1. Measures of association

A common application of *Measures of Association* in statistics is one in which factors or coefficients are used to quantify a relationship between two or more variables (Haug, 2019). In palynology, *Measures of Association* have been used (see Birks, 2019) to assess the similarity of two or more pollen samples, which are essentially datasets comprised of observed frequencies of different taxa.

A.1.1. Pearson product–moment correlation coefficient

Pearson Product–Moment Correlation Coefficient, which is commonly used as an index for the linear relationship between two variables, is indeed an estimate of the product–moment correlation where the respective parameters are substituted by the "sample moments". The underlying assumption of *Pearson Correlation Coefficient* is that the sample being examined has been randomly drawn from a population with a normal distribution. An indiscriminate use of this method in attributing statistical significance to calculated coefficients is problematic since it will render inaccurate results in cases where the conditions of normal distribution are not met and/or where a curved or curvilinear relationship occurs between the two variables for which the relationship is attempting to be determined (Allen and Yen, 2002; Tang et al., 2012).

The formula for *Pearson Correlation Coefficient* is provided below in Eq. 1 (Prentice, 1980), where r_{ij} represents the correlation coefficient between two pollen spectra, i and j, p_{ik} is the proportion of pollen type k in pollen spectrum i, p_{jk} is the proportion of pollen type k in pollen spectrum j, $\overline{p}_i$ indicates mean pollen proportion for pollen spectrum i ($\overline{p}_i = \frac{\sum_k p_{ik}}{n}$), and n is the total number of pollen taxa. The notations are consistent in all formulas below:

$$r_{ij} = \frac{\sum_k (p_{ik} - \overline{p}_i)\left(p_{jk} - \overline{p}_j\right)}{\left[\sum_k (p_{ik} - \overline{p}_i)\right]^{1/2}\left[\sum_k \left(p_{jk} - \overline{p}_j\right)\right]^{1/2}} \quad (1)$$

A.1.2. Spearman rank-order correlation coefficient

The non-parametric *Spearman Rank-order Correlation* is used to measure the degree of association between two multi-step or ordinal level measured variables, where the calculated correlations between two ordered data series and the significance of correlation coefficients are used to test whether there is an association or not (Siegel and Castellan, 1988).

Spearman Rank-order Correlation Coefficient can be calculated by the following formula in Eq. 2, where x_{ik} is rank number for pollen type k in pollen spectrum i and x_{jk} is the rank number for pollen type k in pollen spectrum j (Keen et al., 2014).

$$r_{ij} = 1 - \frac{6\sum_{k}\left(x_{ik} - x_{jk}\right)}{n^3 - n} \tag{2}$$

Although this coefficient is inherently not an appropriate tool for some data types, such as the count data in palynology (although it has been used in previous analyses of pollen spectra, see Keen et al., 2014), we include this approach to test the effectiveness of our evaluation.

A.2. Measures of dissimilarity

In statistics, dissimilarity measures refer to a set of metrics that represent the distance between data points, and how these samples are related or how close they are to each other. In palynology, their application in previous studies has been similar to *Measures of Association* (Prentice, 1980, 1982a, 1982b; Overpeck et al., 1985).

A.2.1. Euclidean distance

Euclidean Distance is a numerical measure of how different two data objects are in a range of $[0, \infty)$. The concept behind this measure is to find the distance between two points in a *Euclidean Space* by Cartesian coordinates of those points using the Pythagorean theorem (Gower, 1985). Proximity to 0 indicates two objects that are alike and vice versa, i.e. $d_{ii} = 0$, $d_{ij} > 0$ (Gordon, 1999). The formula is provided below in Eq. 3 where d_{ij} represents the similarity between two pollen spectra, i and j (Birks and Berglund, 1979).

$$d_{ij} = \sqrt{\sum_{k}\left(p_{ik} - p_{jk}\right)^2} \tag{3}$$

A.2.2. Chi-square distance

According to *The Concise Encyclopedia of Statistics*, the *Euclidean Distance* between the frequencies of objects is called *the Chi-square Distance,* which is a dissimilarity measure that is often used in palynology (Dodge, 2008; Birks and Deacon, 1973). Its use is mainly proposed in the context of *Principle Component Analysis* as a relative distance to the total abundance within each pollen dataset (Heegaard et al., 2006).

$$d_{ij} = \sqrt{\sum_{k}\frac{\left(p_{ik} - p_{jk}\right)^2}{p_{ik} + p_{jk}}} \tag{4}$$

Because our calculations for this study will be based on absolute frequencies, *Chi-square Distance* is an appropriate approach within our context.

A.3. Ecological diversity measures

Diversity indices that compare different regions, taxa and trophic levels are frequently used to describe general properties of communities (Morris et al., 2014; Li et al., 2018; Cursach et al., 2020). Like ecologists, palynologists have been using these indices as a measure of ecological diversity (Keen et al., 2014). One of our main reasons for considering diversity indices in our study is because the compatibility of the distribution of the pollen sample to the overall population (representativeness) is one of the foundational assumptions of sampling theory. As such, monitoring variations in diversity indices for samples of different sizes can be considered as an indicator of sample size adequacy.

A.3.1. Richness (S)

As a simple and popular approach to represent diversity in a certain habitat, richness is defined as the total number of species or attributes present (Whittaker, 1972).

A.3.2. Berger–Parker (BP) dominance

Since species abundance and dominance are also critical to understanding diversity, Wolfgang H. Berger and Frances L. Parker proposed to use the maximum proportional abundance value to represent diversity (Berger and Parker, 1970). The approach is represented in Eq. 5, where P indicates the proportion of a certain pollen type in a pollen spectrum.

$$BP = P_{max} \tag{5}$$

A.3.3. Shannon Diversity Index (alternately, Shannon-Wiener Index)

The foundations of Shannon's Diversity Index lie in statistical physics (Boltzmann, 1872), modified a century later in information theory by Claude Shannon (Shannon, 1948). When the method is adapted to ecology as an indicator of diversity, interpretation of the index focuses on the uncertainty level of the system. The theory behind *Shannon's Diversity Index* states that in a diverse and evenly distributed system, an individual can belong to any of the taxa included in the system, and the prediction of its identity will carry much uncertainty (Magurran, 1988). On the other hand, in a less diverse system, prediction of an individual's identity will be easier.

The formula for *Shannon's Diversity Index* is given below in Eq. 6 where p_i is the proportion of pollen type *i*.

$$H' = -\sum p_i ln(p_i) \tag{6}$$

A.3.4. Simpson diversity (D1)

Another commonly used index in measuring diversity is the *Simpson Diversity Index* (Simpson, 1949). Although Simpson's original paper did not contain the index itself, nor consider it in an ecological context (Gorelick, 2006), his successors used it to interpret ecological diversity (Magurran, 1988; Morris et al., 2014). Traditional calculation of Simpson's diversity is shown below in Eq. 7. The index value represents the probability of two randomly chosen individuals belonging to different taxa.

$$D_1 = 1 - \sum p_i^2 \tag{7}$$

A.3.5. Simpson dominance (D2)

Simply the inverse of Simpson's Diversity, Simpson's Dominance (D_2) also has a widespread use in measuring diversity (Simpson, 1949). In both Simpson indices in Eqs. 7 and 8, a higher index value indicates higher diversity.

$$D_2 = \frac{1}{\sum p_i^2} \tag{8}$$

A.3.6. Simpson evenness

Derived from diversity measures of diversity, dominance, and richness, evenness can be considered as the degree by which individuals are split among species with low values indicating that one or a few species dominate, and high values indicating that relatively equal numbers of individuals belong to each species (Morris et al., 2014).

$$E = \frac{D_2}{S} \tag{9}$$

A.3.7. Rarefaction

The fundamentals of the rarefaction theory are owed to a number of studies (Heck et al., 1975; Simberloff, 1978, 1979; Tipper, 1979; Gart et al., 1982; Siegel and German, 1982). To represent the palynological richness of an investigated area, H. J. B. Birks and J. M. Line recommended the use of rarefaction analysis to estimate average palynological richness (Birks and Line, 1992), and Weng used this method to estimate the minimum pollen count (Weng et al., 2006). The original formula is given below in Eq. 10, where $E(T_n)$ is the expected number of pollen types in a sample of *n* individual grains selected at random without replacement from a count of *N* grains containing *T* taxa, namely the estimated number of pollen types that would have been found if only *n* grains had been counted. In the formula, *N* is the minimum pollen count, *T* is the number of pollen types in the original pollen count, N_i is the number of individual grains assigned to pollen type (taxon) *i* in the original pollen count, and *n* is the total number of grains counted in the sample.

$$E(T_n) = \sum_{i=1}^{T} \left(1 - \left[\frac{(N - N_i!)(N - n)!}{(N - N_i - n)!N!}\right]\right) \tag{10}$$

A.4. Reliability measurements

A practical test for the reliability (in our case representativeness) of a dataset is to take two samples and assess their reliability coefficients. This provides an objective and unbiased measure of the degree of reliability (also replicability) between two samples (Allen and Yen, 2002). The first attempt to establish some form of reliability of palynological samples belongs to Barkley (Barkley, 1934) who used *Pearson Correlation* and *Spearman-Brown's Prophecy* formula to obtain the reliability coefficient. This coefficient can either be obtained for a given sample size, or it can be used to determine the required size of a sample for a given reliability level.

A.4.1. Spearman-Brown's prophecy

In measurement theory, the *Spearman-Brown* approach is one of many to define and interpret test reliability – i.e. a survey-based data collection method for collecting people's evaluations regarding a phenomenon with a predefined scale. For a reliable test, a high correlation between test based observed scores and the test's *true* scores are sought, where the definition of the true scores is based on the *True Score Theory* (Allen and Yen, 2002). According to the theory, scores gathered from parallel test halves (i.e. splitting the data into two sub-sets) should be correlated if only the entire test is reliable.

Spearman-Brown's assessment of the reliability of the whole test using the formula for *Pearson Correlation* is meaningful, only if the sample meets two important criteria: that it is randomly selected and normally distributed (Allen and Yen, 2002). Barkley's adaptation of the formula is provided in Eq. 11, where *X* is the number of pollen grains of each type found in the first half, *Y* is the number of pollen grains of each type found in the second half and, *M* is obtained by dividing the sum of each half by the number of types of pollen grains found during the count of a slide – essentially an average.

$$r = \frac{MXY - (MX^{*}MY)}{\sqrt{MX^{2} - (MX)^{2}}\sqrt{MY^{2} - (MY)^{2}}} \tag{11}$$

Barkley and his followers later applied this approach to the *Prophecy Coefficient of Reliability* (r_p) for use in palynology (Barkley, 1934; and later Djamali and Cilleros, 2020) by adapting the formula in Eq. 12 to reflect $\overline{N}$ which is the necessary number of the pollen count to obtain a given r_p.

$$r_p = \frac{\overline{N}r}{1 + r(\overline{N} - 1)} \tag{12}$$

A.4.2. Split-half reliability

In measurement theory, one approach to evaluate internal consistency is the *Split-Half Reliability* estimate (Allen and Yen, 2002). This entails dividing the measurement into two halves (i.e., splitting the total pollen count data into two halves), and then searching for similarity between the two subsets of data. If the halves are inherently similar, the *Prophecy Coefficient of Reliability* can be calculated using the *Spearman-Brown* formula provided in Eq. 13.

$$r_p = \frac{2r}{1 + r} \tag{13}$$

In our study, we used *Split-Half* as an alternate method for verification while constructing our novel approach.

A.5. Pollen data analyzed in this study

Akgöl Lake: Karlıoğlu Kılıç N., Yilmaz Dagdeviren R., Acar D., Kucukdemirci M., Makaroglu O., Karaoz M. O., Sahin Altun D., Tutar A., Horuz A., Gurel A., Islam Z., Cagatay M. N., 2020. Akgöl'de (Sakarya) Fosil Polen Analizine Dayalı Paleovejetasyon Araştırmaları: İlk Bulgular. Journal of Geography-Cografya Dergisi, 40: 219–230. (Extended abstract, in English).

Aktaş Lake: Karlıoğlu Kılıç, N., Caner, H., Erginal, A.E., Ersin, S., H. Haluk Selim, H. Kaya, 2018. Environmental changes based on multi-proxy analysis of core sediments in Lake Aktaş, Turkey: Preliminary results, Quaternary International, 486: 89–97. doi.org/10.1016/j.quaint.2018.02.004.

Caspian Sea: core IGS-1 from Shirshov Institute of Oceanology, counted by S. Leroy, unpublished.

Coldharbour Moss, Peak District National Park: Heather moorland, radiocarbon dated counted by Jessie Woodbridge, unpublished.

Gnadensee in Lake Constance: core ARG6 from University of Constance, counted by S. Leroy, unpublished.

Issyk-Kul: surface samples taken by S. Krivonogov, counted by S. Leroy, unpublished.

Karataş Lake: Şenkul, Ç., Kalıpçı, E., 2019. Güneybatı Anadolu'dan Yeni Bir Paleoekolojik Değerlendirme: Karataş Gölü ve Çevresinin Geç Holosen Paleovejetasyon Değişimleri ve İklim ile İlişkileri. Coğrafya Dergisi, 38: 35–47. doi.org/10.26650/JGEOG2019-0006

Northeast Black Sea: Marret, F., Bradley, L. R., Tarasov, P. E., Ivanova, E. V., Zenina, M. A., & Murdmaa, I. O. (2019). The Holocene history of the NE Black Sea and surrounding areas: An integrated record of marine and terrestrial palaeoenvironmental change. The Holocene, 29(4), 648–661. https://doi.org/10.1177/0959683618824769

Tuzla Lake: Şenkul, Ç., Memiş, T., Eastwood, W.J., Doğan, U., 2018. Mid-to late-Holocene paleovegetation change in vicinity of Lake Tuzla (Kayseri), Central Anatolia, Turkey, Quaternary International, 486: 98–106. https://doi.org/10.1016/j.quaint.2018.05.026.

References

Adam, D.P., 1974. Palynological applications of princip al component and cluster analyses. J. Res. US Geol. Surv. 2 (6), 727–741.

Addison-Smith, B., Wraith, D., Davies, J.M., 2020. Standardising pollen monitoring: quantifying confidence intervals for measurements of airborne pollen concentration. Aerobiologia 36, 605–615.

Alf, C., Lohr, S., 2007. Sampling assumptions in introductory statistics classes. Am. Stat. 61 (1), 71–77.

Allen, M.J., Yen, W.M., 2002. Introduction to Measurement Theory. Waveland Press.

Barkley, F.A., 1934. The statistical theory of pollen analysis. Ecology 15 (3), 283–289.

Berger, W.H., Parker, F., 1970. Diversity of planktonic foraminifera in deep-sea sediments. Science 168, 1345–1347.

Birks, H., 2019. Contributions of Quaternary botany to modern ecology and biogeography. Plant Ecol. Div. 12 (3–4), 189–385.

Birks, H.J.B., Berglund, B.E., 1979. Holocene pollen stratigraphy of Southern Sweden: a reappraisal using numerical methods. Boreas 82, 57–279.

Birks, H.J.B., Berglund, B.E., 2018. One hundred years of Quaternary pollen analysis 1916–2016. Veg. Hist. Archaeobotany 27 (2), 271–309.

Birks, H.J.B., Birks, H.H., 1980. Quaternary Palaeoecology. Botany School, University of Cambridge.

Birks, H.J.B., Deacon, J., 1973. A numerical analysis of the past and present flora of the British Isles. New Phytol. 72, 877–902.

Birks, H.J.B., Gordon, A.D., 1984. Numerical Methods in Quaternary Pollen Analysis. Academic Press, pp. 277–312.

Birks, H.J.B., Line, J.M., 1992. The use of rarefaction analysis for estimating palynological richness from quaternary pollen-analytical data. The Holocene 2 (1), 1–10.

Boltzmann, L., 1872. Weitere Studien über das Warmegleichgewicht unter Gasmolekülen (Further Studies on the Thermal Equilibrium of Gas Molecules, The Kinetic Theory of Gases). Sitzungsberichte Akademie der Wissenschaften 66, 275–370.

Bonny, A.P., 1972. A method for determining absolute pollen frequencies in lake sediments. New Phytol. 71, 339–405.

Bowman, P.W., 1931. Study of a peat bog near the Matamek River, Quebec, Canada, by the method of pollen analysis. Ecology 12 (4), 694–708.

Brush, G.S., Brush, L.M., 1994. In: Traverse, A. (Ed.), Transport and Deposition of Pollen in an Estuary: Signature of the Landscape. Sedimentation of Organic Particles. Cambridge University Press, pp. 33–46.

Chevalier, M., Davis, B.A.S., Heiri, O., Seppä, H., Chase, B.M., Gajewski, K., Lacourse, T., Telford, R.J., Finsinger, W., Guiot, J., Kühl, N., Maezumi, 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.

Cleal, C., Pardoe, H.S., Berry, C.M., Cascales-Miñana, B., Davis, B.A.S., Diez, J.B., Filipova-Marinova, M.V., Giesecke, T., Hilton, J., Ivanov, D., Kustatscher, E., Leroy, S.A.G., McElwain, J.C., Opluštil, S., Popa, M.E., Seyfullah, L.J., Stolle, E., Thomas, B.A., Uhl, D., 2021. Palaeobotanical experiences of plant diversity in deep time. 1: how well can we identify past plant diversity in the fossil record? Palaeogeogr. Palaeoclimatol. Palaeoecol. 576, 10481 https://doi.org/10.1016/j.palaeo.2021.110481.

Cursach, J., Juan, R., Gomez-Martinez, C., Cardona, C., Capó, M., Lázaro, A., 2020. The role of landscape composition and heterogeneity on the taxonomicaland functional diversity of Mediterranean plant communities in agricultural landscapes. PloS One 15 (9), 1–18.

Dale, M.B., Walker, D.A., 1970. Information analysis of pollen diagrams 1. Pollen Spores 12 (2), 1–37.

Djamali, M., Cilleros, K., 2020. Statistically significant minimum pollen count in Quaternary pollen analysis; the case of pollen-rich lake sediments. Rev. Paleobot. Palynol. 275.

Dodge, Y., 2008. The Concise Encyclopedia of Statistics. Springer, New York, NY.

Erdtman, G., 1931. Pollen statistics: a new research method in paleo-ecology. Science 73 (1893), 399–401.

Erdtman, G., 1952. Pollen Morphology and Plant Taxonomy – Angiosperms. Almquist and Wiksell, Stockholm, p. 539.
Erdtman, G., 1954. An Introduction to Pollen Analysis. Chronica Botanica, Waltham, MA, U.S.A.
Erdtman, G., 1957. Pollen and Spore Morphology / Plant Taxonomy. Gimnospermae, Pteridophyta, Bryophyta (Illustrations) (An Introduction to Palynology II.), 451. Almquist and Wiksell, Stockholm.
Faegri, K., Iversen, J., 1964. Textbook of Pollen Analysis, 2nd edition. Munksgaard, Copenhagen, Denmark.
Faegri, K., Iversen, J., 1989. Textbook of Pollen Analysis, 4th edition. John Wiley & Sons Ltd., Norway.
Gart, J.J., Siegel, A.F., German, R.Z., 1982. Rarefaction and taxonomic diversity. Biometrics 38, 235–241.
Gordon, A.D., 1999. Classification. CRC Press.
Gordon, A.D., Birks, H.J.B., 1974. Numerical methods in Quaternary paleoecology II. Comparison of pollen diagrams. New Phytol. 73 (1), 221–249.
Gorelick, R., 2006. Combining richness and abundance into a single diversity index using matrix analogues of Shanon's and Simpson's indices. Ecography 29-4, 525–530.
Gower, J.C., 1985. Properties of Euclidean and non-Euclidean distance matrices. Linear Algebra Appl. 67, 81–97.
Haug, M.G., 2019. Measure of Association. Encyclopedia Britannica.
Heck, K.L., van Belle, C., Simberloff, D., 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. Ecology 56, 1459–1461.
Heegaard, E., Lotter, A.F., Birks, H.J.B., 2006. Aquatic biota and the detection of climate change: are there consistent aquaticecotons. J. Paleolimnol. 35, 507–518.
Holmes, P.L., 1994. Thre sorting of spore and pollen by water: Experimental and field evidence. In: Traverse, A. (Ed.), Sedimantation of Organic Particles. Cambridge University Press, Cambridge, pp. 9–33.
Hooghiemstra, H., Lézine, A.-M., Leroy, S.A.G., Dupont, L., Marret, F., 2006. Late Quaternary palynology in marine sediments: a synthesis of the understanding of pollen distribution patterns in the NW African setting. Quat. Int. 148 (1), 29–44.
Iversen, J., 1946. Geologisk datering af en senglacial boplads ved Bromine. Aarboger Nordisk Oldkyndighed Historie 198.
Karlıoğlu Kılıç, N., Caner, H., Erginal, A.E., Ersin, S., Haluk Selim, H., Kaya, H., 2018. Environmental changes based on multi-proxy analysis of core sediments in Lake Aktaş, Turkey: preliminary results. Quat. Int. 486, 89–97. https://doi.org/10.1016/j.quaint.2018.02.004.
Karlıoğlu Kılıç, N., Yilmaz Dagdeviren, R., Acar, D., Kucukdemirci, M., Makaroglu, O., Karaoz, M.O., Sahin Altun, D., Tutar, A., Horuz, A., Gurel, A., Islam, Z., Cagatay, M. N., 2020. Akgöl'de (Sakarya) Fosil Polen Analizine Dayalı Paleovejetasyon Araştırmaları: İlk Bulgular. J. Geogr.-Cografya Dergisi 40, 219–230 (Extended abstract, in English).
Keen, H.F., Gosling, W.D., Hanke, F., Miller, C.S., Montoya, E., Valencia, B.G., Williams, J.J., 2014. A statistical sub-sampling tool for extracting vegetation community and diversity information from pollen assemblage data. Paleogeogr. Paleoclimatol. Paleoecol. 408, 48–59. https://doi.org/10.1016/j.palaeo.2014.05.001.
Lewis, I.F., Cocke, E.C., 1929. Pollen analysis of dismal swamp peat. J. Elisha Mitchell Sci. Soc. 45, 37–58.
Li, S., Su, P., Zhang, H., Zhou, Z., Xie, T., Shi, R., Gou, W., 2018. Distribution patterns of desert plan diversity and relationship to soil properties in the Heihe River Basin, China. Ecosphere 9 (7).
Magurran, A.E., 1988. Ecological Diversity and Its Measurement. Princeton University Press, p. 175.
Maher, L.J., 1972. Nomograms for computing 0.95 confidence limits of pollen data. Rev. Palaeobot. Palynol. 13 (2), 85–93.
Manten, A.A., 1966. Half a century of modern palynology. Earth Sci. Rev. 2, 277–316.
Marret, F., Bradley, L.R., Tarasov, P.E., Ivanova, E.V., Zenina, M.A., Murdmaa, I.O., 2019. The Holocene history of the NE Black Sea and surrounding areas: an integrated record of marine and terrestrial palaeoenvironmental change. The Holocene 29 (4), 648–661. https://doi.org/10.1177/0959683618824769.
Moore, P.D., Webb, J.A., Collinson, M.E., 1991. Pollen Analysis. Blackwell Scientific Publications, Oxford.
Morris, E.K., Caruso, T., Buscot, F., Fischer, M., et al., 2014. Choosing and using diversity indices: insights for ecological application from the German biodiversity exploratories. Ecol. Evol. 4, 3514–3524. https://doi.org/10.1002/ece3.1155.
Mosimann, J.E., 1965. Statistical methods for the pollen analyst: Multinomial and negative multinomial techniques. In: Kummel, B., Raup, D. (Eds.), Handbook of Paleontological Techniques. Freeman, San Francisco, California, pp. 636–673.
Nelson Jr., A.A., 1980. Research design: measurement, reliability, and validity. Am. J. Hosp. Pharm. 37 (6), 851–857 (PMID: 7395882).
Overpeck, J.T., Webb, T., Prentice, I.C., 1985. Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. Quatern. Res. 23 (01), 87–108.
Pardoe, H., Cleal, C., Berry, C., Cascales-Miñana, B., Davis, B., Diez, J., Filipova-Marinova, M., Giesecke, T., Hilton, J., Ivanov, D., Kustatscher, E., Leroy, S., Mcelwain, J., Opluštil, S., Popa, M., Seyfullah, L., Stolle, E., Thomas, B., Uhl, D., 2021. Palaeobotanical experiences of plant diversity in deep time. II: how to measure and analyse past plant biodiversity. Palaeogeogr. Palaeoclimatol. Palaeoecol. 580, 110618 https://doi.org/10.1016/j.palaeo.2021.110618.
Prentice, I.C., 1980. Multidimensional scaling as a research tool in Quaternary palynology: a review of theory and methods. Rev. Palaeobot. Palynol. 31, 71–104.
Prentice, I.C., 1982a. Calibration of pollen spectra in terms of species abundance. In: Berglund, B.E. Palaeohydrological Changes in the Temperate Zone in the Last 5,000 years: IGCP 158, Subproject B, Lake and Mire Environments, Project Guide 3: 25–51. Department Quaternary Geology, University of Lund, Sweden.
Prentice, I.C., 1982b. Multivariate methods for the presentation and analysis of data. In: Berglund, B.E. Palaeohydrological Changes in the Temperate Zone in the Last 15,000 years: IGCP 158, Subproject B, Lake and Mire Environments, Project Guide, 3. Department Quaternary Geology, University of Lund, Sweden, pp. 53–77.
Reille, M., 1992. Pollen et spores d'Europe et d'Afrique du Nord. Laboratoire de botanique historique et de palynology, Marseille.
Sears, P.B., 1930. A record of post-glacial climate in Northern Ohio. Ohio J. Sci. 30 (4), 205–217.
Şenkul, Ç., Kalıpçı, E., 2019. Güneybatı Anadolu'dan Yeni Bir Paleoekolojik Değerlendirme: Karataş Gölü ve Çevresinin Geç Holosen Paleovejetasyon Değişimleri ve İklim ile İlişkileri. Coğrafya Dergisi 38, 35–47. https://doi.org/10.26650/JGEOG2019-0006.
Şenkul, Ç., Memiş, T., Eastwood, W.J., Doğan, U., 2018. Mid-to late-Holocene paleovegetation change in vicinity of Lake Tuzla (Kayseri), Central Anatolia, Turkey. Quaternary International 486, 98–106. https://doi.org/10.1016/j.quaint.2018.05.026.
Shannon, C., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (379–423), 623–656.
Siegel, S., Castellan, N.J.J., 1988. The Spearman Rank-Order Correlation Coefficient, Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill International Editions, Boston, pp. 235–244.
Siegel, A.F., German, H.Z., 1982. Rarefaction and taxonomic diversity. Biometrics 38, 23541.
Simberloff, D., 1978. Use of rarefaction and related methods in ecology. In: American Society for Testing and Materials ASTM STP, 652, pp. 150–165.
Simberloff, D., 1979. Rarefaction as a distribution-free method of expressing and estimating diversity. Stat. Ecol. Ser. 6, 159–176.
Simpson, E.H., 1949. Measurement of diversity. Nature 163, 688.
Tang, W., He, H., Tu, X.M., 2012. Applied Categorical and Count Data Analysis. CRC Press.
Tipper, J.C., 1979. Rarefaction and rarefiction - the use and abuse of a method in paleoecology. Paleobiology 5, 423–431.
Walker, D., Wilson, S.R., 1978. A statistical alternative to the zoning of pollen diagrams. J. Biogeogr. 5 (1), 1–21.
Webb III, T., Brayson, R.A., 1972. Late and postglacial climatic change in the northern midwest, USA: quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis. Quatern. Res. 2 (1), 70–115.
Weng, C., Hooghiemstra, H., Duivenvoorden, J.F., 2006. Challenges in estimating past plant diversity from fossil pollen data: statistical assessment, problems, and possible solutions. Divers. Distrib. 12 (3), 310–318.
Woodbridge, et al., 2021. What drives biodiversity patterns? Using long-term multi-disciplinary data to discern centennial-scale change. J. Ecol. https://doi.org/10.1111/1365-2745.13565.
Whittaker, R.H., 1972. Evolution and measurement of species diversity. Taxon 21 (2–3), 213–251. https://doi.org/10.2307/1218190.